IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF TEXAS
AUSTIN DIVISION

INTELLECTUAL VENTURES II LLC,

Plaintiff,

v.

TESLA, INC.,

Defendant.

Civil Action No. 1:24-cv-884-ADA

JURY TRIAL DEMANDED

TESLA, INC.'S OPENING CLAIM CONSTRUCTION BRIEF

TABLE OF CONTENTS

I. INTRODUCTION 1

II. OVERVIEW OF ASSERTED PATENTS 2

A. The ’639 Patent 2

B. The ’805 Patent 2

C. The ’395 Patent 3

D. The ’416 Patent 5

III. LEGAL STANDARDS 5

A. Claim Construction 5

B. Indefiniteness 6

IV. LEVEL OF ORDINARY SKILL IN THE ART 6

V. AGREED CONSTRUCTIONS 7

A. ’743 Patent 7

B. ’158 Patent 7

VI. DISPUTED TERMS 8

A. “said target feature information” (’639 patent, claim 1) / “said feature information” (’639 patent, claim 7) 8

B. “may at least partially describe a potential destination” (’805 patent, claim 1) 12

C. “potential geometric objects for plausibility using a matching algorithm so as to identify one or more plausible objects” (’805 patent, claim 1) 16

D. “said processing” (’395 patent, claim 1) 18

E. “broadcast information” (’416 patent, claim 1) 24

VII. CONCLUSION 27

TABLE OF AUTHORITIES

Cases **Page(s)**

24/7 Customer, Inc. v. LivePerson, Inc.,
235 F. Supp. 3d 1102 (N.D. Cal. 2016) 11

Arthur A. Collins, Inc. v. N. Telecom Ltd.,
216 F.3d 1042 (Fed. Cir. 2000).....26

Berkheimer v. HP Inc.,
881 F.3d 1360 (Fed. Cir. 2018).....12

Bushnell Hawthorne, LLC v. Cisco Sys., Inc.,
813 F. App'x 522 (Fed. Cir. 2020) 6, 18

Cellular Commc'ns Equip. LLC v. AT&T, Inc.,
No. 2:15-cv-576, 2016 WL 7364266 (E.D. Tex. Dec. 19, 2016) 11

Chef Am., Inc. v. Lamb-Weston, Inc.,
358 F.3d 1371 (Fed. Cir. 2004).....9

In re Downing,
745 F. App'x 988 (Fed. Cir. 2018) 19

eCardless Bancorp, Ltd. v. PayPal Inc.,
No. 24-cv-01054, 2024 WL 4804975 (N.D. Cal. Nov. 14, 2024) 19

Fairfield Indus., Inc. v. Wireless Seismic, Inc.,
No. 4:14-cv-2972, 2015 WL 1034275 (S.D. Tex. Mar. 10, 2015) 14

Halliburton Energy Servs., Inc. v. M-I LLC,
514 F.3d 1244 (Fed. Cir. 2008).....18

Infinity Computer Prods., Inc. v. Oki Data Americas, Inc.,
987 F.3d 1053 (Fed. Cir. 2021).....6

Intellectual Ventures I LLC v. General Motors Co.,
No. 6:21-cv-1088-ADA, Dkt. 83, Dkt. 84 (W.D. Tex. Dec. 1, 2022) 7

Interval Licensing LLC v. AOL, Inc.,
766 F.3d 1364 (Fed. Cir. 2014).....6, 12, 16

Kumar v. Ovonic Battery Co., Inc.,
351 F.3d 1364 (Fed. Cir. 2003).....26

Mirror Imaging LLC v. Wells Fargo Bank, N.A.,
No. 6:21-cv-0465-ADA, Dkt. 28 (W.D. Tex. Feb. 22, 2022).....13

Nautilus, Inc. v. Biosig Instruments, Inc.,
572 U.S. 898 (2014).....6

Novo Indus., L.P. v. Micro Molds Corp.,
350 F.3d 1348 (Fed. Cir. 2003).....9, 11

Phillips v. AWH Corp.,
415 F.3d 1303 (Fed. Cir. 2005).....25

Rillito River LLC v. Bamboo Indus. LLC,
No. 2:17-cv-00181, 2018 WL 4350095 (E.D. Cal. Sept. 11, 2018) 15

Standard Oil Co. v. Am. Cyanamid Co.,
774 F.2d 448 (Fed. Cir. 1985).....13

Synchronoss Techs., Inc. v. Dropbox, Inc.,
987 F.3d 1358 (Fed. Cir. 2021).....9

Zadro Prods., Inc. v. SDI Techs., Inc.,
No. 17-cv-1406-WCB, 2019 WL 10252726 (D. Del. June 19, 2019).....13

Other Authorities

MPEP § 2173.05(e) (9th ed., Nov. 2024) 20

I. INTRODUCTION

Intellectual Ventures II LLC ("IV") asserts twelve patents against Defendant Tesla, Inc. ("Tesla"), six of which include terms that require construction at this stage: U.S. Patent Nos. 6,894,639 ("the '639 patent"); 7,181,743 ("the '743 patent"); 7,336,805 ("the '805 patent"); 8,898,395 ("the '395 patent"); 9,232,158 ("the '158 patent"); and 10,136,416 ("the '416 patent") (collectively, the "Asserted Patents"). Tesla proposes constructions that remain faithful to the proper scope of the Asserted Patents' claims. For example, Tesla's construction of "broadcast information" from claim 1 of the '416 patent simply applies the plain meaning of the term, which is "information that is received via a broadcast transmission to all UEs within a base station's coverage area." That plain meaning is also consistent with the intrinsic record and how relevant extrinsic evidence defines the term.

Where claim language is objectively uncertain, Tesla argues indefiniteness. First, claim 1 of the '639 patent recites "***said*** target feature information," but there is no other "target feature information" in the claim. Moreover, this limitation could reasonably be interpreted as either "***a*** target feature information" or as "said target ***specific*** feature information," but the intrinsic record does not specify which one (if either) is correct. Because there are multiple potential interpretations, the term lacks objective certainty and is indefinite.

Second, the "***said*** processing" limitation in claim 1 of the '395 patent also lacks antecedent basis and fails to inform a person of ordinary skill in the art ("POSITA") with reasonable certainty about the claim scope. Claim 1 recites multiple structures which could perhaps perform "said processing" and it also recites multiple groups of instructions which could perhaps be the subject of "said processing." Moreover, the timing of "said processing" is objectively uncertain. Because the intrinsic record does not inform a POSITA of which of multiple potential interpretations of "said processing" (if any) is claimed, claim 1 lacks objective certainty and is indefinite.

Third, claim 1 of the '805 patent recites terms of degree whose scope is objectively uncertain in light of the intrinsic evidence—*i.e.*, "may ***at least partially describe*** a potential destination" and "***potential geometric objects for plausibility*** using a matching algorithm so as to identify one or more ***plausible objects***." While terms of degree are not *per se* indefinite, here, the claimed terms lack any objective certainty and both are therefore indefinite.

Thus, Tesla respectfully requests the Court adopt its proposed constructions.

II. OVERVIEW OF ASSERTED PATENTS[1]

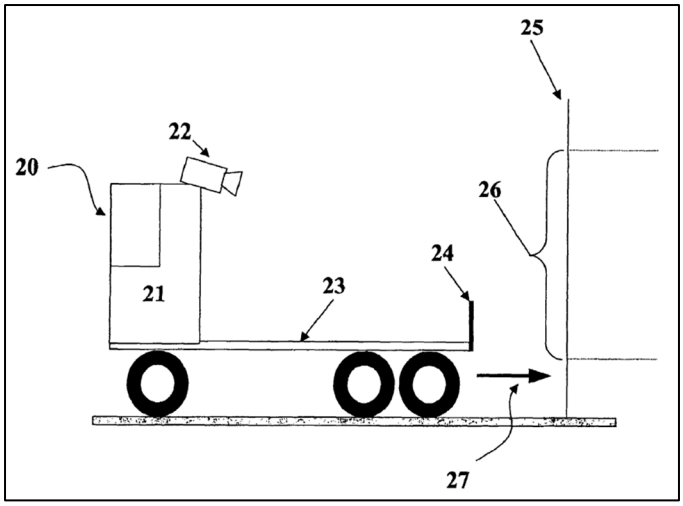
A. The '639 Patent

The '639 patent, titled "Generalized Hebbian Learning for Principal Component Analysis and Automatic Target Recognition, Systems and Method," was filed in December 1991 and issued on May 17, 2005. *See* Ex. 1 ('639 patent). The '639 patent is directed to a method of distinguishing targets from background clutter. *Id.* at Abstract. More specifically, the '639 patent describes calculation of data statistics from input data provided by sensors and uses the data statistics to identify and select "feature information" that is used to distinguish targets from background clutter. *Id.*

B. The '805 Patent

The '805 patent, titled "Docking Assistant," issued on February 26, 2008. *See* Ex. 2 ('805 patent). The '805 patent purports to describe a method for assisting a driver docking a truck (*e.g.*, in a parking bay or loading platform) using image analysis. *Id.* at Abstract. The vehicle uses a camera (*e.g.*, image sensor 22 on vehicle 20 in Figure 2) to obtain image data. *Id.* at 2:56–57.

[1] Tesla provides an overview for only those patents with disputed claim terms (the '639, '805, '395, and '416 patents), but not patents with agreed constructions (the '743 and '158 patents).



'805 patent at Fig. 2.

The obtained images are analyzed to locate objects, such as a cargo door 26 in Figure 2. *Id.* at 7:13–29. The patented system purports to accomplish this by breaking down the image data into "edge segments" and storing relationships between the edge segments in a "mathematical tree structure." *Id.* at 3:37–39; 4:43–45. The edge segments "are then analyzed to check for the presence of a geometric object that is similar to a geometric form which, typically, at least partially describes a potential destination." *Id.* at 3:40–43. The detected geometric objects are then "analyzed for plausibility using a matching algorithm." *Id.* at 2:41–43, 3:3–8, 3:47–50. Plausible objects undergo additional "acceptance analysis." *Id.* at 3:41–54. Accepted objects are then used to compute an "optimized travel path" to assist in vehicle guidance. *Id.* at 3:58–64.

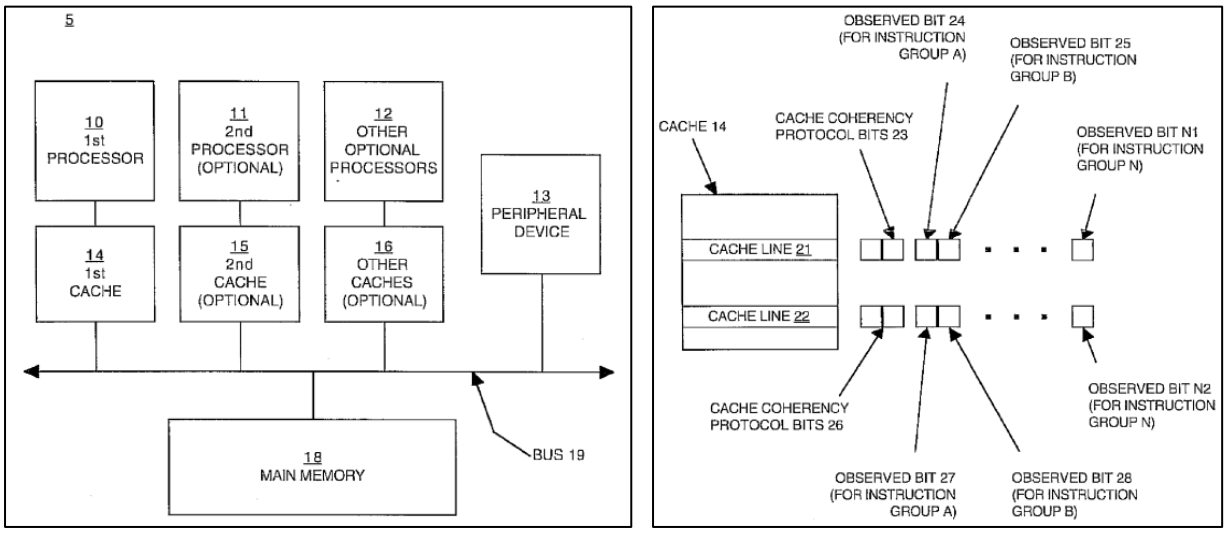
C. The '395 Patent

The '395 patent, titled "Memory Management for Cache Consistency," issued on November 25, 2014. *See* Ex. 3 ('395 patent). It is a continuation of an application originally filed on April 7, 2005 (now U.S. Patent No. 7,376,798).

The '395 patent purports to describe management of cache memory in a multiprocessor architecture in which groups of instructions are executed atomically by a processor. *Id.* at Abstract;

id. at 2:8–9 (Summary of the Invention); *see also id.* at 2:1–7 (alleging prior attempts "are limited to a single memory operation per instruction group" and purporting to give a solution to "overcome this disadvantage"); *id.* at 4:11–12. Each group can include instructions that, when executed by the processor, cause a cache line associated with that processor to be accessed, which in turn causes an indicator associated with that group of instructions in the cache line to be updated to reflect that access. *Id.* at 2:9–15. "The cache line is indicated as having been accessed until execution of the group of instructions is ended." *Id.* at 2:15–16.

With reference to Figures 1 and 2, processor 10 and external agents 11-13 may each be executing a group of instructions. *Id.* at 4:3–5. An instruction group is handled atomically (*e.g.*, it is treated as a unit) and it may seek access to the cache associated with the processor executing that instruction or to the cache of an external agent. *Id.* at 4:11–15.



'395 patent at Fig. 1 (left), Fig. 2 (right).

When a given processor, *e.g.*, processor 10, begins to execute a group of instructions, it allocates an observed bit for that group of instructions to each cache line in cache 14. *Id.* at 5:28–32. If cache line 21 is accessed because of an instruction in instruction group A being executed by processor 10, then observed bit 24 is set. *Id.* at 5:32–37. Similarly, if cache line 21 is also

accessed because of an instruction in instruction group B being executed by processor 10, then observed bit 25 is also set.

When execution of an instruction group by a processor, *e.g.*, processor 10, is ended, then the observed bits associated with that instruction group are cleared. For example, when instruction group A "is committed, rolled back, or aborted," then the observed bit 24 associated with instruction group A in cache line 21 is cleared. *Id.* at 5:41–47. Additionally, if an external agent 11–13 snoops cache line 21 while observed bits 24 and 25 remain set, then instruction group A (corresponding to observed bit 24) and instruction group B (corresponding to observed bit 25) are rolled back, and the observed bits 24 and 25 are cleared. *Id.* at 5:49–61.

D. The '416 Patent

The '416 patent is titled "Communicating on a Shared Channel in a Wireless Network" and issued on November 20, 2018. *See* Ex. 4 ('416 patent). It relates to cellular communications. *Id.* at 1:15–21. Specifically, the '416 patent purports to reduce wasting cellular resources on unused signaling channels. *Id*. at 3:58–61. It does so by using bits in the Broadcast Channel ("BCH") to indicate whether a particular timeslot within a frame is needed for signaling, and if not, providing an indication that the channel can be used as a shared channel. *Id*. at 9:38–58.

III. LEGAL STANDARDS

A. Claim Construction

In view of the Court's familiarity with claim construction, Tesla will forgo a lengthy recitation of law.[2] Tesla cites particularly relevant authority inline within specific arguments, and otherwise reserves the right to respond to any authority cited in IV's responsive brief.

[2] *See* Judge Albright's "Standing Order Governing Proceedings (OGP) 4.1–Patent Cases" (Apr. 14, 2022) at Section IX ("The Court has familiarity with the law of claim construction and

B. Indefiniteness

"A patent is invalid for indefiniteness if its claims, read in light of the patent's specification and prosecution history, fail to inform, with reasonable certainty, those skilled in the art about the scope of the invention." *Nautilus, Inc. v. Biosig Instruments, Inc.*, 572 U.S. 898, 898–99 (2014). Claims are indefinite when, for example, "the claim language 'might mean several different things and no informed and confident choice is available among the contending definitions.'" *Interval Licensing LLC v. AOL, Inc.*, 766 F.3d 1364, 1371 (Fed. Cir. 2014) (quoting *Nautilus*, 572 U.S. at 911 n.8). Seemingly clear terms also may be indefinite based on inconsistent or confusing usage. *See Infinity Computer Prods., Inc. v. Oki Data Americas, Inc.*, 987 F.3d 1053, 1062 (Fed. Cir. 2021) ("We recognize that, in a vacuum, it might seem odd to hold 'computer' indefinite. . . . Yet the indefiniteness here does not reside in the term 'passive link' or 'computer' on its own but rather in the relationship between the two in the context of these claims."). Further, a claim could be indefinite if a term does not have proper antecedent basis when a POSITA, "faced with the claims and the specification, could not, with reasonable certainty, discern the meaning of the claim term." *Bushnell Hawthorne, LLC v. Cisco Sys., Inc.*, 813 F. App'x 522, 526–27 (Fed. Cir. 2020).

IV. LEVEL OF ORDINARY SKILL IN THE ART

Tesla identifies the level of a POSITA for each patent with a disputed claim term:

- **'639 Patent**: A POSITA in December 1991 would have been someone knowledgeable and familiar with the image processing systems pertinent to the '639 patent. That person would have a bachelor's degree in computer science, computer engineering, electrical engineering, or equivalent training, and approximately four years of experience working in the field of image processing. Lack of work experience can be remedied by additional education, and vice versa.

- **'805 Patent**: A POSITA in June 2004 would have had a bachelor's degree in electrical engineering, computer engineering, computer science, or a related

---

encourages the parties to forego lengthy recitations of legal authorities and to instead focus on the substantive issues unique to each case.").

subject, and four years of work experience in image processing, automated vehicle control/navigation, or a related field. Less experience may be necessary with additional education (*e.g.*, a master's degree), and likewise, less education may be necessary with additional work experience (*e.g.*, 5–7 years).

- **'395 Patent**: A POSITA in April 2005 would have had a bachelor's degree in electrical or computer engineering, computer science, or a related subject, and two years of professional experience "in the data processing arts" (Ex. 3 ('395 patent) at 3:4–7) and multiprocessor speculative execution or a related field. Less experience may be necessary with additional education, and likewise, less education may be necessary with additional work experience.

- **'416 Patent**: A POSITA in September 2007 would have had a bachelor's degree in electrical engineering, computer engineering, or an equivalent, and approximately two years of professional experience relating to wireless communication technology. Lack of professional experience can be remedied by additional education, and vice versa.

V. AGREED CONSTRUCTIONS

The parties agreed to the following constructions and request the Court adopt these constructions:

A. '743 Patent

Term	Agreed Construction
Preamble (claim 10)	Limiting

B. '158 Patent

Term	Agreed Construction[3]
"integration time" (claims 1, 8)	"the time the image sensor collects and integrates signal from the scene"
"an image capture device" (claim 1)	"a device including a plurality of sensors, where two or more of the plurality of sensors each capture an overlapping portion of the same scene"

[3] This Court previously construed these two terms from the '158 Patent. *See Intellectual Ventures I LLC v. General Motors Co.*, No. 6:21-cv-1088-ADA, Dkt. 83 at 2, Dkt. 84 at 46–52 (W.D. Tex. Dec. 1, 2022). The parties agree the Court should adopt its prior constructions.

VI. DISPUTED TERMS[4]

A. "said target feature information" ('639 patent, claim 1) / "said feature information" ('639 patent, claim 7)

Tesla's Proposed Construction	IV's Proposed Construction
Indefinite	Plain and ordinary meaning. Not indefinite.

The term "said target feature information" is indefinite because it lacks antecedent basis and can potentially be interpreted in multiple ways. The claim language and specification provide multiple, equally plausible, ways to address the lack of antecedent basis, and the prosecution history only adds confusion. When Tesla asked IV to identify the antecedent basis, IV refused and instead doubled down on its "plain and ordinary" construction. Ex. 5 (12-18-2024 Email fr. Downing to Ellis re: claim construction).

There can be no dispute the term "said target feature information" lacks antecedent basis. The term refers to "said target feature information" even though the claims do not previously introduce "target feature information." *See* Ex. 1 ('639 patent) at claim 1. Thus, the term must be construed to correct the absence of an antecedent basis.

But it is unclear how that can be accomplished because there are at least two ways to remedy the absence of antecedent basis and neither is more likely correct than the other. First, "***said*** target feature information" could refer to "***a*** target feature information," which means the word "said" was drafted erroneously. This rewrite would introduce a concept ("target feature information") to the claim that is not found in any of the other limitations. Second, "***said*** target feature information" could refer to "said target ***specific*** feature information," which would mean "specific" was omitted erroneously. This second rewrite would make all the "feature information"

[4] The disputed terms are addressed in the order specified in Section IX of the Court's OGP.

the same throughout the claim, but it would change the meaning of the term entirely. Neither option is more likely correct than the other in view of the intrinsic record and, thus, neither is appropriate. *See Novo Indus., L.P. v. Micro Molds Corp.*, 350 F.3d 1348, 1354–58 (Fed. Cir. 2003) (finding claim indefinite and reversing district court's attempt to correct error in claim because "district court can correct a patent only if (1) the correction is not subject to reasonable debate based on consideration of the claim language and the specification and (2) the prosecution history does not suggest a different interpretation of the claims"); *see also Synchronoss Techs., Inc. v. Dropbox, Inc.*, 987 F.3d 1358, 1366–67 (Fed. Cir. 2021) (affirming indefiniteness finding because patentee's "proposal would require rewriting the claims, but it is not our function to rewrite claims to preserve their validity" (internal quotations and citations omitted)); *see also Chef Am., Inc. v. Lamb-Weston, Inc.*, 358 F.3d 1371, 1374 (Fed. Cir. 2004) ("Thus, in accord with our settled practice we construe the claim as written, not as the patentees wish they had written it.").

As to the first option, the claim language and specification support replacing "***said*** target feature information" with "***a*** target feature information." This edit would introduce the concept of "target feature information" (whereas the claim previously referred only to "target ***specific*** feature information"). This edit is consistent with the specification which teaches target "feature information" that may not be target specific. When discussing the background of the invention, the specification discusses "feature information" that varies from target to target (*i.e.*, "target ***specific*** feature information") and feature information that may *not* vary target to target (*i.e.*, "target feature information"):

> With respect to the complexity of pattern recognition and pattern classification techniques, many of these techniques rely on feature information to differentiate preferred targets from background clutter. Coordination and selection of this feature information is an ongoing problem, as the ***desired feature information itself (e.g., contrast-based and texture-based) often varies from target to target (e.g., size, stationary, make-up, composition, orientation)***, target to background (*e.g.*,

> location of target), data set to data set (*e.g.*, lighting, time-of-day, ambient temperature, and context of scene, sensor characteristics), and data source to data source (*e.g.*, one-dimensional or multi-dimensional data, such as digitized infrared imagery, digitized TV imagery, digitized Infra-red imagery, speech samples, or radar samples).

Ex. 1 ('639 patent) at 1:29–43 (emphasis added). Similarly, the specification elsewhere refers to both "target ***specific*** feature information" and more general "feature information" that may not be target specific. *Compare* Ex. 1 ('639 patent) at 5:50–55 ("The advantages of this method and system and the corresponding hardware and software embodiments are that they generate and select the ***target specific feature information*** for the classification of preferred targets in the context of particular backgrounds."), *with id.* at 7:15–20 ("As shown in FIG. 10, the classification system inputs the data in module 2, extracts ***feature information*** from the inputted data in module 4 and inputs the ***feature information*** into the classification module 6 . . . ."). Accordingly, the specification distinguishes between "target feature information" and "target ***specific*** feature information," and the two are not interchangeable.

But that is not the end of the analysis. Other claim language suggests the second option: that "said target feature information" should be edited as "said target ***specific*** feature information." Claim 9 recites a similar scope to claim 1, except that claim 9 is an apparatus claim. Indeed, claim 9 includes a set of apparatus that perform functions similar to those in claim 1. Among those, claim 9 includes a generator limitation that performs the same "generating" function as in limitation 1(c), except the generator limitation in claim 9 refers to "said target ***specific*** feature information" instead of "said target feature information" as in claim 1. *See* Ex. 1 ('639 patent) at cl. 9 ("(c) a generator to generate said target specific feature information from said data statistics outputed from said calculator"). This suggests that, like claim 9, claim 1 may have been intended to refer to "target ***specific*** feature information," but does not.

The prosecution history adds to the confusion. When the application was filed, the "generating" limitation recited "said target ***specific*** feature information," but later in prosecution the word "specific" was omitted. *Compare* Ex. 6 ('639 patent Prosecution History, Dec. 18, 1991 claim submission) at 3, *with id.* ('639 patent Prosecution History, May 28, 2004 claim amendments) at 8. And to the extent IV argues "specific" was inadvertently omitted, the same change was made in claim 18 (unasserted), but not in other claims (*e.g.*, claim 16), so the omission of "specific" across different claims undermines any argument the change was inadvertent. *Id.* at 6, 10.

Therefore, a POSITA would not know if "***said*** target feature information" is intended to (i) introduce "target feature information" (another type of information distinct from and broader than the previously introduced "target ***specific*** feature information"), or (ii) refer back to the previously introduced "target ***specific*** feature information." *E.g.*, *24/7 Customer, Inc. v. LivePerson, Inc.*, 235 F. Supp. 3d 1102, 1108 (N.D. Cal. 2016) (finding claim indefinite because "claim 1 recites elements containing different variations of the term 'data'" so the "Court agrees with LivePerson that it is ambiguous whether these variations of 'data' are 'equivalents, data subsets of each other, or completely different sets of data'"). Even if the inclusion of "said" (or removal of "specific") was done in error, the Court cannot rewrite the claim because there are now multiple ways to interpret "said target feature information." *See Novo*, 350 F.3d at 1354–58; *Cellular Commc'ns Equip. LLC v. AT&T, Inc.*, No. 2:15-cv-576, 2016 WL 7364266, at *10–*11 (E.D. Tex. Dec. 19, 2016) (finding claim indefinite and refusing to correct antecedent basis issue because correction open to reasonable debate). Thus, these claim terms are indefinite.

B. "may at least partially describe a potential destination" ('805 patent, claim 1)

Tesla's Proposed Construction	IV's Proposed Construction
Indefinite	Plain and ordinary meaning.

The dispute here turns on whether the specification provides enough clarity to interpret the term of degree: "may at least partially describe a potential destination." It does not. The specification is silent on how to measure or determine when a potential destination is "at least partially describe[d]" by a geometric object. The specification's examples either add more subjectivity or ignore the standard altogether.

The term is found in the "analyzing" limitation of claim 1, which requires analyzing the "edge segments" in image data to identify a "geometrical form that may at least partially describe a potential destination":

> [A]nalyzing the plurality of edge segments for the presence of a geometric object associated with a geometrical form that ***may at least partially describe a potential destination*** so as to identify one or more potential geometric objects; . . . .

The limitation includes several relative terms ("at least," "partially," and "potential"), which individually or combined together result in what is indisputably a term of degree.

Although a term of degree is not necessarily indefinite, courts routinely find them indefinite when the specification provides no objective guidance for measuring or otherwise determining the term of degree. "Although absolute or mathematical precision is not required, it is not enough, as some of the language in our prior cases may have suggested, to identify *some standard* for measuring the scope of the phrase." *Interval Licensing LLC v. AOL, Inc.*, 766 F.3d 1364, 1370–71 (Fed. Cir. 2014) (internal quotations and citations omitted) (emphasis in original). Instead, "[t]he claims, when read in light of the specification and the prosecution history, must provide objective boundaries for those of skill in the art." *Id.* at 1371; *see also Berkheimer v. HP Inc.*, 881

F.3d 1360, 1364 (Fed. Cir. 2018) ("Our case law is clear that the objective boundaries requirement applies to terms of degree."). In fact, courts have found "at least partially" to be indefinite when there is no clear objective standard disclosed in the specification. In *Standard Oil*, the Federal Circuit upheld the district court's ruling that the phrase "at least partially soluble" was indefinite. *Standard Oil Co. v. Am. Cyanamid Co.*, 774 F.2d 448, 451–53 (Fed. Cir. 1985). This Court similarly found "configured to at least partially link two or more interfaces" was indefinite in *Mirror Imaging LLC v. Wells Fargo Bank, N.A*. *See* Ex. 7 (No. 6:21-cv-0465-ADA, Dkt. 28 at 4 (W.D. Tex. Feb. 22, 2022)).

Indeed, the District of Delaware took issue with "at least in part" for precisely the reasons "at least partially" is indefinite here: because the specification provides no guidance to define the term. In *Zadro Prods., Inc. v. SDI Techs., Inc.*, the court ruled "at least in part generally concentrically aligned" was indefinite because the specification "provides no guidance as to what amount of generally concentric alignment would satisfy the requirement that at least part of the light emitting regions or diodes be generally concentrically aligned with the first reflective mirror plate." No. 17-cv-1406, 2019 WL 10252726, at *1–*3 (D. Del. June 19, 2019). The court noted "the addition of the phrase 'at least in part' must be interpreted as intended to broaden the reach the claim language beyond what it would have been had only the phrase 'generally concentrically aligned' been used." *Id.* at *3. But the claim was indefinite because "there is no guidance in the claim language or the specification as to the limits on how broadly that language in the claims should be read." *Id.*

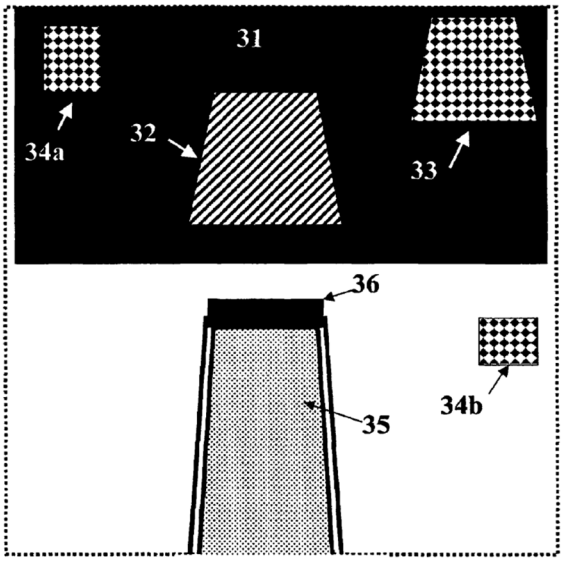
So too here. The specification provides no objective, reasonably certain way to measure when a "geometric object" "may at least partially describe a potential destination." This phrase is never described with any detail or used in a relevant example. The specification discusses the

"analyzing" limitation, but that discussion does not provide an objective standard for determining whether something "at least partially describe[s] a potential destination." Rather, the discussion uses other terms of degree, such as "substantially," "roughly," and "approximately," to describe the claimed "analy[sis]" (in this example, analyzing whether a docking station "substantially corresponds" to a rectangle):

> In a subsequent step, these edge segments are then analyzed to check for the presence of a geometric object that ***is similar to*** a geometric form which, ***typically, at least partially*** describes a potential destination. If the typical destination is a docking station for trucks at a warehouse, for example, then the ***typical geometric form substantially corresponds to a rectangle having roughly identical side lengths of approximately*** 2.5 m. Within the framework of the inventive method, the detected geometric objects that correspond to the typical geometric form are then analyzed for plausibility using a matching algorithm, these objects, which are classified as plausible, undergoing an additional acceptance analysis . . . .

Ex. 2 ('805 patent) at 3:39–52 (emphasis added). Simply substituting different relative words ("similar to," "substantially," or "roughly") for the claimed term of degree ("at least partially") is not sufficient. *See, e.g.*, *Fairfield Indus., Inc. v. Wireless Seismic, Inc.*, No. 4:14-cv-2972, 2015 WL 1034275, at *14–*16 (S.D. Tex. Mar. 10, 2015) (finding indefiniteness and noting patentee's original proposed construction "merely substituted one ambiguous term, 'to the extent reasonably practical,' for another 'substantially prevent'").

The specification provides an example of a determination of a potential destination, but the example further confuses the issue. Figure 3 illustrates a loading platform 35 for a truck docking at a cargo door 32 of a warehouse:



The specification glosses over the technique for determining that geometric objects in this image "may at least partially describe[s] a potential destination," instead saying this step is "assumed":

> Due to the elevated position of image sensor (22), a cargo door (26, 32) having a rectangular geometry is produced as a trapezoidal image in the image data. ***It is assumed exemplarily that, after processing of the image data in image-processing unit (11) and of the tree traversal algorithm in unit (12), objects 32, 33, 34 a and 34 b are identified as geometric objects typically corresponding to a potential destination***. It is assumed that object 32 corresponds to the image of cargo door (26) sought as a destination.

Ex. 2 ('805 patent) at 7:22–31 (emphasis added). It cannot simply be the case that the image is analyzed for real objects, because at least items 34a and 34b (which the specification says is assumed to "at least partially describe a potential destination") are not images of real objects. *Id.* at 7:33–37. There is no explanation as to what criteria the specification uses to make its assumptions. Thus, the specification falls short of providing an objective standard for measuring when a geometric object "at least partially describe[s] a potential destination."

Because the specification does not provide an objective basis for a POSITA to determine what degree of "describing" would satisfy the requirement of "at least partially describ[ing] a potential destination," claim 1 is indefinite. *See, e.g.*, *Rillito River LLC v. Bamboo Indus. LLC*,

No. 2:17-cv-00181, 2018 WL 4350095, at *9–*11 (E.D. Cal. Sept. 11, 2018) (ruling "partially surround" was indefinite because "patent offers no details as to the degree of 'partially surround'" and "claim term is insufficiently defined to offer one ordinarily skilled in the art a means by which to calculate 'partially surround'").

C. "potential geometric objects for plausibility using a matching algorithm so as to identify one or more plausible objects" ('805 patent, claim 1)

Tesla's Proposed Construction	IV's Proposed Construction
Indefinite	Plain and ordinary meaning.

Like the other term from the '805 patent, the dispute here turns on whether the specification provides enough clarity to determine or measure the term of degree: "plausibility" and "plausible objects." Again, it does not. The relevant language requires "plausibility using a matching algorithm . . . to identify one or more plausible objects." The specification discusses "matching algorithms" but fails to explain when the outcome of the matching algorithm is "plausible." Thus, claim 1 is indefinite.

The relevant limitation from claim 1 requires "analyzing the one or more potential geometric objects for plausibility using a matching algorithm so as to identify one or more plausible objects," thereby including terms of degree ("plausibility" and "plausible"). As explained above, the law requires the specification to provide a standard for measuring terms of degree. *See Interval Licensing*, 766 F.3d at 1370–71. Here, though, the specification is silent about the basis to determine when a "potential geometric object" is "plausible." For example, the specification is unclear about what constitutes a sufficient match from the "matching algorithm" for the "potential geometric objects" to be "plausible."

The matching algorithm referenced in the specification is silent on when such matching algorithm produces a "plausible object." The specification explains the "IPC algorithm" described

in a paper by Paul Besl is "particularly advantageously suited" for the claimed "matching algorithm." Ex. 2 ('805 patent) at 5:25–29. But the Besl paper discusses how to process collected data so the data can be analyzed for a match—it never explains when the outcome of the matching algorithm is "plausible" or how to determine "plausibility." *See generally* Ex. 8 (Besl Paper, TESLAIV000023924). Beside the IPC algorithm, the specification only mentions "matching algorithm" at a high level without ever providing details for when it produces "plausible objects." *See, e.g.*, Ex. 2 ('805 patent) at 3:3–8 ("In addition, the image-processing unit includes a comparator unit (13) for analyzing the detected geometric objects which correspond to the typical geometric form, to ***check for plausibility using a matching algorithm*** for making a comparison with object patterns stored in a memory unit (17)." (emphasis added)).

The specification also refers to an "acceptance analysis," but that discussion similarly fails to provide objective criteria for determining what is "plausible." For example, the specification discusses rejecting potential destinations based on either shape or position but does not explain how to determine whether a particular shape or position should result in rejecting a potential destination:

> In one especially advantageous type of acceptance analysis, in those cases in which the surrounding field is recorded by the sensor from an elevated position, those objects which ***do not exhibit any distorted images, as compared to their usual geometric shapes***, are rejected as potential destinations. For example, the distortion of rectangular or square geometric shapes of potential destinations is manifested as a trapezoid in an image formation. In another effective type of acceptance analysis, based on knowledge of the customary position of destinations, ***objects not corresponding to these specifications in the image data*** are rejected. For example, if an object is included in the image data whose position relative to the imaging sensor is known, then, on the basis of its image formation in the image data, an artificial horizon may be generated, on whose basis the relative position of the recognized objects is ascertained. In this manner, ***objects, whose position deviates from the usual position of the destination***, are excluded from the further processing.

Ex. 2 ('805 patent) at 5:64–6:15 (emphasis added). The disclosure of rejecting a destination because the objects "do not exhibit any distorted images," the objects do "not correspond[] to the specification," or the position "deviates" from the usual position does not (a) provide a standard to determine distortion, correspondence, or deviation; or (b) any guidance on how much distortion, correspondence, or deviation is required to reject a destination such that the object is not "plausible." This example does not add clarity to the term at issue.

The specification mentions "plausible objects" only twice but neither passage helps. The first instance parrots the claim language. Ex. 2 ('805 patent) at 2:43–47. The second instance fares no better: "By already observing such elementary rules when processing the hierarchical tree structure, the subsequent outlay for processing may be limited already at this stage of the inventive method to relatively few ***plausible objects which are actually similar to a potential destination***." Ex. 2 ('805 patent) at 5:8–13. Adding terms of degree, like "actually similar," falls short of providing an objective standard for identifying "plausible objects."

Because the specification fails to provide an objective standard to determine when the matching algorithm produces "plausible objects," claim 1 of the '805 patent is indefinite.

D. "said processing" ('395 patent, claim 1)

Tesla's Proposed Construction	IV's Proposed Construction
Indefinite	Plain and ordinary meaning. Not indefinite.

The limitation "said processing" lacks antecedent basis and thus requires modification, but there is no clear correction for that lack of antecedent basis. Thus, the term is indefinite. *See Halliburton Energy Servs., Inc. v. M-I LLC*, 514 F.3d 1244, 1249 (Fed. Cir. 2008) (claim is "indefinite if a term does not have proper antecedent basis where such basis is not otherwise present by implication or the meaning is not reasonably ascertainable"); *Bushnell Hawthorne*, 813

F. App'x at 526–27 (claim is indefinite when scope could not be reasonably ascertained given lack of antecedent basis); *In re Downing*, 745 F. App'x 988, 996 (Fed. Cir. 2018) (holding "the lack of an antecedent basis does not render a claim indefinite" only "as long as the claim apprises [a POSITA] of its scope, and therefore, serves the notice function required by § 112 ¶ 2"); *eCardless Bancorp, Ltd. v. PayPal Inc.*, No. 24-cv-01054, 2024 WL 4804975, at *12 (N.D. Cal. Nov. 14, 2024) ("Antecedent basis is necessary so [POSITA] can ascertain the scope of a patent and its claimed inventions with 'reasonable certainty.'" (citing *Nautilus*, 572 U.S. at 909)).

The claim language causes further confusion as it introduces: (1) different groups of instructions that could perhaps be the subject of "said processing"; (2) multiple structures that could perhaps perform "said processing"; and (3) different times at which "said processing" could occur. Specifically, "said processing" *might* refer to an action on "an instruction in a first group of instructions," "an instruction in a second group of instructions," or an instruction in "a third group of instructions." *See* Ex. 9 (Herlihy Decl.) ¶ 43. Further, "said processing" *might* refer to an action taken by the claimed "processor" or by the claimed "agent other than said processor." *See* Ex. 9 (Herlihy Decl.) ¶ 43. Finally, there is added objective uncertainty as to the timing of "said processing," which a POSITA would recognize *might* be performed immediately or, alternatively, the cache line *might* be marked for a later rollback. *Compare* Ex. 3 ('395 patent) at 5:49–56 (describing "one embodiment" in which "the instruction groups associated with the observed bits that are set are forced to roll back and then reissue"), *with id.* at 8:54–62 (describing other "embodiments in accordance with the present invention" where "a protocol . . . causes all active instruction groups to roll back and reissue ***after the cache line state transition***" (emphasis added)); *see also* Ex. 9 (Herlihy Decl.) ¶ 47 (explaining a POSITA would understand, *e.g.*, that the rollback *might* suspended for a brief duration to give a conflicting instruction group a chance to

complete). Indeed, the timing uncertainty further compounds the uncertainty over which structure performs "said processing." Ex. 9 (Herlihy Decl.) ¶ 47.

As seen in claim 1 (below), the word "processing" appears twice. The first instance recites "said processing" as part of a wherein clause (limitation 1[c][2]) following the "executing a third group of instructions" step (shown in **green**),[5] and the second instance is in a "processing" method step (limitation 1[d]) for "said first group and said second group of instructions" (shown in **blue**):

> 1[pre] A method of managing memory in a computer system, said method comprising:
>
> 1[a] setting a first bit of an indicator associated with a cache line in a cache memory if said cache line has been accessed in response to a processor executing an instruction in a first group of instructions;
>
> 1[b] setting a second bit of said indicator while said first bit remains set if said cache line has also been accessed in response to said processor executing an instruction in a second group of instructions;
>
> 1[c][1] executing a third group of instructions that causes said cache line to be accessed, wherein said third group of instructions is executed by an agent other than said processor,
>
> 1[c][2] wherein ***said processing*** comprises rolling back said first group of instructions provided said first bit is set and rolling back said second group of instructions provided said second bit is set before allowing an instruction in said third group to access said cache line, and otherwise granting said access; and
>
> 1[d] ***processing*** said first group and said second group of instructions according to a value of said indicator.

The first point of objective uncertainty is what is processed by "said processing." The recitation of "said processing" in limitation 1[c][2], which is within the method step for executing the ***third*** group of instructions (limitation 1[c][1]), suggests the action is being taken on the ***third***

[5] The use of the article "said" rather than "a" is significant. The Manual of Patent Examining Procedure dictates how the articles "a" and "the" / "said" are used in patent claims. Specifically, the first time a word or phrase is mentioned, the article "a" should be used, while thereafter the article "the" or "said" should be used to refer back to the antecedent basis. MPEP § 2173.05(e) (9th ed., Nov. 2024).

group of instructions. *See* Ex. 9 (Herlihy Decl.) ¶ 45. However, the remainder of the limitation indicates otherwise as it recites "rolling back said ***first*** group of instructions . . . and rolling back said ***second*** group of instructions." The specification does not help to resolve this uncertainty because instructions from each of the first, the second, and the third groups of instructions are all potentially being processed in the computer system. *E.g.*, Ex. 3 ('395 patent) at 6:62–65. Accordingly, it is not reasonably ascertainable to a POSITA from the full specification which group of instructions is subject to "said processing" in claim 1. *See* Ex. 9 (Herlihy Decl.) ¶ 45.

Further compounding this objective uncertainty, the claim language is also unclear as to which structure performs "said processing," as the claim recites both "a processor" and "an agent other than said processor" before reciting the "said processing" limitation. Indeed, "said processing" is recited in limitation 1[c][2], which is a sub-step of the third method step. The claim requires the third method step (limitation 1[c][1]) is performed "by an agent ***other than said processor***." Ex. 3 ('395 patent) at cl. 1 (emphasis added). A POSITA would therefore understand that "said processing" in the sub-step of the third method step is likewise performed by that same "agent ***other than said processor***." *Id.*; Ex. 9 (Herlihy Decl.) ¶ 44. However, the remainder of the limitation concerns the first and the second groups of instructions, which are executed by ***the processor*** in the earlier-recited method steps (limitations 1[a] and 1[b]). *See* Ex. 3 ('395 patent) at cl. 1; Ex. 9 (Herlihy Decl.) ¶ 44. Furthermore, the specification is silent as to which structure performs "said processing" that "comprises rolling back said first group of instructions . . . and rolling back said second group of instructions." *See, e.g.*, Ex. 3 ('395 patent) at 5:54–56 (describing the result—"then the instruction groups associated with the observed bits that are set are forced to roll back"—without identifying the responsible structure), *id.* at 5:58–61 (same), *id.* at 6:6–11 (same), *id.* at 7:59–62 (same, with reference to step 43 in Figure 4), *id.* at 6:65–7:3

(describing "some or all of the processes . . . are executed by one or more of the agents **10-13**"); *see also id.* at 6:12–16 (describing *when*, but not *how*, rollbacks could occur), *id.* at 6:24–29 (same). Accordingly, it is not reasonably ascertainable to a POSITA whether "said processing" in claim 1 is performed by the recited "processor" or the recited "agent other than said processor." *See* Ex. 9 (Herlihy Decl.) ¶ 44.

Claim 2 also adds objective uncertainty. The language of dependent claim 2 refers to "said processing," but it does not resolve the uncertainties in independent claim 1:

> 2. The method of claim 1 wherein ***said processing*** further comprises:
>
> determining a state of said cache line, wherein said state is specified according to a cache coherency protocol comprising at least a modified state, a shared state, and an invalid state; and
>
> rolling back said first group of instructions provided said first bit is set and rolling back said second group of instructions provided said second bit is set, and then granting said access provided said cache line is in other than said shared state, and otherwise granting said access only when said cache line is in said shared state.

Rather, it remains unclear which of the two recited structures performs "said processing," and which of the three recited groups of instructions is the subject of "said processing," making this claim indefinite. Specifically, the state of the cache line can be determined when processing any of the first, the second, or the third groups of instructions by, *e.g.*, checking cache coherency protocol bits. *See* Ex. 9 (Herlihy Decl.) ¶ 46; *see also* Ex. 3 ('395 patent) at 6:20–24 ("In one embodiment, . . . the cache coherency protocol bits **23** are also checked to determine the state (*e.g.*, the MESI state) of cache line **21**.").

Finally, the prosecution history does not provide reasonable clarity. A "processing" step was added by amendment as the *second limitation* of dependent claim 34:

> 34. (Currently Amended) The method of Claim 33 further comprising:
>
> executing a third group of instructions that causes said cache line to be accessed;
>
> ~~using a value of said indicator to determine whether said cache line has been read and generating a result thereof;~~ and
>
> processing ~~handling~~ said first group and said second group of instructions according to a value of said indicator ~~said result~~.

Ex. 10 ('395 patent Prosecution History, June 17, 2010 Response to Non-Final Rejection) at 3 (highlight added). Claim 35 (which depended from claim 34) was also amended to recite a "said processing" limitation, which then derived its antecedent basis from amended claim 34:

> 35. (Currently Amended) The method of Claim 34 wherein said third group of instructions is executed by an agent other than said processor, wherein said processing ~~handling~~ comprises rolling back said first group of instructions provided said first bit is set and rolling back said second group of instructions provided said second bit is set before allowing an instruction in said third group to access said cache line ~~granting said access provided said indicator indicates said cache line has been read~~, and otherwise granting said access ~~without rolling back said first group and said second group of instructions~~.

Id. at 4 (highlight added). However, in a later amendment, the text of amended claim 35 was added to the *first limitation* of amended claim 34, and that text was then added to independent claim 33 thereby introducing an antecedent basis issue by changing the order of limitations:

> executing a third group of instructions that causes said cache line to be accessed, wherein said third group of instructions is executed by an agent other than said processor, wherein said processing comprises rolling back said first group of instructions provided said first bit is set and rolling back said second group of instructions provided said second bit is set before allowing an instruction in said third group to access said cache line, and otherwise granting said access; and
>
> processing said first group and said second group of instructions according to a value of said indicator.

Ex. 11 ('395 patent Prosecution History, June 30, 2014 Response to Decision on Appeal) at 2 (excerpt of amended claim 33, now claim 1) (highlights added). Further, by rewriting "said processing" to now be part of the step of "executing a third group of instructions . . . by an agent other than said processor," the Applicant compounded the objective uncertainties discussed above.

When Tesla explained the issue and asked IV to clarify the meaning of "***said*** processing," IV refused and instead doubled down on its "plain and ordinary" construction. *E.g.*, Ex. 5 (12-18-2024 Email fr. Downing to Ellis re: claim construction).

The term "said processing" is indefinite because: (1) it lacks an antecedent basis, (2) the specification provides multiple ways to interpret this claim language, and (3) a POSITA cannot reasonably resolve the objective uncertainty.

E. "broadcast information" ('416 patent, claim 1)

Tesla's Proposed Construction	IV's Proposed Construction
Plain meaning, which is "information that is received via a broadcast transmission to all UEs within a base station's coverage area."	Plain and ordinary meaning.

Tesla's proposed construction has three parts: (1) the receipt of information ("information that is received"); (2) through a broadcast ("via a broadcast transmission"); (3) by devices within

a coverage area ("to all UEs within a base station's coverage area"). Each of these parts is within the plain meaning of the phrase—Tesla construes the term only to assist the factfinder. IV disputes Tesla's construction under the guise of "plain and ordinary meaning" but has never explained the nature of its dispute.

Tesla's proposed construction is consistent with the claim language. The full limitation refers to "receiv[ing]" information through a "broadcast": "a receiver, the receiver configured to ***receive broadcast information*** including a plurality of bits, wherein each of the plurality of bits indicates a respective time interval that physical shared channel resources are utilized to signal channels." Ex. 4 ('416 patent) at cl. 1 (emphasis added). Although the claim does not explicitly state the information is received by "all UEs within a base station's coverage area," that is the plain meaning of "broadcast." As an example, a common usage of "broadcast" is a radio broadcast or television broadcast, which is a communication that is not directed or otherwise addressed to a *particular* user or device, but rather received by all radio or television consumers within an applicable coverage area.

This interpretation is also consistent with the intrinsic record. For example, the '416 patent explains that "broadcast information" is information "relevant to that cell, such as the network identity[.]" *Id*. at 3:20–23. The '416 patent distinguishes "broadcast" messages sent to "***all*** UEs in a cell" from communications "targeted" to particular UEs. *Id.* at 11:16–26 ("[T]he indicator bits may be either ***broadcast to all*** UEs in a cell, ***or transmitted to a targeted group*** of one or more UEs in a cell.") (emphasis added).

Other aspects of the intrinsic record similarly require Tesla's interpretation. U.S. Patent Nos. 6,111,865 ("the '865 patent"), 8,493,909 ("the '909 patent"), and 8,446,849 ("the '849 patent") were all considered during prosecution and are thus part of the intrinsic record. *Phillips*

v. AWH Corp., 415 F.3d 1303, 1317 (Fed. Cir. 2005) (explaining intrinsic record "includes the prior art cited during the examination of the patent."); *see also Kumar v. Ovonic Battery Co., Inc.*, 351 F.3d 1364, 1368 (Fed. Cir. 2003) (adopting definition provided by prior art in intrinsic record); *Arthur A. Collins, Inc. v. N. Telecom Ltd.*, 216 F.3d 1042, 1044–45 (Fed. Cir. 2000) (construing term based on its usage in prior art cited in patent). Each uses "broadcast" to refer to communications to all receivers within a range. For example, in the '909 patent, a "broadcast" is a message sent "all the UEs" in a cell. Ex. 12 ('909 patent) at 4:37–41 ("The MAC-b (MAC-broadcast) entity handles only the radio-cell-relevant system information which is broadcast to all the UEs in the relevant radio cell by mapping the logical broadcast control channel BCCH to the transport broadcast channel BCH."). Similarly, when referring to communications with other terminals, the '865 patent uses "broadcast" to refer to communication sent to "all" terminals. Ex. 13 ('865 patent) at 5:51–54 ("The multi-bit quick page message could also be used to indicate that the full paging channel should be performed for a longer duration so that system parameter changes can be broadcast to all the wireless terminals 10."). The '849 patent includes similar usage. Ex. 14 ('849 patent) at 9:48–55 ("Transport Channels comprises a Broadcast Channel (BCH), Downlink Shared Channel (DL-SCH) and a Paging Channel (PCH), the PCH for support, of UE power saving (DRX cycle is indicated by the network to the UE), broadcasted over entire cell . . . .").

Tesla's interpretation of the plain and ordinary meaning is also corroborated by the relevant extrinsic evidence. The '416 patent relies on teachings from cellular standards (referred to as 3GPP standards). *E.g.*, Ex. 4 ('416 patent) at page 2 (citing 3GPP standards as prior art), 1:25–33 (citing 3GPP standards as "Background"), 9:33–38 (referring to the patented system as being "described in a context" of channels "defined in" various 3GPP standards). 3GPP standards

(including, for example, 3GPP TS 25.221, which is cited in the '416 patent at 2:2–3) rely on a standardized vocabulary promulgated by the 3GPP for terms such as "broadcast." *See, e.g.*, Ex. 15 (3GPP TR 21.905 v10.3.0, "3rd Generation Partnership Project; Technical Specification Group Services and Systems Aspects; Vocabulary for 3PP Specification (Release 10)) ("3GPP Vocabulary"). And, like Tesla's proposed interpretation, the 3GPP Vocabulary defines "broadcast" as "denot[ing] unidirectional ***distribution to all users***." *Id.* at TESLAIV000022278 (emphasis added).

Thus, Tesla's interpretation of "broadcast" is consistent with the intrinsic and extrinsic evidence and should be adopted.

VII. CONCLUSION

For these reasons, the Court should adopt (i) the parties' agreed constructions and (ii) Tesla's proposed constructions for the disputed terms in the Asserted Patents.

Dated: January 9, 2025

Respectfully Submitted,

FISH & RICHARDSON P.C.

By: */s/ Daniel R. Gopenko*

Aamir A. Kazi (admitted to W.D. Tex.)
kazi@fr.com
1180 Peachtree Street, 21st Floor
Atlanta, GA 30309
(404) 892-5005

Daniel R. Gopenko (admitted to W.D. Tex.)
gopenko@fr.com
1000 Maine Ave SW, Suite 1000
Washington, DC 20024
(202) 783-5070

David M. Hoffman (TX Bar No. 24046084)
hoffman@fr.com
111 Congress Avenue, Suite 2000
Austin, TX 78701
(512) 472-5070

Matt Colvin (TX Bar No. 24087331)
colvin@fr.com
Michael Ellis (TX Bar No. 24102726)
ellis@fr.com
1717 Main Street, Suite 5000
Dallas, TX 75201
(214) 747-5070

Kathryn Quisenberry (TX Bar No. 24105639)
quisenberry@fr.com
909 Fannin Street, Suite 2100
Houston, TX 77010
(713) 654-5300

TESLA, INC.

Alex Hanna (admitted to W.D. Tex.)
CA Reg. In-House Counsel
alehanna@tesla.com
3000 Hanover Street
Palo Alto, CA 94304
(650) 586-2409

COUNSEL FOR DEFENDANT, TESLA, INC.

IPR2025-00218 Exhibit 2003 Page 32 of 257

CERTIFICATE OF SERVICE

The undersigned hereby certifies that the foregoing document was filed electronically in compliance with Local Rule CV-5(a) on January 9, 2025, and was served via CM/ECF on all counsel of record.

/s/ Daniel R. Gopenko
Daniel R. Gopenko

IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF TEXAS
AUSTIN DIVISION

INTELLECTUAL VENTURES II LLC, Plaintiff, v. TESLA, INC., Defendant.	Civil Action No. 1:24-cv-884-ADA

DECLARATION OF DANIEL R. GOPENKO IN SUPPORT OF TESLA'S OPENING CLAIM CONSTRUCTION BRIEF

I, Daniel R. Gopenko, declare as follows:

1. I am an attorney at the law firm of Fish & Richardson P.C. I make this declaration on behalf of my client, Tesla, Inc. and in support of Tesla's Opening Claim Construction Brief. I have personal knowledge of the matters stated in this declaration and would testify truthfully to them if called upon to do so.

2. Attached hereto as Exhibit 1 is a true and correct highlighted copy of U.S. Patent No. 6,894,639.

3. Attached hereto as Exhibit 2 is a true and correct highlighted copy of U.S. Patent No. 7,336,805.

4. Attached hereto as Exhibit 3 is a true and correct highlighted copy of U.S. Patent No. 8,898,395.

5. Attached hereto as Exhibit 4 is a true and correct highlighted copy of U.S. Patent No. 10,136,416.

6. Attached hereto as Exhibit 5 is a true and correct highlighted copy of an email chain from Downing to Ellis re: claim construction, dated December 18, 2024.

7. Attached hereto as Exhibit 6 is a true and correct highlighted copy of excerpts of the prosecution history of U.S. Patent No. 6,894,639 (Dec. 18, 1991 Submitted Claims and May 28, 2004 Claim Amendments).

8. Attached hereto as Exhibit 7 is a true and correct highlighted copy of the Claim Construction Order dated February 22, 2022 in *Mirror Imaging v. Capital One, N.A., et al*, Case No. 6:21-cv-0465-ADA, Dkt. 28 (W.D. Tex.).

9. Attached hereto as Exhibit 8 is a true and correct copy of Besl, *A Method for Registration of 3-D Shapes*, IEEE Transactions of Pattern Analysis and Machine Intelligence, vol 14, No. 2, February 1992, bates labelled TESLAIV000023924.

10. Attached hereto as Exhibit 9 is a true and correct copy of the declaration of Dr. Maurice Herlihy, dated January 9, 2025.

11. Attached hereto as Exhibit 10 is a true and correct highlighted copy of excerpts of the prosecution history of U.S. Patent No. 8,898,395 (June 17, 2010 Response to Non-Final Rejection).

12. Attached hereto as Exhibit 11 is a true and correct highlighted copy of excerpts of the prosecution history of U.S. Patent No. 8,898,395 (June 30, 2014 Response to Decision on Appeal).

13. Attached hereto as Exhibit 12 is a true and correct highlighted copy of U.S. Patent No. 8,493,909.

14. Attached hereto as Exhibit 13 is a true and correct highlighted copy of U.S. Patent No. 6,111,865.

15. Attached hereto as Exhibit 14 is a true and correct highlighted copy of U.S. Patent No. 8,446,849.

16. Attached hereto as Exhibit 15 is a true and correct highlighted copy of excerpts of 3GPP TR 21.905 v10.3.0, bates labelled TESLAIV000022271.

I declare under penalty of perjury under the laws of the United States that the foregoing is true and correct. Executed at Washington, DC this 9th day of January 2025.

/s/ Daniel R. Gopenko
Daniel R. Gopenko

EXHIBIT 1

US006894639B1

(12) **United States Patent**
Katz

(10) **Patent No.: US 6,894,639 B1**
(45) **Date of Patent: May 17, 2005**

(54) **GENERALIZED HEBBIAN LEARNING FOR PRINCIPAL COMPONENT ANALYSIS AND AUTOMATIC TARGET RECOGNITION, SYSTEMS AND METHOD**

(75) Inventor: **Alan Jerry Katz**, Dallas, TX (US)

(73) Assignee: **Raytheon Company**, Waltham, MA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **07/810,005**

(22) Filed: **Dec. 18, 1991**

(51) **Int. Cl.**[7] **G01S 13/00**
(52) **U.S. Cl.** **342/90**; 342/27; 342/89; 342/159; 342/175; 342/195; 706/15
(58) **Field of Search** 342/27, 28, 61, 342/62–81, 89–103, 159–164, 175, 192, 193–197, 22; 367/87, 93–98, 135, 136; 704/200, 231, 232–257; 706/15–44; 382/103

(56) **References Cited**

U.S. PATENT DOCUMENTS

3,913,099	A	*	10/1975	Wehner et al.	342/192
3,967,282	A	*	6/1976	Young et al.	342/22
3,992,710	A	*	11/1976	Gabriele et al.	342/90
4,001,820	A	*	1/1977	Rosenbaum et al.	342/192
4,389,647	A	*	6/1983	Fanuele et al.	342/192
4,484,193	A	*	11/1984	Bellew	342/98
4,490,718	A	*	12/1984	Opitz et al.	342/192
4,641,138	A	*	2/1987	Opitz	342/192
4,829,306	A	*	5/1989	Gjessing et al.	342/159
4,847,817	A	*	7/1989	Au et al.	367/135
4,897,660	A	*	1/1990	Gold et al.	342/192
4,907,001	A	*	3/1990	Harmuth	342/159
4,995,088	A	*	2/1991	Farhat	706/40
5,040,214	A	*	8/1991	Grossberg et al.	704/240
5,078,952	A	*	1/1992	Gozani et al.	706/20
5,376,940	A	*	12/1994	Abatzoglou	342/192
5,943,661	A	*	8/1999	Katz	706/16

FOREIGN PATENT DOCUMENTS

WO	WO91/02323	*	2/1991	G06F/15/80

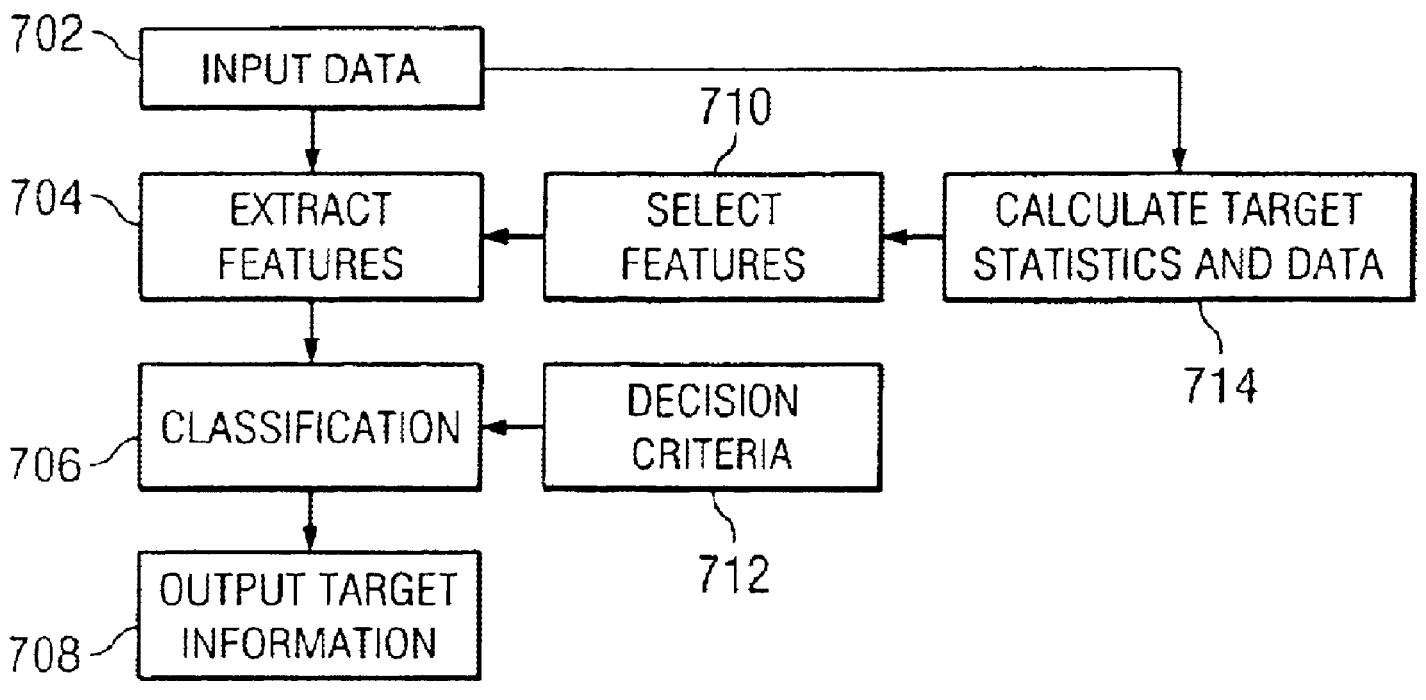
* cited by examiner

Primary Examiner—Bernarr E. Gregory
(74) *Attorney, Agent, or Firm*—Baker Botts L.L.P.

(57) **ABSTRACT**

This application discloses a method for distinguishing targets from clutter, comprising the steps inputting data, calculating data statistics from said data and using said data statistics to select target specific feature information to distinguish specific targets from background clutter, generating said target specific feature information from said data statistics, extracting said target specific feature information from said data, using said target specific feature information to distinguish specific targets from background clutter, and outputting target and background clutter information. Classification systems, including hardware and software embodiments, are also disclosed.

20 Claims, 7 Drawing Sheets



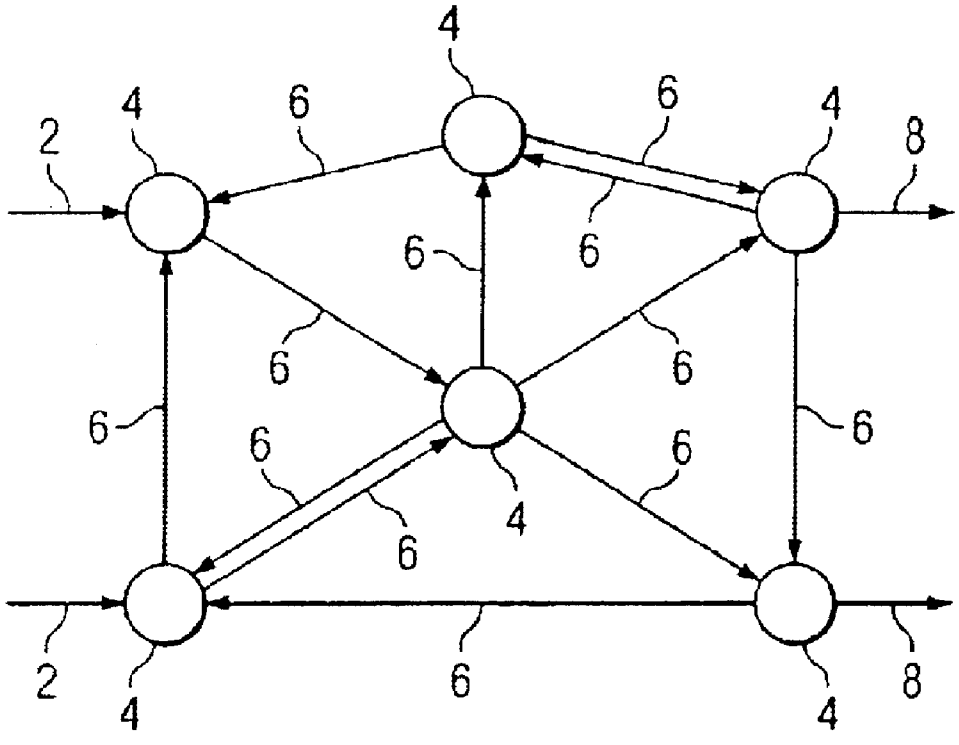


FIG. 1
(PRIOR ART)

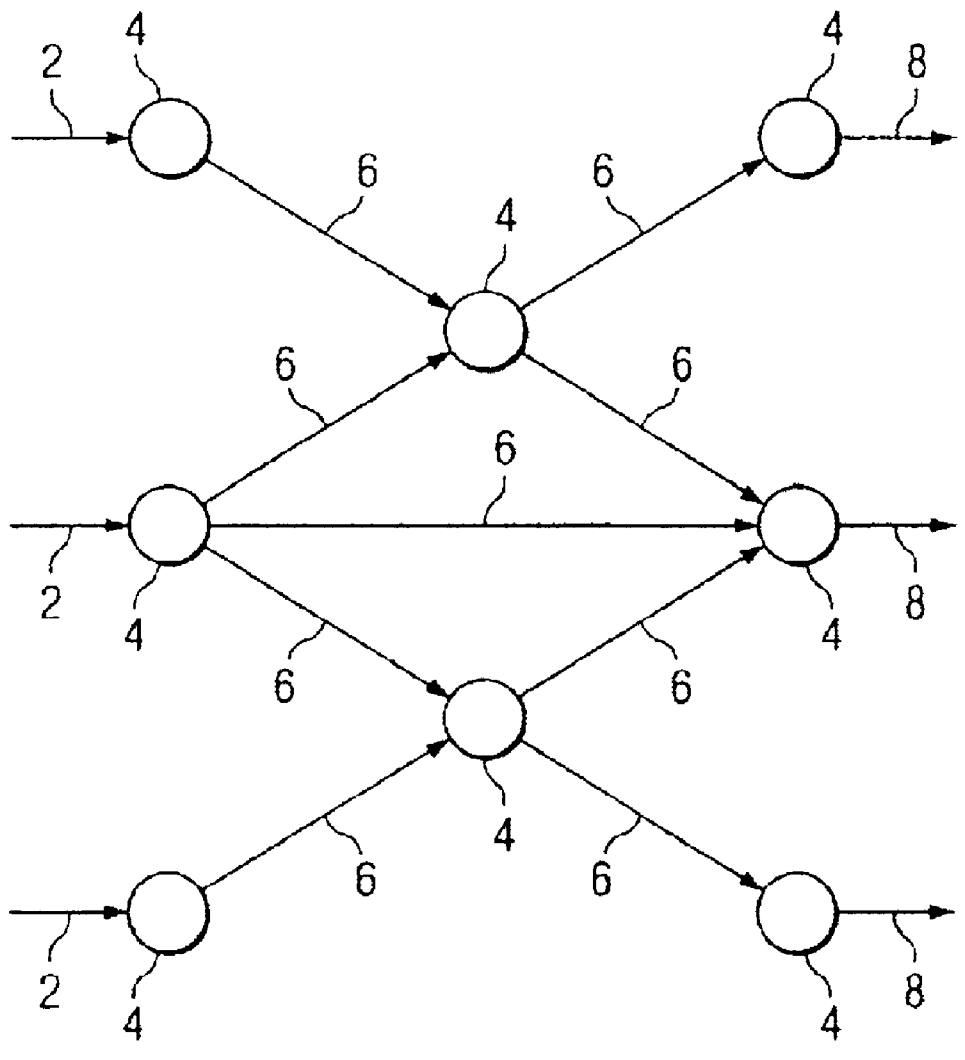


FIG. 2
(PRIOR ART)

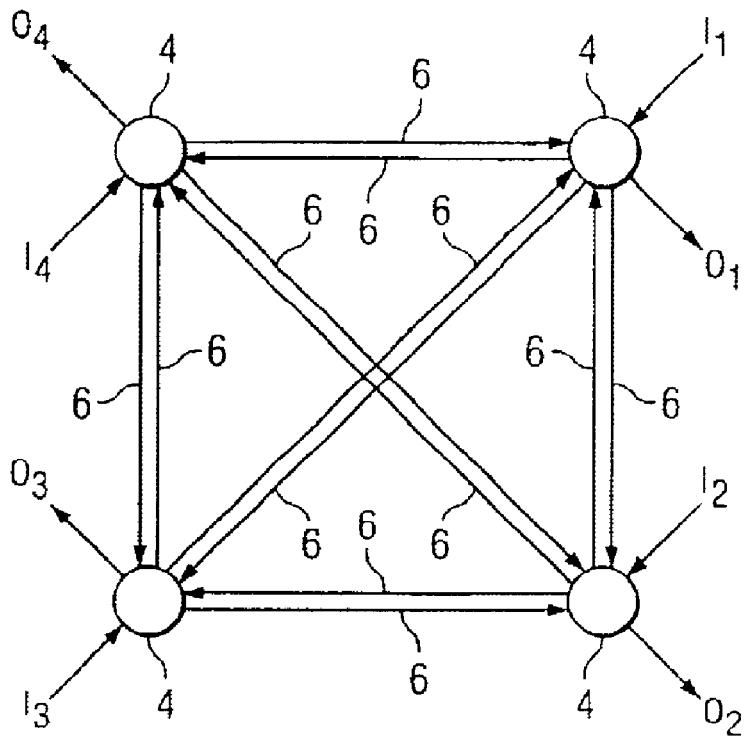


FIG. 3
(PRIOR ART)

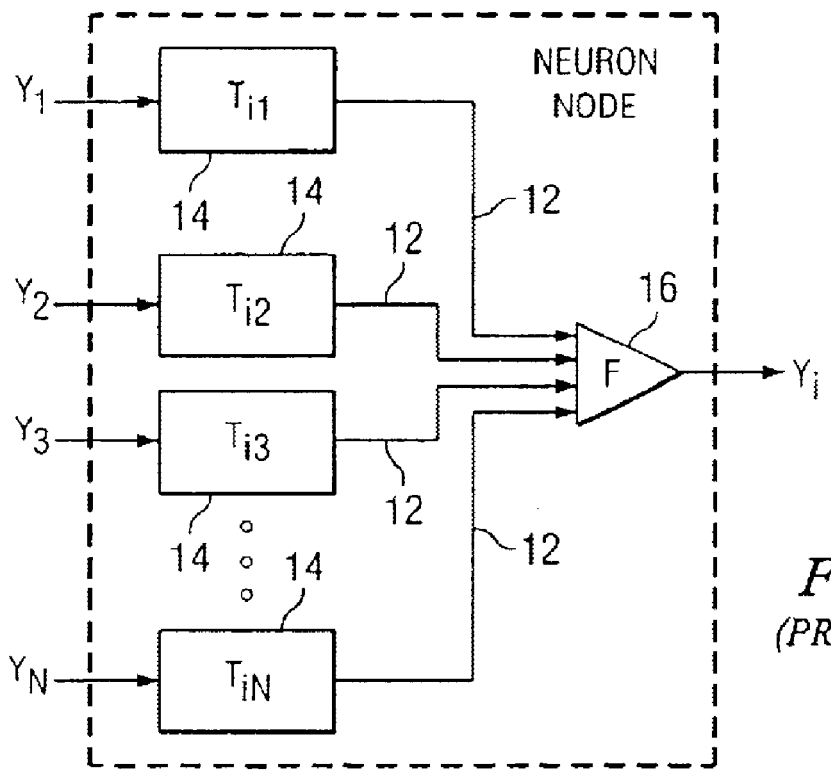


FIG. 4
(PRIOR ART)

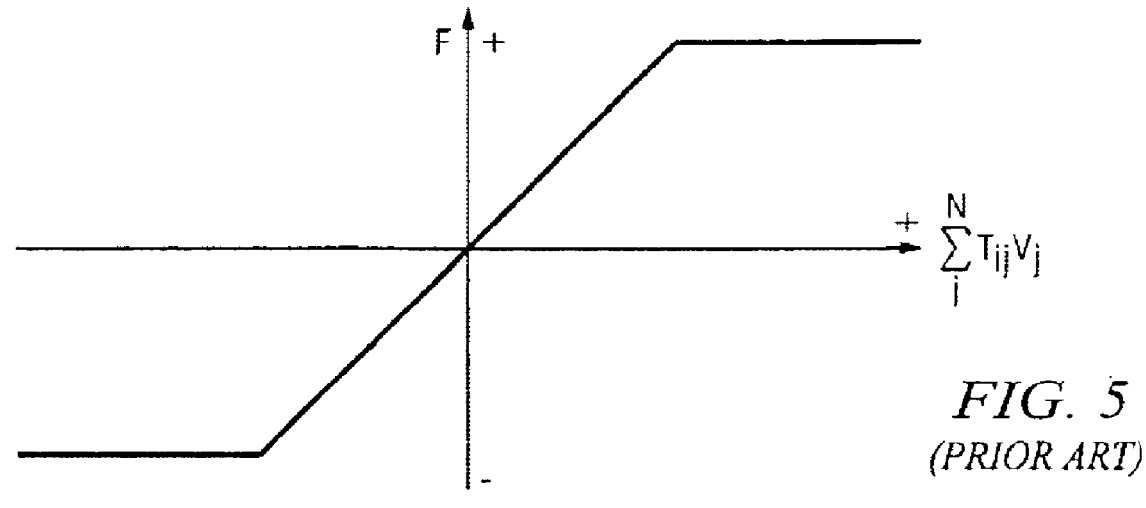


FIG. 5
(PRIOR ART)

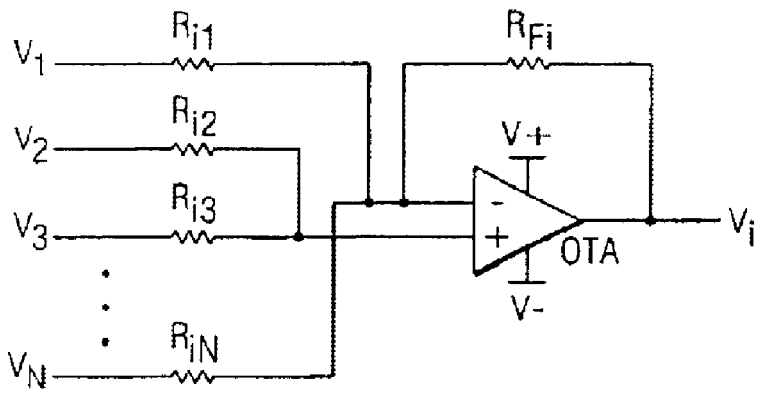


FIG. 6
(PRIOR ART)

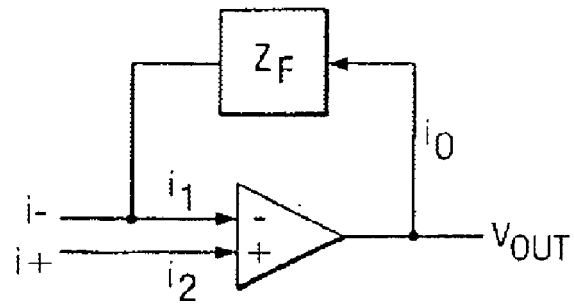


FIG. 7
(PRIOR ART)

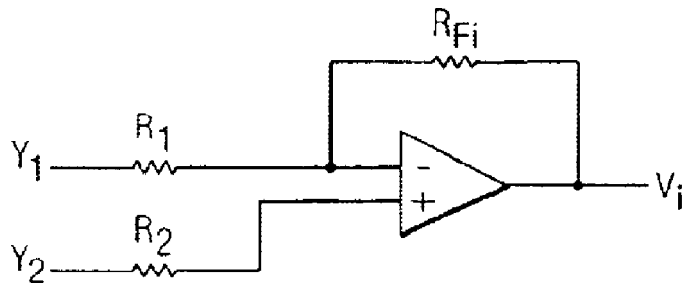


FIG. 8
(PRIOR ART)

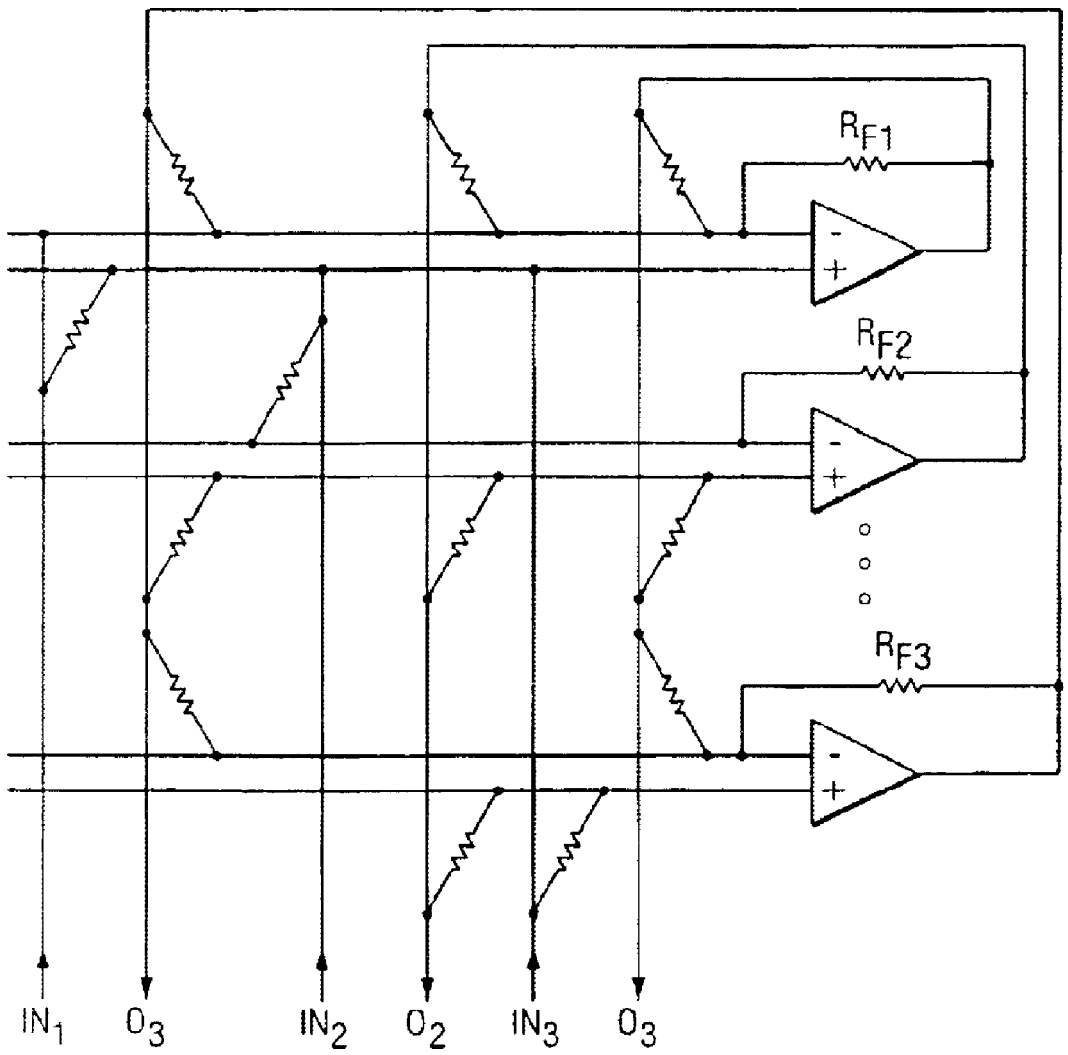


FIG. 9
(PRIOR ART)

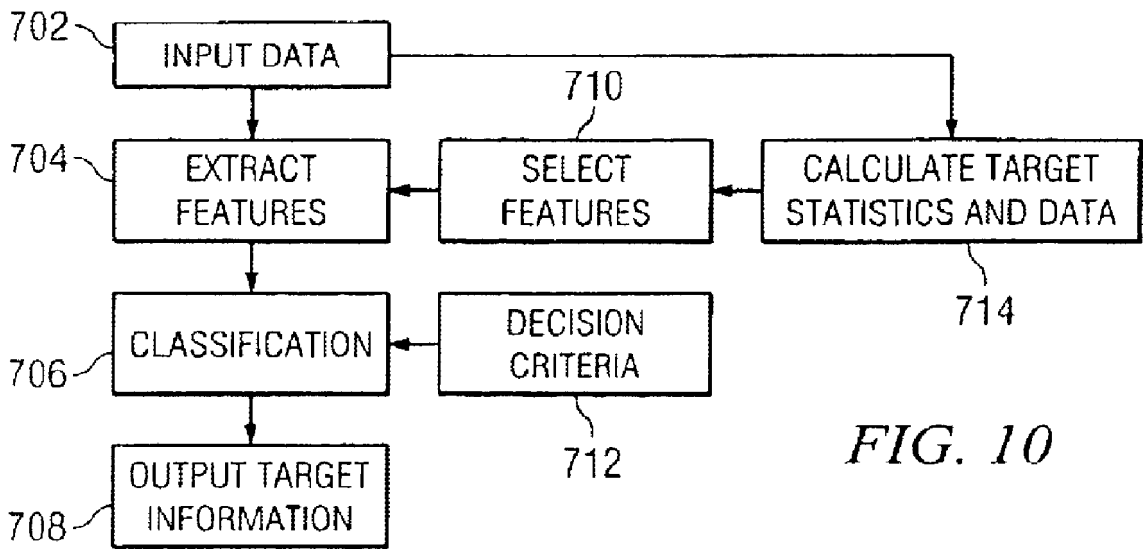

702
INPUT DATA
710
704
EXTRACT FEATURES
SELECT FEATURES
CALCULATE TARGET STATISTICS AND DATA
714
706
CLASSIFICATION
DECISION CRITERIA
712
708
OUTPUT TARGET INFORMATION

FIG. 10

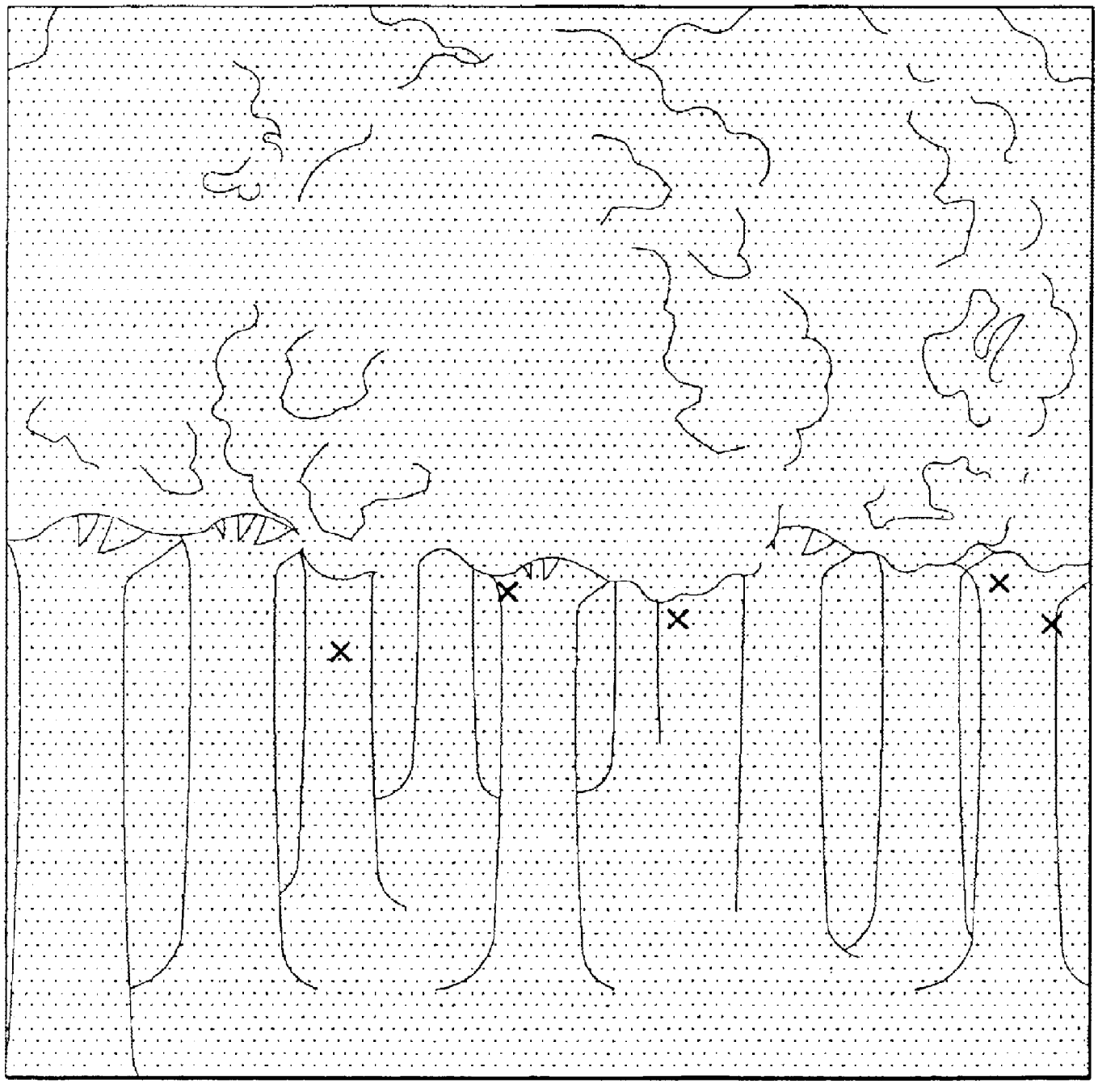


FIG. 11

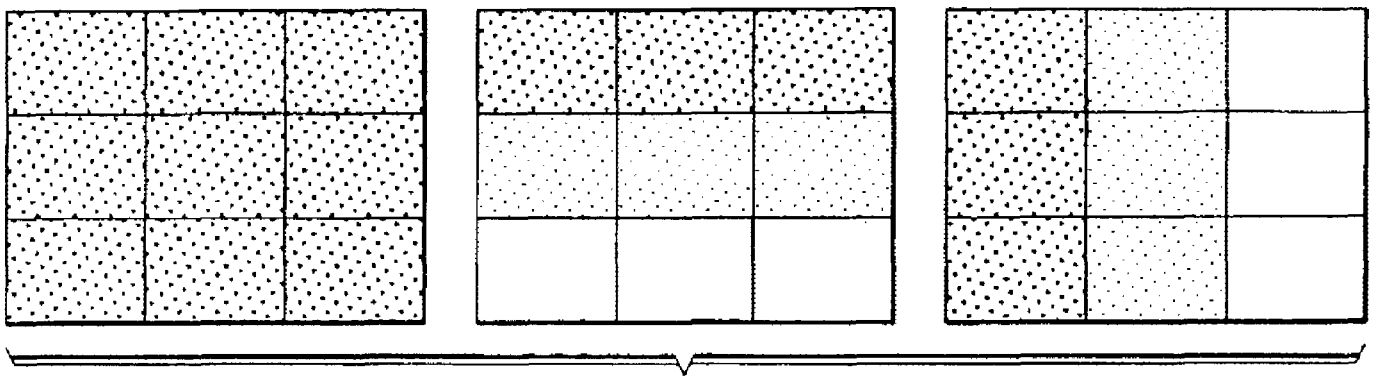


FIG. 12a

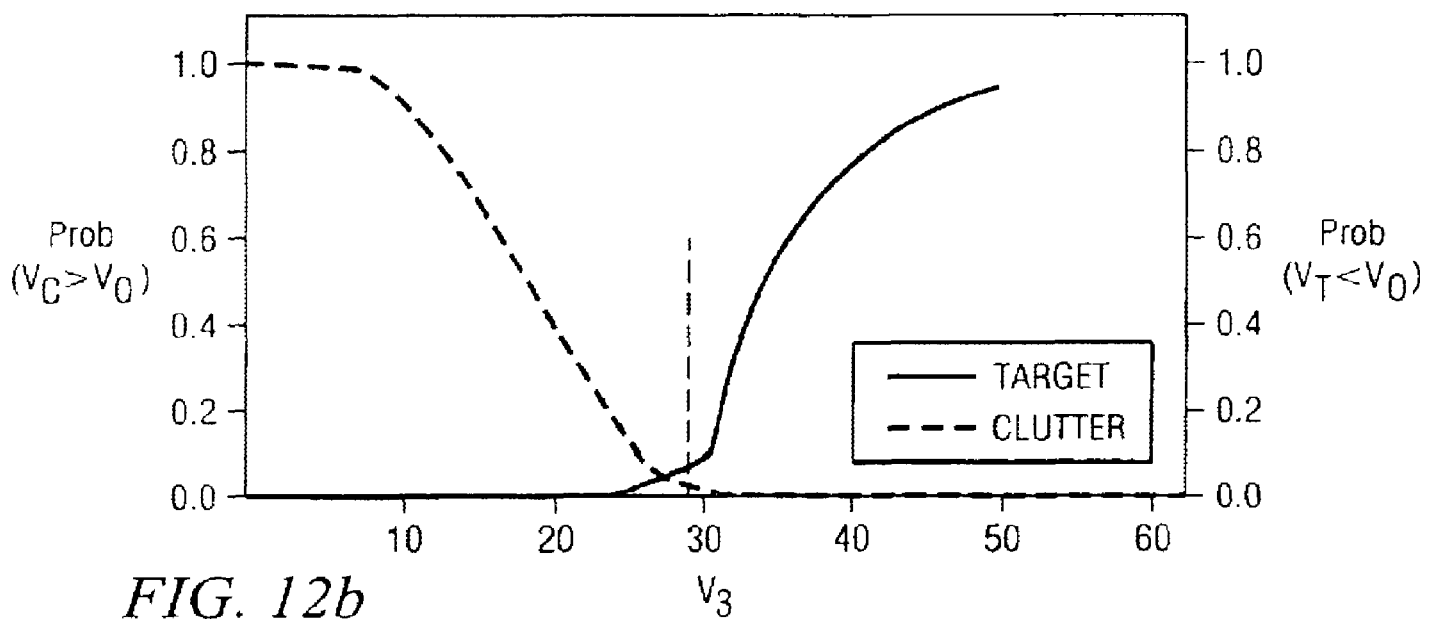


FIG. 12b

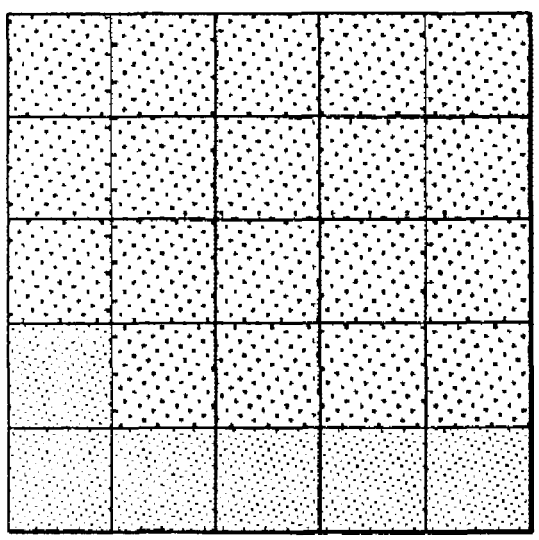


FIG. 13a

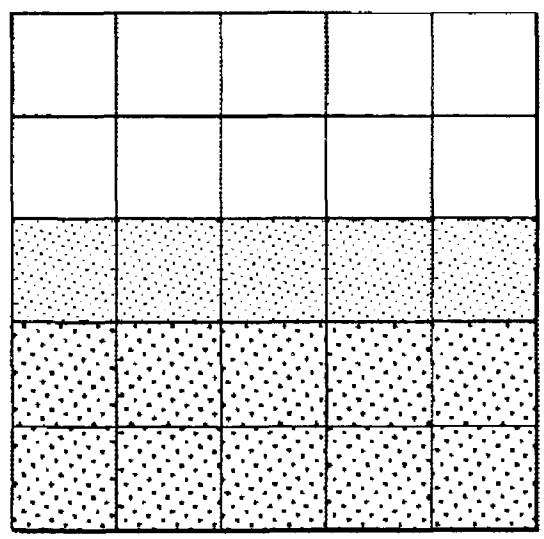


FIG. 13b

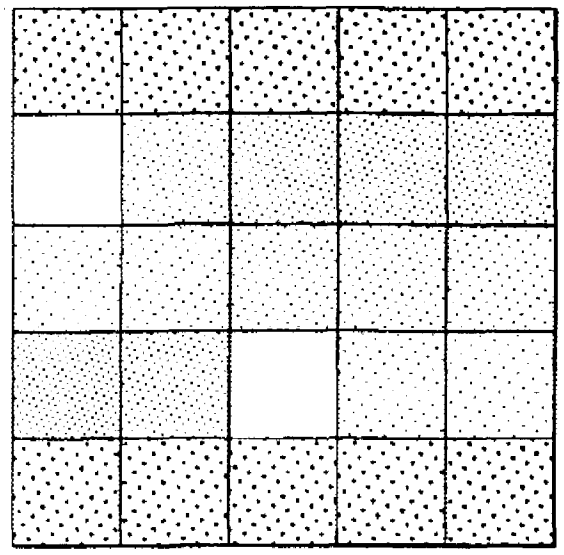


FIG. 13c

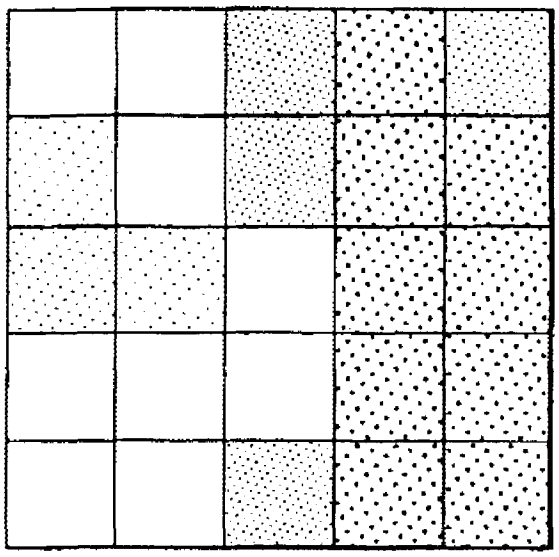


FIG. 13d

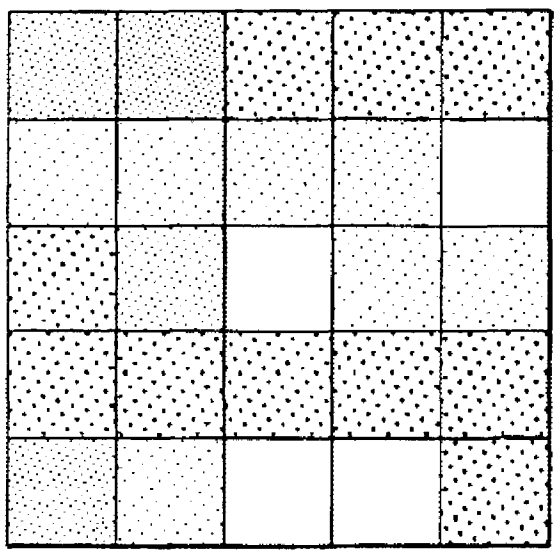


FIG. 13e

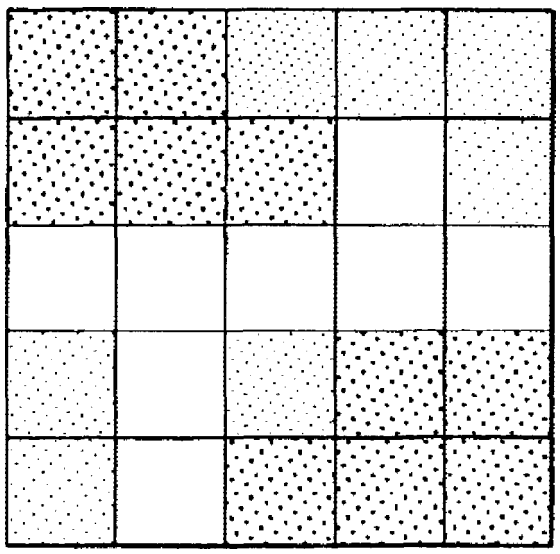


FIG. 13f

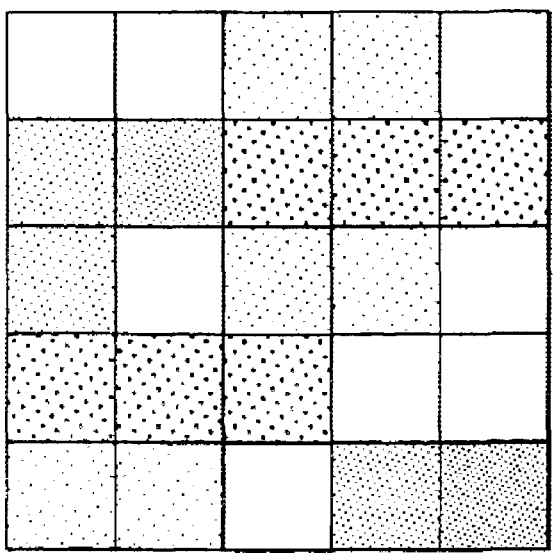


FIG. 13g

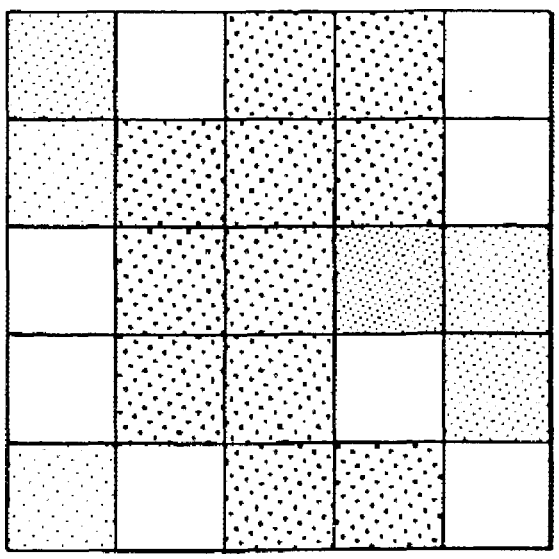


FIG. 13h

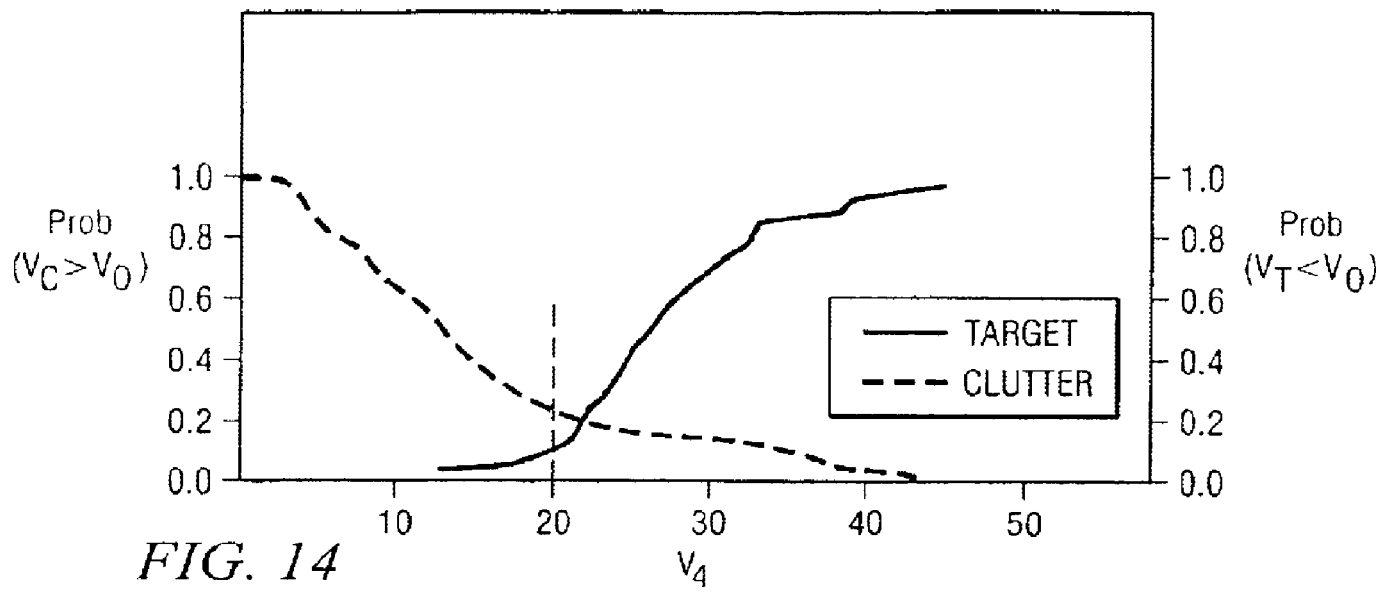


FIG. 14

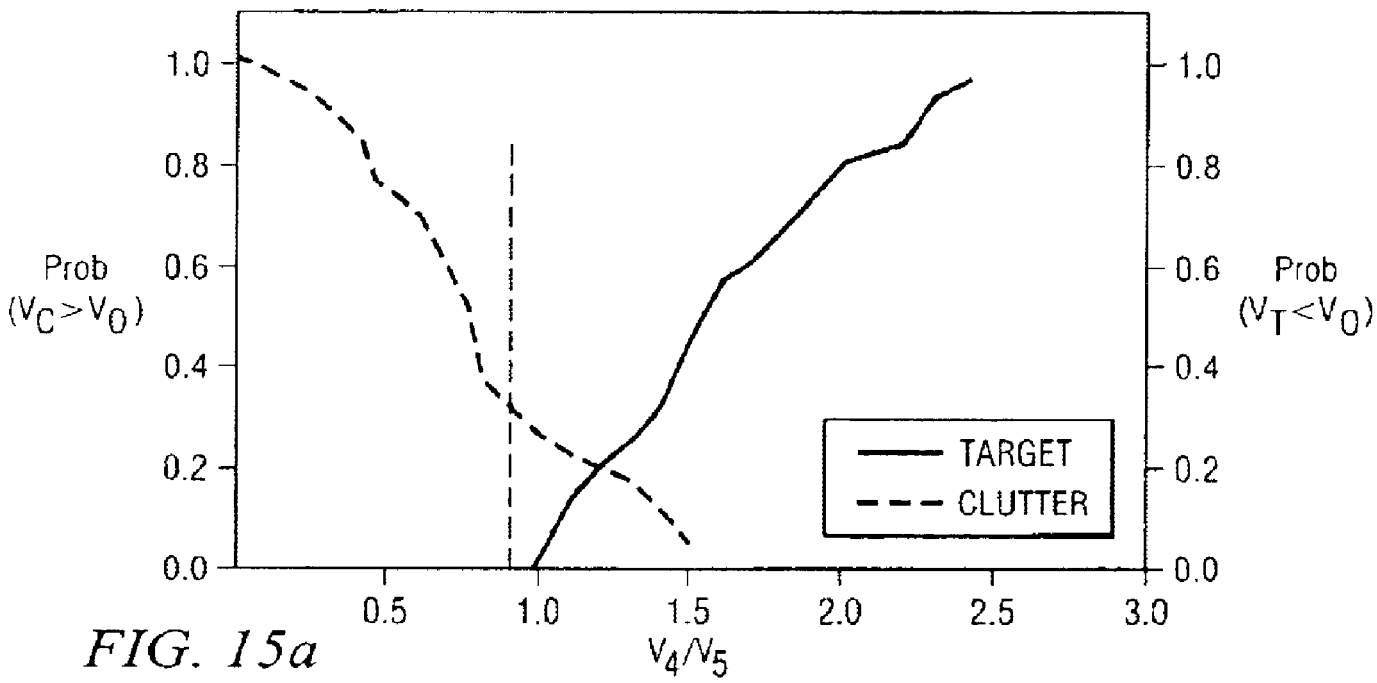


FIG. 15a

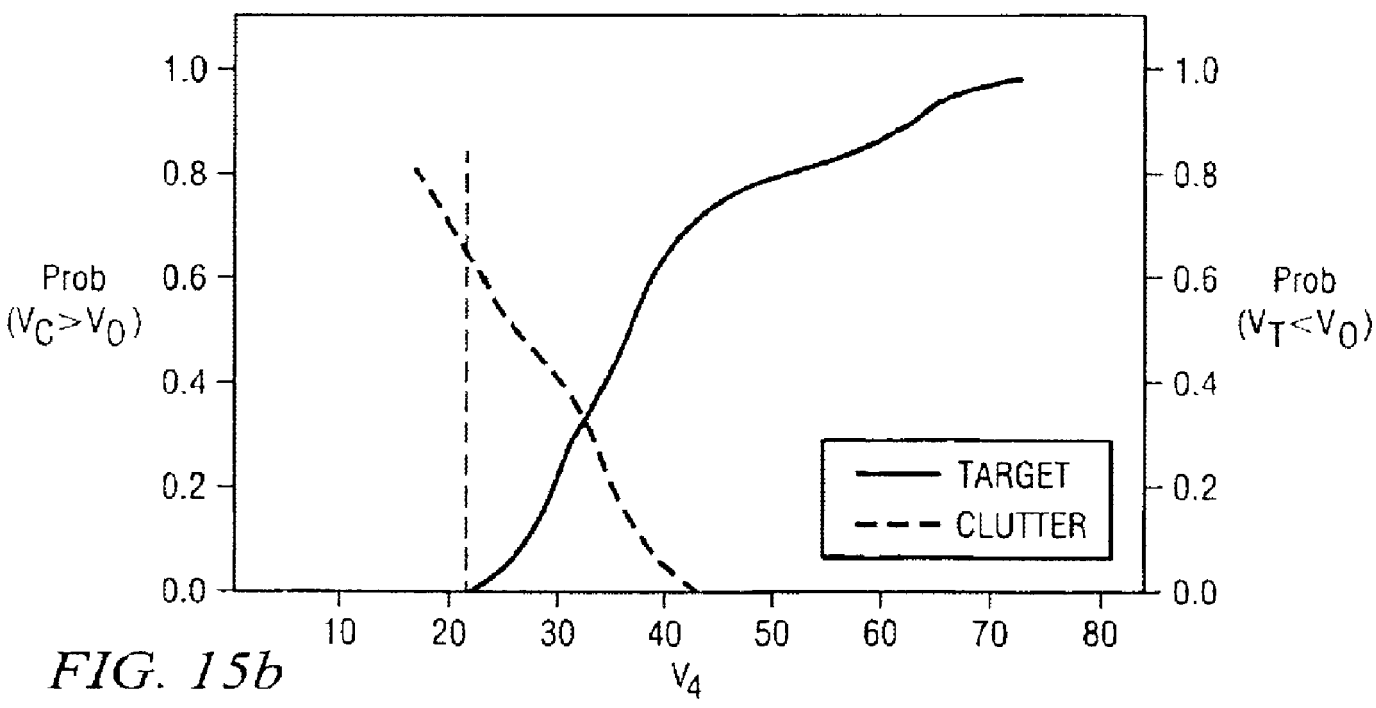


FIG. 15b

GENERALIZED HEBBIAN LEARNING FOR PRINCIPAL COMPONENT ANALYSIS AND AUTOMATIC TARGET RECOGNITION, SYSTEMS AND METHOD

(C) Copyright, Texas Instruments Incorporated 1991. A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

FIELD OF THE INVENTION

The present invention relates generally to products and methods pertaining to pattern recognition and classification procedures, parallel computer architectures, learning techniques, and neural networks.

BACKGROUND OF THE INVENTION

Target recognition and pattern classification involves the evaluation of new observations on the basis of past observations to distinguish targets or desired patterns from background clutter. This task is complicated by the complex nature of real world environments. Moreover, the computational overhead of many practical classification problems strain serial computer resources.

With respect to the complexity of pattern recognition and pattern classification techniques, many of these techniques rely on feature information to differentiate preferred targets from background clutter. Coordination and selection of this feature information is an ongoing problem, as the desired feature information itself (e.g., contrast-based and texture-based) often varies from target to target (e.g., size, stationary, make-up, composition, orientation), target to background (e.g. location of target), data set to data set (e.g., lighting, time-of-day, ambient temperature, and context of scene, sensor characteristics), and data source to data source (e.g., one-dimensional or multi-dimensional data, such as digitized infrared imagery, digitized TV imagery, digitized Infra-red imagery, speech samples, or radar samples).

Recently, researchers have focused on the perception of features in the visual field by biological visual systems and the organization of these biological visual systems. In short, this research has focused on what features are perceived and how are the receptive fields (e.g. neurons sensitive to particular features) organized. This work contributes to the understanding of biological vision systems and lays the groundwork for artificial vision systems in such areas as robotics and automatic target recognition.

In particular, D. J. Hubel and T. N. Wiesel in their article Hubel, D. J. and Wiesel, T. N, "Receptive Fields, Binocular Interaction, and Functional Architecture in The Cat's Visual Cortex." *Journal of Physiology*. vol. 160, pp. 106–154 (1962) and D. Marr in his book Marr, D. *Vision* W. H. Freeman and co. San Francisco 1982 are credited with initiating research in this field. Recent work, however, in artificial neural networks suggests mechanisms and optimization strategies that explain the formation of receptive fields and their organization in mammalian vision systems. This work includes Miller, K. d., Keller, J. B. and Stryker, M. P. "Ocular Dominance Column Development: Analysis and Simulation." *Science*, vol. 245, pp. 605–615 (1989), Durbin, R. and Michison, G. "A Dimension Reduction Framework For Understanding Cortical Maps." *Nature*, vol. 343, pp. 644–647 (1990), Linsker, R. "Self Self-organization In a Perceptual Network." Computer vol. 21, pp. 105–117 (1988), and Kammen, D. M. and Yuille, A. L. "Spontaneous Symmetry-breaking Energy Function." *Biological Cybernetics*, vol. 59, pp. 23–31 (1988). Linsker demonstrated how Hebbian learning algorithms, which change synaptic connections according to the degree of correlation between neuronal inputs and outputs, give rise to layers of center-surround and orientation-selective cell, even if the input to the initial layer is random white Gaussian noise. Kammen and Yuille show that orientation-selective receptive fields can also develop from a symmetry-breaking mechanism. Under certain conditions, the receptive fields perform a principal component analysis of the input data, as was shown in Oja, E. J. "A Simplified Neuron Model As A Principal Component Analyzer." *Mathematics and Biology*, vol. 15, pp. 267–273 (1982). Similarly, the article by Brown, T. H., Kairiss, E. W., and Keenan, C. L., "Hebbian synapses: Biophysical mechanisms and algorithms," Annual Review of Neuroscience, vol. 13, pp. 475–511 (1990) suggested Hebbian learning occurs in nature.

With respect to the computational requirements, neural networks provide parallel computational implementations. These networks embody an approach to pattern recognition and classification based on learning. Example patterns are used to train these networks to isolate distinctions between the particular patterns and background clutter for proper classification.

With respect to neural networks, the architecture of a neural network can be simply represented by a data-dependency graph, such as that shown in FIG. **1**. As shown, this data-dependency graph includes a set of active nodes **4** and a set of passive communicative links **6**. The graph nodes **4** represent artificial neurons and the passive communication links **6** define unidirectional communication paths between the artificial neurons **4** (graph nodes **4**). Additional links **2** are used as network inputs, which indicate which neurons will receive input signals from outside the network. Similarly, additional links **8** are used as network outputs, which indicate which neurons will transmit output signals to outside the network. A communication link **6** effectively makes the output state of one neuron at the tail of the link available to the neuron at the head of the link. For example, if the neuron states were made available as electrical voltages or currents, the communication links could be provided as simple conductive wires.

Referring to FIG. **2**, some architectures are feed-forward, wherein signals are inputed through communication links **2** which direct the input signals, flow through the communication links **6** to the neurons **4** in the neural network in a layer-by-layer fashion, and outputted through the communications links **8** which direct the output signals. No circular path of feedback exists within this type of network. Other networks are fully interconnected, wherein all neurons are connected to all other neurons. FIG. **3** shows an example of this type of network, wherein all neurons **4** are connected to all other neurons **4** through corresponding communication links **6** with input and output communication links as shown. In general, each neuron in a neural network architecture has a high fan-in and a high fan-out.

With respect to the neuron function itself, neuron states are generally defined by non-linear local functions. The output value of a neuron is defined by a transfer function, which depends only upon the neurons' current internal state, a set of local parameters called synapses, and the value of signals received from other neurons. FIG. **4** shows a simplified schematic of a neuron, along with its network links

12, synaptic parameters **14**, and typical input-output transfer function **16**. Input voltages V_1, V_2, V_3, . . . V_N are received from other neurons and output voltage V_i is outputed to other neurons. The neuron transfer function **16** is usually described as a non-linear function of the sum-of-products of the synapse parameters **14** with their associated input signal values. FIG. **5** shows a typical non-linear function that must be provided by each neuron that operates according to equation 1, which is defined as follows

$$V_i = F\left(\sum_{j=1}^{N} \sigma_{ij} V_i\right)$$

In defining the transfer function described by equation 1, V_i is the output value (or state) of neuron i,

$$\sum ij$$

is the value of the synapse parameter that modifies the effect of the output of neuron j on the state of neuron i, and F is a function of the sum-of-products of the synapses with neuron states. With respect to FIG. **5**, F is a linear function of the sum-of-products input near the zero sum region, while saturating to positive and negative values when the sum exceeds a preset positive or negative value.

With respect to training and using the neural network, the synapses affect the transfer function by modulating the strength of signals received from other neurons in the network. In virtually all models, the behavior of the network as a whole is changed by altering the values of the synaptic parameters. Once the parameters are altered, inputs can be applied to the network, and the processed output taken from the designated set of neurons. The time during which the synapse parameter values are altered is generally called the Training Phase. The time during which input signals flow through the network while the synapse parameters remain fixed is called the Recall or Relaxation Phase.

With respect to neuron circuit embodiments, FIG. **6** shows the schematic diagram of an artificial neuron implemented using resistors and an operational transconductance amplifier. The resistors R_{ij} are associated primarily with the synapses in equation 1, while the operational transconductance amplifier (OTA) and its feedback component R_6 is used to provide the desired transfer function F_i.

With respect to the operation of an OTA, as shown in FIG. **6**, an ideal OTA generates an output voltage proportional to the difference of the currents entering its inputs, multiplied by a very large gain factor. FIG. **7** shows an ideal OTA. The defining equation **7** for this circuit may be derived by the analysis, as follows

$$V_{OUT}=A(i_2-i_1)\ A>>>1$$

$$i_1=i^-+i_0$$

$$V_{out}=A(i^+-i^--i_0)$$

$$i_o = \frac{V_{out}}{Z_F}$$

Thus,

$$V_{out} = \frac{Z_F A}{A + Z_F}(i^+ - i^-)$$

If $A>>Z_F$,

$$V_{out}=Z_F(i^+-i^-)$$

The gain factor A is infinite in the idea case, and very large values of A (Over 1 million) can be obtained in practice using off-the-shelf OTA components. An ideal OTA will have a zero input resistance at its inputs, so that the operation of the circuit may be understood by assuming equation 7, and by assuming that the voltage at either input referenced to ground is always zero. Electrical components placed between the output of an OTA and its negative input induce negative current feedback in the circuit while components placed between the OTA output and its positive input induce positive current feedback. Simple resistive feedback causes the OTA to generate an output voltage proportional to the net current difference at its' positive and negative input terminals, as shown in FIG. **8** and the corresponding equations 8, 9 and 10, which are as follows

$$V_i = R_{FI}\frac{V_2}{R_2} - \frac{V_1}{R_1}\ V_1 = \sum_{j=1}^{2} \sigma_{ij} V_j\ \sigma_{ij} = \frac{R_{FI}}{R_j}$$

Using these assumptions, the operation of the neuron circuit can be explained. The resistors attached to the input of the amplifier shown in FIG. **6** serve to produce input currents proportional to the product of the input voltages V_j and the conductances of the resistors R_{ij}. The input resistors thus act as the synaptic parameters in FIG. **4**, while the neuron state values are represented as the output voltages of the OTAs that implement each neuron function. In general, synapse parameters can be bipolar. That is, these parameters can take on positive and negative values. However, it is not simple in general to implement negative conductances that physically represent mathematically negative synaptic parameters. Fortunately, the differencing nature of the OTA can be used to provide the effect of negative conductances using positive conductance elements. This is accomplished by attaching resistors that provide the absolute (positive) value of the desired conductance to the negative input of the OTA. Since the subtraction of a positive current in the OTA is equivalent to the addition of a negative current, negative synapse parameters can be equally well implemented using standard resistive elements. The feedback resistor R_6 defines the slope of the OTA transfer function in the linear region, while the power supply voltages determine the limiting values of the transfer function in regions of strong negative and positive bias. The circuit shown in FIG. **6** provides the basic functional requirements for a typical neural network neuron. Using a more complex feedback impedance in place of R_6 can provide more complex transfer functions when necessary.

FIG. **9** shows a schematic of a general neural network based on the use of resistor elements for the synapse parameters and OTAs for the neurons. A non-intersecting grid of wires is used to allow synapse resistors to be connected from OTA outputs to OTA inputs. Pairs of grid lines are used to allow each synapse resistor to be connected from an OTA output to either a positive or negative OTA input terminal. This functionality may be implemented in a software domain as well. For instance, the software algo-

rithm shown in Table 1 produces a dynamic change in neuron state similar to the dynamic changes expected from the OTA array shown in FIG. **9**. The synaptic parameters (Resistor values) are assumed to be fixed and pre-stored in a two dimensional matrix R. The OTA feedback resistors are assumed to be the same for all neurons. The neuron states are assumed to be initially preset to certain values and are initially defined by a one dimensional matrix V. All parameters are assumed to be double precision floating point. While only the relaxation model is shown, other models exist and are well known in the art. The training algorithm varies according to the specific neural network type.

SUMMARY OF THE INVENTION

This application discloses a method for distinguishing targets from background clutter, comprising the steps of inputting data (e.g., one-dimensional or multidimensional data, such as digitized infrared imagery, digitized TV imagery, speech samples, or radar samples), calculating data statistics from the data and using said data statistics to select target specific feature information (e.g., contrast-based and texture-based) to distinguish specific targets from background clutter, generating the target specific feature information from the data statistics, extracting the target specific feature information from the data, using the target specific feature information to distinguish specific targets from background clutter (e.g., parametric and non-parametric techniques and one-class, two-class, or multi-class classification procedures), and outputting target and background clutter information. In particular, this method uses the Hebbian Learning to distinguish and take advantage of second order correlation information.

The invention may also be incorporated into a classification system, comprising a data input (e.g., one-dimensional or multi-dimensional data, such as digitized infrared imagery, digitized TV imagery, speech samples, or radar samples), a calculator to calculate data statistics from the data input, a selector to use the data statistics to select target specific feature information (e.g., contrast-based and texture-based) to distinguish specific targets from background clutter, a generator to generate the target specific feature information from the data statistics, a feature extractor to extract target specific feature information from the data input, and a classifier to use the target specific feature information to distinguish targets from background clutter and output target and background clutter data (e.g., parametric and non-parametric techniques and one-class, two-class, or multi-class classification procedures).

The advantages of this method and system and the corresponding hardware and software embodiments are that they generate and select the target specific feature information for the classification of preferred targets in the context of particular backgrounds, and thereby increases the robustness of the overall classification process. In particular, this method and system uses a variation of the Hebbian learning algorithm and recently discovered relationships between neural network Hebbian learning algorithms and statistically-based principal component analysis to generate a set of distinguishing target characteristics pertinent to the classification and separation of targets from background clutter in automatic target recognition applications.

In addition, advantages also include the tailoring the features (in the form of principal component filters) to reflect the characteristics of a specific target or target class which leads to separability of the targets based on a relatively small number of features, which reduces the size of the feature space and the corresponding size of the training set required for adequate generalization. Similarly, using only target training examples for feature generation removes biases that arise from nonrepresentative clutter training sets and signal-to-noise ratios aid in identifying features that are more robust over the training data. Likewise, hierarchies of feature filters covering relevant length scales in the image set provide further discrimination of the object classes and indicate any scale invariant properties of the objects. Finally, carrying out filter convolutions is a local operation and can be parallel over the image to yield fast implementation times.

Although our focus is on automatic target recognition, the described learning and classification methods extend to a variety of pattern recognition domains. The invention may be implemented in software and hardware configurations as well. Parallel computer archetectures, such as neural networks, are used extensively.

BRIEF DESCRIPTION OF THE DRAWINGS

Further features and advantages will become apparent from the following and more particular description of the various embodiments of the invention, as illustrated in the accompanying drawing, wherein:

FIG. **1** shows a data-dependency graph of a typical neural neural network.

FIG. **2** shows an example of a feed-forward neural network.

FIG. **3** shows an example of a fully connected neural network.

FIG. **4** shows a simplified schematic of a neuron, along with its network links, synaptic parameters, and typical input-output transfer function.

FIG. **5** shows a typical non-linear function provided by each neuron operating according to defining equations.

FIG. **6** shows the schematic diagram of an artificial neuron implemented using resistors and an operational transconductance amplifier.

FIG. **7** shows an ideal operational transconductance amplifier.

FIG. **8** is a 2-input operational transconductance amplifier, which shows how equation 1 is provided.

FIG. **9** shows a schematic of a general neural network based on the use of resistor elements for the synapse parameters and OTAs for the neurons.

FIG. **10** shows a preferred system configuration.

FIG. **11** is a TV image from data set I, which contains the five targets (marked with an "x"). that were used for training.

FIG. **12***a* shows 3×3 pixel filters determined by learning with the generalized Hebbian algorithm, which are the first three principal components for the training target regions from data set I.

FIG. **12***b* is a graph showing the separability of target and clutter objects based on variance values derived from the third filter in FIG. **12***a*.

FIGS. **13***a*, **13***b*, **13***c*, **13***d*, **13***e*, **13***f*, **13***g*, and **13***h* show the **5**×**5** pixel filters derived from training data set II, which represent top eight principal components ordered by decreasing eigenvalue.

FIG. **14** is a graph showing variances derived from Filter **5** in FIG. **13***e* in conjunction with the probability of detection and the false alarm rate.

FIG. **15***a* is a graph showing the ratio of variances derived from Filters **4** and **5** in FIGS. **13***d* and **13***e*.

FIG. **15***b* is a graph showing the variance derived from Filter **4** in a second set of 5×5 pixel filters described in the detailed description in conjunction with the probability of detection and the false alarm rate.

Table 1 shows a software algorithm that produces a dynamic change in neuron state similar to the dynamic changes expected from the OTA array shown in FIG. **9**.

Table 2 is a description of the two data sets discussed in text.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

FIG. **10** shows a preferred classification system configuration. As shown in FIG. **10**, the classification system inputs the data in module **2**, extracts feature information from the inputted data in module **4** and inputs the feature information into the classification module **6** which uses decision criteria **12** to classify the target from background clutter and outputs this information in module **8**. The feature selection module **10** uses target statistics and data which are extracted from the input data in module **14** to select and generate features specific to a given target class and inputs this information into module **4** that extracts features. These distinguishing target characteristics help the classifier separate targets from background clutter, which makes the classification system more robust.

In particular, the preferred system shown in FIG. **10** accepts primarily long-wavelength infrared (IR) and television (TV) imagery to be analyzed. As discussed, infra, the data inputs are primarily derived from two data sets that characterize two target scenarios: targets that are partially occluded (e.g., targets along tree lines) and targets that resemble background clutter. Various procedures are used to extract feature information, which will be discussed infra in conjunction with the feature selection process. The preferred system shown in FIG. **10** uses binary tree structure to classify inputted feature information, wherein the parameters defined by the eigenvalues associated with the principal components, discussed infra in conjunction with the feature selection process, constitute the decision criteria.

With respect to the feature selection process, the preferred embodiment uses target statistics and a neural network that implements a generalized Hebbian learning algorithm to select and generate features specific to a given target class. The generalized Hebbian learning algorithm employed, which was discussed in the Sanger, T. D. "Optimal Unsupervised Learning." *Neural Networks* vol. 2, pp. 459–473 (1989), selects features for automatic target recognition from the inputted imagery. The generalized Hebbian Learning algorithm generates principal components or receptive features that resemble those found in Linsker's work, discussed supra. Inter alia, Sanger shows that the GHA determines the principal components of the data set in order of decreasing eigenvalue, which is important because the higher eigenvalue components are generally the most important or prominent features.

The generalized Hebbian learning algorithm trains a one-layer neural network, where the input nodes define arrays of pixel intensity values from image data and the output nodes index the principal components. Sanger describes the procedure with the following equation:

$$c_{ij}(T+I) = c_{ij}(T) + \gamma(T)\left[Y_i(T)x_j(T) - Y_i(T)\sum_{k<i} c_{kj}(T)y_k(T)\right]$$

where c_{ij} is the weight or connection strength between the j^{th} input neuron and the i^{th} output neuron, x_j is the j^{th} component of the input vector, y_i is the i^{th} component of the output vector, and Y(T) is a learning parameter that decreases with time such that

$$\lim_{T\to\infty} \gamma(T) = 0 \text{ and } \sum_{T=0}^{T=\infty} \gamma(T) = \infty$$

The second term on the rhs of equation 1 is the Hebbian term and the third term ensures that the algorithm learns successive eigenvectors (which are the principal components) of the co-variance matrix of the input vectors ordered by decreasing eigenvalue. This decomposition of the co-variance matrix in terms of eigenvectors is the well-known Karhunen-Loeve transform. Local operations affect equation 1, which distinguishes the relationship described in equation 1 from other relationships characterized by the Karhunen-Loeve transform and underscores the importance of equation 1 for training neural networks. Whereas Sanger applies equation 1 to image coding, texture segmentation, and the generation of receptive fields, principal components have been used to characterize image texture, as discussed in Ade, F. "Characterization Of Textures By Eigenfilters." *Signal Processing*, vol. 5, pp. 451–457 (1983) and are used in the preferred embodiment.

The preferred embodiment uses equation 1 to develop receptive fields for identifying a specific target object. The extracted characteristics of the target object are embedded in the second-order image statistics. Inputs to the network are r×s arrays of pixel values, which are rastered into r×s component vectors, from image subregions that contain the target of interest. The resulting principal components are directions in the r×s dimensional input space with maximum variance. Projections of input vectors along the principal component directions are maximally distinguishable. Eigenvalues corresponding to the principal components determined from equation 1 provide a measure of the variance in the principal component directions. Since vectors in the input space are made up of pixel intensities, the principal components generally correspond to prominent intensity patterns or features in the object of interest. The preferred embodiment trains on several examples of the target object to smooth out noise present in individual examples and to generate principal components that signify features common to different occurrences of the object.

The generated principal components are arrayed in r×s matrices to produce receptive fields or filters that are convolved with the original image data during classification. The preferred embodiment convolves these filters such that the spatial sampling of pixel intensities used to construct the input vectors is preserved. Next, every r×s array of pixels contained in the image subregion of interest is multiplied by the generated filters and then compute the variances of the resulting convolutions. Variances (these are related to the eigenvalues of the principal components) or ratios of the variances (these provide a measure of the relative content of two patterns) compose the parameter sets used for classification. The preferred embodiment uses range information, if available, to scan the image to properly adjust the size of the box or window circumscribing the subregion of interest to reflect the target size.

The spatial sampling density entering into the construction of the input vectors is an important parameter in the generation of the receptive fields. This parameter corresponds to the synaptic connection density in Linsker's Hebbian algorithm, discussed supra, which only yields principal components if the connection density is held fixed. The preferred embodiment uses the spatial sampling density to determine the scale of the feature. Please note that the spatial sampling density must be appropriately scaled to ensure that the same feature scale is measured in all cases, if the target object occurs at different ranges in the image data, as in the inputted data reviewed in the preferred embodiment. As the spatial sampling density for targets at nearer ranges are reduced, the preferred embodiment averages over the shorter length scales to avoid aliasing effects. Thus, the preferred embodiment assumes that the target at the longest range fits into a u×v pixel box. Input vectors for this case are formed from intensity values of r×s blocks of pixels (where these blocks are smaller than the box size) extracted from the box circumscribing the target object. The preferred embodiment composes input vectors from 2r×2 s blocks of pixels for targets at half the initial range and extracts the intensity value from every second pixel. The preferred embodiment continues in this fashion in closer ranges.

The preferred embodiment also scales the spatial sampling density for a given target sample to generate a hierarchy of receptive fields at different scales. The relevant scales are set by the smallest scale detectable (effectively the resolution) for the target seen at longest range and the size of the target object. This hierarchy characterizes the target object in terms of what features become relevant at different length scales. For example, self-similar objects, which are objects that look the same at any scale, are expected to have an invariant feature set as a function of scale. Scaling of the spatial sampling density resembles renormalization group analysis that assumes system behavior is governed by how the physical operators scale, discussed in Wilson, K. G. "The Renormalization Group: Critical Phenomena And The Kondo Problem." *Review of Modern Physics* vol. 47, PP. 773–840 (1975).

The inventor used test data sets with military targets to evaluate the effectiveness of the preferred system implementation, described supra. Table 2 shows a description of the two data sets used. Images in data set I were TV whereas those in data set II were long-wavelength IR. Target objects for both data sets were of several different types, so the preferred embodiment needed to find receptive fields common to all types. Different orientations of the target objects in these data sets did not appear to generate problems. Apparently, targets from data set I were positioned in the field-of-view at long enough range that there was little sensitivity to orientation, and targets from data set II were sufficiently spherically symmetric to neutralize orientation effects. Preprocessing procedures were used to normalize linearly the pixel intensities so they fell in the range from 0 to 255.

The inventor defined clutter objects for both data sets as any region in the image passed by the screener and not a target fell into the clutter class, which is consistent with conventional screeners. The inventor then used this definition to measure the capability of the principal component features to distinguish targets from objects similar in appearance.

FIG. **11** is a TV image from data set I, which contains the five targets (marked with an "x") that were used for training. For data set I, the experiments used 3×3 pixel filters using five targets contained in a single image and tested filter performance on 9 additional images. Targets in the training image represented different types of the target class. Target heights in data set I ranged from 4 pixels to 32 pixels.

FIG. **12***a* displays the 3×3 pixel filters determined by learning with the generalized Hebbian algorithm, which are the first three principal components for the training target regions from data set I. The first filter emphasizes regions that have strong grayshade contrast with the background environment; most of the targets show strong contrast with background. The second and third features highlight regions with strong horizontal and vertical grayshade gradients, respectively; target regions all have sharp transitions in grayshade, from pixels within the targets to pixels outside the targets.

FIG. **12***b* shows the separability of target and clutter objects based on variance values derived from the third filter in FIG. **12***a*. Abscissa values V_3 are the variances derived from Filter **3** in FIG. **12***a*. The left-hand curve is the probability for a clutter region to have a variance V_c larger than variance V_o; the right-hand curve is the probability for a target region to have a variance V_t smaller than variance V_o. The dotted line indicates the lower-bound on V_t determined from the training data. Please note the dotted line lies close to the point where the two probability curves cross, which marks the demarcation point in a Bayes classification approach. In particular, FIG. **12***b* indicates that most of the background clutter has a much smaller vertical gradient content than the targets; tree lines and roads and horizons in the image set extend mostly parallel to the horizontal edges of the image and provide little vertical gradient content. The inventor also examined the same images with a feature set based on contrast and edge measures and found that between **8** and **16** of these features were required to achieve the same level of separability found in FIG. **12***b* and that the discriminant surface, which divided targets and clutter, was highly nonlinear.

The experiments generated eight 5×5 pixel filters, which are shown in FIGS. **13***a*, **13***b*, **13***c*, **13***d*, **13***e*, **13***f*, **13***g*, and **13***h* for target set II from four examples in a single image and tested filter performance on 23 images. The 5×5 pixel filters were derived from training data for data set II. The filters shown represent the top eight principal components and are ordered by decreasing eigenvalue. Only filters **4** and **5** were used for classification. SIN denotes the signal to noise ratio defined infra. In addition, the 5×5 pixel filters shown in FIGS. **13***a*, **13***b*, **13***c*, **13***d*, **13***e*, **13***f*, **13***g*, and **13***h* represent top eight principal components ordered by decreasing eigenvalue. Target heights ranged from 15 pixels in the far range to 175 pixels in the near range. Data set II was in several ways more challenging than the first data set because the noise level was higher and the target texture was very similar to the texture of the clutter. The filters were ranked by a signal-to-noise ratio (SIN), as shown in FIGS. **13***a*, **13***b*, **13***c*, **13***d*, **13***e*, **13***f*, **13***g*, and **13***h*, where S is the mean of the variances from the four examples in the training set and N is the variance of the variances from the same training examples. S/N is a natural criterion, since the target characteristics sought are both prominent and invariant over the data set, which leads to large values of S/N.

FIG. **14** is a graph showing variances derived from Filter **5** in FIG. **13***e* in conjunction with the probability of detection and the false alarm rate. V_5 denotes variances derived from Filter **5** in FIG. **13***e*. The dotted line is the lower bound on V_t determined from the training data. With the lower bound, the probability of detection, P_d, is 0.89 and the false alarm rate, P_{fa} is 0.27. As shown in FIG. **14**, filter **5** had the largest S/N ratio among the eight generated and alone provides

considerable separation of targets and clutter. FIG. **14** also shows a lower cutoff for the target region, as the training examples establish a lower bound on the target region in the feature space due to the higher degree of noise and corresponding smaller variance values associated with these examples. The use of noisy examples, (and, hence, examples more easily confused with clutter) helps estimate the true position (as determined from an infinitely large data set) of the classification discriminant surface from examples that lie close to the boundary between target and clutter than from examples that lie further away.

To achieve further separability of targets from clutter, the inventor also examined ratios of the variance values from the filters and found the ratios of the variance outputs from filters **4** and **5** in FIGS. **13***d* and **13***e* further distinguish targets and clutter, as shown in FIG. **15***a*. In particular, FIG. **15***a* is a graph showing the ratio of variances derived from Filters **4** and **5** in FIGS. **13***d* and **13***e*. V_c, V_t, and V_o indicate the ratio of variances from Filters **4** and **5**. The dotted line is determined from training data. The clutter curve uses only false alarms from the results in FIG. **14**, whereas the target curve uses all the target points. Using the dotted line to discriminate targets and clutter (together with results from FIG. **14**), the detection probability is defined to be 0.89 and the false alarm rate is defined to be 0.10, which is a significant improvement over 0.27.

In addition, the experiments used the training image to generate a second set of eight 5×5 pixel filters at a larger length scale by halving the spatial sampling density of the input vectors to the neural network. The filter with the second largest S/N ratio was the same as filter **5** in FIG. **13**. This same filter further reduces the number of clutter points mistakenly identified as targets, as shown in FIG. **15***b*, which suggests some self-similarity in the target characteristics and indicates the persistence of a particular feature over a factor of two in length scales. In particular, FIG. **15***b* is a graph showing the variance derived from Filter **4** in a second set of 5×5 pixel filters described in the detailed description in conjunction with the probability of detection and the false alarm rate. V_4 is the variance derived from Filter **4** in a second set of 5×5 pixel filters, described infra. The clutter curve uses only false alarms from the results in FIG. **15***a*. All target points, however, are used. The dotted line is the discriminator between targets and clutter set by the training data. If the results shown in FIGS. **14**, **15***a*, and **15***b* are used in sequence, the detection probability is determined to be 0.89 and the false alarm rate is determined to be 0.07, which is another significant improvement in performance. As a group, the three filters and their corresponding variance outputs establish a binary classification tree for distinguishing targets from background clutter. The final probability of detection (ratio of the number of targets detected to total number of targets) was 0.89, whereas the false alarm rate (ratio of the number of clutter points mistakenly classified as targets to total number of clutter points) was 0.07.

The values and ranges disclosed herein are given for a preferred embodiment of the present invention, of course these values may vary as the parameters are changed. Therefore it is to be understood that many changes in detail may be made as a matter of engineering choices without departing from the spirit and scope of the invention, as defined by the appended claims. In particular, alternate procedures that share the same functionally as the procedures discussed, supra, such as alternate classification techniques or data inputs, are plainly within the scope of the invention.

I claim:

1. A method for distinguishing targets from background clutter, comprising the steps of

(a) inputting data;

(b) calculating data statistics from said data and using said data statistics to select target specific feature information to distinguish specific targets from background clutter;

(c) generating said target feature information from said data statistics;

(d) extracting said target specific feature information from said data;

(e) using said target specific feature information to distinguish specific targets from background clutter; and

(f) outputting target and background clutter information.

2. The method of claim **1**, wherein said data is one-dimensional, two-dimensional, or multi-dimensional.

3. The method of claim **1**, wherein said data is digitized imagery, speech signals, or radar signals.

4. The method of claim **1**, wherein parallel hardware architectures are employed to use said feature information to distinguish particular targets from background clutter.

5. The method of claim **4**, wherein said parallel hardware architectures comprise neural networks.

6. The method of claim **1**, wherein said feature information is contrast-based or texture-based.

7. The method of claim **1**, wherein said feature information is further used to distinguish two target classes or a plurality of target classes from one another.

8. The method of claim **1**, wherein Hebbian learning techniques are used to select said target specific feature information.

9. A classification system, comprising

(a) a calculator to calculate data statistics from a data input;

(b) a selector to use said data statistics outputed from said calculator to select target specific feature information to distinguish specific targets from background clutter, wherein said target specific feature information uniquely describes a target for recognition purposes;

(c) a generator to generate said target specific feature information from said data statistics outputed from said calculator;

(d) a feature extractor to extract target specific feature information from said data input; and

(e) a classifier to use said target specific feature information extracted from said data input by said feature extractor to distinguish targets from background clutter and output target and background clutter data.

10. The classification system of claim **9**, wherein said data input is one-dimensional, two-dimensional, or multidimensional.

11. The classification system of claim **9**, wherein said data input is digitized imagery, speech signals or radar signals.

12. The classification system of claim **9**, wherein parallel hardware architectures are employed in the classifier to use said feature information to distinguish particular targets from background clutter.

13. The classification system of claim **12**, wherein said parallel hardware architectures comprise neural networks.

14. The classification system of claim **9**, wherein said feature information is contrast-based or texture-based.

15. The classification system of claim **9**, wherein said selector employs Hebbian learning techniques to select said target specific feature information.

16. A classification hardware system, comprising

(a) a calculator to calculate data statistics from a data input;

(b) a selector to use said data statistics outputed from said calculator to select target specific feature information to distinguish specific targets from background clutter, wherein said target specific feature information uniquely describes a target for recognition purposes;

(c) a generator to generate said target specific feature information from said data statistics outputed from said calculator,

(d) a feature extractor to extract target specific feature information from said data input; and

(e) a classifier to use said target specific feature information extracted from said data input by said feature extractor to distinguish targets from background clutter and output target and background clutter data.

17. The classification hardware system of claim 16, wherein said selector employs Hebbian learning techniques to select said target specific feature information.

18. A classification software system, comprising

(a) calculation software to calculate data statistics from a data input;

(b) selection software to use said data statistics outputed from said calculator to select target specific feature information to distinguish specific targets from background clutter, wherein said target feature information uniquely describes a target for recognition purposes;

(c) generator software to generate said target specific feature information from said data statistics outputed from said calculator;

(d) feature extraction software to extract target specific feature information from said data input;

(e) classification software to use said target specific feature information extracted from said data input by said feature extractor to distinguish targets from background clutter and output target and background clutter data; and

(f) wherein the calculation software, selection software, generator software, feature extraction software, and classification software reside on at least one computer-readable medium.

19. The classification software system of claim 18, wherein said selector employs Hebbian learning techniques to select said target specific feature information.

20. A method for distinguishing targets from background clutter, comprising the steps of

(a) inputting data having targets and background clutter;

(b) applying neural network learning procedures to said data to construct a set of operators;

(c) applying said set of operators to data to construct a set of features;

(d) choosing optimal features from said set of features to distinguish targets from clutter; and

(e) using said optimal features to distinguishing targets from background clutter.

* * * * *

EXHIBIT 2

US007336805B2

(12) **United States Patent**
Gehring et al.

(10) **Patent No.: US 7,336,805 B2**
(45) **Date of Patent: Feb. 26, 2008**

(54) **DOCKING ASSISTANT**

(75) Inventors: **Ottmar Gehring**, Kemen (DE); **Harro Heilmann**, Ostfildern (DE); **Frederic Holzmann**, Stuttgart (DE); **Andreas Schwarzhaupt**, Landau (DE); **Gernot Spiegelberg**, Heimsheim (DE); **Armin Sulzmann**, Oftersheim (DE)

(73) Assignee: **DaimlerChrysler AG**, Stuttgart (DE)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 439 days.

(21) Appl. No.: **11/154,772**

(22) Filed: **Jun. 16, 2005**

(65) **Prior Publication Data**

US 2005/0281436 A1 Dec. 22, 2005

(30) **Foreign Application Priority Data**

Jun. 16, 2004 (DE) 10 2004 028 763

(51) **Int. Cl.**

G06K 9/00	(2006.01)
G06K 9/34	(2006.01)
G06K 9/36	(2006.01)
H04N 7/00	(2006.01)
H04N 7/18	(2006.01)
B60Q 1/48	(2006.01)
G08G 1/123	(2006.01)
G05D 1/00	(2006.01)

(52) **U.S. Cl.** **382/104**; 382/103; 382/113; 382/173; 382/181; 382/291; 348/116; 348/118; 348/148; 340/932.2; 340/933; 340/958; 340/995.25; 340/995.28; 701/1; 701/23; 701/300

(58) **Field of Classification Search** 382/103, 382/104, 113, 173, 181, 199, 203, 291; 340/932.2, 340/933, 935, 937, 958, 995.25–995.28; 348/113, 116, 118–120, 148–149; 701/1, 701/23–28, 207–212, 300–302

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,906,940	A *	3/1990	Greene et al.	382/100
4,931,937	A *	6/1990	Kakinami et al.	701/300

(Continued)

FOREIGN PATENT DOCUMENTS

DE	297 23 648 U1	6/1999
DE	101 41 464	3/2004
DE	103 23 915	2/2005
JP	2001-343212	12/2001
JP	2002-172988	6/2002
JP	2002-172989	6/2002

OTHER PUBLICATIONS

Robert Sedgewick, Algorithms in C, Chapter 4, pp. 35 to 49, Chapter 5, pp. 187-189, Chapter 5.6, pp. 230-249, Addison-Wesley Pub. Comp. Inc. 1998.

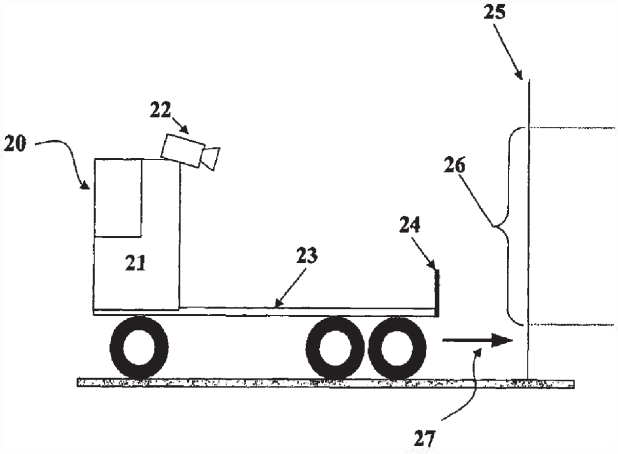
Paul J. Besl, A Method for Registration of 3-D Shapes, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14, No. 2, Feb. 1992, pp. 239-256.

Primary Examiner—Bhavesh M Mehta
Assistant Examiner—Manav Seth
(74) *Attorney, Agent, or Firm*—Davidson, Davidson & Kappel, LLC

(57) **ABSTRACT**

Many day-to-day driving situations require that an operator of a motor vehicle guide the motor vehicle along a specific course and bring the vehicle to a stop at a specific location, for example in a parking bay or at a loading platform. To assist a vehicle operator in such situations, a method and a suitable device for implementing this method, include detecting the potential target objects in the image data of an image sensor and identifying the potential target objects as potential destinations in a multi-stage exclusionary method, whereupon a trajectory describing an optimized travel path is computed at least in relation to the most proximate destination. By using the multi-stage exclusionary method according to the present invention, it is possible to reliably identify potential destinations in complex image scenarios solely on the basis of their geometric form, even when the destinations have not been encoded by specific symbols.

19 Claims, 3 Drawing Sheets



U.S. PATENT DOCUMENTS

4,970,653	A *	11/1990	Kenue	701/301
5,220,508	A *	6/1993	Ninomiya et al.	701/207
5,245,422	A *	9/1993	Borcherts et al.	348/119
5,351,044	A *	9/1994	Mathur et al.	340/901
5,386,285	A *	1/1995	Asayama	340/435
5,487,116	A *	1/1996	Nakano et al.	382/104
5,517,412	A *	5/1996	Unoura	701/23
5,555,312	A *	9/1996	Shima et al.	382/104
5,555,555	A *	9/1996	Sato et al.	382/104
5,612,686	A *	3/1997	Takano et al.	340/903
5,646,614	A *	7/1997	Abersfelder et al.	340/932.2
5,680,313	A *	10/1997	Whittaker et al.	701/300
5,790,403	A *	8/1998	Nakayama	701/28
5,832,116	A *	11/1998	Rezzouk	382/199
5,844,505	A *	12/1998	Van Ryzin	340/988
5,991,427	A *	11/1999	Kakinami et al.	382/104
6,172,601	B1 *	1/2001	Wada et al.	340/436
6,507,660	B1 *	1/2003	Wirtz et al.	382/103
6,744,380	B2 *	6/2004	Imanishi et al.	340/937
6,794,987	B2 *	9/2004	Schiffmann et al.	340/435
6,894,606	B2 *	5/2005	Forbes et al.	340/435
6,952,488	B2 *	10/2005	Kelly et al.	382/104
7,116,246	B2 *	10/2006	Winter et al.	340/932.2
7,209,221	B2 *	4/2007	Breed et al.	356/5.02
2002/0130953	A1	9/2002	Riconda et al.	348/115
2004/0056950	A1	3/2004	Takeda	348/92
2005/0002558	A1	1/2005	Franke et al.	382/154

* cited by examiner

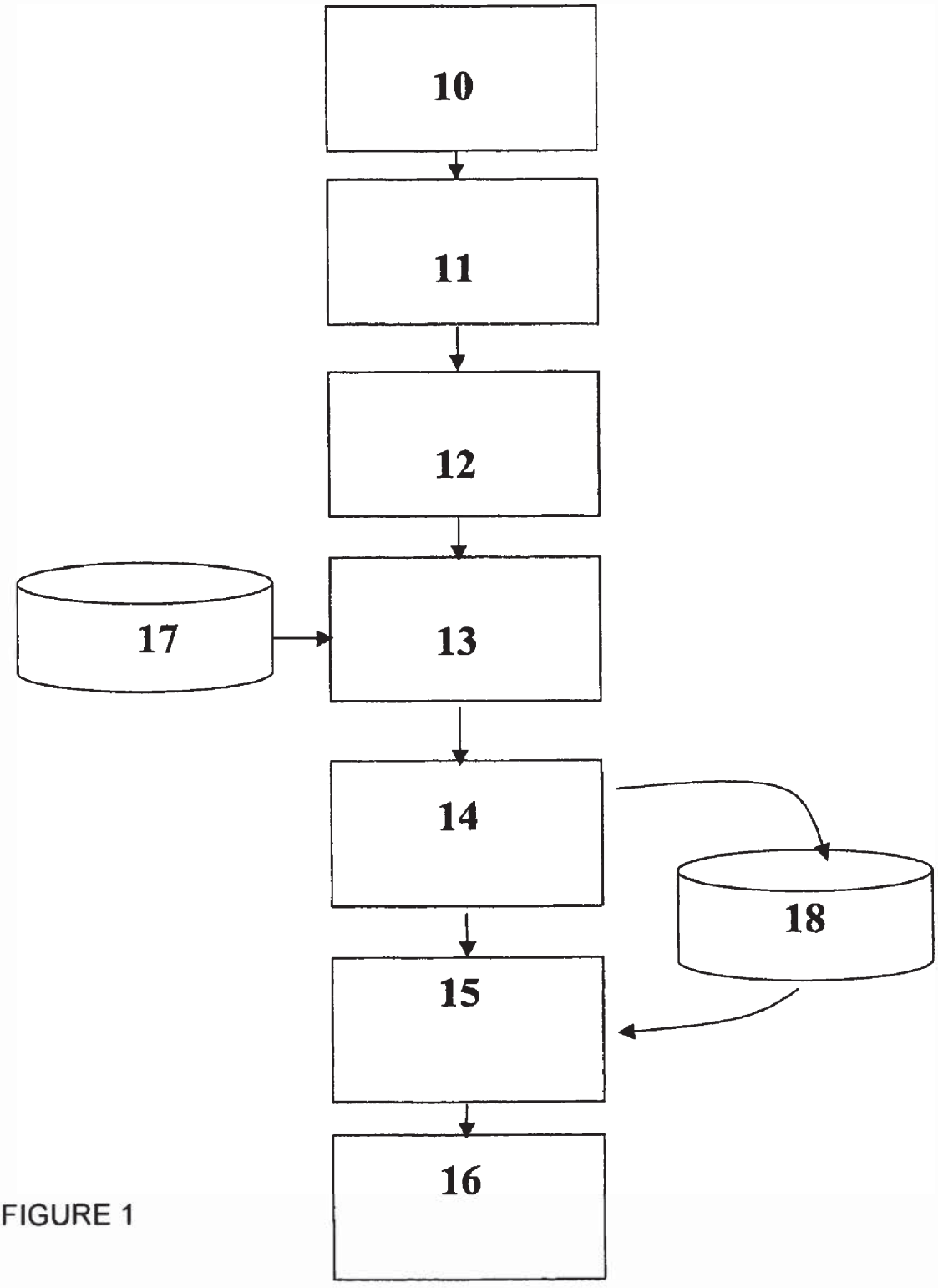


FIGURE 1

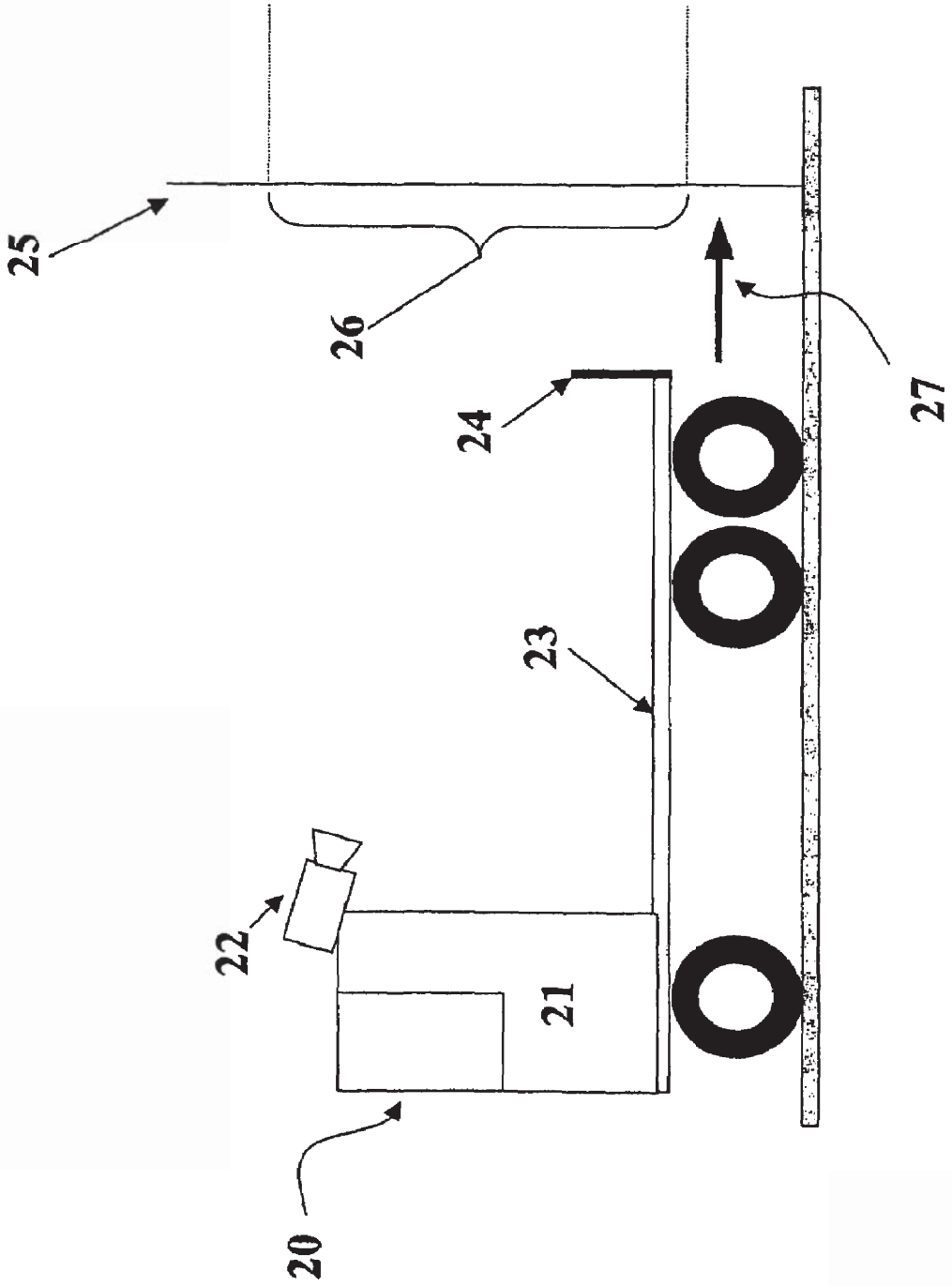


Figure 2

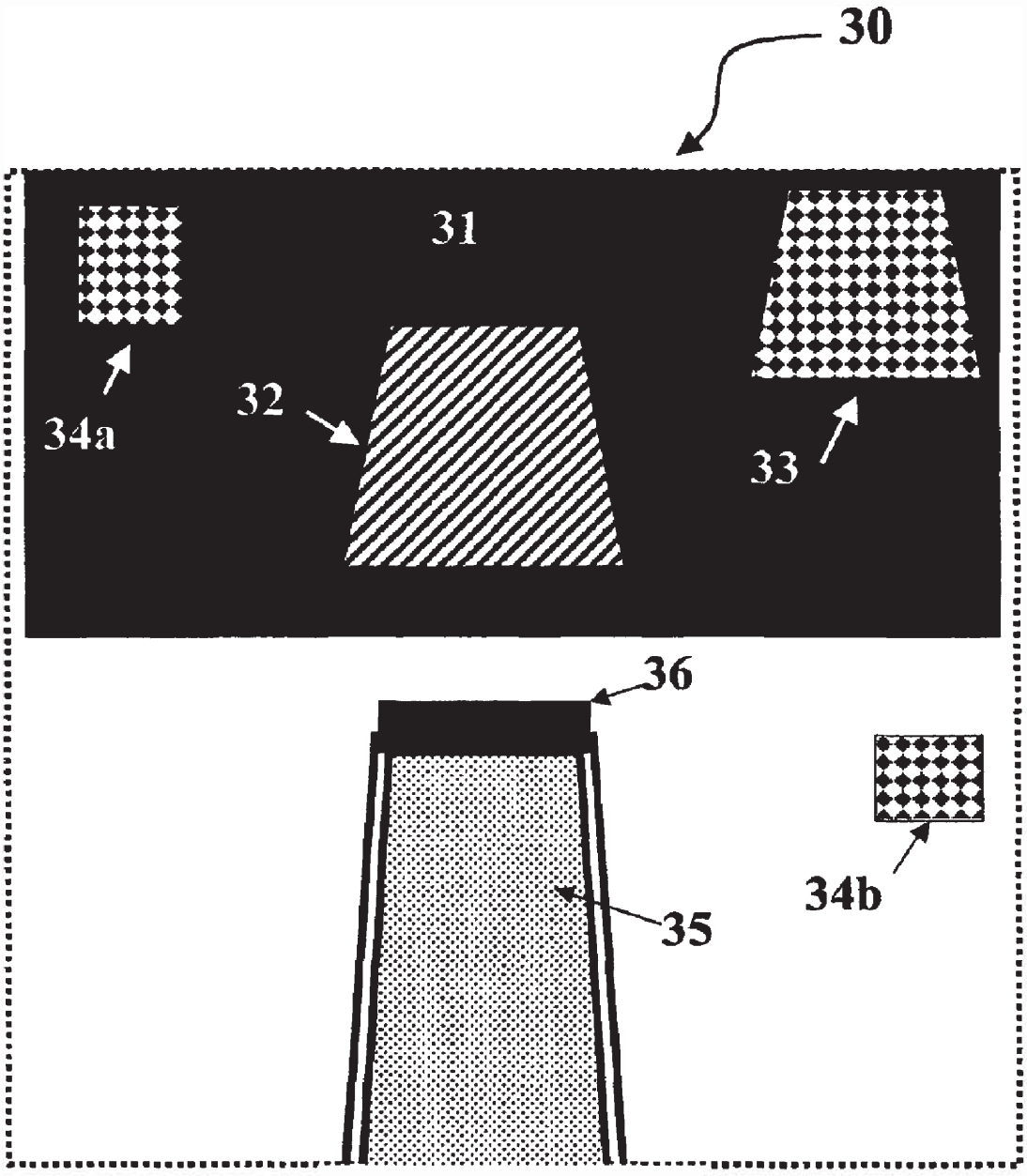


Figure 3

DOCKING ASSISTANT

Priority is claimed to German Patent Application No. DE 10 2004 028 763.5, filed on Jun. 16, 2004, the entire disclosure of which is incorporated by reference herein.

The present invention is directed to a method for assisting vehicle guidance on the basis of image data, particularly when maneuvering trucks toward docking stations, as well as to a device suited for implementing the method.

Many day-to-day driving situations require that an operator of a motor vehicle guide the motor vehicle along a specific course and bring the vehicle to a stop at a specific location, for example in a parking bay or at a loading platform. To assist a vehicle operator in such situations, the Japanese Patent Application JP 2001-343212 A describes a camera-based system for the guided entry into a parking bay whose boundaries are marked on the road surface. The system utilizes the fact that parking bays marked on the road surface are typically bounded on the right and left by clearly visible lines (lane marking signatures). In the image data acquired by the camera integrated in the vehicle, these visual signatures (boundary lines) are identified within an image-processing unit, and their orientation is measured. Since the visual signatures are parallel straight lines, they are also reproduced as straight lines in the camera image data, so that their angular deviation from the x- and y-axis of the camera image can easily be determined. From the angular deviations of both straight-line sections relative to each other and the knowledge of the distance that separates them, the vehicle's distance to the parking bay and its orientation relative to the same can be calculated in a geometrically simple manner. The image data are presented to the driver of the vehicle on a display, directional arrows being superposed on the display to indicate how far and in which direction the vehicle needs to be controlled in order to reach the parking bay.

Correspondingly, the Japanese Patent Applications JP 2002-172988 A and JP 2002-172989 A describe how, by using the image-processing system known from the Japanese Patent Application JP 2001-343212 A, an at least partially autonomous vehicle guidance into the parking bay can be carried out, in that the lane required for parking is precalculated. However, the evaluation of the image data for purposes of positional determination requires that the signatures (boundary lines) be clearly visible to enable their angular deviation in the image data to be ascertained. In particular, a correct positional determination requires that the starting points of the visual signatures on the lane be clearly ascertainable. In reality, however, this is not always possible, due to soiling of and wear to the lane markings.

A camera-based position detection and lane control system for motor vehicles that is rugged with respect to obscuration and soiling of the visual signatures, is described in the post-published German Patent Application DE 10323915.4. It discusses determining the position of the motor vehicle relative to a visual signature, which is used to mark the destination, by matching a template to camera image data acquired from the motor vehicle's surrounding field. This requires superposing the visual signature in the field surrounding the vehicle on a template of a visual signature stored in a memory. When the existing coordinate systems are known, the position of the motor vehicle relative to the visual signature can be directly inferred, in particular, from the compression and rotation parameters of this template matching. By using template matching for the problem at hand, one very advantageously utilizes the fact that this method is highly reliable, even when the visual signature in the image data is not fully visible due to obscuration or soiling. As do the other known related art methods for assisting the driver of a motor vehicle in approaching a destination, the system described in the German Patent Application DE 10323915.4 also requires that the destination be provided with a specific, previously known signature. In order for the visual signature to be uniquely identifiable in the image data, it must be conceived in a way that makes it clearly distinguishable from other visual signatures actually existing in the surrounding vicinity or from those that are only seemingly present.

SUMMARY OF THE INVENTION

An object of the present invention is to provide a method for assisting vehicle guidance on the basis of image data, particularly when maneuvering trucks toward docking stations, as well as to provide a device suited for implementing the method, which functions without the need for affixing specific visual signatures at the destination.

The present invention provides a method for assisting guidance of a motor vehicle on the basis of image data, particularly when maneuvering trucks toward docking stations, in which image data are acquired by an imaging sensor from the surrounding field of the motor vehicle; from the acquired image data, the positional parameters of at least one potential destination relative to the motor vehicle being extracted, and, as a consequence thereof, a trajectory describing an optimized travel path being calculated in order to assist a subsequent vehicle guidance for at least one of the potential destinations. According to the method, to extract the relative positional parameters of the at least one potential destination, the image data undergo an edge detection and edge segmentation, in order to break down the image data into individual edge segments, whose interrelationships are stored in a mathematical tree structure. Subsequently, these edge segments are then analyzed to check for the presence of a geometric object that is similar to a geometric form which, typically, at least partially describes a potential destination. The detected geometric objects that correspond to the typical geometric form are analyzed for plausibility using a matching algorithm. These plausible objects undergo an additional acceptance analysis to the effect that, based on the knowledge of the imaging properties of the imaging sensor in relation to its surrounding field, the shape of the image formation of the objects in the image data is analyzed. In addition, at least that object which is accepted in this manner and which corresponds to the most proximate destination is stored, along with the corresponding relative positional data, in an object list, and, to this end, at least one trajectory describing an optimized travel path is computed.

The present invention also provides a device for assisting a motor vehicle guidance on the basis of image data, in particular for maneuvering trucks toward docking stations. The device includes an imaging sensor (**10**, **22**) for acquiring image data from the field surrounding the motor vehicle; an image-processing unit (**11**) for extracting positional parameters of at least one potential destination relative to the motor vehicle from the image data; a processing unit (**15**) for computing at least one trajectory that describes the optimized travel path to one of the potential designations; and a system (**16**) for assisting in the vehicle guidance to one of the destinations. The image-processing unit (**11**) includes an edge detector and segmenter which extracts the relative positional parameters of the at least one potential destination from the image data. Connected downstream of the edge

detector and segmenter is a unit (**12**) for locating objects in the image data that have geometric shapes that typically correspond at least partially to the potential destination. In addition, the image-processing unit includes a comparator unit (**13**) for analyzing the detected geometric objects which correspond to the typical geometric form, to check for plausibility using a matching algorithm for making a comparison with object patterns stored in a memory unit (**17**). A unit for acceptance analysis (**14**) is provided, which, based on the comparison of selected geometric objects, analyzes the shape of the image formation of the objects in the image data, based on the knowledge of the imaging properties of the imaging sensor in relation to its surrounding field. The device communicates with a data memory (**18**) in which at least that geometric object detected in the image data which corresponds to the most proximate destination is stored, along with the corresponding relative positional data, in an object list, and then transfers these data to a processing unit (**15**) communicating herewith to calculate a trajectory describing an optimized travel path, this processing unit making these data available to a subsequent driver assistance system (**16**).

In the system according to the present invention for assisting a motor vehicle guidance on the basis of image data, in particular for maneuvering trucks toward docking stations, including a method and a device suited for implementing this method, image data are acquired by an imaging sensor from the surrounding field of a vehicle and, from these data, the positional parameters of at least one potential destination relative to the motor vehicle are extracted. This results in the calculation of a trajectory describing an optimized travel path in order to assist a subsequent vehicle guidance for at least one of the potential destinations. In this connection, along the lines of the present invention, to extract the relative positional parameters of the at least one potential destination, the image data undergo an edge detection and edge segmentation. To that end, the image data are broken down into individual edge segments, and their interrelationships are stored in a mathematical tree structure. In a subsequent step, these edge segments are then analyzed to check for the presence of a geometric object that is similar to a geometric form which, typically, at least partially describes a potential destination. If the typical destination is a docking station for trucks at a warehouse, for example, then the typical geometric form substantially corresponds to a rectangle having roughly identical side lengths of approximately 2.5 m. Within the framework of the inventive method, the detected geometric objects that correspond to the typical geometric form are then analyzed for plausibility using a matching algorithm, these objects, which are classified as plausible, undergoing an additional acceptance analysis to the effect that, based on the knowledge of the imaging properties of the imaging sensor in relation to its surrounding field, the shape of the image formation of the objects in the image data is analyzed (thus, an image of a typical rectangular geometric form viewed from an elevated location is formed in the image data as a trapezoid that tapers toward the top). In a final method step, at least that object which is determined within the scope of the acceptance analysis and which corresponds to the most proximate destination is stored, along with the corresponding relative positional data, in an object list, and, to this end, at least one trajectory describing an optimized travel path is computed.

In this manner, a driver assistance is able to be devised by using the device according to the present invention and on the basis of the method according to the present invention, which computes a trajectory describing an optimized travel path to at least the most proximate destination, based solely on the knowledge of the geometric appearance of the destination, without having to affix specific symbols or markings.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is elucidated in the following on the basis of advantageous exemplary embodiments and with the aid of the drawings, in which:

FIG. **1**: shows a block diagram describing the method sequence;

FIG. **2**: shows schematically, a camera system according to the present invention mounted on a truck; and

FIG. **3**: shows exemplarily, image data recorded using the camera system from FIG. **2**.

DETAILED DESCRIPTION

As becomes clear from FIG. **1**, the method according to the present invention is essentially sequential, in a first step, an image sensor (**10**) recording image data from the field surrounding a motor vehicle. In this connection, it is generally a question of a camera sensor which records image data in the visible light spectrum. However, it is equally conceivable that the image sensor (**20**) functions within an essentially invisible wavelength range, in particular in the infrared or in the ultraviolet wavelength range. Using such an image sensor makes it advantageously possible for the field surrounding the motor vehicle to be recorded to be actively illuminated by headlights which radiate light in this wavelength range, while objects or people in the area are exposed to a nonglare-type illumination. On the other hand, within the scope of the present invention, a millimeter wave radar or a lidar may be used as image sensor (**10**).

The image data generated by image sensor (**10**) are further processed in an image-processing unit (**11**), this unit including, in particular, an edge detector and segmenter, with whose assistance the image data are processed in such a way that, at least for one potential destination, the relative positional parameters are able to be extracted from the image data. In this context, based on the knowledge of their position and location, the extracted edges and segments are able to be advantageously stored in a hierarchically organized tree structure in the image data. Many different proven methods for creating and organizing such a hierarchical tree structure are available to one skilled in the art; reference is made here exemplarily to the comprehensive compilation and discussion of widely varying tree structures by Robert Sedgewick (R. Sedgewick, *Algorithms in C*, Chapter 4, Addision-Wesley Pub. Comp. Inc., 1990). Based on the knowledge of the geometric appearance of a destination, this hierarchical tree structure may, at this point, be processed in a unit (**12**) downstream from the image-processing unit (**11**), in order to find geometric shapes that typically correspond to the potential destination. This investigation to check for the existence of a geometric object in the image data that typically corresponds to a potential destination is performed on the basis of the hierarchical tree structure, to this end, a tree traversal algorithm known from the related art being able to be effectively used. In the case of a tree traversal algorithm, the individual branches are systematically processed, beginning with the start nodes of the tree (R. Sedgewick, *Algorithms in C*, Chapter 5.6, Addision-Wesley Pub. Comp, Inc., 1998). In this connection, based on the knowledge of the typical geometric shape of a destination, the edges and segments stored in parameters in the branches

of the hierarchical tree are analyzed to check if they are able to be pieced together to form geometric objects which describe suitable, typical destinations. If the destination has a square shape, for example, then only those edge elements or segments, whose adjoining lateral surfaces substantially form a right angle and whose distances to the respective opposing side pairs are approximately of the same value, are grouped together to form an object. By already observing such elementary rules when processing the hierarchical tree structure, the subsequent outlay for processing may be limited already at this stage of the inventive method to relatively few plausible objects which are actually similar to a potential destination. Selection rules of this kind are generally able to be very well defined within the framework of the tree traversal algorithm, since typical destinations are mostly man-made architectural objects, which, as a rule, have a rectangular, in particular square, or also, however, round shape.

The geometric objects pieced together by the tree traversal algorithm, in connection with the rules derived from the geometric shape of the destination, undergo an additional acceptance analysis. In this connection, the image data are compared in a comparator unit (**13**), with the aid of a matching algorithm, to object patterns describing the destination stored in a memory (**17**). As a matching algorithm, the IPC algorithm (Besl, P. J., McKay, N. D., *A Method for Registration of* 3-*D Shapes*, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 14, no. 2, 1992, pp. 224-231) is particularly advantageously suited. The iterative IPC algorithm makes it possible for the objects determined with the aid of the tree transversal algorithm to be scaled and rotated in a way that minimizes the quadratic error with respect to the object's deviation from the ideal object pattern of the destination. The distance to the potential destination is also able to be easily estimated from the parameters resulting from the scaling and orientation (in particular rotation). At this point, the iterative sequence of the IPC algorithm is briefly explained by way of example:

P denotes the position of the object pattern in space. e_i denotes the deviation between the stored object pattern and the segments ascertained by the tree traversal algorithm and grouped into an object. The iteration steps bear the designation n. For subsequent correction factor c_j, it must, therefore, follow that $P_{n+1}=P_n-c_j$. The Jacobi matrix J_{ij} corresponding thereto is defined in this connection as

$$J_{ij}=\frac{\partial e_i}{\partial c_j}.$$

By linear approximation, it follows that Jij*cj*=e_i. Within the scope of the iteration, the optimal vector, which describes the necessary scaling and rotation of the object determined by the tree traversal algorithm, must satisfy the equation

$$\|J\cdot c_j-e_i\|^2$$

In this connection, this IPC algorithm is stabilized by the iteration steps and exhibits an overcontrolled convergence.

The objects, which are scaled and oriented (rotated) in comparator unit (**13**) by the matching algorithm, subsequently undergo a further acceptance analysis in unit (**15**). In one especially advantageous type of acceptance analysis, in those cases in which the surrounding field is recorded by the sensor from an elevated position, those objects which do not exhibit any distorted images, as compared to their usual geometric shapes, are rejected as potential destinations. For example, the distortion of rectangular or square geometric shapes of potential destinations is manifested as a trapezoid in an image formation. In another effective type of acceptance analysis, based on knowledge of the customary position of destinations, objects not corresponding to these specifications in the image data are rejected. For example, if an object is included in the image data whose position relative to the imaging sensor is known, then, on the basis of its image formation in the image data, an artificial horizon may be generated, on whose basis the relative position of the recognized objects is ascertained. In this manner, objects, whose position deviates from the usual position of the destination, are excluded from the further processing.

Of the geometric objects corresponding to a potential destination that remain following the acceptance analysis, at least the parameters (in particular position and distance) of that object which corresponds to the most proximate destination are stored in a memory unit (**18**). At least for this object, the trajectory describing an optimized travel path to the potential destination may then be computed in a processing unit (**15**) which is linked to memory unit (**18**). The steering properties and capabilities of the motor vehicle are considered in an especially advantageous manner in the calculation of a trajectory describing an optimized travel path. To that end, the device according to the present invention should have a memory in which the data required for this purpose are stored. On the other hand, it would also be conceivable to provide the device with an input unit for inputting the parameters describing the steering properties and capabilities of the motor vehicle. This makes it easily possible to adapt the computing program to various designs (different body designs or trailers). In this computing program, within the framework of the calculation of the trajectory describing an optimized travel path, those potential destinations are also rejected which are not able to be reached in consideration of the given steering properties and capabilities of the motor vehicle.

The thus calculated trajectories may then be made available to subsequent driver assistance systems (**16**). In this connection, the method according to the present invention is very advantageously conceived in such a way that the driver of the motor vehicle is informed about the position of at least one of the potential destinations and, in particular, about the course of the trajectory which is computed for that purpose and which describes an optimized travel path. This information may be provided by showing the ideal trajectory on a display; ideally, in this connection, the trajectory of the planned lane being superposed symbolically on camera image data representing the driver's field of view.

As a matter of course, the vehicle guidance may subsequently be carried out automatically or at least semi-automatically, on the basis of the computed trajectory to the nearest destination. It would be conceivable, for example, when working with a semi-automatic or manual vehicle guidance, for a deviation of the vehicle from the optimal trajectory describing the travel path to be indicated by audible or visual signaling means. This makes it possible, in a simple manner, for the driver of the vehicle to be assisted in observing the lateral guidance.

In the course of the lateral guidance, it is particularly advantageous for the trajectory to be constantly readapted with the aid of a Kalman filter and on the basis of a continuous image analysis (temporal updating), in order to substantially eliminate interference effects in the image data.

It is also intended, during the vehicle guidance, for the position of the objects stored as potential destinations in the object list to be continuously monitored, so that, in the event that the originally most proximate destination can no longer be easily reached due to circumstances, that potential destination is selected as a new, most proximate destination, which, on the basis of the active position of the motor vehicle, may best be reached in consideration of its steering properties and capabilities. If a new destination is thus ascertained to be the nearest destination to the vehicle, then, on the basis of the active position, a trajectory describing an optimized travel path to this destination must be computed.

FIG. **2** shows exemplarily a typical situation of a truck (**20**) docking at a cargo door (**26**, **32**) of a warehouse (**25**, **31**). In this case, imaging sensor (**22**) is mounted above driver's cabin (**21**) of truck (**20**). Image sensor (**22**) is aligned in such a way that, in travel direction (**27**) of truck (**20**), it is able to record its loading platform (**23**, **35**) as well as its loading platform end region (**24**, **36**), as well as warehouse (**25**, **31**) and cargo door (**26**, **32**) situated therein. Image data (**30**) resulting from a corresponding image data acquisition are shown schematically in FIG. **3**. Due to the elevated position of image sensor (**22**), a cargo door (**26**, **32**) having a rectangular geometry is produced as a trapezoidal image in the image data. It is assumed exemplarily that, after processing of the image data in image-processing unit (**11**) and of the tree traversal algorithm in unit (**12**), objects **32**, **33**, **34***a* and **34***b* are identified as geometric objects typically corresponding to a potential destination. It is assumed that object **32** corresponds to the image of cargo door (**26**) sought as a destination. Object **33** corresponds to another rectangular structure, for example a window of warehouse (**25**, **31**). Apparent objects **34***a* and **34***b* do not represent images of real objects, but rather result from interference in the image data which coincidentally produces edge sections and segments of a kind that results in the tree traversal algorithm piecing them together to form a potential destination object. These objects would then be scaled and oriented in comparator unit (**13**) by a matching algorithm. However, based on the knowledge of the camera position and camera imaging geometry, objects **34***a* and **34***b* would be eliminated within the framework of the downstream unit for acceptance analysis, since they do not have a trapezoidal shape. In addition, knowing its position relative to the camera position, the image formation of loading platform end region (**36**) may be retrieved during the acceptance analysis to produce an artifical horizon. On the basis of this artificial horizon, object **33** may then also be eliminated, since it is situated in a position (far above the horizon) that is abnormal for the destination (cargo door). After executing the individual method steps according to the present invention, solely object **32** remains as a potential destination in this example, so that its parameters may be stored in destination memory (**18**) and be retrieved for computing the trajectory describing an optimized travel path.

What is claimed is:

1. A method for assisting guidance of a motor vehicle on the basis of image data, the method comprising:

acquired image data using an imaging sensor from a surrounding field of the motor vehicle;

extracting from the acquired image data positional parameters of at least one potential destination relative to the motor vehicle; and

calculating at least one trajectory describing an optimized travel path using the positional parameters so as to assist a subsequent vehicle guidance for at least one of the potential destinations,

wherein the extracting includes:

performing an edge detection and edge segmentation on the image data so as to break down the image data into a plurality of edge segments and storing interrelationships of the plurality of edge segments in a mathematical tree structure;

analyzing the plurality of edge segments for the presence of a geometric object associated with a geometrical form that may at least partially describe a potential destination so as to identify one or more potential geometric objects;

analyzing the one or more potential geometric objects for plausibility using a matching algorithm so as to identify one or more plausible objects;

performing an additional acceptance analysis so as to identify at least one accepted object, the additional acceptance analysis including analyzing a shape of the image formation of each object in the image data based on knowledge of at least one imaging property of the imaging sensor relative to the surrounding field; and

storing at least one accepted object that corresponds to a most proximate destination together with corresponding relative positional data in an object list.

2. The method as recited in claim **1**, further comprising processing the mathematical tree structure using a tree traversal algorithm.

3. The method as recited in claim **1**, wherein the potential geometric objects include at least one of rectangles and squares.

4. The method as recited in claim **1**, wherein the matching algorithm includes an IPC algorithm.

5. The method as recited in claim **1**, wherein, when the image sensor is disposed in an elevated position an image of a rectangular or square potential destination is formed as a trapezoid in the image data, and the acceptance analysis rejects as potential destinations those objects that do not exhibit a distorted image a typical geometric shape.

6. The method as recited in claim **1**, wherein the calculation of the at least one trajectory includes considering a steering property and a capability of the motor vehicle.

7. The method as recited in claim **6**, further comprising rejecting a potential destination that is not reachable given the steering property and capability of the motor vehicle.

8. The method as recited in claim **1**, further comprising informing a driver of the motor vehicle about a position of at least one potential destination and about the at least one trajectory.

9. The method as recited in claim **1**, further comprising performing an at least semi-automatic vehicle guidance to the most proximate destination using the at least one trajectory.

10. The method as recited in claim **9**, further comprising using at least one of an audible and a visual signaling during the vehicle guidance to assist a driver of the vehicle in observing the trajectory.

11. The method as recited in claim **9**, further comprising:

continuously monitoring a position of the stored at least one accepted object during the vehicle guidance;

selecting a new most proximate destination went the most proximate destination is no longer easily reachable, wherein the new most proximate destination is selected based on an active position of the motor vehicle and in determining the at least one accepted object that is best reachable, considering a steering properties and a capability of the vehicle; and

computing a new trajectory describing a new optimized travel path based on the active position.

12. The method as recited in claim 1, wherein the motor vehicle is a truck and wherein the surrounding field includes a docking station.

13. A device for assisting guidance of a motor vehicle using image data, the device comprising:

an imaging sensor configured to acquire image data from a field surrounding the motor vehicle;

an image-processing unit configured to extract positional parameters of at least one potential destination relative to the motor vehicle from the image data;

a processing unit configured to compute at least one trajectory describing an optimized travel path to a potential designation; and

a vehicle guidance system configured to assist in guiding the vehicle to the potential destination,

wherein the image-processing unit includes:

an edge detector and segmenter for extracting relative positional parameters of the at least one potential destination from the image data;

a locating unit disposed downstream from the image processing unit and configured to locate corresponding objects in the image data having geometric shapes corresponding at least partially to the at least potential destination;

a memory unit for storing object patterns;

a comparator unit configured to analyze the corresponding geometric objects, to check for plausibility using a matching algorithm for comparing the object patterns stored in the memory unit;

a data memory storing the geometric object detected in the image data corresponding to the most proximate destination is stored, along with the corresponding relative positional data, in an object list;

an acceptance analysis unit configured to analyze a shape of an image formation of the objects in the image data based on the knowledge of the imaging properties of the imaging sensor in relation to its surrounding field, to communicate with the data memory, and to transfer the most proximate destination to the processing unit.

14. The device as recited in claim 13, wherein the imaging sensor includes at least one of a camera, a laser scanner and a millimeter wave radar.

15. The device as recited in claim 13, wherein the imaging sensor is disposed at an elevated position on the vehicle from where the imaging sensor to observe a part of the field surrounding the motor vehicle relevant to the driving task.

16. The device as recited in claim 15, wherein the imaging sensor is mounted in such a way that an image of at least one portion of the motor vehicle is included in the image data.

17. The device as recited in claim 13, wherein the vehicle guidance system communicates with at least one of a steering, an acceleration and a braking equipment of the motor vehicle so as to be able to act on the equipment at least semi-autonomously.

18. The device as recited in claim 13, wherein the vehicle guidance system includes at least one of a visual and an audible output for communicating with a driver of the motor vehicle.

19. The device as recited in claim 13, wherein the motor vehicle is a truck and the surrounding field includes a docking station.

* * * * *

EXHIBIT 3

US008898395B1

(12) **United States Patent**
Rozas

(10) **Patent No.: US 8,898,395 B1**
(45) **Date of Patent: *Nov. 25, 2014**

(54) **MEMORY MANAGEMENT FOR CACHE CONSISTENCY**

(76) Inventor: **Guillermo J. Rozas**, Los Gatos, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 466 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **12/121,531**

(22) Filed: **May 15, 2008**

Related U.S. Application Data

(63) Continuation of application No. 11/102,127, filed on Apr. 7, 2005, now Pat. No. 7,376,798.

(51) **Int. Cl.**
G06F 12/00 (2006.01)
G06F 12/08 (2006.01)

(52) **U.S. Cl.**
CPC ***G06F 12/0815*** (2013.01); *G06F 12/0806* (2013.01)
USPC **711/144**; 711/141; 711/143; 711/145; 711/130; 711/133; 710/22; 714/11; 714/12

(58) **Field of Classification Search**
None
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,086,429	A *	2/1992	Gray et al.	714/13
5,428,761	A	6/1995	Herlihy et al.	
5,701,432	A	12/1997	Wong et al.	
5,974,438	A	10/1999	Neufeld	
6,006,299	A	12/1999	Wang et al.	
6,189,074	B1	2/2001	Pedneau	
6,263,407	B1	7/2001	Arimilli et al.	
6,304,944	B1	10/2001	Pedneau	
6,625,694	B2	9/2003	Masri et al.	
6,658,536	B1	12/2003	Arimilli et al.	
6,738,864	B2	5/2004	Chauvel	
6,779,085	B2	8/2004	Chauvel	
6,839,813	B2	1/2005	Chauvel	
6,938,130	B2	8/2005	Jacobson et al.	
2006/0112226	A1 *	5/2006	Hady et al.	711/130

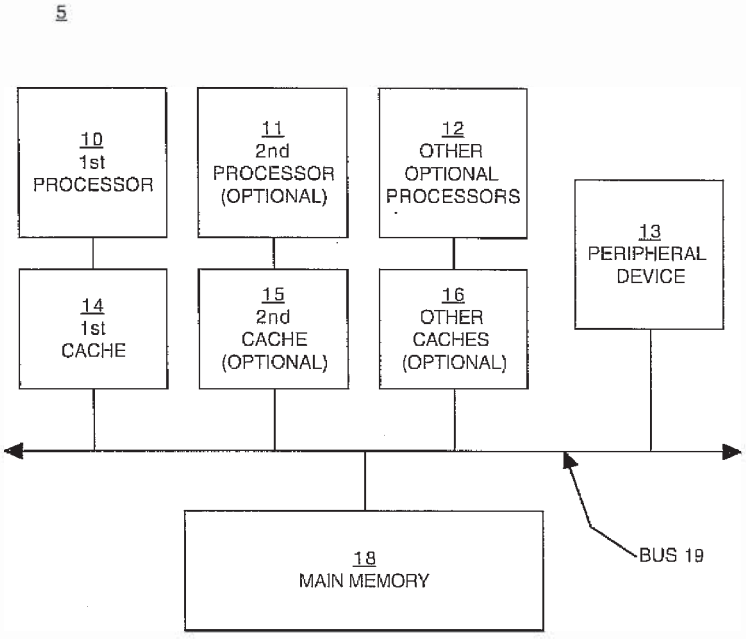
* cited by examiner

Primary Examiner — Matthew Bradley
Assistant Examiner — Alan Otto

(57) **ABSTRACT**

Methods and systems for maintaining cache consistency are described. A group of instructions is executed. The group of instructions can include multiple memory operations, and also includes an instruction that when executed causes a cache line to be accessed. In response to execution of that instruction, an indicator associated with the group of instructions is updated to indicate that the cache line has been accessed. The cache line is indicated as having been accessed until execution of the group of instructions is ended.

20 Claims, 5 Drawing Sheets



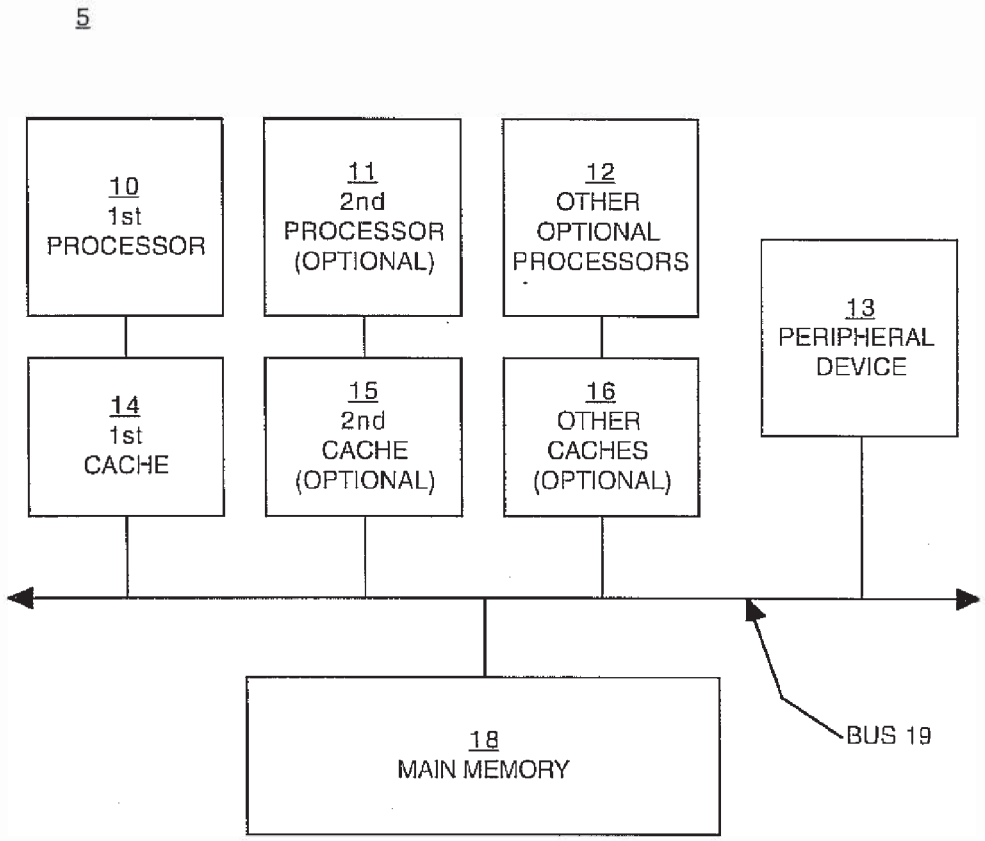

5
10
1st
PROCESSOR
11
2nd
PROCESSOR
(OPTIONAL)
12
OTHER
OPTIONAL
PROCESSORS
13
PERIPHERAL
DEVICE
14
1st
CACHE
15
2nd
CACHE
(OPTIONAL)
16
OTHER
CACHES
(OPTIONAL)
18
MAIN MEMORY
BUS 19

Figure 1

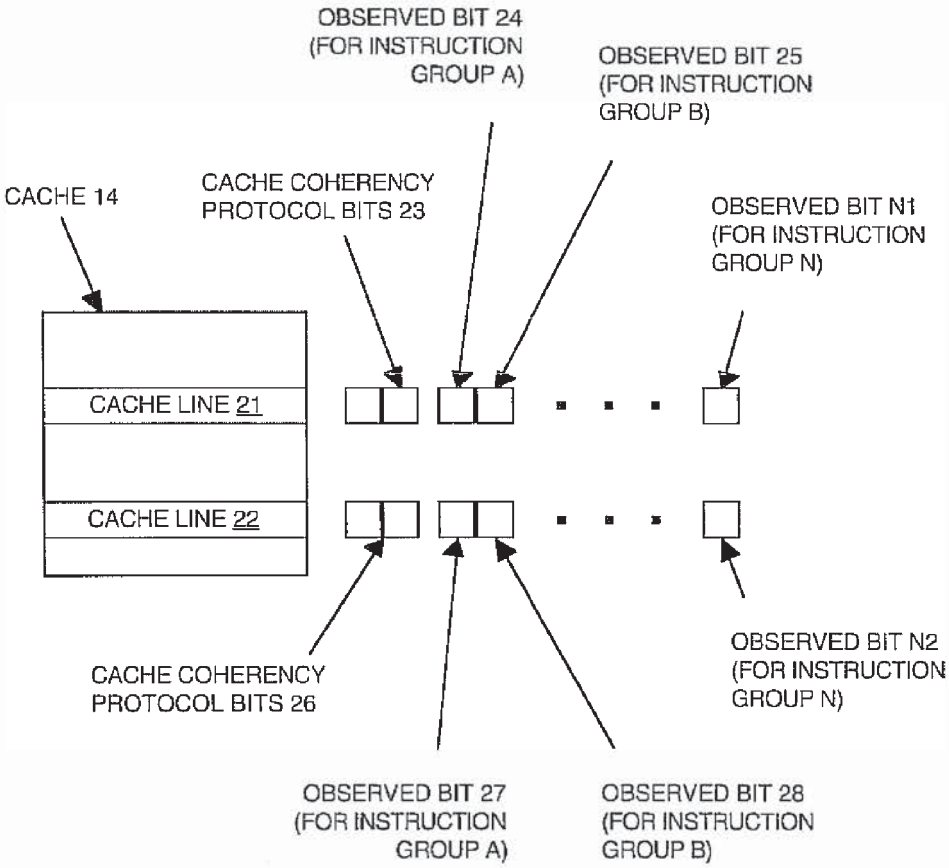

OBSERVED BIT 24
(FOR INSTRUCTION
GROUP A)
OBSERVED BIT 25
(FOR INSTRUCTION
GROUP B)
CACHE 14
CACHE COHERENCY
PROTOCOL BITS 23
OBSERVED BIT N1
(FOR INSTRUCTION
GROUP N)
CACHE LINE 21
CACHE LINE 22
CACHE COHERENCY
PROTOCOL BITS 26
OBSERVED BIT N2
(FOR INSTRUCTION
GROUP N)
OBSERVED BIT 27
(FOR INSTRUCTION
GROUP A)
OBSERVED BIT 28
(FOR INSTRUCTION
GROUP B)

Figure 2

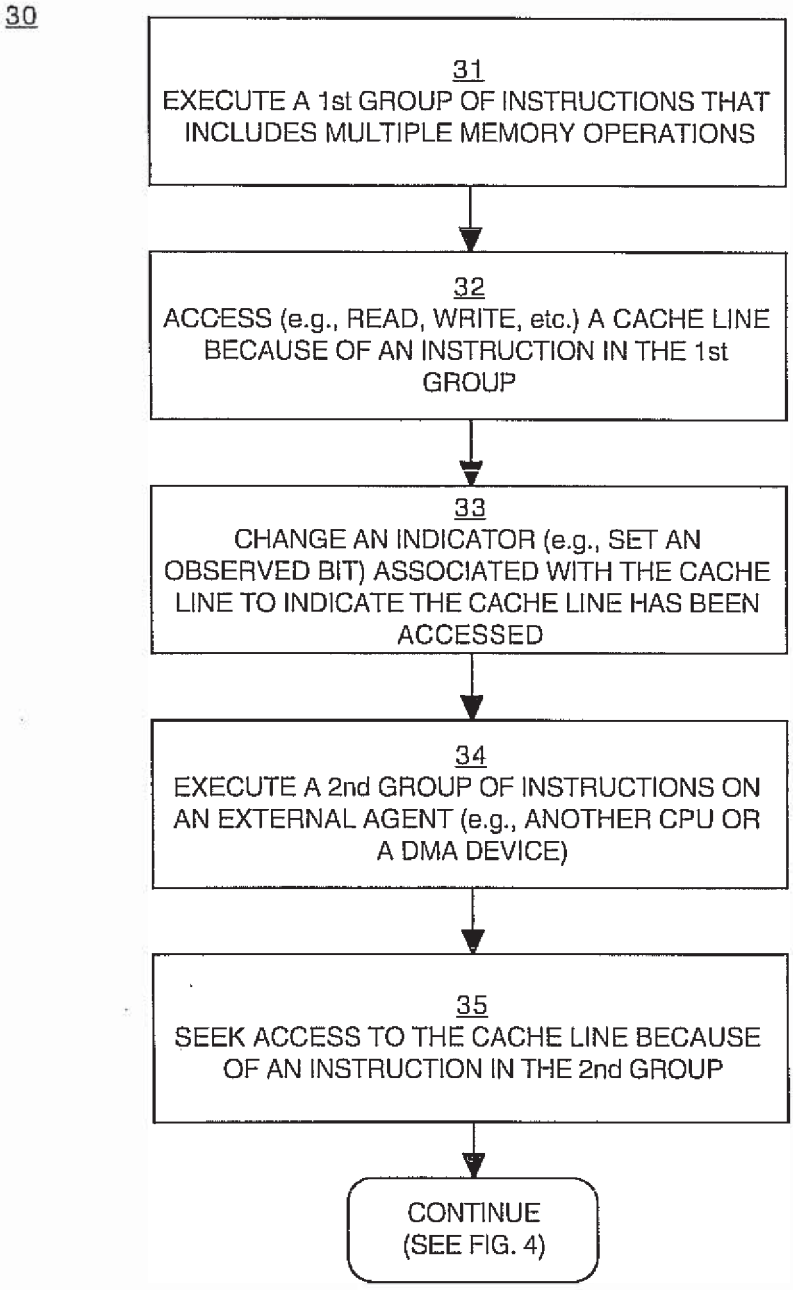

30
31
EXECUTE A 1st GROUP OF INSTRUCTIONS THAT INCLUDES MULTIPLE MEMORY OPERATIONS
32
ACCESS (e.g., READ, WRITE, etc.) A CACHE LINE BECAUSE OF AN INSTRUCTION IN THE 1st GROUP
33
CHANGE AN INDICATOR (e.g., SET AN OBSERVED BIT) ASSOCIATED WITH THE CACHE LINE TO INDICATE THE CACHE LINE HAS BEEN ACCESSED
34
EXECUTE A 2nd GROUP OF INSTRUCTIONS ON AN EXTERNAL AGENT (e.g., ANOTHER CPU OR A DMA DEVICE)
35
SEEK ACCESS TO THE CACHE LINE BECAUSE OF AN INSTRUCTION IN THE 2nd GROUP
CONTINUE
(SEE FIG. 4)

Figure 3

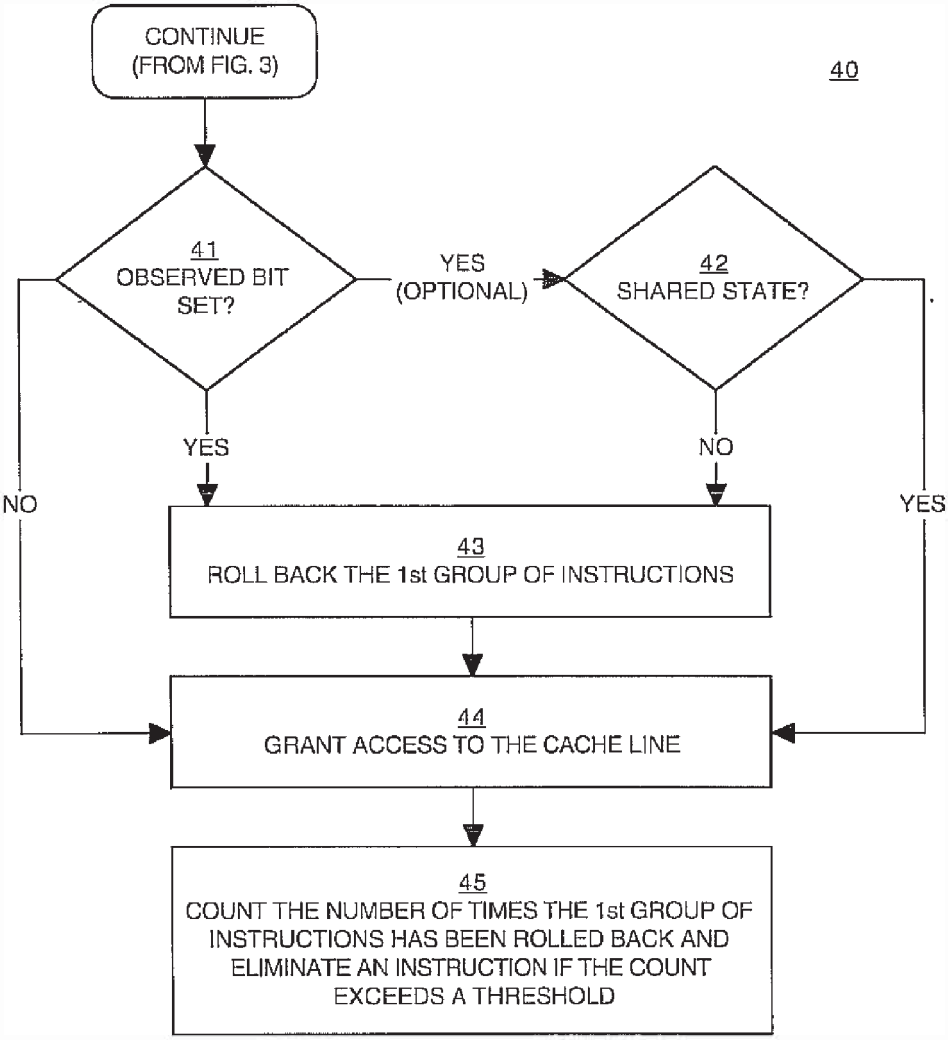

CONTINUE
(FROM FIG. 3)
40
41
OBSERVED BIT SET?
YES
(OPTIONAL)
42
SHARED STATE?
YES
NO
NO
YES
43
ROLL BACK THE 1st GROUP OF INSTRUCTIONS
44
GRANT ACCESS TO THE CACHE LINE
45
COUNT THE NUMBER OF TIMES THE 1st GROUP OF INSTRUCTIONS HAS BEEN ROLLED BACK AND ELIMINATE AN INSTRUCTION IF THE COUNT EXCEEDS A THRESHOLD

Figure 4

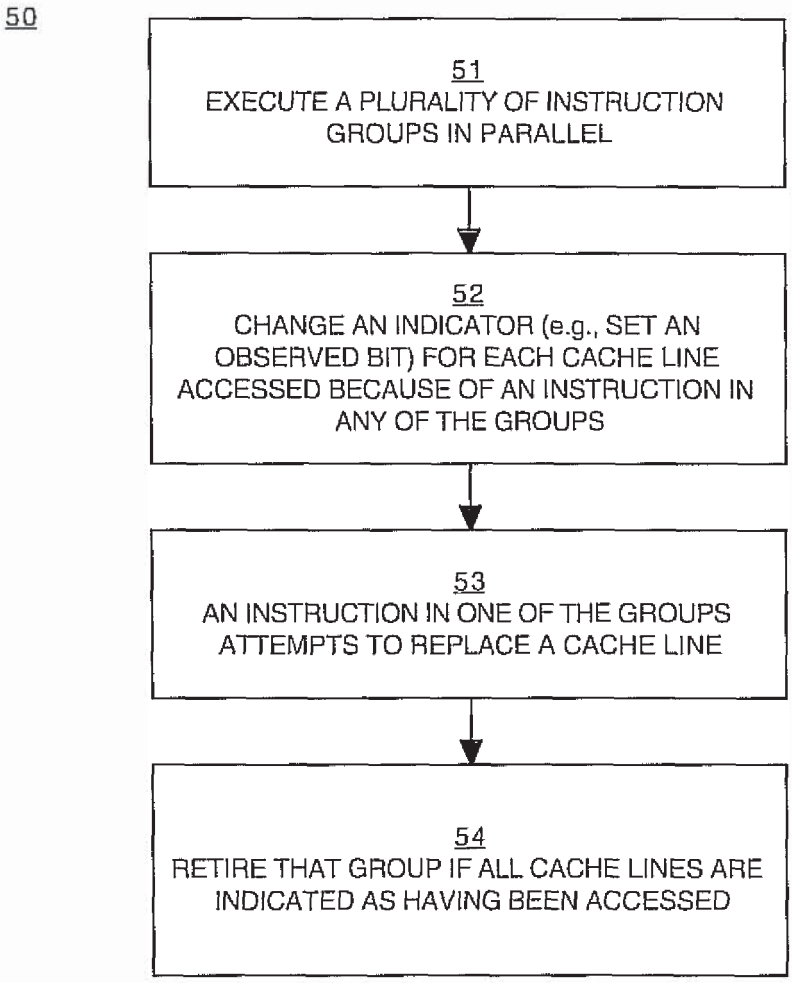

50
51
EXECUTE A PLURALITY OF INSTRUCTION GROUPS IN PARALLEL
52
CHANGE AN INDICATOR (e.g., SET AN OBSERVED BIT) FOR EACH CACHE LINE ACCESSED BECAUSE OF AN INSTRUCTION IN ANY OF THE GROUPS
53
AN INSTRUCTION IN ONE OF THE GROUPS ATTEMPTS TO REPLACE A CACHE LINE
54
RETIRE THAT GROUP IF ALL CACHE LINES ARE INDICATED AS HAVING BEEN ACCESSED

Figure 5

MEMORY MANAGEMENT FOR CACHE CONSISTENCY

RELATED UNITED STATES PATENT APPLICATION

This application is a Continuation Application of the co-pending, commonly-owned U.S. patent application Ser. No. 11/102,127, filed Apr. 7, 2005, by G. J. Rozas, and entitled "Memory Management Methods and Systems That Support Cache Consistency."

BACKGROUND OF THE INVENTION

1. Field of the Invention

Embodiments of the present invention relate to computer system memory, in particular the management of cache memory.

2. Related Art

With direct memory access (DMA), an input/output (I/O) system can issue read requests and writes directly to main memory without passing through the central processing unit (CPU). However, if the I/O system uses DMA to write to main memory, and changes data cached previously by the CPU, then the CPU will not receive the new data unless the CPU fetches the data from main memory. Also, for DMA reads, the CPU cache may contain more recent data than main memory, and so the I/O system will not receive the new data unless it reads the cache instead of main memory. Multiprocessor systems, particularly systems referred to as shared-memory simultaneous multiprocessor (SMP) architectures, have to deal with similar types of scenarios. The MESI (Modified, Exclusive, Shared, Invalid) protocol is a popular cache consistency (coherency) protocol that addresses these issues. A modification of the MESI protocol is the MOESI (Modified, Owned, Exclusive, Shared, Invalid) protocol. These protocols are known in the art.

Constraints are also applied to memory operations to prevent processors and DMA systems from reordering memory operations at will. If, for example, each processor could reorder memory operations for optimization, code sequences that work on a single processor would not work on multiprocessor systems. One type of constraint can be referred to as sequential consistency. With sequential consistency, the legal orders of memory operations are those that are indistinguishable from strict interleaving of the operations from each thread of control. For example, for two threads, the operations of one thread can be interleaved with those of the other thread, but the order of operations within each thread is preserved.

There are different classes of high-level processor architectures with regard to operation reordering, optimization and speculation. One such class can be referred to as a lumped in-order architecture, and another such class as a lumped out-of-order architecture. With lumped in-order architectures, instructions are lumped into instruction groups that can be committed and rolled back atomically. Different instruction groups are committed sequentially and in order, but within an instruction group, arbitrary reordering and optimization can occur. Full speculation is possible within an instruction group, but speculation, reordering and optimization across instruction groups is limited. With lumped out-of-order architectures, instructions are lumped into instruction groups that can be committed and rolled back atomically. Different groups can execute out of order but are committed in order.

SUMMARY OF THE INVENTION

A method or system that can maintain sequential consistency for lumped (in-order or out-of-order) architectures would be advantageous. Prior attempts to achieve a sequentially consistent, lumped architecture are limited with respect to the number of memory operations that can be included in an instruction group. Specifically, such attempts are limited to a single memory operation per instruction group. Embodiments in accordance with the present invention overcome this disadvantage.

According to one embodiment of the present invention, a group of instructions is executed. The group of instructions can include multiple memory operations, and also includes an instruction that when executed causes a cache line to be accessed (e.g., read, write, store, load, etc.). In response to execution of that instruction, an indicator associated with the group of instructions is updated to indicate that the cache line has been accessed. The cache line is indicated as having been accessed until execution of the group of instructions is ended. If an external agent (e.g., another processor or a DMA system) snoops the cache and the cache line (or any other cache line in the cache, for that matter) is indicated as having been accessed, then the instruction group is rolled back and reissued. If an external agent (e.g., another processor or a DMA system) snoops the cache and no cache line is indicated as having been accessed, then the snoop can be processed using a conventional cache coherency protocol (MESI, for example).

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and form a part of this specification, illustrate embodiments of the present invention and, together with the description, serve to explain the principles of the invention. The drawings referred to in this description should not be understood as being drawn to scale except if specifically noted.

FIG. **1** is a block diagram of a portion of a computer system upon which embodiments of the present invention can be implemented.

FIG. **2** is a block diagram of a cache and showing cache tags that are used according to one embodiment of the present invention.

FIG. **3** is a flowchart of a method for managing memory in accordance with one embodiment of the present invention.

FIG. **4** is a flowchart that continues from the flowchart of FIG. **3**.

FIG. **5** is a flowchart of a method for replacing a cache line in accordance with one embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

Reference will now be made in detail to the various embodiments of the invention, examples of which are illustrated in the accompanying drawings. While the invention will be described in conjunction with these embodiments, it will be understood that they are not intended to limit the invention to these embodiments. On the contrary, the invention is intended to cover alternatives, modifications and equivalents, which may be included within the spirit and scope of the invention as defined by the appended claims. Furthermore, in the following detailed description of the present invention, numerous specific details are set forth in order to provide a thorough understanding of the present invention. However, it will be recognized by one of ordinary skill in the art that the present invention may be practiced without these specific details. In other instances, well-known methods, procedures, components, and circuits have not been described in detail as not to unnecessarily obscure aspects of the present invention.

Some portions of the detailed descriptions that follow are presented in terms of procedures, logic blocks, processing, and other symbolic representations of operations on data bits within a computer memory. These descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. A procedure, logic block, process, etc., is here, and generally, conceived to be a self-consistent sequence of steps or instructions leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated in a computer system. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, bytes, values, elements, symbols, characters, terms, numbers, or the like.

It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the following discussions, it is appreciated that throughout the present invention, discussions utilizing terms such as "executing," "changing," "using," "accessing," "handling," "rolling back," "maintaining," "eliminating," "setting," "clearing," "incrementing," "decrementing," "associating," "granting," "determining" or the like, refer to the action and processes (e.g., flowcharts **30**, **40** and **50** of FIGS. **3**, **4** and **5**, respectively) of a computer system or similar intelligent electronic computing device (generally, a controller), that manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

FIG. **1** is a block diagram of a portion of a computer system **5** upon which embodiments of the present invention can be implemented. Computer system **5** can include elements in addition to those illustrated in FIG. **1**. Generally speaking, embodiments in accordance with the present invention pertain to a computer system (or similar electronic device) that includes a processor, a main memory, a cache associated with the processor, and one or more external agents that can access the main memory and the processor's cache.

In the example of FIG. **1**, computer system **5** includes a first processor **10**, an optional second processor **11**, and other optional processors **12** that are coupled to main memory **18** via bus **19**. In the present embodiment, computer system **5** also includes a peripheral device **13** coupled to main memory **18** via bus **19**. For ease of discussion, the processors **10**, **11** and **12** and the peripheral device **13** may be collectively referred to herein as agents **10-13**. Relative to processor **10**, processors **11** and **12** and peripheral device **13** may be referred to as external agents; relative to processor **11**, processors **10** and **12** and peripheral device **13** may be referred to as external agents; and so on.

In one embodiment, peripheral device **13** is capable of direct memory access (DMA). In general, DMA refers to a hardware circuit or software agent that can transfer data directly to or from main memory **18**.

Associated with the processors **10-12** are first cache **14**, optional second cache **15**, and other optional caches **16**, respectively. In one embodiment, each of the processors **10-12** has their own cache. The caches may be physically incorporated into the processors **10-12**, or they may be physically external to the processors **10-12**. In the present embodiment, each of the processors **10-12** can access each of the caches **14**, **15** and **16** via bus **19**.

In general, the agents **10-13** may each be executing a group of instructions according to a lumped in-order architecture or according to a lumped out-of-order architecture. The group of instructions can each include one or more memory operations (e.g., read, load, store, write, lock acquisition, etc.). The group of instructions can be demarcated by commit points (e.g., a commit ending one group of instructions can mark the beginning of the next group of instructions, which in turn ends at another commit). An instruction group is handled atomically (e.g., it is treated as a unit). An instruction (or instructions) in the group of instructions being executed by one of the agents **10-13** may seek access to a cache associated with any of the other agents **10-13**.

In one embodiment, computer system **5** implements a cache coherency protocol such as, but not limited to, MESI (Modified, Exclusive, Shared, Invalid). According to the MESI protocol, a cache line is identified as being in one of the modified, exclusive, shared or invalid states. In the modified state, the cache line includes more recent information than main memory **18** (only the information in the cache line is valid because main memory **18** is not up to date), and no other cache holds the information in the cache line. In the exclusive state, the information in the cache line is not in any other cache and the information in the cache line is unmodified, and accordingly main memory **18** is up to date. In the shared state, the information in the cache line may be in one or more other caches, and main memory **18** is up to date. In the invalid state, the cache line does not contain valid information. MESI is consistent with sequential consistency.

FIG. **2** is a block diagram of cache **14** according to one embodiment of the present invention. Cache **14** includes a number of cache lines exemplified by cache lines **21** and **22**.

In the embodiment of FIG. **2**, a cache line tag is associated with each of the cache lines in cache **14**. In one such embodiment, each cache line tag includes a number of cache coherency protocol bits. For example, cache coherency protocol bits **23** are associated with cache line **21**, and cache coherency protocol bits **26** are associated with cache line **22**. The cache coherency protocol bits **23** and **26** can be used to indicate the state of the cache lines **21** and **22**, respectively, according to a cache coherency protocol such as the MESI protocol.

Furthermore, in the present embodiment, each cache line tag includes a number of additional bits referred to herein as observed bits for ease of discussion. For example, observed bits **24**, **25**, . . . , N**1** are associated with cache line **21**, and observed bits **27**, **28**, . . . , N**2** are associated with cache line **22**. The observed bits for cache **14** are associated with groups of instructions executed by the agent (e.g., processor **10**) that is associated with cache **14**. In one embodiment, each cache line tag is extended by one observed bit for each group of instructions that may be executed by processor **10** (FIG. **1**) at the same time. If, for example, processor **10** can execute five (5) instruction groups in parallel, then the cache line tags are each extended by 5 observed bits.

In the example of FIG. **2**, observed bits **24** and **27** (associated with cache line **21**) are associated with instruction group A, and observed bits **25** and **28** (associated with cache line **22**) are associated with instruction group B.

In one embodiment, an observed bit is set (e.g., set to a value of one) for a cache line in a cache if execution of an instruction in the group of instructions executed by the agent associated with that cache causes that cache line to be accessed (e.g., load, store, read, or write). For example, in one embodiment, observed bit **24** is set if an instruction in instruction group A, when executed by processor **10**, causes cache

line **21** to be accessed, because cache line **21** is in cache **14** associated with processor **10** (FIG. **1**). In such an embodiment, observed bit **27** (also associated with instruction group A) is not set unless cache line **22** is also accessed because of instruction group A.

In general, with reference to FIGS. **1** and **2**, a group of instructions executed by processor **10** may result in multiple cache lines in cache **14** being accessed. For each cache line in cache **14** accessed because of a group of instructions executed by processor **10**, the observed bit corresponding to that group of instructions is set. However, an observed bit associated with a cache line in cache **14** is not set if an instruction executed on an external agent **11-13** causes access (e.g., a snoop or interrogation, or perhaps a read) to that cache line. For example, in one embodiment, an instruction executed by an agent other than processor **10** that causes cache line **21** to be accessed will not set an observed bit for cache line **21**.

In one embodiment, when execution of a group of instructions is ended (e.g., the instruction group is committed, rolled back, or aborted), each observed bit set by that group of instructions is cleared (e.g., unset, or set to a value of zero). For example, if instruction group A is committed (or retired), rolled back or aborted, any of the observed bits **24** and **27** that are set are cleared. From another perspective, the observed bits associated with instruction group A, if set, remain set until the execution of instruction group A is ended.

To summarize, in one embodiment, when an agent (e.g., processor **10**) begins to execute a group of instructions, it allocates an observed bit for that group of instructions to each cache line in the cache associated with the agent (e.g., cache **14** associated with processor **10**). If a cache line (e.g., cache line **21**) has been accessed because of an instruction in an instruction group (e.g., instruction group A) being executed by processor **10**, then the observed bit for that cache line and corresponding to instruction group A (e.g., observed bit **24**) is set. If another cache line (e.g., cache line **22**) has also been accessed because of an instruction in instruction group A, then the observed bit for that cache line and corresponding to instruction group A (e.g., observed bit **27**) is set. When execution of an instruction group is ended, the observed bits associated with that instruction group are cleared, and thus the observed bits are available for subsequent instruction groups executed by the agent. For example, if instruction group A is committed, rolled back or aborted, then observed bits **24** and **27** are cleared.

In one embodiment, the observed bits are used in the following manner. Continuing with reference to FIGS. **1** and **2**, if an external agent (e.g., another processor **11** or **12** or a DMA agent such as peripheral device **13**) executes an instruction that causes access to (or a request for access to) cache line **21**, for example, and any of the observed bits **24**, **25**, . . . , N**1** have been set, then the instruction groups associated with the observed bits that are set are forced to roll back and then reissue. Also, any of the observed bits **24**, **25**, . . . , N**1** associated with cache line **21** that have been set are cleared. For example, if observed bits **24** and **25** have been set, then instruction group A (which corresponds to observed bit **24**) and instruction group B (which corresponds to observed bit **25**) are rolled back, and observed bits **24** and **25** are cleared. The rollback also clears the observed bits associated with other cache lines in cache **14** if those observed bits correspond to the groups of instructions that are rolled back. For example, if the instruction groups A and B are rolled back as mentioned above, then observed bits **27** and **28** are also cleared if those observed bits are set. The cache state transition occurs at this point, before the instruction groups A and B are reissued and are able to again cause cache line **21** (and perhaps cache line **22**) to be accessed.

In general, if an external agent **11-13** accesses, or requests access to, a cache line in cache **14**, and an observed bit associated with that cache line has been set, then all the observed bits that have been set for that cache line are cleared, all the instruction groups corresponding to the observed bits that are cleared are rolled back, and the observed bits associated with the instruction groups that are rolled back are also cleared for other cache lines in cache **14**.

In one embodiment, the rollbacks just described occur only if an external agent **11-13** accesses, or requests access to, a cache line in cache **14**, an observed bit associated with that cache line has been set, and the cache line is not in the shared state. Again consider an example in which an external agent (e.g., another processor **11** or **12** or a DMA agent such as peripheral device **13**) executes an instruction that seeks access to cache line **21**. In one embodiment, in addition to checking the observed bits **24**, **25**, . . . , N**1** associated with cache line **21**, the cache coherency protocol bits **23** are also checked to determine the state (e.g., the MESI state) of cache line **21**. In a MESI embodiment, for example, if cache line **21** is in the shared state, and if the access from the external agent is for sharing (e.g., it is not exclusive), then access to cache line **21** can be granted without rolling back the instruction groups associated with the observed bits **24**, **25**, . . . , N**1** that are set.

In one embodiment, observed bits that are set are not forgotten (e.g., cleared) except in the case of a rollback, commit or abort. In such an embodiment, a cache line that has at least one of its observed bits set is not replaceable, although the cache line and its set observed bit(s) can be moved to another location. This can lead to a situation in which a cache line is needed, but no cache lines are available (that is, all cache lines have at least one of their respective bits set, so none of the cache lines can be replaced or moved). In one embodiment, if such a situation arises, the group of instructions that needs the cache line, as well as following instruction groups, is rolled back and then reissued.

To prevent instruction groups from being repeatedly rolled back, in one embodiment, the number of times each instruction group is rolled back is counted and compared to a threshold. In such an embodiment, if the number of times an instruction group is rolled back exceeds the threshold, then the instruction group is rolled back for the last time. If it fails to execute to completion in the next attempt, the instruction group is reformed into a new instruction group that is reduced in size (e.g., one or more instructions is removed from the instruction group), and the counter for the new instruction group is set to zero. In the limiting case, the instruction group is eventually reduced to a single instruction that will be able to complete execution.

FIGS. **3** and **4** are flowcharts **30** and **40** of a method for managing memory in accordance with one embodiment of the present invention (the flowchart of FIG. **4** is a continuation of the flowchart of FIG. **3**). Although specific steps are disclosed in flowcharts **30** and **40**, such steps are exemplary. That is, embodiments of the present invention are well-suited to performing various other steps or variations of the steps recited in flowcharts **30** and **40**. It is appreciated that the steps in flowcharts **30** and **40** may be performed in an order different than presented, and that not all of the steps in flowcharts **30** and **40** may be performed. Some or all of the processes and steps of flowcharts **30** and **40** can be realized, in one embodiment, as a series of instructions that reside in hardware or

software (e.g., in main memory **18** of FIG. **1**) and are executed by one or more of the agents **10-13** in a computer system **5** (FIG. **1**).

In step **31** of FIG. **3**, a first group of instructions (e.g., instruction group A) is executed (e.g., by processor **10** of FIG. **1**). In one embodiment, the first group of instructions includes multiple memory operations (e.g., reads, writes, loads, stores, lock acquisition, etc.). At least one of the instructions in the first group of instructions, when executed, causes a cache line (e.g., cache line **21** of FIG. **2**) in a cache associated with processor **10** (e.g., cache **14**) to be accessed (e.g., read, written, etc.).

In step **32** of FIG. **3**, and continuing with reference to FIG. **2**, that instruction is executed, causing cache line **21** to be accessed.

In step **33** of FIG. **3**, with reference also to FIG. **2**, an indicator associated with the first group of instructions is changed, in order to indicate that cache line **21** has been accessed because of the first group of instructions. In one embodiment, the indicator is also associated with cache line **21**. In one embodiment, a bit (e.g., observed bit **24**) associated with cache line **21** and the first group of instructions is set. In one embodiment, cache line **21** is indicated as having been accessed (that is, for example, the observed bit remains set) until execution of the first group of instructions is ended (e.g., because of a roll back, commit or abort).

In step **34** of FIG. **3**, with reference also to FIG. **1**, a second group of instructions is executed by an agent external to processor **10** (e.g., processor **11** or **12** or a DMA agent such as peripheral device **13**). At least one of the instructions in the second group of instructions, when executed, causes (or will cause) the cache line of interest (e.g., cache line **21** of FIG. **2**) to be accessed.

In step **35** of FIG. **3**, access to the cache line of interest (e.g., cache line **21** of FIG. **2**) is sought because of the second group of instructions. As will be seen, the first group of instructions is then handled depending on whether or not cache line **21** has been accessed because of the first group of instructions and is identified as having been accessed (e.g., by an observed bit). That is, it is possible that cache line **21** is accessed because of the second group of instructions before it is accessed because of the first group of instructions.

In step **41** of FIG. **4**, with reference also to FIG. **2**, a determination is made with regard to whether or not the cache line of interest (e.g., cache line **21**) has been accessed because of the first group of instructions. In one embodiment, if the observed bit (e.g., observed bit **24**) associated with cache line **21** and the first instruction group is set, then flowchart **40** proceeds to step **43**, and otherwise flowchart **44** proceeds to step **44**. In another embodiment, if the observed bit (e.g., observed bit **24**) associated with cache line **21** and the first instruction group is set, then flowchart **40** proceeds to step **42**.

In step **42** of FIG. **4**, and with reference to FIG. **2**, a determination is made with regard to whether or not cache line **21** is in the shared state (e.g., according to the MESI protocol). If cache line **21** is in the shared state, then flowchart **40** proceeds to step **44**, and otherwise flowchart **40** proceeds to step **43**.

In step **43** of FIG. **4**, with reference also to FIG. **2**, the first group of instructions is atomically rolled back, and the observed bits **24**, **25**, . . . , N1 associated with cache line **21** are cleared. Other observed bits, associated with other cache lines in cache **14** and corresponding to the first group of instructions, are also cleared if they are set. For example, if observed bit **27** (which also corresponds to the first group of instructions) is set, then it is also cleared by the rollback of the first group of instructions. Also, other groups of instructions associated with any of the observed bits of cache **14** that are cleared are also rolled back. For example, if observed bit **25** is set, then it is cleared and the instruction group associated with observed bit **25** (e.g., instruction group B) is rolled back, which in turn results in observed bit **28** (which also corresponds to instruction group B) of cache line **22** being cleared.

In step **44** of FIG. **4**, access to cache line **21** (FIG. **2**) is granted.

In step **45** of FIG. **4**, in one embodiment, the number of times that the first group of instructions has been rolled back is counted. If the number of times the first group of instructions has been rolled back exceeds a threshold, then one or more instructions are eliminated from the first instruction group. Specifically, in one embodiment, the first group of instructions is reissued, and if the next execution attempt results in another rollback, then the first instruction group is reduced by one or more instructions.

FIG. **5** is a flowchart **50** of a method for replacing a cache line in accordance with one embodiment of the present invention. Although specific steps are disclosed in flowchart **50**, such steps are exemplary. That is, embodiments of the present invention are well-suited to performing various other steps or variations of the steps recited in flowchart **50**. It is appreciated that the steps in flowchart **50** may be performed in an order different than presented, and that not all of the steps in flowchart **50** may be performed. Some or all of the processes and steps of flowchart **50** can be realized, in one embodiment, as a series of instructions that reside in hardware or software (e.g., in main memory **18** of FIG. **1**) and are executed by one or more of the agents **10-13** in a computer system **5** (FIG. **1**).

In step **51** of FIG. **5**, a number of instruction groups are executed in parallel using processor **10** (FIG. **1**), for example.

In step **52** of FIG. **5**, with reference also to FIGS. **1** and **2**, each cache line in the cache associated with processor **10** (e.g., cache **14**) that is accessed because of the groups of instructions being executed in step **51** is indicated as being accessed. For example, as previously described herein, an observed bit can be set for each cache line accessed because of the groups of instructions of step **51**.

In step **53** of FIG. **5**, one of the instructions in one of the instruction groups attempts to replace a cache line.

In step **54**, in one embodiment, if all of the cache lines in the cache of interest (e.g., cache **14**) are indicated as having been accessed (e.g., at least one observed bit is set for each of the cache lines in cache **14**, then the instruction group that has the instruction that attempted the cache line replacement (the instruction group of step **53**) is rolled back.

In summary, methods and systems that can maintain sequential consistency for lumped (in-order or out-of-order) architectures are described. Significantly, embodiments in accordance with the present invention permit sequentially consistent, lumped architectures in which multiple memory operations can be included in instruction groups.

Embodiments in accordance with the present invention can incorporate variations to the features described above. For systems with large speculation ability (e.g., potentially 20 or more active instruction groups), the cost (e.g., the storage cost) of the associated number of observed bits can be reduced in various ways. For example, a protocol can be established in which any external snoop causes all active instruction groups to roll back and reissue after the cache line state transition. In effect, such a protocol acts as if there is a single observed bit that is always set, except during the period of inactivity after rollback, and so no additional storage is needed for the observed bit. Alternatively, instead of an observed bit per cache line per instruction group, there can be an observed bit per instruction group.

Furthermore, new hardware structures can be added to track which cache lines have been accessed, without the storage of data. These hardware structures can be direct-mapped, set-associative, or fully-associative and of varying capacity.

Also, instead of having an observed bit per cache line per instruction group, a smaller memory structure (e.g., another cache, or a memory structure within main memory) that covers the whole cache address space and that has an observed bit per entry and per instruction group can be used. Each entry covers some portion of the cache address space, in such a way that every cache line address maps uniquely to an entry, but an entry maps to many cache lines. For example, some subset of the address bits of a cache line can be used to identify an entry. The subset of the address bits are the address summary. When accessing a cache line, the entry corresponding to the address summary for the cache line is used to set the observed bits. Similarly, on a snoop request, the address is "summarized," and the entry corresponding to the summary is examined for the presence of observed bits. The observed bits are cleared when an instruction group is committed, rolled back, or aborted. To illustrate, consider a cache complex where each cache holds 32 bytes, so that, for example, the bottom five bits (bits **4-0**) of an address indicate bytes within the same cache line. Furthermore, consider a 32-entry observed bit structure, where bits **9-5** of the cache line address are used to select an entry, and this entry holds the observed bits. Then, any two addresses that have the same bits **9-5** will use the same entry in the summary structure; that is, they collide. This can mean that when a snoop comes, an instruction group that did not really conflict may be rolled back, because there is a conflict/collision in the summary structure even though there would not have been a conflict/collision in the cache. Thus, there is a tradeoff between the size of the summary structure and the overhead incurred by the "spurious" rollbacks due to collisions.

Embodiments in accordance with the present invention are thus described. While the present invention has been described in particular embodiments, it should be appreciated that the present invention should not be construed as limited by such embodiments, but rather construed according to the below claims.

What is claimed is:

1. A method of managing memory in a computer system, said method comprising:

setting a first bit of an indicator associated with a cache line in a cache memory if said cache line has been accessed in response to a processor executing an instruction in a first group of instructions;

setting a second bit of said indicator while said first bit remains set if said cache line has also been accessed in response to said processor executing an instruction in a second group of instructions;

executing a third group of instructions that causes said cache line to be accessed, wherein said third group of instructions is executed by an agent other than said processor, wherein said processing comprises rolling back said first group of instructions provided said first bit is set and rolling back said second group of instructions provided said second bit is set before allowing an instruction in said third group to access said cache line, and otherwise granting said access; and

processing said first group and said second group of instructions according to a value of said indicator.

2. The method of claim **1** wherein said processing further comprises:

determining a state of said cache line, wherein said state is specified according to a cache coherency protocol comprising at least a modified state, a shared state, and an invalid state; and

rolling back said first group of instructions provided said first bit is set and rolling back said second group of instructions provided said second bit is set, and then granting said access provided said cache line is in other than said shared state, and otherwise granting said access only when said cache line is in said shared state.

3. The method of claim **1** further comprising maintaining a first count of how many times said first group of instructions has been rolled back and a second count of how many times said second group of instructions has been rolled back.

4. The method of claim **3** further comprising eliminating one or more instructions from said first group of instructions if said first count exceeds a threshold and eliminating one or more instructions from said second group of instructions if said second count exceeds said threshold.

5. The method of claim **1** further comprising clearing said first bit when said execution of said first group of instructions is ended, wherein said second bit remains set until execution of said second group of instructions is ended.

6. The method of claim **1** further comprising:

executing a plurality of groups of instructions in parallel;

changing an indicator associated with each cache line in said cache that has been accessed because of an instruction in any of said plurality of groups of instructions; and

provided all cache lines in said cache memory are indicated as having been accessed, rolling back a group of instructions that includes an instruction that when executed attempts to replace information in a cache line of said cache memory, and otherwise continuing said executing of said plurality of groups of instructions.

7. A computer system comprising:

a processor;

a cache memory for use by said processor; and

a memory unit coupled to said processor and having stored therein instructions, said instructions comprising:

instructions to execute a first group of instructions using said processor, wherein said first group of instructions includes an instruction that causes a cache line of said cache memory to be read;

instructions to set a first bit of an indicator associated with said cache line to indicate that said cache line has been read;

instructions to execute a second group of instructions using said processor, wherein said second group of instructions includes an instruction that causes said cache line to be read;

instructions to set a second bit of said indicator while said first bit remains set to indicate said cache line has been accessed by both said first group of instructions and said second group of instructions;

instructions to execute a third group of instructions that causes said cache line to be accessed, wherein said third group of instructions is for execution by an agent other than said processor, wherein said instructions further comprise instructions to roll back said first group of instructions provided said first bit is set and to roll back said second group of instructions provided said second bit is set before allowing an instruction in said third group to access said cache line, wherein otherwise said access is granted; and

instructions to process said first group and said second group of instructions according to a value of said indicator.

8. The computer system of claim 7 wherein said instructions further comprise:

instructions to determine a state of said cache line, wherein said state is specified according to a cache coherency protocol comprising at least a modified state, a shared state, and an invalid state; and

instructions to roll back said first group of instructions provided said first bit is set and to roll back said second group of instructions provided said second bit is set, wherein said access is granted provided said cache line is in other than said shared state, and wherein otherwise said access is granted only when said cache line is in said shared state.

9. The computer system of claim 8 wherein said instructions further comprise instructions to maintain a first count of how many times said first group of instructions has been rolled back and a second count of how many times said second group of instructions has been rolled back.

10. The computer system of claim 9 wherein said instructions further comprise instructions to eliminate one or more instructions from said first group of instructions if said first count exceeds a threshold and to eliminate one or more instructions from said second group of instructions if said second count exceeds said threshold.

11. The computer system of claim 7 wherein said instructions further comprise instructions to clear said first bit when said execution of said first group of instructions is ended, wherein said second bit remains set until execution of said second group of instructions is ended.

12. The computer system of claim 7 wherein said instructions further comprise:

instructions to execute a plurality of groups of instructions in parallel using said processor;

instructions to change an indicator associated with each cache line in said cache memory that has been accessed because of an instruction in any of said plurality of groups of instructions; and

instructions to roll back a group of instructions that includes an instruction that when executed attempts to replace information in a cache line of said cache memory provided all cache lines in said cache memory are indicated as having been accessed, wherein otherwise said plurality of groups of instructions continue to be executed.

13. A method of managing shared memory in a computer system, said method comprising:

executing a first group of instructions with a processor;

associating a first state with a cache line in a cache comprising a plurality of cache lines, wherein said first state is specified according to a cache coherency protocol comprising at least a modified state, a shared state, and an invalid state;

associating a second state with said cache line during said executing of said first group of instructions, wherein said second state indicates that said cache line has been accessed because of said first group of instructions;

executing a second group of instructions with said processor;

associating a third state with said cache line during said executing of said second group of instructions, wherein said third state indicates that said cache line has been accessed because of said second group of instructions;

executing a third group of instructions that causes said cache line to be accessed; and

rolling back said first group of instructions and said second group of instructions before granting said access.

14. The method of claim 13 wherein said first group of instructions includes multiple memory operations and said second group of instructions includes multiple memory operations.

15. The method of claim 13 wherein said rolling back is only performed provided said cache line is in other than said shared state.

16. The method of claim 13 further comprising maintaining a first count of how many times said first group of instructions is rolled back and a second count of how many times said second group of instructions is rolled back.

17. The method of claim 16 further comprising eliminating one or more instructions from said first group of instructions if said first count exceeds a threshold and eliminating one or more instructions from said second group of instructions if said second count exceeds said threshold.

18. The method of claim 13 wherein said second state is indicated by setting a first bit associated with said cache line, wherein said first bit is cleared when said execution of said first group of instructions is ended, and wherein said third state is indicated by setting a second bit associated with said cache line, wherein said second bit is cleared when said execution of said second group of instructions is ended.

19. A computer system comprising:

a processor;

a cache memory coupled to said processor and having a plurality of cache lines; and

a memory unit coupled to said processor, wherein said memory unit has a plurality of indicator bits per cache line, wherein a first indicator bit associated with a cache line of said cache memory is set to indicate that said cache line has been accessed responsive to said processor executing a first group of instructions and wherein said first indicator bit is cleared when execution of said first group of instructions is ended, wherein a second indicator bit associated with said cache line is set to indicate that said cache line has been accessed responsive to said processor executing a second group of instructions in parallel with said first group and wherein said second indicator bit is cleared when execution of said second group of instructions is ended;

wherein further, in response to an agent other than said processor executing a third group of instructions that causes said cache line to be accessed, said first group of instructions is rolled back provided said first bit is set and said second group of instructions is rolled back provided said second bit is set before allowing an instruction in said third group to access said cache line, wherein otherwise said access is granted.

20. The computer system of claim 19 wherein a subset of address bits of said cache line is used to identify an entry in said memory unit, wherein an indicator bit of said entry is set in response to said cache line being accessed.

* * * * *

EXHIBIT 4

US010136416B2

(12) **United States Patent**
Beale

(10) **Patent No.: US 10,136,416 B2**
(45) **Date of Patent: Nov. 20, 2018**

(54) **COMMUNICATING ON A SHARED CHANNEL IN A WIRELESS NETWORK**

(71) Applicant: **Intellectual Ventures Holding 81 LLC**, Las Vegas, NV (US)

(72) Inventor: **Martin Warwick Beale**, Bristol (GB)

(73) Assignee: **Intellectual Ventures Holding 81 LLC**, Las Vegas, NV (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **15/611,309**

(22) Filed: **Jun. 1, 2017**

(65) **Prior Publication Data**

US 2018/0027539 A1 Jan. 25, 2018

Related U.S. Application Data

(63) Continuation of application No. 13/212,867, filed on Aug. 18, 2011, now Pat. No. 9,674,818, which is a continuation of application No. 11/863,205, filed on Sep. 27, 2007, now Pat. No. 8,027,291.

(51) **Int. Cl.**
H04W 72/04 (2009.01)
H04W 48/10 (2009.01)

(52) **U.S. Cl.**
CPC ***H04W 72/0406*** (2013.01); *H04W 48/10* (2013.01)

(58) **Field of Classification Search**
CPC H04J 3/16; H04J 3/1688; H04J 3/17; H04J 3/24
USPC 370/322, 326, 328–331, 335–338, 370/342–345, 348
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

6,111,865	A	8/2000	Butler et al.
6,504,827	B1	1/2003	Kuo et al.
7,072,329	B2	7/2006	Willars et al.
7,088,697	B1	8/2006	Benz et al.
7,113,496	B2	9/2006	Koo et al.
7,318,187	B2	1/2008	Vayanos et al.
8,446,849	B2	5/2013	Damnjanovic et al.
8,493,909	B2	7/2013	Choi
2005/0105502	A1	5/2005	Kim et al.
2005/0117553	A1	6/2005	Wang et al.
2005/0135320	A1	6/2005	Tiedemann et al.
2007/0042786	A1	2/2007	Chillariga et al.
2008/0014969	A1	1/2008	Laroia et al.
2008/0186936	A1	8/2008	Chun et al.

(Continued)

FOREIGN PATENT DOCUMENTS

CN	1089412	A	7/1994
CN	1860808	B	1/2012

(Continued)

OTHER PUBLICATIONS

Office Action, Japanese Patent Application No. 2010-526245, dated Jan. 15, 2013.

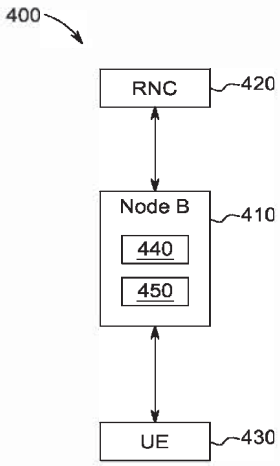
(Continued)

Primary Examiner — Alvin A Zhu
(74) *Attorney, Agent, or Firm* — Volpe and Koenig, P.C.

(57) **ABSTRACT**

User equipment (UE) may receive broadcast information with bits that indicate a respective time interval that physical shared channel resources are utilized to signal channels. The UE may monitor the signaled channels in indicated time intervals. The UE may determine whether signaling information for the UE is present based on an identification of the UE.

12 Claims, 8 Drawing Sheets



(56) **References Cited**

U.S. PATENT DOCUMENTS

2008/0188220	A1	8/2008	DiGirolamo et al.
2008/0311926	A1	12/2008	Fischer et al.
2009/0207774	A1	8/2009	Lee et al.
2010/0234032	A1	9/2010	Chun et al.
2013/0028234	A1	1/2013	DiGirolamo et al.

FOREIGN PATENT DOCUMENTS

GB	2260245	A1	4/1993
JP	11243579	A2	9/1999
JP	2000201196	A2	7/2000
JP	2002532987	T2	10/2002
WO	2005034542	A1	4/2005
WO	2005101879	A1	10/2005
WO	07052187	A1	5/2007

OTHER PUBLICATIONS

Third Generation Partnership Project, "Technical Specification Group Radio Access Network; High Speed Downlink Packet Access (HSDPA); Overall description; Stage 2 (Release 6)," 3GPP TS 25.308 V6.4.0, Mar. 2007.

Third Generation Partnership Project, "Technical Specification Group Radio Access Network; High Speed Downlink Packet Access (HSDPA); Overall description; Stage 2 (Release 7)," 3GPP TS 25.308 V7.3.0, Jun. 2007.

Nokia, "Analysis of HSDPA in Cell_Fach State," R1-062884, 3GPP TSG-RAN WG1 Meeting #46bis, Seoul, South Korea, Oct. 9-13, 2006.

Third Generation Partnership Project, "Technical Specification Group Radio Access Network; Physical Channels and Mapping of Transport Channels onto Physical Channels (TDD) (Release 6)," 3GPP TS 25.221 v6.4.0, Jun. 2005.

Third Generation Partnership Project, "Technical Specification Group Services and System Aspects; Multimedia Broadcast/Multicast Service; Stage 1 (Release 8)," 3GPP TS 22.146 v8.3.0, Jun. 2007.

Non-Final Rejection issued by USPTO, dated Sep. 30, 2010 for U.S. Appl. No. 11/863,205.

Notice of Allowance issued by USPTO, dated Jun. 13, 2011 for U.S. Appl. No. 11/863,205.

International Search Report, PCT Application No. PCT/EP2008/062177, dated Mar. 4, 2009.

Non-Final Rejection, U.S. Appl. No. 13/212,867, dated Jul. 9, 2015.

Non-Final Rejection, U.S. Appl. No. 13/212,867, dated Aug. 8, 2013.

Non-Final Rejection, U.S. Appl. No. 13/212,867, dated May 22, 2014.

Final Rejection, U.S. Appl. No. 13/212,867, dated Mar. 21, 2016.

Final Rejection, U.S. Appl. No. 13/212,867, dated Dec. 17, 2014.

Final Rejection, U.S. Appl. No. 13/212,867, dated Feb. 10, 2014.

Extended European Search Report, European Patent Application No. 17163346, dated Oct. 18, 2017.

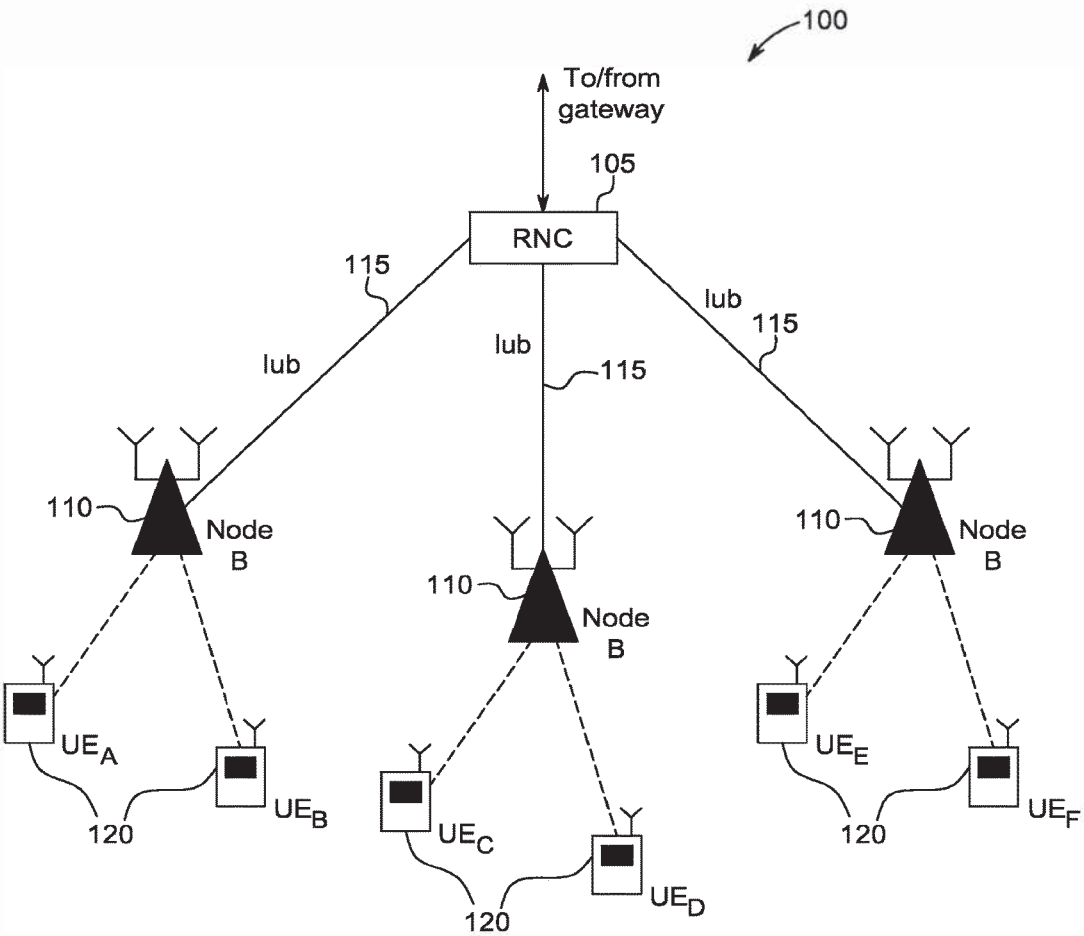


FIG. 1
PRIOR ART

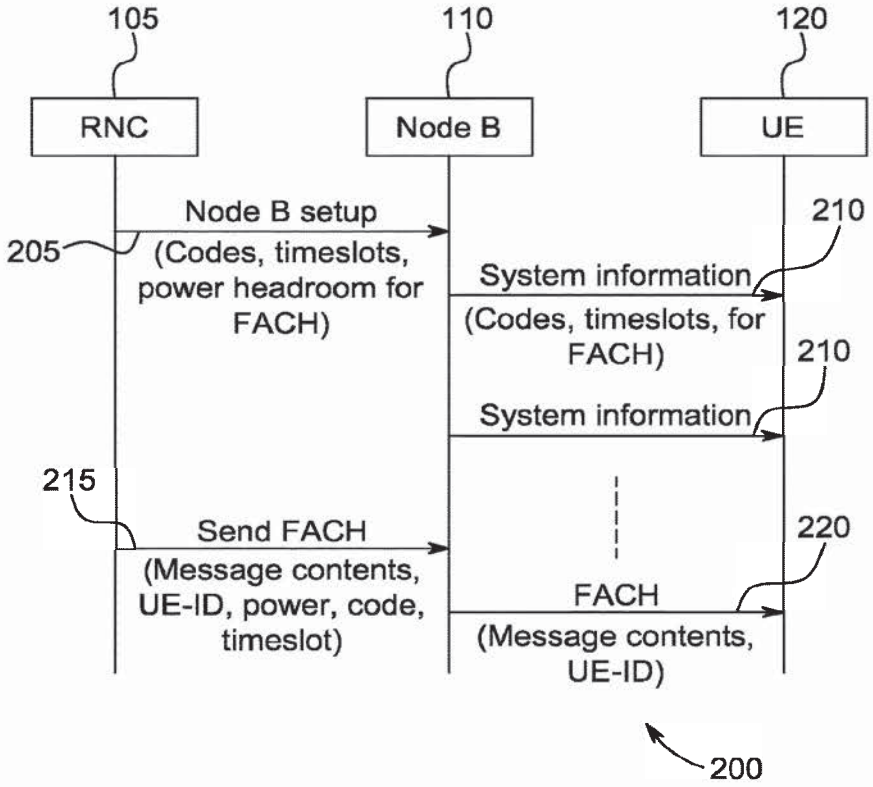


FIG. 2
PRIOR ART

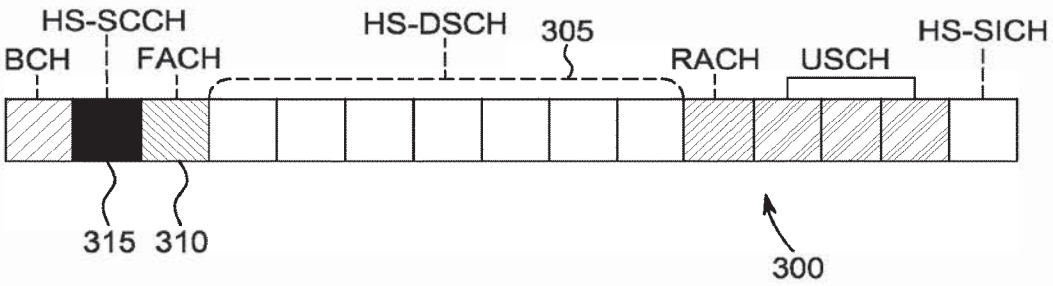


FIG. 3
PRIOR ART

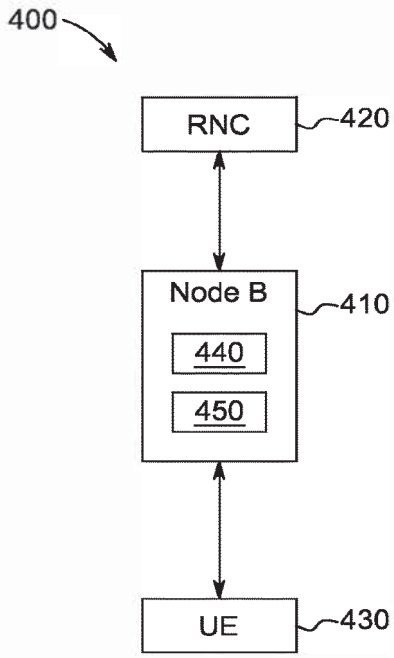

400
RNC
420
Node B
410
440
450
UE
430

FIG. 4

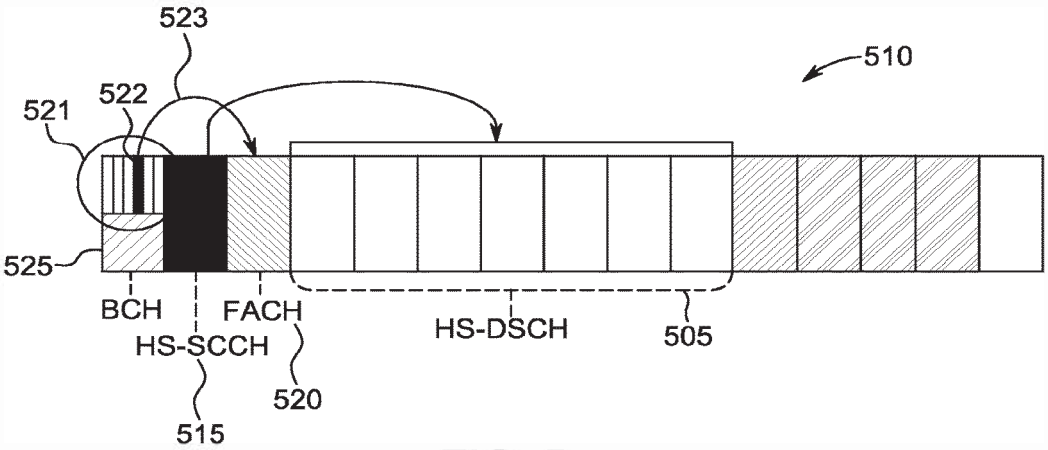

523
510
521
522
525
BCH
FACH
HS-SCCH
HS-DSCH
505
520
515

FIG. 5

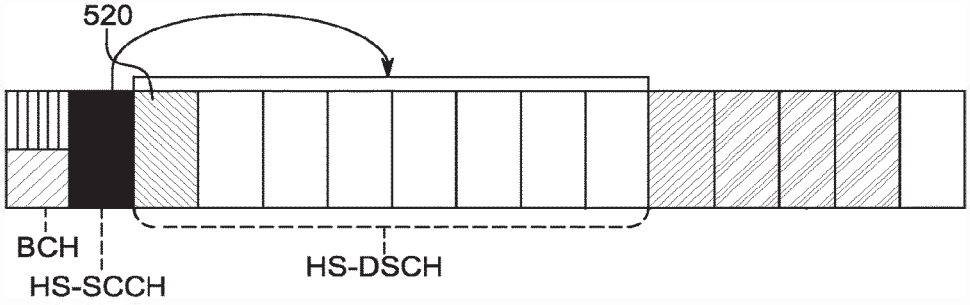

520
BCH
HS-SCCH
HS-DSCH

FIG. 6

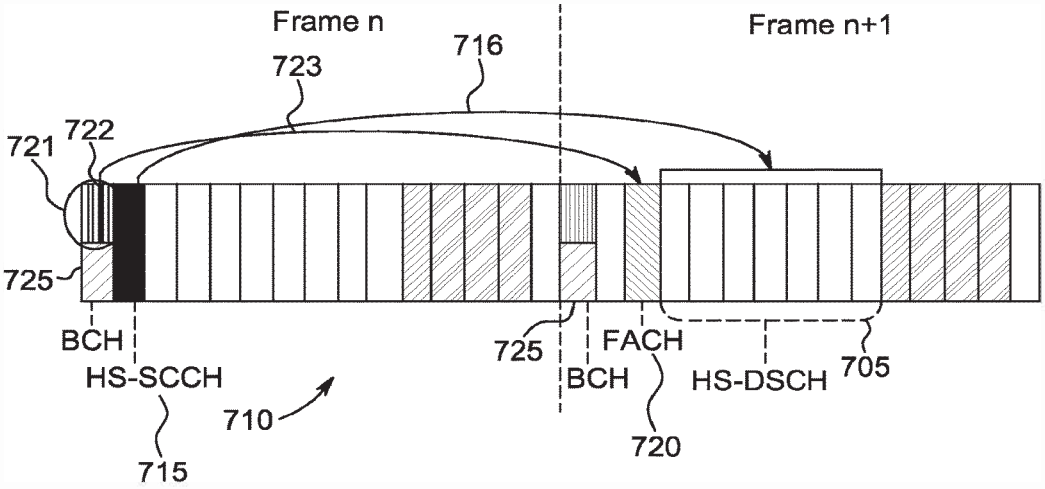

Frame n
Frame n+1
716
723
721
722
725
BCH
HS-SCCH
710
715
725
FACH
BCH
HS-DSCH
705
720

FIG. 7

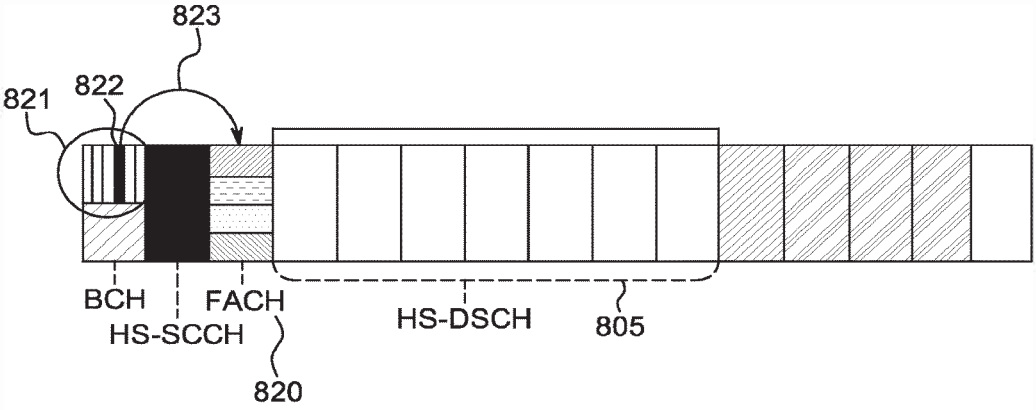

823
821
822
BCH
HS-SCCH
FACH
820
HS-DSCH
805

FIG. 8

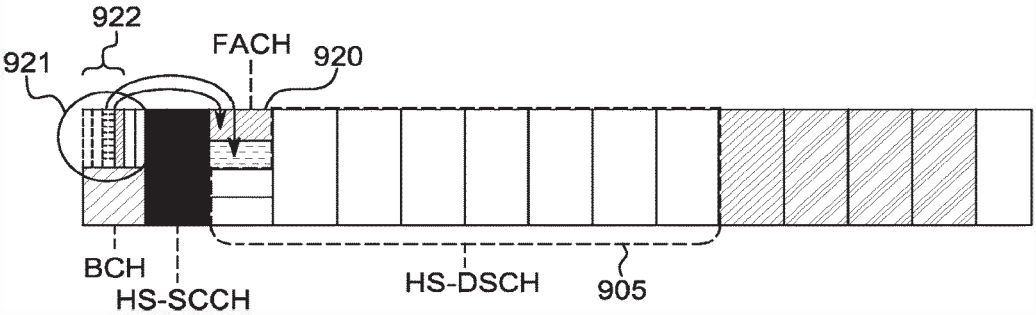

922
921
FACH
920
BCH
HS-SCCH
HS-DSCH
905

FIG. 9

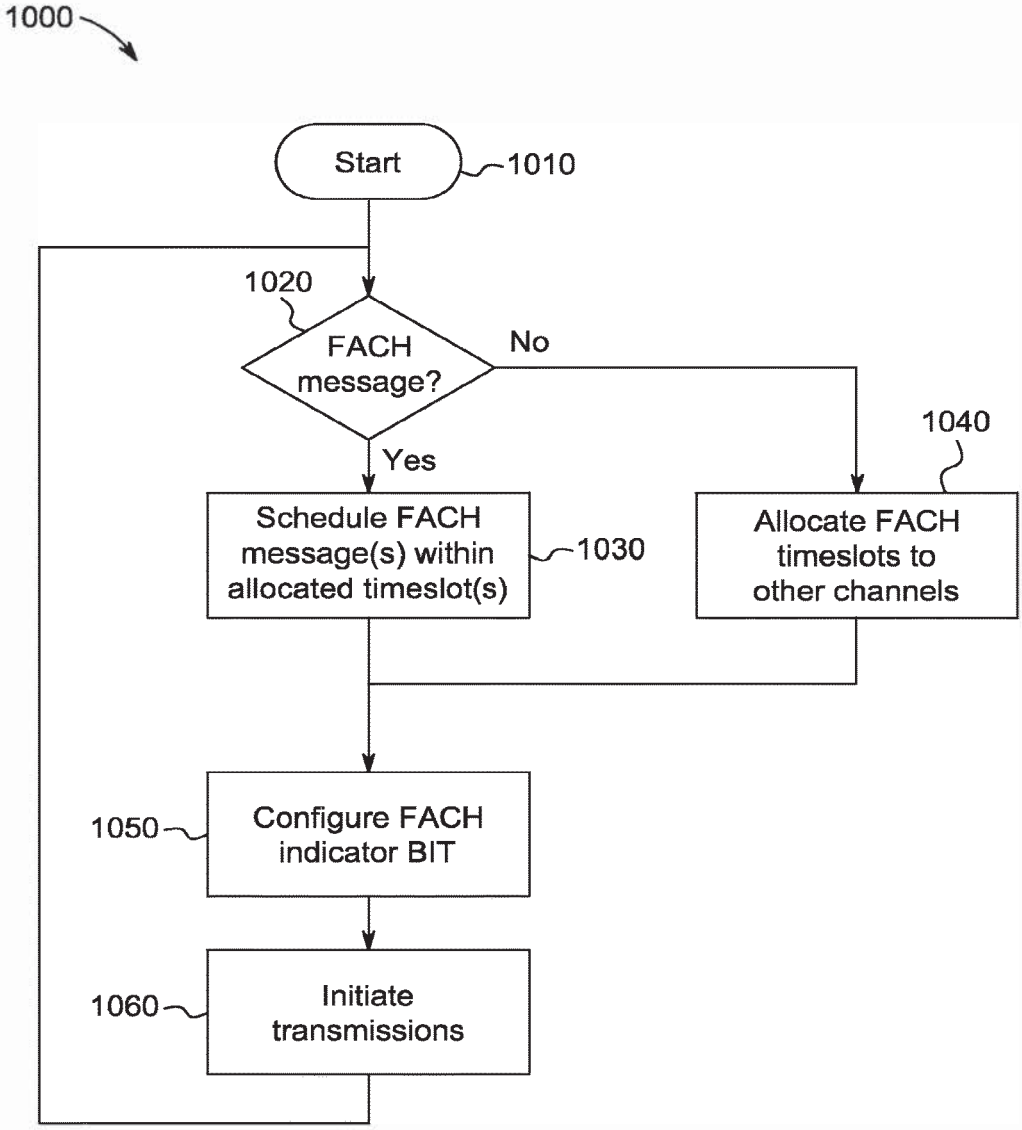

1000
Start
1010
1020
FACH message?
No
Yes
1040
Schedule FACH message(s) within allocated timeslot(s)
1030
Allocate FACH timeslots to other channels
1050
Configure FACH indicator BIT
1060
Initiate transmissions

FIG. 10

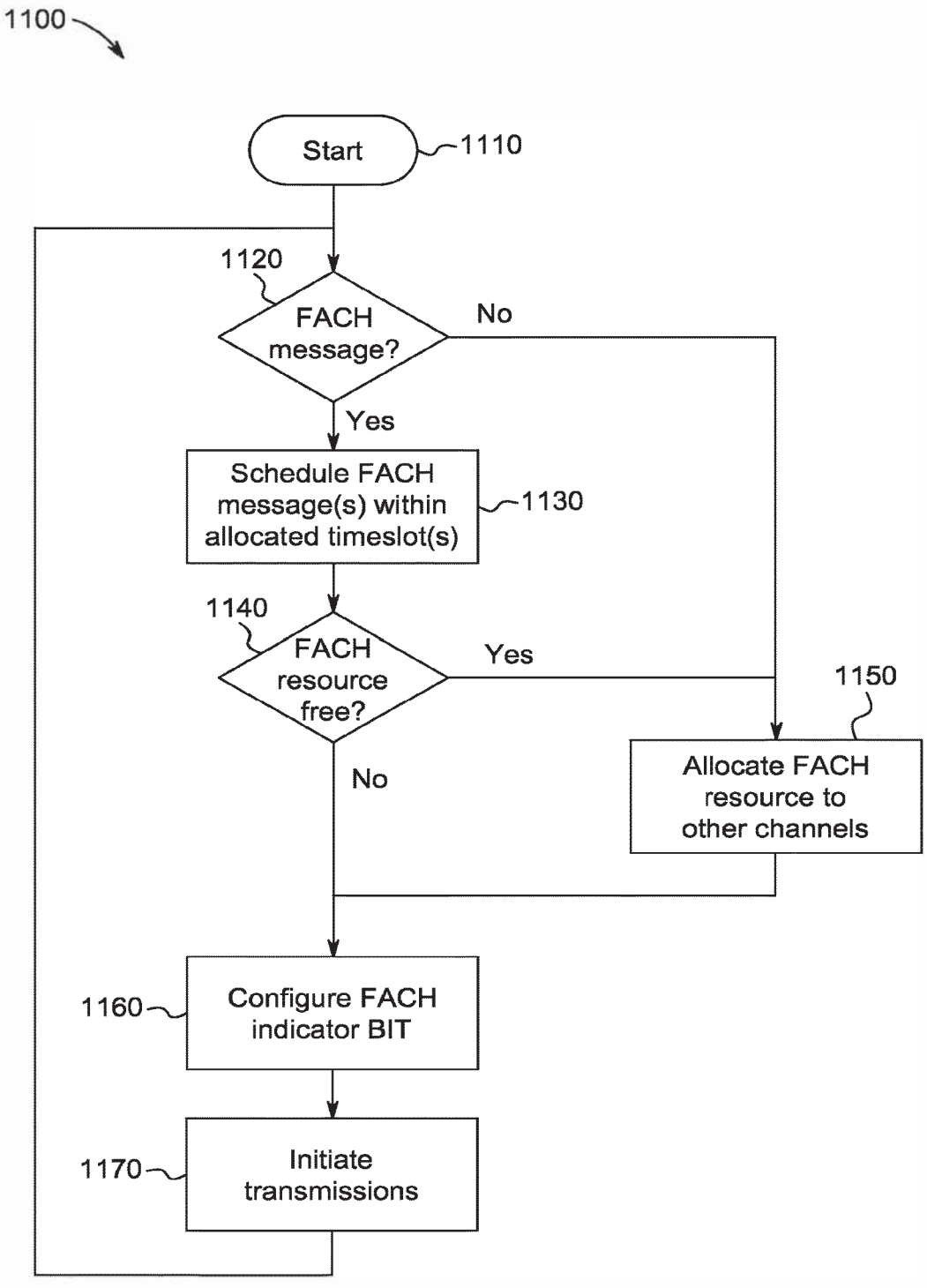

1100
Start
1110
1120
FACH message?
No
Yes
Schedule FACH message(s) within allocated timeslot(s)
1130
1140
FACH resource free?
Yes
No
1150
Allocate FACH resource to other channels
1160
Configure FACH indicator BIT
1170
Initiate transmissions

FIG. 11

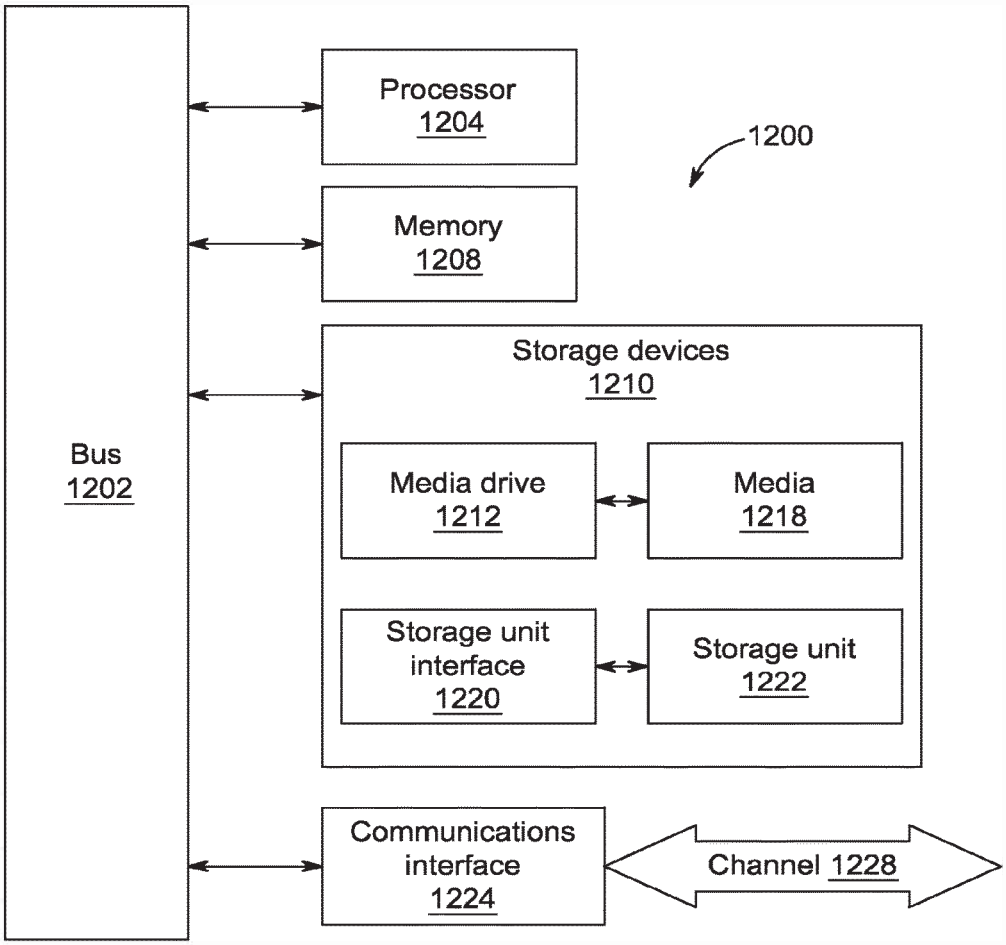

Bus
1202
Processor
1204
1200
Memory
1208
Storage devices
1210
Media drive
1212
Media
1218
Storage unit
interface
1220
Storage unit
1222
Communications
interface
1224
Channel 1228

FIG. 12

COMMUNICATING ON A SHARED CHANNEL IN A WIRELESS NETWORK

CROSS REFERENCE TO RELATED APPLICATION(S)

This application is a continuation of U.S. patent application Ser. No. 13/212,867, filed Aug. 18, 2011, which is a continuation of U.S. patent application Ser. No. 11/863,205, filed Sep. 27, 2007, which issued as U.S. Pat. No. 8,027,291 on Sep. 27, 2011, which are incorporated by reference as if fully set forth.

FIELD OF THE INVENTION

The field of the invention relates, in general, to a method and apparatus for transmitting transport channels over a physical channel of a cellular communication system. In particular, but not exclusively, the field of the invention relates to allocating channel resources within a physical channel of a cellular communication system.

BACKGROUND OF THE INVENTION

Known cellular communication systems, such as a Universal Mobile Telecommunications System (UMTS) Radio Access Network (UTRAN), standardised by the 3rd Generation Partnership Project (3GPP), typically consist of a set of radio network controllers (RNCs), Node B, also known as Node-Bs, and mobile stations (MSs), also known as User Equipment (UEs). FIG. **1** illustrates an example of a known network topology for part of such a cellular communication system **100**.

The RNCs **105** provide a connection with, for example, a Media Gateway (not shown), which acts as a translation unit between, in this case, the UMTS network and, for example, a Public Switched Telephone Network (PSTN). The RNC **105** also performs some of the higher layer processing for the UMTS network, performing functions, such as, setting up and managing radio bearers, radio resource management, supporting mobility, controlling initial access of UEs to the communication system, radio link control (RLC), etc.

The Node Bs **110** typically perform lower layer processing for the network, performing such functions as Medium Access Control (MAC), formatting blocks of data for transmission and physically transmitting transport blocks to UEs.

As can be seen in FIG. **1**, Node Bs **110** are connected to an RNC **105** via an interface (Iub) **115**. This interface between a Node B **110** and an RNC **105** may be a leased line, for example provided by a fixed line telecommunications provider, a microwave link, an Ethernet cable or some other form of communication link. The Node Bs are connected wirelessly to the UEs **120**.

In order to conserve battery life, when a UE **120** is not involved in active connections, it is known for the UE **120** to go into an idle state, whereby the UE powers down its radio frequency circuitry (RF). When a UE **120** is in the idle state, it is important to allow the Node B to initiate a connection to the UE **120**, for example when there is an in-coming call for the UE **120**.

In order to achieve this, it is known for a UE **120** to periodically power up its radio circuitry in order to monitor specific channels in order to determine whether it is required to establish a connection with the network. UMTS provides two services with which a Node B is able to indicate to a specific UE **120** that it is required to establish a connection with the network. One is termed a Paging CHannel (PCH) and the other is termed a Forward Access CHannel (FACH). Details of these and other services are provided in 3GPP TS 25 221 (Universal Mobile Telecommunications System (UMTS); Physical channels and mapping of transport channels onto physical channels (TDD)) and other documents referenced therein.

The PCH is a downlink transport channel that is used to carry control information to a UE when the network does not know the specific location of the UE, i.e. the specific Node B to which the UE is attached.

In the known art, the PCH comprises two blocks: the Paging Indicator CHannel (PICH) and the PCH itself. The PICH comprises multiple indicator bits. Each UE is associated with one of the indicator bits within the PICH. Accordingly, when the UE is paged by the network using the PCH, the network sets the relevant indicator bit in the PICH. When in the idle state, the UE periodically decodes the PICH to see if the indicator bit with which it is associated has been set. If the relevant indicator has been set, the UE then reads the PCH.

Each UE has a unique identifier (UE-ID). When the network pages a UE, as previously mentioned, the network sets the relevant indicator within the PICH for that UE, and transmits the UE-ID for the UE being paged, and the relevant message within the PCH. In this manner, since more than one UE may be associated with an indicator bit in the PICH, the UE-ID enables a UE to determine whether the message is intended for that UE or not. If the PCH contains the UE-ID for the UE, the UE then reads the message, and performs the required actions. Thus, in the known art, the indicator bits are used in a paging channel to inform the UE whether it needs to turn its radio on for reading the PCH (i.e. as a battery saving mode).

The UMTS standard dictates that the PCH (and PICH) are always transmitted at a reference power level.

In a wireless communications system, the communication medium is divided into units of resource. A unit of resource can be a single code (e.g. UMTS FDD), a plurality of codes, a set of codes and timeslots (e.g. UMTS TDD), a set of timeslots (e.g. a TDD system) or a set of tones, tones and symbols or tones, symbols and timeslots (e.g. an OFDM system).

The FACH is a downlink transport channel that is used to carry control information to a UE when the system knows the location cell of the UE, e.g. the specific Node B to which the UE is attached. The FACH allows short messages to be sent from the Node B to the UE. These short messages are typically control type messages that are used, for example, to allocate physical resources to the UE, set up dedicated physical channels, etc.

The FACH is transmitted on a set of physical resources that are pre-defined and broadcast by the Node B on a Broadcast CHannel (BCH). The FACH is controlled by the RNC, which defines the codes and timeslots that are reserved for FACH transmissions. The RNC also reserves an amount of power headroom for the transmission of the FACH. When the FACH is transmitted, the RNC defines the power with which the FACH must be transmitted by the Node B. Power headroom is the amount of power that the RNC reserves for allocation of FACH resources. The power that is not reserved for 'power headroom' is allocated to the Node B to do with as it sees fit (for example scheduling HS-DSCH resource into). Thus, when the RNC allocates power headroom, it informs the Node B that the Node B cannot allocate that power headroom, as the power headroom is reserved by the RNC for the RNC to allocate resource into.

Unlike the PCH (and PICH), the FACH is not required to always be transmitted, and typically is only used when a message is required to be sent to a UE.

FIG. **2** illustrates a high level signal flow example **200** of an implementation for FACH transmissions. An RNC sends a 'NODE B SETUP' message to a Node B, instructing the Node B to reserve a certain set of codes and timeslots for FACH transmissions. This message may also reserve power headroom for use by the RNC. In UMTS, the 'NODE B SETUP' message for configuring the FACH transport channel is the 'COMMON TRANSPORT CHANNEL SETUP' message, sent to the Node B over the Iub interface **115**.

The Node B **110** then periodically transmits a 'SYSTEM INFORMATION' message **210** on the Broadcast CHannel (BCH). This SYSTEM INFORMATION message **210** informs UEs of the physical resources used for FACH transmissions, such as the codes and timeslots in a Code Division Multiple Access (CDMA) system or the sub-carriers and timeslots in an Orthogonal Frequency-Division Multiplexing (OFDM) system. The SYSTEM INFORMATION message **210** typically also contains other broadcast information relevant to that cell, such as the network identity etc.

A UE **120** is able to receive the SYSTEM INFORMATION message **210**, and configure its FACH decoding function based on the information contained within the SYSTEM INFORMATION message **210**.

When the RNC **105** needs to send a message to a UE **120** using the FACH, the RNC sends a 'SEND FACH' message **215** to the Node B. In UMTS, the 'SEND FACH' message is carried using FACH frame protocol (FACH FP) messages over the Iub interface **115**. The SEND FACH message **215** defines the specific code and timeslot to be used for the FACH transmission **220** from the Node B **110** to the UE **120**, as well as the power level to be applied to the FACH transmission, the message contents and the identity of the UE (the UE-ID) that are to be included in the FACH transmission. The Node B **110** subsequently transmits the FACH message **220** to the UE **120**, as defined by the RNC **105** in the SEND FACH message **215**.

The UE **120** decodes the FACH **220** every frame, according to the definition of the FACH provided within the SYSTEM INFORMATION message **210**. If the UE-ID in the FACH message **220** matches the identity that has been assigned to the UE **120**, the UE **120** acts on the message contents in the FACH **220**.

FIG. **3** illustrates an example of a known assignment of physical resources **300** for a single Time Division Duplex (TDD) system, and shows that one timeslot per frame is assigned for FACH usage. As previously mentioned, the physical resources **300** for the FACH are pre-assigned by the RNC **105**, in the NODE B SETUP message **205**, and cannot be reused by other channels. Even when the FACH is lightly used, for example when most UEs are in a connected state, receiving data traffic on a traffic channel, indicated as a High Speed Downlink Shared Channel (HS-DSCH) **305** in FIG. **3**, the FACH timeslot **310** is still required to be reserved.

As will be appreciated by a skilled artisan, the fact that the FACH timeslot **310** is required to be reserved, particularly during light use of the FACH, is an inefficient use of physical resources. Consequently, the inventor of the present invention has recognised and appreciated that it is desirable for the timeslot **310** assigned to the FACH to be able to be used for other purposes when it is not required for the transmission of FACH messages. For the example illustrated in FIG. **3**, the traffic channel resources, namely the HS-DSCH resources **305**, comprise seven timeslots. If the timeslot **310** reserved for FACH transmissions could be utilised for a traffic channel during periods when no FACH transmissions were required to be sent, the traffic channel resources could be increased from seven timeslots to eight timeslots. This would increase the traffic channel resources by fourteen percent (14%) during those periods.

A known solution for the reuse of the FACH timeslot is for the traffic channel resources, which for the example illustrated in FIG. **3** is in the form of the HS-DSCH **305**, to be used for the transmission of FACH messages. In this case, the information that would have been transmitted on the FACH transport channel is instead transmitted on the HS-DSCH transport channel. In this solution the Node B **110**, as opposed to the RNC **105**, controls the FACH, and the FACH timeslot of FIG. **3** becomes an HS-DSCH timeslot. In this manner, in the example of FIG. **3**, the HS-DSCH **305** always comprises eight timeslots, as opposed to seven.

The HS-DSCH **305** is a shared channel that is controlled by the Node B **110**. The content of the HS-DSCH **305** is allocated via Shared Control CHannels for the HS-DSCH (HS-SCCH) **315**, which is a downlink physical channel that carries higher layer control information for the HS-DSCH.

The HS-SCCH **315** contains a UE-ID relating to the UE **120** for which content within the HS-DSCH **305** is intended. The UE-ID ensures that only that UE **120** for which the content of the HS-DSCH **305** is intended decodes the HS-SCCH **315** successfully. The HS-SCCH **315** also instructs the UE of those codes and timeslots that are allocated to it for the HS-DSCH **305** transmission allocated to it, and the coding and modulation of the HS-DSCH **305**, e.g. code rate and modulation order, such as quadrature phase shift keying (QPSK), 16-QAM (Quadrature Amplitude Modulation), etc. Having decoded the HS-SCCH **315**, the UE **120** is able to decode relevant timeslot(s) of the HS-DSCH **305** using the information contained within the HS-SCCH **315**, and retrieve the relevant content.

When the HS-DSCH **305** is used to transmit the FACH, the codes and timeslots that are used for the HS-SCCH **315**, as well as a FACH-ID are broadcast by the Node B, for example within the SYSTEM INFORMATION message **210** illustrated in FIG. **2**. UEs that are required to monitor the FACH subsequently attempt to decode the HS-SCCH **315** using the FACH-ID. If a UE **120** is successful in decoding the HS-SCCH **315** using the FACH-ID, a FACH message is present within the HS-DSCH **315**.

Having decoded the HS-SCCH **315** using the FACH-ID, the UE **120** then decodes the relevant part of the HS-DSCH **305** as allocated by the HS-SCCH **315**, and using the information provided within the HS-SCCH **315**.

The UE-ID of the relevant UE **120** is included with the HS-DSCH **305**. If this matches the UE-ID of the UE **120** decoding the HS-DSCH **305**, the UE **120** acts on the FACH message contents in the HS-DSCH **305**.

When a FACH message is to be sent to a UE **120**, the RNC **105** requests the Node B **110** to send a FACH to a UE **120** with a certain message. The Node B **110** then schedules the transmission of the FACH message in the HS-DSCH **305**, along with any traffic data that is to be sent in the HS-DSCH **305**. The Node B **110** chooses the codes and timeslots that are to be used for the HS-DSCH **305** carrying the FACH message, as well as the coding and modulation to be applied to the HS-DSCH **305**. The Node B then transmits an HS-SCCH **315** that allocates the codes and timeslots on the HS-DSCH **305** for the FACH message, along with the coding and modulation used, and encodes the HS-SCCH **315** using the FACH-ID.

When there are no requests from the RNC **105** for the Node B **110** to transmit a FACH message, the Node B is able to use all of the HS-DSCH **305** resource for traffic data. In this manner, physical resources are used more efficiently, and in particular physical resources are not permanently assigned for transmitting FACH messages, in particular when no, or few, FACH messages are required to be transmitted.

However, a problem with this known technique for reusing the FACH timeslot is that it requires the use of the HS-SCCH **315** and the HS-DSCH **305** channels for transmitting FACH messages, as opposed to just a FACH channel, and as such carries a penalty in terms of power consumption.

FACH messages may be transmitted to a UE **120** when the UE is in an idle state (or alternatively in a connected state). Accordingly, the Node B **110** is unaware of the state of the transmission path between itself and the UE **120**. As a result, normal power control is ineffective since the Node B **110** has no information with which to make an informed decision on the appropriate power level to use. Consequently, the power requirements must initially depend on the number of bits to be transmitted, and the coding rate, without explicit knowledge of the path loss between the UE and Node B.

A number of bits transmitted on the HS-SCCH **315** is comparable to a number of bits that are transmitted on a traditional FACH, in the order of 60 bits (57 bits for a 3.84 Mcps TDD HS-SCCH). The exact number of bits to be transmitted on the FACH depends on the particular implementation. The power requirements for HS-SCCH **315** and FACH at the start of a connection are therefore substantially the same, due to the similar number of bits to be carried on each channel, for example approximately 33% of the Node B transmit power.

A problem with this known technique for reusing the FACH timeslot is that both the HS-SSCH **315** and the HS-DSCH **305** must be used in order to transmit a FACH message, each requiring approximately 33% of the Node B transmit power, albeit in different timeslots. This is in contrast to the traditional method of transmitting FACH messages, where using the dedicated FACH timeslot only required the one timeslot. Thus, the known technique for reusing the FACH timeslot requires an additional timeslot at 33% of the Node B transmit power.

A skilled artisan will appreciate that, although the reuse of the FACH timeslot by mapping the FACH onto HS-DSCH resources provides a considerable improvement in the use of the physical resources, this increase in the power requirements is undesirable.

SUMMARY OF THE INVENTION

Accordingly, the invention seeks to mitigate, alleviate or eliminate one or more of the abovementioned disadvantages, singly or in any combination.

According to a first aspect of the invention, there is provided a cellular communication network node comprising a transmitter operably coupled to signal processing logic and arranged to transmit at least one direct signalling channel indicator bit generated by the signal processing logic. The signal processing logic is arranged, upon receipt of a message to be transmitted, to set at least one direct signalling channel indicator bit to indicate that a direct signalling channel is active in at least one unit of resource. The signal processing logic is further arranged to reallocate the at least one unit of resource for use by at least one other channel when no message is required to be transmitted.

Thus, embodiments of the invention may allow improved use of the communication resource in the communication system, for example by allowing physical resource to be dynamically switched between different channel types.

According to an optional feature of the invention, the signal processing logic may be arranged to generate direct signalling channel resource information comprising a description of the physical resources and transport formats applied to the direct signalling channel(s). Thus, embodiments of the invention may allow the physical resources of the direct signalling channel to be transmitted relatively infrequently to all UEs. Thus, in this manner, a reduction in the signalling load is provided, as compared to the classic case where the physical resources for the direct signalling channel are sent whenever the direct signalling channel is allocated.

According to an optional feature of the invention, the signal processing logic may be arranged, upon receipt of a message to be transmitted, to set the at least one direct signalling channel indicator bit in a first frame, and to transmit the message in a second frame. For example, the second frame may be a next consecutive frame to the first frame. Thus, embodiments of the invention may allow the UE to have a longer time to process the indicator bit, thereby allowing the UE complexity to be reduced and/or to allow the UE to turn its receiver off in the time during which the direct signalling channel is transmitted.

According to an optional feature of the invention, the signal processing logic may allocate a plurality of direct signalling channels to at least one unit of resource. Thus, embodiments of the invention may allow the system to operate with a reduced number of direct signalling channel indicator bits. The bits that are freed up on the indicator channel may then be used for other indication purposes (e.g. paging or E-DCH ACK/NACK signalling).

According to an optional feature of the invention, the signal processing logic may generate a single direct signalling channel indicator bit for transmission, or a plurality of direct signalling channel indicator bits, one for each direct signalling channel, or a number of indicator bits for transmission, where each indicator bit corresponds to a plurality of direct signalling channels. Thus, embodiments of the invention may allow for finer granularity control of the allocated direct signalling channel, for example the network may allocate some of the resource as direct signalling channel instead of either allocating all or none of the resource. This allows for more targeted allocation of resource to direct signalling channels and traffic channels (e.g. HS-DSCH).

According to an optional feature of the invention, the signal processing logic may be arranged to allocate a plurality of direct signalling channels to a plurality of units of resource. Thus, embodiments of the invention may allow a single indicator bit to control more than one direct signalling channel, thus either allowing more direct signalling channels (when the number of indicator bits that can be sent are constrained) or reduced usage of the indicator bits.

According to an optional feature of the invention, the signal processing logic, upon receipt of at least one message to be transmitted, may be arranged to re-allocate at least one unit of resource not required for transmitting the one or more message, for use by at least one channel other than the direct signalling channel. Thus, embodiments of the invention may allow better usage of the resource, for example some of the resource may be assigned to the direct signalling channel and some of the resource may be reassigned to the other channel when it is not needed for the direct signalling

channel. Furthermore, the usage of the resource may be more tailored to the traffic profile, if the direct signalling channel is not greatly required, but there is a lot of traffic on the traffic channel.

According to an optional feature of the invention, the at least one direct signalling channel may be a Forward Access Channel(s) (FACH) within a Universal Mobile Telecommunications System (UMTS) network. For example, the at least one direct signalling channel unit of resource may be re-allocated for use by a High Speed Downlink Shared CHannel (HS-DSCH). Thus, embodiments of the invention may allow direct applicability to an application. In this case, for example, under-utilised FACH resource may be transferred to HS-DSCH resource, thus increasing the potential throughput of HS-DSCH.

According to an optional feature of the invention, the signal processing logic may generate the at least one direct signalling channel indicator bit within a Broadcast CHannel (BCH) for broadcast. For example, the signal processing logic may locate the at least one direct signalling channel indicator bit within a Transport Format Combination Indicator (TFCI) type field within the BCH. Furthermore, according to an optional feature of the invention, the signal processing logic may generate the direct signalling channel indicator bit(s) as part of a data portion of the BCH for broadcast. Thus, embodiments of the invention may allow transmission directly on the broadcast channel, thereby removing a need for a separate indicator channel. Alternatively, when an indicator channel naturally exists (e.g. PICH or E-HICH), the paging or E-DCH acknowledgement capacity of those indicator channels is not reduced, as direct signalling channel indicator bits do not replace indicator bits that might be required for the other functions, such as PICH or E-HICH.

According to an optional feature of the invention, the signal processing logic may generate the at least one direct signalling channel indicator bit within a dedicated direct signalling channel indicator channel or within a dedicated paging indicator channel or within a E-DCH Hybrid ARQ Indicator CHannel (E-HICH) of a Universal Mobile Telecommunications System (UMTS) network. Thus, embodiments of the invention may allow a reduction in usage of the broadcast channel. In this manner, an indicator channel may also be more lightly coded (than a broadcast channel), thereby allowing for lower power/lower complexity decoding of the indicator channel (compared to a case where the indicator bits are sent on the broadcast channel).

In accordance with a second aspect of the invention there is provided a method for transmitting transport channels over a physical channel of a cellular communication system. The method comprises determining whether at least one message is to be transmitted. If it is determined that at least one message is required to be transmitted, scheduling the at least one message for transmission in at least one unit of resource allocated to at least one direct signalling channel, and setting at least one direct signalling channel indicator bit for transmission. If it is determined that no message is to be transmitted, re-allocating at least one unit of resource allocated for use by the at least one direct signalling channel for use by at least one channel other than the direct signalling channel. Thus, embodiments of the invention may provide a more flexible usage of resources, for example direct signalling channel resource that is under-utilised may be re-assigned to other (e.g. traffic) channels.

In accordance with a third aspect of the invention there is provided a computer program product comprising program code for signalling in a wireless communication system from a network element to a wireless communication unit. The computer program product comprising program code for setting, upon receipt of a message to be transmitted, at least one direct signalling channel indicator bit. The program code further comprises code for transmitting a message in at least one unit of resource allocated for use as a direct signalling channel, and re-allocating the at least one unit of resource allocated to the direct signalling channel for use by at least one other channel when no message is required to be transmitted. Thus, embodiments of the invention may allow the system to run on general purpose hardware, thus reducing cost.

In accordance with a fourth aspect of the invention there is provided a wireless communication unit arranged to receive a signalling message allocating a first set of resources in a wireless communication system from a cellular communication network node. The wireless communication unit further comprises logic for receiving a at least one direct signalling channel indicator bit on a communication channel, the at least one direct signalling channel indicator bit allocating a direct signalling channel. The wireless communication unit further comprises logic for modifying the wireless communication unit's first set of resources in response to a status of the at least one direct signalling channel indicator bit.

In the existing 3GPP system, for example, whenever any resource is applied to FACH, resources in the FACH slot and HS-DSCH resources cannot easily be allocated to the same UE, since the amount of resource in the FACH slot and HS-DSCH slot differ and existing signalling restrictions mean that the same amount of physical resource has to be allocated in each timeslot. Thus, embodiments of the invention address this limitation and may allow the system to allocate more resource to a UE.

In this manner, in contrast to normal 3GPP operation where the same number of codes has to be allocated to a wireless communication unit in all of the timeslots that it is allocated due to signalling restrictions, the wireless communication unit is signalled whether FACH is active. Advantageously, the wireless communication unit is able to use this indication to modify, for example, the HS-DSCH resource allocation, as the wireless communication unit knows that if the FACH indicator bit was set then any HS-DSCH resources that are common to FACH resources in the FACH timeslot must be allocated to FACH and not to HS-DSCH.

In accordance with a fifth aspect of the invention there is provided a method for resource allocation in a wireless communication system by a wireless communication unit. The method employed by the wireless communication unit comprises receiving by a wireless communication unit a signalling message allocating a first set of resources in a wireless communication system via at least one direct signalling channel indicator bit on a communication channel, the at least one direct signalling channel indicator bit allocating a direct signalling channel; and modifying the wireless communication unit's first set of resources in response to a status of the at least one direct signalling channel indicator bit.

These and other aspects, features and advantages of the invention will be apparent from, and elucidated with reference to, the embodiments described hereinafter.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** shows an example of a known network topology for part of a cellular

FIG. **2** shows a high level example of a known implementation for FACH channel transmissions.

FIG. **3** shows an example of a known assignment of physical resources.

Embodiments of the invention will now be described, by way of example only, with reference to the accompanying drawing(s), in which:

FIG. **4** shows part of a cellular communication system in accordance with embodiments of the invention.

FIG. **5** and FIG. **6** show an example of resource allocation in accordance with embodiments of the invention.

FIG. **7** shows an example of resource allocation in accordance with embodiments of the invention.

FIG. **8** shows an example of resource allocation in accordance with embodiments of the invention.

FIG. **9** shows an example of resource allocation in accordance with embodiments of the invention.

FIG. **10** shows a flowchart of a method for transmitting transport channels over a physical channel of a cellular communication system in accordance with embodiments of the invention.

FIG. **11** shows a flowchart of a method for transmitting transport channels over a physical channel of a cellular communication system in accordance with embodiments of the invention.

FIG. **12** illustrates a typical computing system that may be employed to implement processing functionality in accordance with embodiments of the invention.

DESCRIPTION OF EMBODIMENTS OF THE INVENTION

The following embodiments of the invention will be described in a context of a Forward Access Channel (FACH) provided as part of a Universal Mobile Telecommunications System (UMTS) communication network, as defined in 3GPP TS 22.146 and 3GPP TS 25.221 and other documents referenced therein. However, it will be appreciated by a skilled artisan that the inventive concept described herein may be applied to any direct signalling channel, such as a paging channel, an allocation channel such as HS-SCCH or applied to alternative comparable services.

The following embodiments of the invention will be described in a context of timeslots, for example with respect to a TDD implementation, but the inventive concept is equally applicable to any communication 'unit of resource', for example for non-TDD implementations. Thus, hereinafter the term timeslot in a TDD context encompasses any communication 'unit of resource' outside of a TDD domain.

In one embodiment of the invention, in contrast to the known art where the indicator bits are used in a paging channel to inform the UE whether it needs to turn its radio on for reading the PCH (i.e. as a battery saving mode), the indicator bits are used to inform the UE whether it should be reading the resources 'reserved' for FACH as a FACH channel, or whether it should be ignoring those resources (for example when they are being used for another channel such as HS-DSCH).

Referring now to FIG. **4**, there is illustrated a part of a cellular communication network **400** according to embodiments of the invention. The cellular communication network **400** comprises a communication network node, which for the illustrated embodiment is in a form of a Node B (NODE B) **410**, operably coupled to a radio network controller (RNC) **420**, for example by way of a leased line provided by a fixed line telecommunications provider, a microwave link, an Ethernet cable or some other form of communication link. The RNC **420** may be further operably coupled to, for example, a Media Gateway (not shown).

The Node B **410** comprises signal processing logic **440** arranged to transmit one or more indicator bits, for example on a broadcast channel. The signal processing logic **440** is further arranged, upon receipt of a FACH message to be transmitted, for example from the RNC **420**, to set at least one of the one or more indicator bits and to transmit the FACH message in at least one of one or more timeslots allocated for use by one or more FACH channels. The signal processing logic **440** is still further arranged to reallocate the one or more timeslots allocated to the one or more FACH channels, for use by one or more channels other than the one or more FACH channels, when no FACH messages are required to be transmitted.

The RNC **420** may, for example by way of a 'NODE B SETUP' message (e.g. 'COMMON TRANSPORT CHANNEL SETUP' in UMTS), instruct the Node B **410** those timeslot(s) that is/are to be allocated to the FACH channel(s), along with a set of channelisation codes that is reserved for the FACH channel(s) (codes used to separate the different channels that may be present on a certain frequency in a Code Division Multiple Access (CDMA) system), power headroom reserved for the RNC, etc. In a context of a 3GPP system, the generic expressions of 'NODE B SETUP" and 'SEND FACH' message may be referred to, for example, as 'COMMON TRANSPORT CHANNEL SETUP' and 'FACH frame protocol'.

Accordingly, the Node B **410** may comprise a memory element **450**, in which the Node B **410** stores the settings provided by the RNC **420**.

Alternatively, the Node B **410** may define which timeslot(s) are to be allocated to the FACH channel, along with a set of channelisation codes that are reserved for the FACH channel(s), power assigned to Node B FACH transmissions etc. The Node B may then inform the RNC of these definitions.

The signal processing logic **440** of the Node B **410** is arranged to transmit, upon receipt of a FACH message transmitted from the RNC **420**, the one or more indicator bits within, for example, the broadcast channel. In this manner, by monitoring the channel carrying indicator bits, e.g. the broadcast channel, the UE **430** is able to determine whether there is a FACH message. If there is no indicator bit set, the UE **430** is able to power down its radio circuitry until the next time the channel carrying indicator bits is scheduled to be transmitted. If there is a FACH message, the UE **430** is able to read the contents of the FACH channel, and determine whether, or not, the message is intended for itself, as described in greater detail below.

Referring now to FIG. **5**, there is illustrated an example of resource allocation in accordance with embodiments of the invention, where a FACH message is required to be transmitted. In particular, FIG. **5** illustrates a plurality of timeslots **510** within a 3.84 Mcps TDD UMTS frame.

The plurality of timeslots **510** are allocated to transport channels, which for the illustrated embodiments may comprise a Broadcast CHannel (BCH) **525**, a High Speed Downlink Shared Channel (HS-DSCH) **505**, a Shared Control Channel for the HS-DSCH (HS-SCCH) **515**, and a Forward Access CHannel (FACH) **520**. In embodiments of the invention, one timeslot also carries (what will hereinafter be referred to as) an INdicator Channel (INCH) **521**, carrying indicator bits. FIG. **5** is labeled in terms of transport channels, but the skilled artisan will appreciate that transport channels are mapped to physical channels before transmission and that for the transport channels illustrated there is a

one-to-one mapping between transport channel and physical channel. A skilled artisan will also appreciate that the timings between channels as shown in FIG. **5**, and subsequent figures, are illustrative and may vary in an actual implementation of the invention.

Furthermore, the indicator bits may be sent on any channel used to send indicator bits, which may exist on a same timeslot as the BCH (but can also or alternatively exist on a timeslot other than that occupied by the broadcast channel (BCH)). For example, such indicator bits may take over some of the indicators that are used on other indicator channels (for example some of the E-HICH indicator bits would not be used for E-HICH purposes (e.g. acknowledging uplink transmissions), but would be used for transmission of FACH indicator bits instead).

Furthermore, the indicator bits may be either broadcast to all UEs in a cell, or transmitted to a targeted group of one or more UEs in a cell. For example, the targeted group of UEs may be a group of UEs with a particular capability or may be a group of UEs geographically located close to the Node B. An example method of targeting a group of UEs would be to mask the indicator bits with a scrambling code that is known to a group of UEs or to transmit the indicator bits at a power level such that only that group of UEs is able to receive the indicator bits reliably.

As previously mentioned, the FACH is a downlink transport channel that is used to carry control information to a UE. The FACH allows short messages to be sent from the network to the UE. These short messages are typically control type messages that are used, for example, to allocate physical resources to the UE, set up dedicated physical channels, etc. In particular, the FACH provides a service for a network to communicate with UEs, without having to set up dedicated physical resources.

As previously mentioned, with respect to FIG. **4**, the RNC **420** may instruct the Node B **410** those timeslot resources that are to be reserved for the FACH, for example by way of a 'NODE B SETUP' message, along with the channelisation codes and transport formats to be applied to the FACH. Furthermore, when the network needs to send a FACH message to a UE, for example UE **430** of FIG. **4**, the RNC **420** sends a message, for example a 'SEND FACH' message, to the Node B **410** comprising the content of the message, and the identity of the UE **430**, such as a UE-ID for which the message is intended. The RNC **420** may also define other criteria such as the specific timeslot resources to be used (from those reserved), specific channelisation codes to be used, etc. Alternatively, the choice of timeslot resources, channelisation codes, etc. may be left to the Node B **410**, and as such the Node B **410** schedules the transmission of FACH messages within the, or each, FACH channel **520**.

The BCH is a downlink transport channel that is used to broadcast system and cell specific information. In particular, the NODE B **410** may use the BCH **525** to transmit resource information relating to common channels (where a common channel is a channel that a plurality of UEs are meant to decode, in contrast to a dedicated physical channel, which is meant to be decoded by a specific UE). The resource information may comprise descriptions of channelisation codes for common physical channels, timeslot resource allocations, and formats for transport channels such as coding and modulation.

For the illustrated embodiments, the INCH **521** further comprises one or more indicator bits, at least one of which comprises a FACH indicator bit **522**. In this manner, the resource information on BCH further comprises description of the at least one indicator bit within the INCH **521**.

The FACH indicator bit is used to indicate whether a FACH channel to which it relates is 'active'. Thus, for the embodiments illustrated in FIG. **5**, the FACH indicator bit **522** is used to indicate whether the FACH is 'active'.

When the NODE B **410** is required to send a FACH message, for example having received a SEND FACH message from the RNC **420**, the NODE B **410** sets the FACH indicator bit **522**, as illustrated in FIG. **5**. The NODE B **410** also transmits the FACH message within the FACH timeslot, along with the identity of the UE, for example the UE-ID, for which the message is intended.

In this manner, a UE **430** attached to the NODE B **410**, is able to determine whether there is a FACH message by monitoring the INCH **521**, and in particular by reading the FACH indicator bit **522**. Upon reading a set FACH indicator bit, the UE **430** determines that the NODE B **430** is transmitting a FACH message, and using the resource information being transmitted in the BCH **525**, is able to decode the FACH message. If the identity of the UE **430** matches the identity transmitted within the FACH message, the UE **430** acts on the message content.

Referring now to FIG. **6** there is illustrated the resource allocation of FIG. **5**, where no FACH message is required to be transmitted. As the NODE B **410** has no FACH message to transmit, the FACH indicator bit **522** in the INCH **521** is not set. In this manner, a UE **420** attached to the NODE B **410**, for example UE **430**, is able to determine that the NODE B **110** is not transmitting a FACH message, by reading the FACH indicator bit **522**, and accordingly the UE does not attempt to decode HS-DSCH resource as FACH. This may allow the UE to conserve battery power and avoids the UE incorrectly decoding HS-DSCH as FACH, thus improving reliability of the FACH transmission.

As will be appreciated by a skilled artisan, since the NODE B **410** is not transmitting a FACH message, the timeslot allocated to the FACH **520** is free. Therefore, in accordance with embodiments of the invention, the NODE B **410** is arranged to reallocate the FACH timeslot **520** for use by a channel other than the FACH, when no FACH message is required to be transmitted.

For the illustrated embodiments, the FACH timeslot **520** is reallocated for use by the HS-DSCH **505**. The HS-DSCH **505** is a shared channel, the content and physical resources of which are allocated via the HS-SCCH **515**, which is a downlink physical channel that carries control information for the HS-DSCH **505**. A skilled artisan will appreciate that, although the HS-SCCH **515** is illustrated herein as occupying an entire timeslot, the HS-SCCH **515** may be multiplexed with other channels.

As previously mentioned, when the NODE B **410** is not transmitting a FACH message, it reallocates the FACH timeslot **520** for use by another channel, which for the illustrated embodiments is the HS-DSCH **505**.

In this manner, when a FACH message is not being transmitted, the resources available to the HS-DSCH **505** can be expanded from, for the illustrated embodiments, seven timeslots to eight, an increase of approximately 14%. Thus, embodiments of the invention provide a more efficient use of physical resources, as compared with traditional techniques for providing services, such as that of FACH in traditional cellular communication systems.

Furthermore, unlike for the known solution for reusing the FACH timeslot described above, as the HS-DSCH **505** and HS-SCCH **515** are not used for transmitting FACH messages, these two channels do not each require a signifi-

cant proportion (e.g. approximately 33%) of the Node B transmit power when a FACH message is to be transmitted. Rather, only the FACH, when a FACH message is being transmitted, requires 33% of the Node B transmit power.

Thus, an embodiment of the invention may provide an advantage of a more efficient use of physical resources, whilst substantially avoiding an increase in power requirements.

For the embodiments illustrated in FIG. **5** and FIG. **6**, the FACH indicator bit **522** is illustrated as indicating **523** a state of the FACH within the same frame. A skilled artisan will appreciate that, for some UE implementations, this may not be feasible due to the processing requirements this would place on the UE.

Referring now to FIG. **7**, there is illustrated an example of resource allocation in accordance with embodiments of the invention. In the same way as for the embodiments illustrated in FIG. **5** and FIG. **6**, timeslots **710** within a frame are allocated to a BCH **725**, HS-DSCH **705**, HS-SCCH **715**, FACH **720**, INCH **721**, etc.

For the embodiments illustrated in FIG. **7**, the INCH **721** comprises a plurality of indicator bits, one of which, for the illustrated embodiments, is in the form of a FACH indicator bit **722**, which indicates the state of the FACH **720** in a subsequent frame of the physical channel. For the illustrated embodiments, the FACH indicator bit **722** in frame 'n' indicates **723** the state of the FACH **720** within the next consecutive frame, frame '**n+1**'. Thus, in FIG. **7**, the FACH indicator bit **722** in frame 'n' is set, and thereby indicates that the FACH **720** is active in frame '**n+1**'. In this manner, the processing requirements placed on a UE **430** are reduced, since the UE **430** is provided more time in which to prepare to decode the FACH **720** if the FACH **720** is active. As illustrated, the HS-SCCH **715** may also relate **716** to the HS-DSCH **705** in the next consecutive frame.

Referring now to FIG. **8**, there is illustrated an example of resource allocation in accordance with embodiments of the invention. For the embodiments illustrated in FIGS. **5** to **7**, a single FACH indicator bit is used to indicate a state of a single FACH, and the FACH is allocated an entire timeslot within the physical channel. A skilled artisan will appreciate that a NODE B may be arranged to provide more than one FACH, and that the plurality of FACHs may be allocated to one or more timeslots.

For the embodiments illustrated in FIG. **8**, four FACH channels are allocated to a single timeslot, and the FACH indicator bit **822** indicates **823** whether at least one of the FACH channels is active. In this manner, if the NODE B is required to transmit one or more FACH messages, the NODE B sets the FACH indicator bit **822**, and transmits the one or more FACH messages in one or more FACH channels **820** within the allocated timeslot. In this manner, a UE is able to determine whether a FACH message is to be transmitted by reading the FACH indicator bit **822** in the INCH **821**. If the FACH indicator bit **822** is set, the UE then decodes the FACH timeslot and reads each of the four FACH channels therein to determine whether the one or more transmitted FACH messages are intended for itself.

If the NODE B is not required to transmit a FACH message, the NODE B sets the FACH indicator bit **822** to 'inactive', and reallocates the FACH timeslot for use by another transport channel, for example the HS-DSCH **805**.

Referring now to FIG. **9**, there is illustrated an example of resource allocation in accordance with embodiments of the invention. As in the embodiment illustrated in FIG. **8**, four FACH channels are allocated to a single timeslot. However, for the embodiments illustrated in FIG. **9**, the INCH **921** comprises a plurality of FACH indicator bits **922**. For the illustrated embodiments, the INCH **921** comprises four FACH indicator bits **922**, each FACH indicator bit **922** relating to one of the FACH channels **920**.

In this manner, when the NODE B receives one or more FACH messages to be transmitted, the NODE B schedules transmission of the one or more FACH messages in one or more of the FACH channels **920**, and sets one or more of the FACH indicator bits **922**, to indicate those FACH channels **920** that are active.

Thus, FIG. **9** illustrates a case where the NODE B is required to transmit two FACH messages. Accordingly, the NODE B has set two of the FACH indicator bits **922**, and transmits the FACH messages within the two corresponding FACH channels **920**. As will be appreciated by a skilled artisan, for the embodiments illustrated in FIG. **9**, only two out of four of the FACH channels **920** within the FACH timeslot are used. Accordingly, and in accordance with an embodiment of the invention, the NODE B re-allocates the timeslot resource for those FACH channels not required for transmitting FACH messages, for use by another transport channel, such as the HS-DSCH **905**. In this manner, even when the NODE B is required to transmit FACH messages, free resource within the FACH allocated timeslot may be reused for other channels.

In the embodiments shown in FIG. **9**, where a timeslot is partially used by FACH, a UE that is allocated HS-DSCH may advantageously modify its resource allocation based on the status of the FACH indicator bit. For example, if the UE is allocated all eight timeslots and all sixteen codes on HS-DSCH **905**, and the FACH indicator bit is set, the UE may understand that in the FACH timeslot **920** it is actually allocated a reduced number of HS-DSCH codes. This is in contrast to the existing operation in 3GPP, where the same number of codes has to be allocated to a UE in all of the timeslots that it is allocated, due to signalling restrictions. However, in accordance with embodiments of the invention, the UE is signaled whether the FACH is active. Advantageously, the UE is able to use this indication to modify the HS-DSCH resource allocation (as the UE understands that if the FACH indicator bit was set, then the UE cannot be allocated all of the resource in the FACH timeslot as that resource is allocated elsewhere as FACH).

For the various embodiments illustrated and described herein, a single timeslot has been allocated for FACH channels. However, a skilled artisan will appreciate that more than one timeslot may be allocated for FACH channels. Accordingly, and in accordance with an embodiment of the invention, where more than one timeslot is allocated for FACH channels, the one or more timeslots may, individually or collectively, and as a whole or partly, be re-allocated for use by other channels when not required for the transmission of FACH messages.

Referring now to FIG. **10**, there is illustrated a flow chart **1000** of a method for transmitting transport channels over a physical channel of a cellular communication system in accordance with embodiments of the invention.

The method starts at block **1010** and moves to block **1020**, where it is determined whether a FACH message is required to be transmitted, for example whether a 'SEND FACH' command has been received from an RNC.

If it is determined that a FACH message is required to be transmitted, the method moves on to block **1030**, where the FACH message(s) to be transmitted is/are scheduled for transmission in one or more timeslots allocated to one or more FACH channels. The method then moves on to block **1050**.

Referring back to block **1020**, if it is determined that no FACH message is required to be transmitted, the method moves on to block **1040**, where the timeslot(s) allocated to the one or more FACH channels are re-allocated for use by one or more channels other than the one or more FACH channels. The method then moves on to block **1050**.

In block **1050**, the FACH indicator bit is configured as required for transmission in, for example a broadcast channel or indicator channel, indicating whether the FACH is active, e.g. a FACH message is to be transmitted, or not. That is to say, if a FACH message is to be transmitted, the FACH indicator bit is set (e.g. transmitted as a logic '1'), whereas if a FACH message is not to be transmitted, the FACH indicator bit is not set (e.g. transmitted as a logic '0' or alternatively the FACH indicator bit is simply not sent at all (for example in a Discontinuous Transmission (DTX) mode)). The method then moves on to block **1060**, where the transmission of the transport channels is initiated. The method then returns to block **1020**, as shown.

A skilled artisan will appreciate that only those blocks pertinent to the invention have been illustrated and described in order not to obscure the inventive concept.

Referring now to FIG. **11**, there is illustrated a flowchart **1100** of a method for transmitting transport channels over a physical channel of a cellular communication system, in accordance with embodiments of the invention.

The method starts at block **1110** and moves to block **1120**, where it is determined whether a FACH message is required to be transmitted. If it is determined that one or more FACH messages are required to be transmitted, the method moves on to block **1130**, where the one or more FACH message(s) is/are scheduled for transmission within the allocated timeslot(s).

The method then moves on to block **1140**, where it is determined whether resources allocated for FACH messages are free for use by other channels. If it is determined that no FACH resources are free in block **1140**, the method moves on to block **1160**. However, if it is determined that FACH resources are free, the method moves on to block **1150**.

Referring back to block **1120**, if it is determined that no FACH messages are required to be transmitted, the method moves on to block **1150**, where resources allocated for FACH messages that are not required for the transmission of FACH messages are re-allocated for use by other channels.

The method then moves to block **1160**, where the, or each, FACH indicator bit is configured as being required within a broadcast channel or indicator channel, indicating whether the FACH is active, e.g. a FACH message is to be transmitted, or not. The method then moves on to block **1170**, where the transmission of the transport channels is initiated. The method then returns to block **1120**.

A skilled artisan will appreciate that only those blocks pertinent to the invention have been illustrated and described in order not to obscure the inventive concept.

For the various embodiments hereinbefore described, the, or each, indicator bit has been described as being provided within an indicator channel. Alternatively the indicator bits may be sent on the BCH or other channel. By way of example, the indicator bit(s) may be sent within a Transport Format Combination Indicator (TFCI) type field within the BCH, wherein the TFCI field is a field that is separately encoded to data traffic to which it relates. As another example, the indicator bits may be sent in the data portion of the BCH or other channel.

In an embodiment of the invention, the indicator bit(s) may be provided within a dedicated indicator channel. The indicator bits may also be transmitted as repetition encoded bits, or alternatively they may be more robustly encoded, such as jointly convolutionally encoded.

Alternatively, indicator bits may be transmitted within an indicator channel along with other indicator bits. For example indicator bits from the PICH (Paging Indicator CHannel) or E-HICH (E-DCH Hybrid ARQ Indicator CHannel) channels may be reassigned as FACH indicator bits.

While the invention has been described in terms of particular embodiments and illustrative figures, those of ordinary skill in the art will recognize that the invention is not limited to the embodiments or figures described. Although embodiments of the present invention are described, in some instances, using UMTS terminology, those skilled in the art will recognize that such terms are also used in a generic sense herein, and that the present invention is not limited to such systems.

Those skilled in the art will recognize that the operations of the various embodiments may be implemented using hardware, software, firmware, or combinations thereof, as appropriate. For example, some processes can be carried out using processors or other digital circuitry under the control of software, firmware, or hard-wired logic. (The term 'logic' herein refers to fixed hardware, programmable logic and/or an appropriate combination thereof, as would be recognized by one skilled in the art to carry out the recited functions.) Software and firmware can be stored on computer-readable media. Some other processes can be implemented using analog circuitry, as is well known to one of ordinary skill in the art. Additionally, memory or other storage, as well as communication components, may be employed in embodiments of the invention.

FIG. **12** illustrates a typical computing system **1200** that may be employed to implement processing functionality in embodiments of the invention. Computing systems of this type may be used in the NodeB (in particular, the scheduler of the NodeB), core network elements, such as the GGSN, and RNCs, for example. Those skilled in the relevant art will also recognize how to implement the invention using other computer systems or architectures. Computing system **1200** may represent, for example, a desktop, laptop or notebook computer, hand-held computing device (PDA, cell phone, palmtop, etc.), mainframe, server, client, or any other type of special or general purpose computing device as may be desirable or appropriate for a given application or environment. Computing system **1200** can include one or more processors, such as a processor **1204**. Processor **1204** can be implemented using a general or special purpose processing engine such as, for example, a microprocessor, microcontroller or other control logic. In this example, processor **1204** is connected to a bus **1202** or other communications medium.

Computing system **1200** can also include a main memory **1208**, such as random access memory (RAM) or other dynamic memory, for storing information and instructions to be executed by processor **1204**. Main memory **1208** also may be used for storing temporary variables or other intermediate information during execution of instructions to be executed by processor **1204**. Computing system **1200** may likewise include a read only memory (ROM) or other static storage device coupled to bus **1202** for storing static information and instructions for processor **1204**.

The computing system **1200** may also include information storage system **1210**, which may include, for example, a media drive **1212** and a removable storage interface **1220**. The media drive **1212** may include a drive or other mechanism to support fixed or removable storage media, such as

a hard disk drive, a floppy disk drive, a magnetic tape drive, an optical disk drive, a compact disc (CD) or digital video disk (DVD) read or write drive (R or RW), or other removable or fixed media drive. Storage media **1218** may include, for example, a hard disk, floppy disk, magnetic tape, optical disk, CD or DVD, or other fixed or removable medium that is read by and written to by a media drive. As these examples illustrate, the storage media **1218** may include a computer-readable storage medium having stored therein particular computer software or data.

In alternative embodiments, information storage system **1210** may include other similar components for allowing computer programs or other instructions or data to be loaded into computing system **1200**. Such components may include, for example, a removable storage unit **1222** and an interface **1220**, such as a program cartridge and cartridge interface, a removable memory (for example, a flash memory or other removable memory module) and memory slot, and other removable storage units **1222** and interfaces **1220** that allow software and data to be transferred from the removable storage unit **1222** to computing system **1200**.

Computing system **1200** can also include a communications interface **1224**. Communications interface **1224** can be used to allow software and data to be transferred between computing system **1200** and external devices. Examples of communications interface **1224** can include a modem, a network interface (such as an Ethernet or other NIC card), a communications port (such as for example, a universal serial bus (USB) port), a PCMCIA slot and card, etc. Software and data transferred via communications interface **1224** are in the form of signals which can be electronic, electromagnetic, optical or other signals capable of being received by communications interface **1224**. These signals are provided to communications interface **1224** via a channel **1228**. This channel **1228** may carry signals and may be implemented using a wireless medium, wire or cable, fiber optics, or other communications medium. Some examples of a channel include a phone line, a cellular phone link, an RF link, a network interface, a local or wide area network, and other communications channels.

In this document, the terms 'computer program product' 'computer-readable medium' and the like may be used generally to refer to media such as, for example, memory **1208**, storage device **1218**, or storage unit **1222**. These and other forms of computer-readable media may store one or more instructions for use by processor **1204**, to cause the processor to perform specified operations. Such instructions, generally referred to as 'computer program code' (which may be grouped in the form of computer programs or other groupings), when executed, enable the computing system **1200** to perform functions of embodiments of the present invention. Note that the code may directly cause the processor to perform specified operations, be compiled to do so, and/or be combined with other software, hardware, and/or firmware elements (e.g., libraries for performing standard functions) to do so.

In an embodiment where the elements are implemented using software, the software may be stored in a computer-readable medium and loaded into computing system **1200** using, for example, removable storage drive **1212** or communications interface **1224**. The control logic (in this example, software instructions or computer program code), when executed by the processor **1204**, causes the processor **1204** to perform the functions of the invention as described herein.

It will be appreciated that, for clarity purposes, the above description has described embodiments of the invention with reference to different functional units and processors. However, it will be apparent that any suitable distribution of functionality between different functional units, processors or domains may be used without detracting from the invention. For example, functionality illustrated to be performed by separate processors or controllers may be performed by the same processor or controller. Hence, references to specific functional units are only to be seen as references to suitable means for providing the described functionality, rather than indicative of a strict logical or physical structure or organization.

Although the invention has been described in connection with embodiments, it is not intended to be limited to the specific form set forth herein. Rather, the scope of the present invention is limited only by the claims Additionally, although a feature may appear to be described in connection with particular embodiments, one skilled in the art would recognize that various features of the described embodiments may be combined in accordance with the invention. In the claims, the term 'comprising' does not exclude the presence of other elements or steps.

Furthermore, although individually listed, a plurality of means, elements or method steps may be implemented by, for example, a single unit or processor. Additionally, although individual features may be included in different claims, these may possibly be advantageously combined, and the inclusion in different claims does not imply that a combination of features is not feasible and/or advantageous. Also, the inclusion of a feature in one category of claims does not imply a limitation to this category, but rather the feature may be equally applicable to other claim categories, as appropriate.

Furthermore, the order of features in the claims does not imply any specific order in which the features must be performed and in particular the order of individual steps in a method claim does not imply that the steps must be performed in this order. Rather, the steps may be performed in any suitable order. In addition, singular references do not exclude a plurality. Thus, references to 'a', 'an', 'first', 'second', etc. do not preclude a plurality.

It will be understood that the method and apparatus for transmitting transport channels over a physical channel of a cellular communication system described above provides, by way of example only, at least the advantages of a more efficient use of physical resources, whilst substantially avoiding an increase in power requirements.

What is claimed is:

1. A user equipment (UE) comprising:

a receiver, the receiver configured to receive broadcast information including a plurality of bits, wherein each of the plurality of bits indicates a respective time interval that physical shared channel resources are utilized to signal channels; and

a processor, the processor communicatively coupled to the receiver and configured to monitor the signaled channels in indicated time intervals, to determine whether signaling information for the UE is present based on an identification of the UE, and to process the signaling information in response to the determination that signaling information for the UE is present.

2. The UE of claim **1**, wherein the respective time interval or each indicated time interval is part of a shared frame that is allocated by a network device.

3. The UE of claim **1**, wherein the broadcast information is received on a broadcast channel.

4. The UE of claim **1**, wherein the respective time interval or each indicated time interval is two or more timeslots.

5. The UE claim 1, wherein the UE powers down radio circuitry outside of the indicated time intervals.

6. The UE of claim 1, wherein the identification of the UE is assigned by a network device to decode the signaling information.

7. A method performed by a user equipment (UE), the method comprising:

receiving, by the UE, broadcast information including a plurality of bits, wherein each of the plurality of bits indicates a respective time interval that physical shared channel resources are utilized to signal channels;

monitoring, by the UE, the signaled channels in indicated time intervals;

determining, by the UE, whether signaling information for the UE is present based on an identification of the UE; and

processing, by the UE, the signaling information in response to the determination that signaling information for the UE is present.

8. The method of claim 7, wherein the respective time interval or each indicated time interval is part of a shared frame that is allocated by a network device.

9. The method of claim 7, wherein the broadcast information is received on a broadcast channel.

10. The method of claim 7, wherein the respective time interval or each indicated time interval is two or more timeslots.

11. The method claim 7, wherein the UE powers down radio circuitry outside of the indicated time intervals.

12. The method of claim 7, wherein the identification of the UE is assigned by a network device to decode the signaling information.

* * * * *

EXHIBIT 5

From: John Downing <JDowning@kasowitz.com>
Sent: Wednesday, December 18, 2024 12:45 PM
To: Michael Ellis; Intellectual Ventures
Cc: Matt Colvin; Bernard Cryan; Daniel Gopenko; Aamir Kazi; Kate Quisenberry; Rodeen Talebi; Ann Vermeersch; msiegmund@cjsjlaw.com
Subject: RE: Intellectual Ventures v Tesla

[This email originated outside of F&R.]

Michael,

We appreciate you reaching out. To clarify, we think that you mean claim 10 for the '743 patent below. We are looking into these issues and will let you know availability to meet and confer tomorrow when we confirm availability with local counsel, who is cc'd. Per the scheduling order, this meet and confer is meant to narrow terms in dispute and exchange revised lists of terms, which both parties have done. Our understanding of this procedure is that it is not intended to require IV to articulate claim construction positions before Defendant's opening brief and outside of the court's schedule for disclosing positions.

For the '639 and '395, out position is that the terms above are not indefinite.

For the '743, we are in the process of reviewing this issue, but cannot commit to having a response by tomorrow.

Regards,

John

John Downing
Kasowitz Benson Torres LLP
333 Twin Dolphin Drive, Suite 200
Redwood Shores, CA 94065
Tel. (650) 453-5426
Fax. (650) 362-9430
JDowning@kasowitz.com

This e-mail and any files transmitted with it are confidential and may be subject to the attorney-client privilege. Use or disclosure of this e-mail or any such files by anyone other than a designated addressee is unauthorized. If you are not an intended recipient, please notify the sender by e-mail and delete this e-mail without making a copy.

From: Michael Ellis <ellis@fr.com>
Sent: Tuesday, December 17, 2024 3:45 PM
To: Intellectual Ventures <IntellectualVentures@kasowitz.com>
Cc: Matt Colvin <colvin@fr.com>; Bernard Cryan <cryan@fr.com>; Michael Ellis <ellis@fr.com>; Daniel Gopenko <gopenko@fr.com>; Aamir Kazi <kazi@fr.com>; Kate Quisenberry <quisenberry@fr.com>; Rodeen Talebi <talebi@fr.com>; Ann Vermeersch <vermeersch@fr.com>
Subject: Intellectual Ventures v Tesla

ALERT: THIS IS AN EXTERNAL EMAIL. DO NOT CLICK ON ANY LINK, ENTER A PASSWORD, OR OPEN AN ATTACHMENT UNLESS YOU KNOW THAT THE MESSAGE CAME FROM A SAFE EMAIL ADDRESS.

Counsel,

The Proposed Scheduling Order requires the parties to meet and confer to narrow terms in dispute and exchange revised list of terms/constructions by December 19. To that end, we are available to meet and confer after 3pm Eastern on Thursday, December 19. Please let us know your availability.

Further, please be prepared to discuss what IV contends is the antecedent basis for the terms "said processing" in claim 1 of the '395 patent, "said target feature information" in claim 1 of the '639 patent, and "said feature information" in claim 7 of the '639 patent.

Please also be prepared to discuss whether or not IV agrees the preamble of claim 1 of the '743 patent is limiting. IV's statement that the plain and ordinary meaning applies is not responsive to whether or not the preamble is limiting.

Finally, to further narrow down the terms at dispute, Tesla no longer pursues an indefiniteness position for the terms "multi-level transmit power control (TPC) command" / "multi-level TPC command" in the 500 and 153 patents.

Thanks,

Michael R Ellis
Associate ■ Fish & Richardson P.C.

O: 214 760 6119 **C:** 202 607 0752 | ellis@fr.com | Bio

****
This email message is for the sole use of the intended recipient(s) and may contain confidential and privileged information. Any unauthorized use or disclosure is prohibited. If you are not the intended recipient, please contact the sender by reply email and destroy all copies of the original message.
****

EXHIBIT 6

SERIAL NUMBER	PATENT DATE	PATENT NUMBER
07/810005		

SERIAL NUMBER	FILING DATE	CLASS	SUBCLASS	GROUP ART UNIT	EXAMINER
07/810,005	12/18/91	~~395~~ 364	546	~~2301~~ 3602	~~Gorn~~ Harper Gerstl

APPLICANTS

ALAN J. KATZ, DALLAS, TX.

CONTINUING DATA********
VERIFIED

NONE

FOREIGN/PCT APPLICATIONS********
VERIFIED

NONE

BEST AVAILABLE COPY

Foreign priority claimed 35 USC 119 conditions met	Verified and Acknowledged	AS FILED	STATE OR COUNTRY	SHEETS DRWGS.	TOTAL CLAIMS	INDEP. CLAIMS	FILING FEE RECEIVED	ATTORNEY'S DOCKET NO.
☐ yes ☒ no / ☐ yes ☒ no	Examiner's initials	→	TX	14	19	4	$760.00	TI-16314

ADDRESS

RICHARD A. STOLTZ
TEXAS INSTRUMENTS, PATENT DEPT.,
MAIL STATION 219,
P. O. BOX 655474
DALLAS, TX 75265

TITLE

GENERALIZED HEBBIAN LEARNING FOR PRINCIPAL COMPONENT ANALYSIS AND AUTOMATIC TARGET RECOGNITION, SYSTEMS AND METHOD

U.S. DEPT. of COMM.-Pat. & TM Office—PTO-436L (rev. 10-78)

PARTS OF APPLICATION FILED SEPARATELY

NOTICE OF ALLOWANCE MAILED	PREPARED FOR ISSUE		CLAIMS ALLOWED	
	Assistant Examiner	Docket Clerk	Total Claims	Print Claim
ISSUE FEE			DRAWING	
Amount Due / Date Paid	Primary Examiner		Sheets Drwg. / Figs. Drwg. / Print Fig.	
	ISSUE CLASSIFICATION		ISSUE BATCH NUMBER	
	Class	Subclass		

WARNING: The information disclosed herein may be restricted. Unauthorized disclosure may be prohibited by the United States Code Title 35, Sections 122, 181 and 368. Possession outside the U.S. Patent & Trademark Office is restricted to authorized employees and contractors only.

Entered or Counted

MAIL ROOM
DEC 18 1991
32
PAT. & TRADE MARK OFF.

APPLICATION FOR U.S. PATENT
TRANSMITTAL FORM

Attorney Docket: TI-16314

December 18, 1991

Commissioner of Patents and Trademarks
Washington, D.C. 20231
Dear Sir:

"EXPRESS MAIL"
Express Mailing Label Number RB101289562
Date of Deposit December 18, 1991
I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Commissioner of Patents and Trademarks, Washington, D. C. 20231.
Donna J. Otto

Transmitted herewith for filing is the patent application of:

Inventor(s): Alan Jerry Katz

For: Object Screener For Automatic Target Recognition Based On Principal Component Features

Enclosed are 14 sheets of informal drawing and assignment of the invention to TEXAS INSTRUMENTS, INCORPORATED.

FEE CALCULATION					FEE
				Basic Fee	$690
	Number	Allowed	Extra	Rate	
Total Claims	19	20	0	20	0
Independent Claims	4	3	1	72	72
				Total Filing Fee	760
				Assignment Recording Fee	20
				TOTAL FEES	$780

Please charge deposit account No. 20-0668 in the amount of the Total Fees set forth above. The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to said deposit account. A duplicate copy of this sheet is enclosed.

CLAIMS

I claim:

1. A method for distinguishing targets from background clutter, comprising the steps of

(a) inputting data;

(b) calculating data statistics from said data and using said data statistics to select target specific feature information to distinguish specific targets from background clutter;

(c) generating said target specific feature information from said data statistics;

(d) extracting said target specific feature information from said data;

(e) using said target specific feature information to distinguish specific targets from background clutter; and

(f) outputting target and background clutter information.

2. The method of claim 1, wherein said data is one-dimensional, two-dimensional, or multi-dimensional.

3. The method of claim 1, wherein said data is digitized imagery, speech signals, or radar signals.

4. The method of claim 1, wherein parallel hardware architectures are employed to use said feature information to distinguish particular targets from background clutter.

5. The method of claim 4, wherein said parallel hardware architectures

comprise neural networks.

6. The method of claim 1, wherein said feature information is contrast-based or texture-based.

7. The method of claim 1, wherein said feature information is further used to distinguish two target classes or a plurality of target classes from one another.

8. The method of claim 1, wherein Hebbian learning techniques are used to select said target specific feature information.

9. A classification system, comprising

(a) a calculator to calculate data statistics from a data input;

(b) a selector to use said data statistics outputted from said calculator to select target specific feature information to distinguish specific targets from background clutter, wherein said target specific feature information uniquely describes a target for recognition purposes;

(c) a generator to generate said target specific feature information from said data statistics outputted from said calculator;

(d) a feature extractor to extract target specific feature information from said data input; and

(e) a classifier to use said target specific feature information extracted from said data input by said feature extractor to distinguish targets from background clutter and output target and background clutter data.

10. The classification system of claim 9, wherein said data input is one-dimensional, two-dimensional, or multidimensional.

11. The classification system of claim 9, wherein said data input is digitized imagery, speech signals or radar signals.

12. The classification system of claim 9, wherein parallel hardware architectures are employed in the classifier to use said feature information to distinguish particular targets from background clutter.

13. The classification system of claim 12, wherein said parallel hardware architectures comprise neural networks.

14. The classification system of claim 9, wherein said feature information is contrast-based or texture-based.

15. The classification system of claim 9, wherein said selector employs Hebbian learning techniques to select said target specific feature information.

16. A classification hardware system, comprising

(a) a calculator to calculate data statistics from a data input;

(b) a selector to use said data statistics outputed from said calculator to select target specific feature information to distinguish specific targets from background clutter, wherein said target specific feature information uniquely describes a target for recognition purposes;

(c) a generator to generate said target specific feature information from said data statistics outputed from said calculator;

(d) a feature extractor to extract target specific feature information from said data input; and

(e) a classifier to use said target specific feature information extracted

from said data input by said feature extractor to distinguish targets from background clutter and output target and background clutter data.

17. The classification hardware system of claim 16, wherein said selector employs Hebbian learning techniques to select said target specific feature information.

18. A classification software system, comprising

(a) calculation software to calculate data statistics from a data input;

(b) selection software to use said data statistics outputed from said calculator to select target specific feature information to distinguish specific targets from background clutter, wherein said target specific feature information uniquely describes a target for recognition purposes;

(c) generator software to generate said target specific feature information from said data statistics outputed from said calculator;

(d) feature extraction software to extract target specific feature information from said data input; and

(e) classification software to use said target specific feature information extracted from said data input by said feature extractor to distinguish targets from background clutter and output target and background clutter data.

19. The classification software system of claim 18, wherein said selector employs Hebbian learning techniques to select said target specific feature information.

ATTORNEY DOCKET
004578.1360

PATENT
07/810,005

OIPE
MAY 28 2004
PATENT & TRADEMARK OFFICE

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:	Alan Jerry Katz
Serial No.:	07/810,005
Filed:	December 18, 1991
Confirmation No.:	2632
Group No.:	3662
Examiner:	Bernarr E. Gregory
Title:	GENERALIZED HEBBIAN LEARNING FOR PRINCIPAL COMPONENT ANALYSIS AND AUTOMATIC TARGET RECOGNITION, SYSTEMS AND METHODS

RECEIVED
JUN 0 7 2004
GROUP 3600

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

"EXPRESS MAIL"

Express Mailing Label Number
EV 324642475 US
I hereby certify that this paper is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated below and is addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Willie Jiles
Willie Jiles

Date: May 28, 2004

Dear Sir:

RESPONSE PURSUANT TO 37 C.F.R. 1.111

In response to the Office Action mailed March 3, 2004, Applicant respectfully requests the Examiner to reconsider the rejection of the claims in view of the following amendments thereto, and the comments as set forth below. Please amend the application as follows:

In the Claims

1. (Original) A method for distinguishing targets from background clutter, comprising the steps of

(a) inputting data;

(b) calculating data statistics from said data and using said data statistics to select target specific feature information to distinguish specific targets from background clutter;

(c) generating said target feature information from said data statistics;

(d) extracting said target specific feature information from said data;

(e) using said target specific feature information to distinguish specific targets from background clutter; and

(f) outputting target and background clutter information.

2. (Original) The method of claim 1, wherein said data is one-dimensional, two-dimensional, or multi-dimensional.

3. (Original) The method of claim 1, wherein said data is digitized imagery, speech signals, or radar signals.

4. (Original) The method of claim 1, wherein parallel hardware architectures are employed to use said feature information to distinguish particular targets from background clutter.

5. (Original) the method of claim 4, wherein said parallel hardware architectures comprise neural networks.

6. (Original) The method of claim 1, wherein said feature information is contrast-based or texture-based.

7. (Original) The method of claim 1, wherein said feature information is further used to distinguish two target classes or a plurality of target classes from one another.

8. (Original) The method of claim 1, wherein Hebbian learning techniques are used to select said target specific feature information.

9. (Original) A classification system, comprising

(a) a calculator to calculate data statistics from a data input;

(b) a selector to use said data statistics outputed from said calculator to select target specific feature information to distinguish specific targets from background clutter, wherein said target specific feature information uniquely describes a target for recognition purposes;

(c) a generator to generate said target specific feature information from said data statistics outputed from said calculator;

(d) a feature extractor to extract target specific feature information from said data input; and

(e) a classifier to use said target specific feature information extracted from said data input by said feature extractor to distinguish targets from background clutter and output target and background clutter data.

10. (Original) The classification system of claim 9, wherein said data input is one-dimensional, two-dimensional, or multidimensional.

11. (Original) The classification system of claim 9, wherein said data input is digitized imagery, speech signals or radar signals.

12. (Original) The classification system of claim 9, wherein parallel hardware architectures are employed in the classifier to use said feature information to distinguish particular targets from background clutter.

13. (Original) The classification system of claim 12, wherein said parallel hardware architectures comprise neural networks.

14. (Original) The classification system of claim 9, wherein said feature information is contrast-based or texture-based.

15. (Original) The classification system of claim 9, wherein said selector employs Hebbian learning techniques to select said target specific feature information.

16. (Original) A classification hardware system, comprising

(a) a calculator to calculate data statistics from a data input;

(b) a selector to use said data statistics outputed from said calculator to select target specific feature information to distinguish specific targets from background clutter, wherein said target specific feature information uniquely describes a target for recognition purposes;

(c) a generator to generate said target specific feature information from said data statistics outputed from said calculator;

(d) a feature extractor to extract target specific feature information from said data input; and

(e) a classifier to use said target specific feature information extracted from said data input by said feature extractor to distinguish targets from background clutter and output target and background clutter data.

17. (Original) The classification hardware system of claim 16, wherein said selector employs Hebbian learning techniques to select said target specific feature information.

18. (Currently Amended) A classification software system, comprising

(a) calculation software to calculate data statistics from a data input;

(b) selection software to use said data statistics outputed from said calculator to select target specific feature information to distinguish specific targets from background clutter, wherein said target feature information uniquely describes a target for recognition purposes;

(c) generator software to generate said target specific feature information from said data statistics outputed from said calculator;

(d) feature extraction software to extract target specific feature information from said data input; ~~and~~

(e) classification software to use said target specific feature information extracted from said data input by said feature extractor to distinguish targets from background clutter and output target and background clutter ~~data.~~ data; and

(f) wherein the calculation software, selection software, generator software, feature extraction software, and classification software reside on at least one computer-readable medium.

19. (Original) The classification software system of claim 18, wherein said selector employs Hebbian learning techniques to select said target specific feature information.

20. (Previously presented) A method for distinguishing targets from background clutter, comprising the steps of

(a) inputting data having targets and background clutter;

(b) applying neural network learning procedures to said data to construct a set of operators;

(c) applying said set of operators to data to construct a set of features;

(d) choosing optimal features from said set of features to distinguish targets from clutter; and

(e) using said optimal features to distinguishing targets from background clutter.

EXHIBIT 7

UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF TEXAS
WACO DIVISION

MIRROR IMAGING, LLC, Plaintiff v. **BANK OF AMERICA, N.A.,** Defendant	**Case No. 6:21-cv-00463-ADA** **JURY TRIAL DEMANDED**
MIRROR IMAGING, LLC, Plaintiff v. **JPMORGAN CHASE BANK, N.A.,** Defendant	**Case No. 6:21-cv-00464-ADA** **JURY TRIAL DEMANDED**
MIRROR IMAGING, LLC, Plaintiff v. **WELLS FARGO BANK, N.A.,** Defendant	**Case No. 6:21-cv-00465-ADA** **JURY TRIAL DEMANDED**

MIRROR IMAGING, LLC, Plaintiff v. **CAPITAL ONE, N.A.,** Defendant	**Case No. 6:21-cv-00517-ADA** **JURY TRIAL DEMANDED**
MIRROR IMAGING, LLC, Plaintiff v. **PNC BANK, N.A.,** Defendant	**Case No. 6:21-cv-00518-ADA** **JURY TRIAL DEMANDED**

CLAIM CONSTRUCTION ORDER

The Court provided preliminary constructions to the parties in advance of the *Markman* hearing scheduled for February 14, 2022. In advance of that hearing, the parties indicated that all parties had met and conferred, and all parties stipulated to the preliminary constructions as provided. In light of the communication from the parties, the *Markman* hearing was cancelled.

SIGNED this 22nd day of February, 2022.

ALAN D ALBRIGHT
UNITED STATES DISTRICT JUDGE

I. Disputed Constructions

Term [Proposing Party] [Patents/Claims]	Plaintiff's Proposed Construction	Defendants' Proposed Construction	Court's Preliminary Constructions
"electronic image[s]" ['275 Patent, Claims 1, 5, 8; '435 Patent: Claims 11, 13; '009 Patent: Claim 14; '447 Patent: Claims 1, 6, 7, 19] [Defendant JPM and PNC]	Plain and ordinary meaning. In the alternative: "a digitized record of a document."	"digital representation of a photograph or photocopy of a document"	**Plain and ordinary meaning**
"associated with a [first/second] entity" ['275 Patent: Claim 1; '435 Patent: Claim 11; '009 Patent: Claim 14; '447 Patent: Claims 1, 19] [Defendants JPM and PNC]	Plain and ordinary meaning.	"owned or operated by a [first/second] entity"	**Plain and ordinary meaning**
"located remotely from" ['275 Patent: Claim 1; '009 Patent: Claim 14; '447 Patent: Claims 1, 19] [Defendants BOA, WF, CO, JPM, and PNC]	Plain and ordinary meaning.	"geographically distanced from"	**Plain and ordinary meaning**
"store at least some of" / "storing at least some of" / "storing the at least some of" [275 Patent: Claims 1, 5; '435 Patent: Claim 11; '009 Patent: Claim 14; '447 Patent: Claims 1, 19] [Defendants BOA, WF, CO, and JPM]	Plain and ordinary meaning.	Indefinite.	**Not indefinite. Plain and ordinary meaning.**

"first entity" ['435 Patent: Claim 11; '009 Patent: Claim 14; '447 Patent: Claims 1, 19] [Defendants JPM and PNC]	Plain and ordinary meaning.	Indefinite.	**Not indefinite. Plain and ordinary meaning.**
"second entity" ['435 Patent: Claim 11; '009 Patent: Claim 14; '447 Patent: Claims 1, 19] [Defendants JPM and PNC]	Plain and ordinary meaning.	Indefinite.	**Not indefinite. Plain and ordinary meaning.**
"wherein the first entity and the second entity are distinct or related" ['435 Patent: Claim 11] [Defendants BOA, WF, CO, JPM, and PNC]	Plain and ordinary meaning.	Indefinite.	**Indefinite**
"configured to at least partially link two or more interfaces" ['447 Patent: Claim 11] [Defendants BOA, WF, CO, JPM, and PNC]	Plain and ordinary meaning.	Indefinite.	**Indefinite**

//

EXHIBIT 8

A Method for Registration of 3-D Shapes

Paul J. Besl, *Member, IEEE*, and Neil D. McKay

***Abstract*—This paper describes a general-purpose, representation-independent method for the accurate and computationally efficient registration of 3-D shapes including free-form curves and surfaces. The method handles the full six degrees of freedom and is based on the iterative closest point (ICP) algorithm, which requires only a procedure to find the closest point on a geometric entity to a given point. The ICP algorithm always converges monotonically to the nearest local minimum of a mean-square distance metric, and experience shows that the rate of convergence is rapid during the first few iterations. Therefore, given an adequate set of initial rotations and translations for a particular class of objects with a certain level of "shape complexity," one can globally minimize the mean-square distance metric over all six degrees of freedom by testing each initial registration. For example, a given "model" shape and a sensed "data" shape that represents a major portion of the model shape can be registered in minutes by testing one initial translation and a relatively small set of rotations to allow for the given level of model complexity. One important application of this method is to register sensed data from unfixtured rigid objects with an ideal geometric model prior to shape inspection. The described method is also useful for deciding fundamental issues such as the congruence (shape equivalence) of different geometric representations as well as for estimating the motion between point sets where the correspondences are not known. Experimental results show the capabilities of the registration algorithm on point sets, curves, and surfaces.**

***Index Terms*— Free-form curve matching, free-form surface matching, motion estimation, pose estimation, quaternions, 3-D registration.**

I. Introduction

GLOBAL AND local shape matching metrics for free-form curves and surfaces as well as point sets were described in [3] in an attempt to formalize and unify the description of a key problem in computer vision: Given 3-D data in a sensor coordinate system, which describes a data shape that may correspond to a model shape, and given a model shape in a model coordinate system in a different geometric shape representation, estimate the optimal rotation and translation that aligns, or registers, the model shape and the data shape minimizing the distance between the shapes and thereby allowing determination of the equivalence of the shapes via a mean-square distance metric. Of key interest to many applications is the following question: Does a segmented region from a range image match a subset of B-spline surfaces on a computer-aided-design (CAD) model? This paper provides a solution to this free-form surface matching problem as defined in [3] and [5] as a special case of a simple, general, unified approach, which generalizes to n dimensions and provides solutions to 1) the point-set matching problem without correspondence and 2) the free-form curve matching problem. The algorithm requires no extracted features, no curve or surface derivatives, and no preprocessing of 3-D data, except for the removal of statistical outliers.

The main application of the proposed method as described here is to register digitized data from unfixtured rigid objects with an idealized geometric model prior to shape inspection. When inspecting shapes using high-accuracy noncontact measurement devices [4] over a shallow depth of field, the uncertainty in different sensed points does not vary by much. Therefore, for purposes of simplicity and relevance to inspection applications based on thousands of digitized points, the case of unequal uncertainty among points is not considered. Similarly, the removal of statistical outliers is considered a preprocessing step, is probably best implemented as such, and will also not be addressed. In the context of inspection applications, the assumption that a high-accuracy noncontact measurement device does not generate bad data is reasonable since some sensors have the ability to reject highly uncertain measurements.

The proposed shape registration algorithm can be used with the following representations of geometric data:

1) Point sets
2) line segment sets (polylines)
3) implicit curves: $\vec{g}(x, y, z) = 0$
4) parametric curves: $(x(u), y(u), z(u))$
5) triangle sets (faceted surfaces)
6) implicit surfaces: $g(x, y, z) = 0$
7) parametric surfaces: $(x(u, v), y(u, v), z(u, v))$.

This covers most applications that would utilize a method to register 3-D shapes. Other representations are handled by providing a procedure for evaluating the closest point on the given shape to a given digitized point.

This paper is structured as follows: Several relevant papers from the literature are first reviewed. Next, the mathematical preliminaries of computing the closest point on a shape to a given point are covered for the geometric representations mentioned above. Then, the iterative closest point (ICP) algorithm is stated, and a theorem is proven concerning its monotonic convergence property. The issue of the initial registration states is addressed next. Finally, experimental results for point sets, curves, and surfaces are presented to demonstrate the capabilities of the ICP registration algorithm.

II. Literature Review

Relatively little work has been published in the area of registration (pose estimation, alignment, motion estimation) of 3-D

Manuscript received October 30, 1990; revised May 6, 1991.

The authors are with the Computer Science Department, General Motors Research Laboratories, Warren, MI 48090-9055.

IEEE Log Number 9102686.

0162-8828/92$03.00 © 1992 IEEE

Case 1:24-cv-00884-ADA Document 26-9 Filed 01/09/25 Page 3 of 19

free-form shapes. Most of the existing literature addressing global shape matching or registration have addressed limited classes of shapes, namely, 1) polyhedral models, 2) piecewise-(super)quadric models [2], [27], and 3) point sets with known correspondence. The reader may consult [6] and [14] for pre-1985 work in these areas. For a sampling of other more recent related work not discussed below, see [8], [10], [12], [13], [19], [20], [24], [26], [34], [35], [37], [39], [44], [46], [48], [53], [58], [59].

Historically, free-form shape matching using 3-D data was done earliest by Faugeras and his group at INRIA [18], where they demonstrated effective matching with a Renault auto part (steering knuckle) in the early 1980's. This work popularized the use of quaternions for least squares registration of corresponding 3-D point sets in the computer vision community. The alternative use of the singular value decomposition (SVD) algorithm [23], [1], [49] was not as widely known in this time frame. The primary limitation of this work was that it relied on the probable existence of reasonably large planar regions within a free-form shape.

Schwartz and Sharir [50] developed a solution to the free-form space curve matching problem without feature extraction in late 1985. They used a nonquaternion approach to computing the least squares rotation matrix. The method works well with reasonable quality curve data but has difficulty with very noisy curves because the method uses arclength sampling of the curves to obtain corresponding point sets.

Haralick *et al.* [28] addressed the 3-D point-set pose estimation problem using robust methods combined with the least squares SVD registration approach, which provided a robust statistical alternative to the least squares quaternion or SVD point set matching. This algorithm is able to handle statistical outliers and could theoretically be substituted for our quaternion-based algorithm as long as the determinant of the orthonormal matrix is strictly a positive one. A recent conference proceedings [47] contains new contributions on this subject.

Horn [31] derived an alternative formulation of Faugeras's method [18] of least squares quaternion matching that uses the maximum eigenvalue of a 4×4 matrix instead of the minimum eigenvalue. Horn [30] and Brou [11] also developed the extended Gaussian image (EGI) methods allowing the matching of convex and restricted sets of nonconvex shapes based on surface normal histograms.

Taubin [55] has done some interesting work in the area of implicit algebraic nonplanar 3-D curve and surface estimation with applications to position estimation without feature extraction. He describes a method of approximating data points with implicit algebraic forms up to the tenth degree using an approximate distance metric. Global shapes (not occluded shapes) can be identified based on generalized eigenvalues, and the registration transformation can be recovered. The method is shown to be useful for complete planar curve and space curve shapes, but it is unclear that the effectiveness generalizes well to more complicated surfaces, such as terrain data or a human face. Taubin has stated that the numerical methods of the approximate distance fit tend to break down above the tenth degree. He later [56] extended his work in shape description by investigating shape matching based on generalized shape polynomials. This demonstrated some interesting theoretical results but remains to be demonstrated for practical use on complex surfaces.

Szeliski [54] also describes a method for estimating motion from sparse range data without correspondence between the points and without feature extraction. His primary goal was to create a method for estimating the motion of the observer between two range image frames of the same terrain. Given the set of points from one frame, he applies a smoothness assumption to create a smoothing spline approximation of the points. Then, a conventional steepest descent algorithm is used to rotate and translate the second data set so that it minimizes the sum of the covariance-weighted z differences between the points and the surface. His approach is based on a regular xy-grid structure, and true 3-D point-to-surface distances are not computed. The steepest-descent approach is a slower alternative to reaching the local minima than our proposed ICP algorithm described below. Szeliski uses optimal Bayesian mathematics to allow him to downweight noisier values at longer ranges from a simulated range finder. For navigation range imaging sensors, the uncertainty in data points vary significantly from the foreground to the background. For high-accuracy sensors with shallow depths of field, the uncertainty variation between points is orders of magnitude less and is of much less concern. Szeliski provides experimental results for synthetic terrain data and a block. The terrain data motion test was a simple translation along one axis: a 1-D correlation problem. His block test did involve six degrees of freedom, but the block is a very simple shape. Overall, this work presents some interesting ideas, but the experimental results are unconvincing for applications.

Horn and Harris [33] also addressed the problem of estimating the exact rigid-body motion of the observer given sequentially digitized range image frames of the same terrain. They describe a range rate constraint equation and an elevation rate constraint equation. The result is a noniterative least squares method that provides a six-degree-of-freedom motion estimate as long as the motion between frames of data is relatively small. This method is much quicker than the one proposed by Szeliski, but it is not clear that this method generalizes to arbitrary rotations and translations of a shape.

Kamgar-Parsi *et al.* [36] also describe a method for the registration of multiple overlapping range images without distinctive feature extraction. This method works very well using the level sets of 2.5-D range data but is essentially restricted to the three degrees of freedom in the plane since the work was addressed toward piecing together terrain map data.

Li [38] addressed free-form surface matching with arbitrary rotations and translations. His method forms an attributed relational graph of fundamental surface regions for data and model shapes and then performs graph matching using an inexact approach that allows for variability in attributes as well as in graph adjacency relationships. This seems to be a reasonable approach but relies on extraction of derivative-based quantities. Experimental results are shown for a coffee cup and the Renault auto part; see also Wong *et al.* [60] for other related work using attributed graphs for 3-D matching.

TESLAIV000023925

The work of Gilbert and Foo [21] and Gilbert *et al.* [22] is related in that it addresses the computation of distance between two object shapes. Such methods could be the basis for similar shape matching techniques as are described below. The major inconvenience with their method, though, is that object shapes must be decomposed into convex subbodies, which is a problem that, in general, is not trivial for arbitrary CAD models or for digitized 3-D data.

III. Mathematical Preliminaries

In this section, methods for computing the closest point to a given point on the various geometric representations listed above are described. First, the basic geometric entities are covered, followed by parametric entities, and, finally, implicit entities. The reader might consult Mortenson [42] on some of the items below for additional information.

The Euclidean distance $d(\vec{r}_1, \vec{r}_2)$ between the two points $\vec{r}_1 = (x_1, y_1, z_1)$ and $\vec{r}_2 = (x_2, y_2, z_2)$ is $d(\vec{r}_1, \vec{r}_2) = \|\vec{r}_1 - \vec{r}_2\| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$. Let A be a point set with N_a points denoted $\vec{a}_i$: $A = \{\vec{a}_i\}$ for $i = 1, \ldots, N_a$. The distance between the point $\vec{p}$ and the point set A is

$$d(\vec{p}, A) = \min_{i \in \{1, \ldots, N_a\}} d(\vec{p}, \vec{a}_i). \tag{1}$$

The closest point $\vec{a}_j$ of A satisfies the equality $d(\vec{p}, \vec{a}_j) = d(\vec{p}, A)$.

Let l be the line segment connecting the two points $\vec{r}_1$ and $\vec{r}_2$. The distance between the point $\vec{r}$ and the line segment l is

$$d(\vec{p}, l) = \min_{u+v=1} \|u\vec{r}_1 + v\vec{r}_2 - \vec{p}\| \tag{2}$$

where $u \in [0, 1]$ and $v \in [0, 1]$. The required closed-form computations are straightforward. Let L be the set of N_l line segments denoted l_i, and let $L = \{l_i\}$ for $i = 1, \ldots, N_l$. The distance between the point $\vec{p}$ and the line segment set L is

$$d(\vec{p}, L) = \min_{i \in \{1, \ldots, N_l\}} d(\vec{p}, l_i). \tag{3}$$

The closest point $\vec{y}_j$ on the line segment set L satisfies the equality $d(\vec{p}, \vec{y}_j) = d(\vec{p}, L)$.

Let t be the triangle defined by the three points $\vec{r}_1 = (x_1, y_1, z_1)$, $\vec{r}_2 = (x_2, y_2, z_2)$, and $\vec{r}_3 = (x_3, y_3, z_3)$. The distance between the point $\vec{p}$ and the triangle t is

$$d(\vec{p}, t) = \min_{u+v+w=1} \|u\vec{r}_1 + v\vec{r}_2 + w\vec{r}_3 - \vec{p}\| \tag{4}$$

where $u \in [0, 1]$, $v \in [0, 1]$, and $w \in [0, 1]$. The required closed-form computations are again straightforward. Let T be the set of N_t triangles denoted t_i, and let $T = \{t_i\}$ for $i = 1, \ldots, N_t$. The distance between the point $\vec{p}$ and the triangle set T is given by

$$d(\vec{p}, T) = \min_{i \in \{1, \ldots, N_t\}} d(\vec{p}, t_i). \tag{5}$$

The closest point $\vec{y}_j$ on the triangle set T satisfies the equality $d(\vec{p}, \vec{y}_j) = d(\vec{p}, T)$.

A. Point to Parametric Entity Distance

In this section, a parametric curve and a parametric surface are treated as a single parametric entity $\vec{r}(\vec{u})$, where $\vec{u} = u \in \mathcal{R}^1$ should be substituted for parametric curves, and $\vec{u} = (u, v) \in \mathcal{R}^2$ should be substituted for parametric surfaces ($\mathcal{R}$ denotes the real line). The evaluation domain for a curve is an interval, but the evaluation domain for a surface can be any arbitrary closed-connected region in the plane. For more information on parametric entities, such as Bezier and B-spline curves and surfaces, see [9], [15]–[17], [42], [52].

The distance from a given point $\vec{p}$ to a parametric entity E is

$$d(\vec{p}, E) = \min_{\vec{r}(\vec{u}) \in E} d(\vec{p}, \vec{r}(\vec{u})). \tag{6}$$

The computations to compute the distance are not closed form and are relatively involved. One method for computing point-to-curve and point-to-surface distances is described below. Sets of parametric entities are again straightforward once the distance metric for an individual entity is implemented. Let F be the set of N_e parametric entities denoted E_i, and let $F = \{E_i\}$ for $i = 1, N_e$. The distance between a point $\vec{p}$ and the parametric entity set F is

$$d(\vec{p}, F) = \min_{i \in \{1, \ldots, N_e\}} d(\vec{p}, \vec{E}_i). \tag{7}$$

The closest point $\vec{y}_j$ on the parametric entity set F satisfies the equality $d(\vec{p}, \vec{y}_j) = d(\vec{p}, F)$.

Our first step towards computing the distance from a point to a parametric entity is creating a simplex-based approximation (line segments or triangles). For a parametric space curve $C = \{\vec{r}(u)\}$, one can compute a polyline $L(C, \delta)$ such that the piecewise-linear approximation never deviates from the space curve by more than a prespecified distance δ. By tagging each point of the polyline with its corresponding u argument values of the parametric curve, one can obtain an estimate of the u_a argument value of the closest point from the line segment set.

Similarly, for a parametric surface $S = \{\vec{r}(u, v)\}$, one can compute a triangle set $T(S, \delta)$ such that the piecewise-triangular approximation never deviates from the surface by more than a prespecified distance δ. By tagging each triangle vertex with the corresponding (u, v) argument values of the parametric surface, one can obtain an estimate (u_a, v_a) of the argument values of the closest point from the triangle set. As a result of these curve and surface procedures, one can assume that an initial value $\vec{u}_a$ is available such that $\vec{r}(\vec{u}_a)$ is very close to the closest point on the parametric entity.

The point-to-parametric-entity distance problem is ideal for employing a pure Newton's minimization approach when a reliable starting point $\vec{u}_a$ is available. The scalar objective function to be minimized is

$$f(\vec{u}) = \|\vec{r}(\vec{u}) - \vec{p}\|^2. \tag{8}$$

Let $\nabla = [\partial/\partial\vec{u}]^t$ be the vector differential gradient operator (where t implies vector transpose). The minimum of f occurs when $\nabla f = 0$. When the parametric entity is a surface, the 2-D gradient vector is $\nabla f = [f_u, f_v]^t$, and the 2-D Hessian matrix

TESLAIV000023926

Case 1:24-cv-00884-ADA Document 26-9 Filed 01/09/25 Page 5 of 19

is

$$\nabla\nabla^t(f) = \begin{bmatrix} f_{uu} & f_{uv} \\ f_{uv} & f_{vv} \end{bmatrix} \tag{9}$$

where the partial derivatives of the objective function are given by

$$f_u(\vec{u}) = 2\vec{r}_u^t(\vec{u})(\vec{r}(\vec{u}) - \vec{p}) \tag{10}$$
$$f_v(\vec{u}) = 2\vec{r}_v^t(\vec{u})(\vec{r}(\vec{u}) - \vec{p}) \tag{11}$$
$$f_{uu}(\vec{u}) = 2\vec{r}_{uu}^t(\vec{u})(\vec{r}(\vec{u}) - \vec{p}) + 2\vec{r}_u^t(\vec{u})\vec{r}_u(\vec{u}) \tag{12}$$
$$f_{vv}(\vec{u}) = 2\vec{r}_{vv}^t(\vec{u})(\vec{r}(\vec{u}) - \vec{p}) + 2\vec{r}_v^t(\vec{u})\vec{r}_v(\vec{u}) \tag{13}$$
$$f_{uv}(\vec{u}) = 2\vec{r}_{uv}^t(\vec{u})(\vec{r}(\vec{u}) - \vec{p}) + 2\vec{r}_u^t(\vec{u})\vec{r}_v(\vec{u}). \tag{14}$$

The curve case requires only computation of f_u and f_{uu}. The Newton's update formula for either entity is

$$\vec{u}_{k+1} = \vec{u}_k - [\nabla\nabla^t(f)(\vec{u}_k)]^{-1}\nabla f(\vec{u}_k) \tag{15}$$

where $\vec{u}_0 = \vec{u}_a$. When using the starting point selection method described above based on a simplex approximation with a reasonably small δ, Newton's method for computing the closest point generally converges in one to five iterations and typically in three. The computational cost of Newton's method is very low in contrast with finding good starting points.

B. Point to Implicit Entity Distance

An implicit geometric entity is defined as the zero set of a possibly vector-valued multivariate function $\vec{g}(\vec{r}) = 0$. The distance from a given point $\vec{p}$ to an implicit entity I is

$$d(\vec{p}, I) = \min_{\vec{g}(\vec{r})=0} d(\vec{p}, \vec{r}) = \min_{\vec{g}(\vec{r})=0} \|\vec{r} - \vec{p}\|. \tag{16}$$

The calculations to compute this distance are also not closed form and are relatively involved. One method for computing point-to-curve and point-to-surface distances is outlined below. Sets of implicit entities are straightforward once the distance metric for an individual entity is implemented. Let J be the set of N_I parametric entities denoted I_k and $J = \{I_k\}$ for $k = 1, N_I$. The distance between a point $\vec{p}$ and the implicit entity set J is

$$d(\vec{p}, J) = \min_{k\in\{1,\ldots,N_I\}} d(\vec{p}, \vec{I}_k). \tag{17}$$

The closest point $\vec{y}_j$ on the implicit entity I_j satisfies the equality $d(\vec{p}, \vec{y}_j) = d(\vec{p}, J)$.

Our first step towards computing the distance from a point to an implicit entity is creating a simplex-based approximation (line segments or triangles) as was done for parametric entities [7]. Computing the point-to-line set or point-to-triangle set distance yields an approximate closest point $\vec{r}_a$, which can be used to compute the exact distance.

The implicit entity distance problem is quite different from the parametric entity case where unconstrained optimization suffices. To find the closest point on an implicit entity defined by $\vec{g}(\vec{r}) = 0$ to a given point $\vec{p}$, one must solve a constrained optimization problem to minimize a quadratic objective function subject to a nonlinear constraint

$$\text{Minimize } f(\vec{r}) = \|\vec{r} - \vec{p}\|^2 \text{ where } \vec{g}(\vec{r}) = 0. \tag{18}$$

One approach to this problem is to form the augmented Lagrange multiplier system of equations [40]:

$$\begin{aligned} \nabla f(\vec{r}) + \vec{\lambda}^t\nabla\vec{g}(\vec{r}) &= 0 \\ \vec{g}(\vec{r}) &= 0 \end{aligned} \tag{19}$$

where $\nabla = [\partial/\partial\vec{r}]^t$ and solve this system of nonlinear equations via numerical methods. The number of equations and unknowns for the nonlinear system is three for planar curves, four for surfaces, and five for implicitly defined space curves. Continuation methods [41] can be used to solve this problem for algebraic entities even without a good starting point, but a good starting point will allow the use of faster methods, such as the multidimensional Newton's root finding method. From a numerical point of view, the parametric methods are much easier to deal with. From an applied point of view, no industrial CAD systems store free-form curves or surfaces in implicit form. For this reason, implicit surfaces of interest are dealt with in our implemented system either via special case mathematics (e.g., spheres) or via a parametric form. Of course, if there were an application where it was necessary to handle free-form implicit entities in their implicit form [51], the above algorithm could be implemented.

Taubin [55] uses an approximate distance algorithm that implies a simple update formula for surfaces and planar curves when $g(\vec{r}_0)$ is nearly zero:

$$\vec{r}_{k+1} = \vec{r}_k - \frac{\nabla g(\vec{r}_k) g(\vec{r}_k)}{\|\nabla g(\vec{r}_k)\|^2}. \tag{20}$$

This method is only exact if the infinite line with the direction $\nabla g(\vec{r})$ at the starting point $\vec{r}_0$ intersects the implicit entity at a point where the normal vector has that same direction. This is not true in general, and the approximation is generally worse the further the point is from the implicit entity. Therefore, this result cannot be used if precise distance results are required.

C. Corresponding Point Set Registration

All closest point (minimum distance) algorithms have been mentioned in forms that generalize to n dimensions. One more necessary procedure for yielding the least squares rotation and translation is reviewed. For our purposes, the quaternion-based algorithm is preferred over the singular value decomposition (SVD) method in two and three dimensions since reflections are not desired. The SVD approach, based on the cross-covariance matrix of two point distributions, does, however, generalize easily to n dimensions and would be our method of choice for $n > 3$ in any n-dimensional applications. The basic solution of Horn [31] is described below, although the method of Faugeras [18] is equivalent. Our summary stresses the role of the SVD cross-covariance matrix, which is an important relationship not discussed in other work.

TESLAIV000023927

The unit quaternion is a four vector $\vec{q}_R = [q_0 q_1 q_2 q_3]^t$, where $q_0 \geq 0$, and $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$. The 3×3 rotation matrix generated by a unit rotation quaternion is found at the bottom of this page. Let $\vec{q}_T = [q_4 q_5 q_6]^t$ be a translation vector. The complete registration state vector $\vec{q}$ is denoted $\vec{q} = [\vec{q}_R | \vec{q}_T]^t$. Let $P = \{\vec{p}_i\}$ be a measured data point set to be aligned with a model point set $X = \{\vec{x}_i\}$, where $N_x = N_p$ and where each point $\vec{p}_i$ corresponds to the point $\vec{x}_i$ with the same index. The mean square objective function to be minimized is

$$f(\vec{q}) = \frac{1}{N_p} \sum_{i=1}^{N_p} \|\vec{x}_i - \boldsymbol{R}(\vec{q}_R)\vec{p}_i - \vec{q}_T\|^2. \tag{22}$$

The "center of mass" $\vec{\mu}_p$ of the measured point set P and the center of mass $\vec{\mu}_x$ for the X point set are given by

$$\vec{\mu}_p = \frac{1}{N_p} \sum_{i=1}^{N_p} \vec{p}_i \quad \text{and} \quad \vec{\mu}_x = \frac{1}{N_x} \sum_{i=1}^{N_x} \vec{x}_i. \tag{23}$$

The cross-covariance matrix Σ_{px} of the sets P and X is given by

$$\Sigma_{px} = \frac{1}{N_p} \sum_{i=1}^{N_p} [(\vec{p}_i - \vec{\mu}_p)(\vec{x}_i - \vec{\mu}_x)^t] = \frac{1}{N_p} \sum_{i=1}^{N_p} [\vec{p}_i \vec{x}_i^t] - \vec{\mu}_p \vec{\mu}_x^t. \tag{24}$$

The cyclic components of the anti-symmetric matrix $A_{ij} = (\Sigma_{px} - \Sigma_{px}^T)_{ij}$ are used to form the column vector $\Delta = [A_{23} \quad A_{31} \quad A_{12}]^T$. This vector is then used to form the symmetric 4×4 matrix $Q(\Sigma_{px})$

$$Q(\Sigma_{px}) = \begin{bmatrix} \operatorname{tr}(\Sigma_{px}) & \Delta^T \\ \Delta & \Sigma_{px} + \Sigma_{px}^T - \operatorname{tr}(\Sigma_{px}) \boldsymbol{I}_3 \end{bmatrix} \tag{25}$$

where $\boldsymbol{I}_3$ is the 3×3 identity matrix. The unit eigenvector $\vec{q}_R = [q_0 \quad q_1 \quad q_2 \quad q_3]^t$ corresponding to the maximum eigenvalue of the matrix $Q(\Sigma_{px})$ is selected as the optimal rotation. The optimal translation vector is given by

$$\vec{q}_T = \vec{\mu}_x - \boldsymbol{R}(\vec{q}_R)\vec{\mu}_p. \tag{26}$$

This least squares quaternion operation is $O(N_p)$ and is denoted as

$$(\vec{q}, d_{ms}) = \mathcal{Q}(P, X) \tag{27}$$

where d_{ms} is the mean square point matching error. The notation $\vec{q}(P)$ is used to denote the point set P after transformation by the registration vector $\vec{q}$.

IV. The Iterative Closest Point Algorithm

Now that the methods for computing the closest point on a geometric shape to a given point and for computing a least squares registration vector have been outlined, the ICP algorithm can be described in terms of an abstract geometric shape X whose internal representation must be known to execute the algorithm but is not of concern for this discussion. Thus, all that follows applies equally well to 1) sets of points, 2) sets of line segments, 3) sets of parametric curves, 4) sets of implicit curves, 5) sets of triangles, 6) sets of parametric surfaces, and 7) sets of implicit surfaces.

In the description of the algorithm, a "data" shape P is moved (registered, positioned) to be in best alignment with a "model" shape X. The data and the model shape may be represented in any of the allowable forms. For our purposes, the data shape must be decomposed into a point set if it is not already in point set form. Fortunately, this is easy; the points to be used for triangle and line sets are the vertices and the endpoints, and if the data shape comes in a surface or curve form, then the vertices and endpoints of the triangle/line approximation (as described above) are used. The number of points in the data shape will be denoted N_p. Let N_x be the number of points, line segments, or triangles involved in the model shape. As described above, the curve and surface closest-point evaluators implemented in our system require a framework of lines or triangles to yield the initial parameter values for the Newton's iteration; therefore, the number N_x is still relevant for these smooth entities but varies according to the accuracy of the approximation.

The distance metric d between an individual data point $\vec{p}$ and a model shape X will be denoted

$$d(\vec{p}, X) = \min_{\vec{x} \in X} \|\vec{x} - \vec{p}\|. \tag{28}$$

The closest point in X that yields the minimum distance is denoted $\vec{y}$ such that $d(\vec{p}, \vec{y}) = d(\vec{p}, X)$, where $\vec{y} \in X$. Note that computing the closest point is $O(N_x)$ worst case with expected cost $\log(N_x)$. When the closest point computation (from $\vec{p}$ to X) is performed for each point in P, that process is worst case $O(N_p N_x)$. Let Y denote the resulting set of closest points, and let $\mathcal{C}$ be the closest point operator:

$$Y = \mathcal{C}(P, X). \tag{29}$$

Given the resultant corresponding point set Y, the least squares registration is computed as described above:

$$(\vec{q}, d) = \mathcal{Q}(P, Y). \tag{30}$$

The positions of the data shape point set are then updated via $P = \vec{q}(P)$.

$$\boldsymbol{R} = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1 q_2 - q_0 q_3) & 2(q_1 q_3 + q_0 q_2) \\ 2(q_1 q_2 + q_0 q_3) & q_0^2 + q_2^2 - q_1^2 - q_3^2 & 2(q_2 q_3 - q_0 q_1) \\ 2(q_1 q_3 - q_0 q_2) & 2(q_2 q_3 + q_0 q_1) & q_0^2 + q_3^2 - q_1^2 - q_2^2 \end{bmatrix} \tag{21}$$

TESLAIV000023928

A. ICP Algorithm Statement

The ICP algorithm can now be stated:

- The point set P with N_p points $\{\vec{p}_i\}$ from the data shape and the model shape X (with N_x supporting geometric primitives: points, lines, or triangles) are given.
- The iteration is initialized by setting $P_0 = P$, $\vec{q}_0 = [1, 0, 0, 0, 0, 0, 0]^t$ and $k = 0$. The registration vectors are defined relative to the initial data set P_0 so that the final registration represents the complete transformation. Steps 1, 2, 3, and 4 are applied until convergence within a tolerance τ. The computational cost of each operation is given in brackets.
  a. Compute the closest points: $Y_k = \mathcal{C}(P_k, X)$ (cost: $0(N_p N_x)$ worst case, $0(N_p \log N_x)$ average).
  b. Compute the registration: $(\vec{q}_k, d_k) = \mathcal{Q}(P_0, Y_k)$ (cost: $O(N_p)$).
  c. Apply the registration: $P_{k+1} = \vec{q}_k(P_0)$ (cost: $O(N_p)$).
  d. Terminate the iteration when the change in mean-square error falls below a preset threshold $\tau > 0$ specifying the desired precision of the registration: $d_k - d_{k+1} < \tau$.

If a dimensionless threshold is desired, one can replace τ with $\tau\sqrt{tr(\Sigma_x)}$, where the square root of the trace of the covariance of the model shape indicates the rough size of the model shape.

B. Convergence Theorem

A convergence theorem for the ICP algorithm is now stated and proved. The key ideas are that 1) least squares registration generically reduces the average distance between corresponding points during each iteration, whereas 2) the closest point determination generically reduces the distance for each point individually. Of course, this individual distance reduction also reduces the average distance because the average of a set of smaller positive numbers is smaller. We offer a more elaborate explanation in the proof below.

Theorem: The iterative closest point algorithm always converges monotonically to a local minimum with respect to the mean-square distance objective function.

Proof: Given $P_k = \{\vec{p}_{ik}\} = \vec{q}_k(P_0)$ and X, compute the set of closest points $Y_k = \{\vec{y}_{ik}\}$ as prescribed above given the internal geometric representation of X. The mean squared error e_k of that correspondence is given by

$$e_k = \frac{1}{N_p}\sum_{i=1}^{N_p} \|\vec{y}_{ik} - \vec{p}_{ik}\|^2. \tag{31}$$

The $\mathcal{Q}$ operator is applied to get $\vec{q}_k$ and d_k from that correspondence:

$$d_k = \frac{1}{N_p}\sum_{i=1}^{N_p} \|\vec{y}_{ik} - \boldsymbol{R}(\vec{q}_{kR})\vec{p}_{i0} - \vec{q}_{kT}\|^2. \tag{32}$$

It is always the case that $d_k \leq e_k$. Suppose that $d_k > e_k$. If this were so, then the identity transformation on the point set would yield a smaller mean square error than the least squares registration, which cannot possibly be the case. Next, let the least squares registration $\vec{q}_k$ be applied to the point set P_0, yielding the point set P_{k+1}. If the previous correspondence to the set of points Y_k were maintained, then the mean square error is still d_k, that is

$$d_k = \frac{1}{N_p}\sum_{i=1}^{N_p} \|\vec{y}_{ik} - \vec{p}_{i,k+1}\|^2. \tag{33}$$

However, during the application of the subsequent closest point operator, a new point set Y_{k+1} is obtained: $Y_{k+1} = \mathcal{C}(P_{k+1}, X)$. It is clear that

$$\|\vec{y}_{i,k+1} - \vec{p}_{i,k+1}\| \leq \|\vec{y}_{ik} - \vec{p}_{i,k+1}\| \text{ for each } i = 1, N_p \tag{34}$$

because the point $\vec{y}_{ik}$ was the closest point prior to transformation by $\vec{q}_k$ and resides at some new distance relative to $\vec{p}_{i,k+1}$. If $\vec{y}_{i,k+1}$ were further from $\vec{p}_{i,k+1}$ than $\vec{y}_{ik}$, then this would directly contradict the basic operation of the $\mathcal{C}$ closest point operator. Therefore, the mean square errors e_k and d_k must obey the following inequality:

$$0 \leq d_{k+1} \leq e_{k+1} \leq d_k \leq e_k \text{ for all } k. \tag{35}$$

The lower bound occurs, of course, since mean-square errors cannot be negative. Because the mean-square error sequence is nonincreasing and bounded below, the algorithm as stated above must converge monotonically to a minimum value. Q.E.D.

Experimentally, we find fast convergence during the first few iterations that slows down as it approaches the local minimum. Even at this slow pace, somewhere between 30 and 50 iterations yields excellent results: $d_k \approx 0.1\%$ of model shape size. The convergence can be accelerated using a simple additional operation described in the next section.

C. An Accelerated ICP Algorithm

The accelerated ICP algorithm uses a minor variation on the basic line search methods of multivariate unconstrained minimization [45]. As the iterative closest point algorithm proceeds, a sequence of registration vectors is generated: $\vec{q}_1$, $\vec{q}_2$, $\vec{q}_3$, $\vec{q}_4$, $\vec{q}_5$, $\vec{q}_6$, ..., which traces out a path in the registration state space from the identity transformation toward a locally optimal shape match. Consider the difference vector sequence defined by

$$\Delta\vec{q}_k = \vec{q}_k - \vec{q}_{k-1} \tag{36}$$

which defines a direction in the registration state space. Let the angle in 7 space between the two last directions be denoted

$$\theta_k = \cos^{-1}\left(\frac{\Delta\vec{q}_k{}^t \Delta\vec{q}_{k-1}}{\|\Delta\vec{q}_k\|\|\Delta\vec{q}_{k-1}\|}\right) \tag{37}$$

and let $\delta\theta$ be a sufficiently small angular tolerance (e.g., 10°). If

$$\theta_k < \delta\theta \text{ and } \theta_{k-1} < \delta\theta \tag{38}$$

then there is good direction alignment for the last three registration state vectors: $\vec{q}_k$, $\vec{q}_{k-1}$, and $\vec{q}_{k-2}$. Let d_k, d_{k-1},

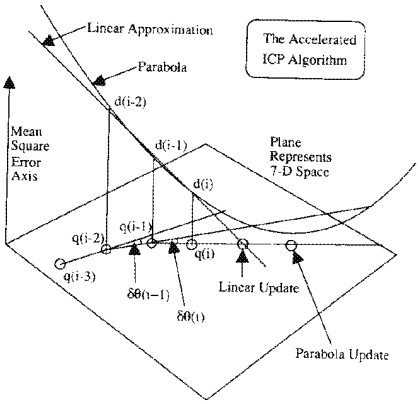


Fig. 1. Consistent direction allows acceleration of the ICP algorithm.

and d_{k-2} be the associated mean square errors, and let v_k, v_{k-1}, and v_{k-2} be associated approximate arc length argument values:

$$v_k = 0, v_{k-1} = -\|\Delta\vec{q}_k\|, v_{k-2} = -\|\Delta\vec{q}_{k-1}\| + v_{k-1}. \quad (39)$$

See Fig. 1 for a picture of the situation. Next, a linear approximation and a parabolic interpolant to the last three data points are computed:

$$d_1(v) = c_1 v + b_1 \; d_2(v) = a_2 v^2 + b_2 v + c_2 \quad (40)$$

which gives us a possible linear update, based on the zero crossing of the line, and a possible parabola update, based on the extremum point of the parabola:

$$v_1 = -b_1/a_1 > 0 \; v_2 = -b_2/2a_2. \quad (41)$$

To be on the safe side, we adopt a maximum allowable value $v_{\max}$. The following logic is used to perform an attempted update:

1) If $0 < v_2 < v_1 < v_{\max}$ or $0 < v_2 < v_{\max} < v_1$, use the parabola-based updated registration vector: $\vec{q}'_k = \vec{q}_k + v_2\Delta\vec{q}_k/\|\Delta\vec{q}_k\|$ instead of the usual vector $\vec{q}_k$ when performing the update on the point set, i.e., $P_{k+1} = \vec{q}'_k(P_0)$.
2) If $0 < v_1 < v_2 < v_{\max}$ or $0 < v_1 < v_{\max} < v_2$ or $v_2 < 0$ and $0 < v_1 < v_{\max}$, use the line-based updated registration vector $\vec{q}'_k = \vec{q}_k + v_1\Delta\vec{q}_k/\|\Delta\vec{q}_k\|$ instead of the usual vector $\vec{q}_k$.
3) If both $v_1 > v_{\max}$ and $v_2 > v_{\max}$, use the maximum allowable update $\vec{q}'_k = \vec{q}_k + v_{\max}\Delta\vec{q}_k/\|\Delta\vec{q}_k\|$ instead of the usual vector $\vec{q}_k$.

We have found experimentally that setting $v_{\max} = 25\|\Delta\vec{q}_k\|$ adaptively has provided a good sanity check on the updates allowing the iterative closest point algorithm to move to the local minimum with a given degree of precision in many fewer steps. A nominal run of more than 50 basic ICP iterations for a given value of τ is typically accelerated to 15 or 20 iterations.

If the updated registration vector were somehow to overshoot the minimum enough to yield a worse mean square error, it would be advantageous to construct a new parabola using the new registration with the last two steps and move to the appropriate minimum. This has not been necessary in our experience. To be rigorous, one can simply ignore the suggested update if it causes a worse mean square error.

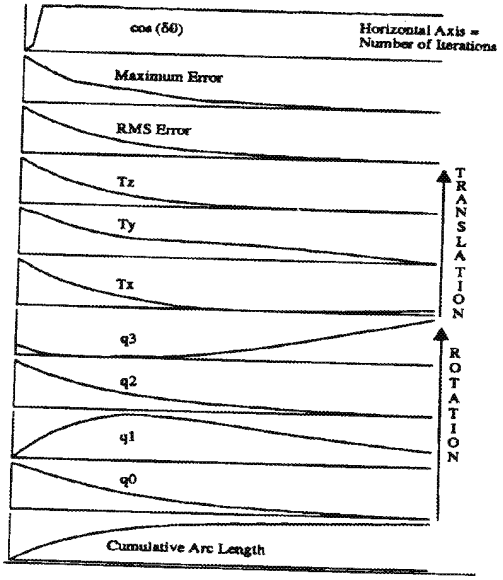


Fig. 2. Various quantities plotted against iteration count for the basic nonaccelerated ICP algorithm.

To give a quantitative example comparison, the registration values, RMS error, maximum error, angular change, and cumulative arc length values were recorded during 50 iterations of both the basic and accelerated ICP algorithms during the same free-form surface matching test. The results for the basic ICP algorithm are shown in Fig. 2. Note the smooth character of all the curves. The most important feature is that the $\cos(\delta\theta)$ plot indicates a consistent direction of updates for all but the first few iterations. In contrast, the accelerated ICP algorithm shows the desirable jumpy behavior as seen in Fig. 3. In addition, note how most quantities get close to their final values after the first acceleration step and very close after two. The acceleration steps occur whenever a V-shaped dip occurs in the plot of $\cos(\delta\theta)$ versus the iteration count.

D. Alternative Minimization Approaches

The ICP algorithm allows us to move from a given starting point to a local minima in 7 space relatively quickly in comparison with other possible alternatives. Each iteration requires only one evaluation of the closest point operator: the most expensive computation. Any optimization method that does not use explicit vector gradient estimates, such as Powell's direction set method, the Nelder-Mead downhill simplex method, or simulated annealling, requires literally *hundreds* to *tens of thousands* of closest point evaluations. These numbers are based on tests done to simulate the action of the least squares registration step involved in one ICP iteration but using instead Powell's direction set method and the Nelder-Mead method from [45].

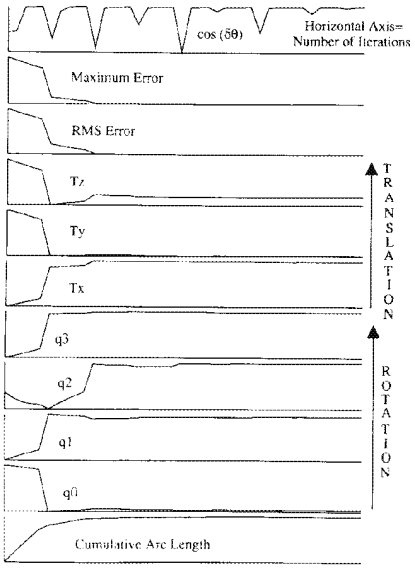


Fig. 3. Various quantities plotted against iteration count for the accelerated ICP algorithm.

Any optimization method that uses explicit vector gradients, such as steepest descent, conjugate gradient, and variable metric schemes, will require at least seven closest point evaluations for each numerical gradient evaluation. Therefore, such a method would have to converge in three or four iterations to be competitive with the accelerated ICP method. Such generic methods generally require many more than three iterations with the number of required closest point evaluations running well over 100 even in ideal circumstances. If a pure numerical Hessian-based Newton's method were used, the numerical gradient and Hessian computations would require at least 13 closest point evaluations per iteration, implying that the iteration would have to converge in two iterations to be competitive with the accelerated ICP algorithm. A pure Newton's method might require only three iterations if the initial point were already well into the region of attraction surrounding a local minimum, but the initial iterations would not be handled well by Newton's method.

Whenever an accelerated parabolic update takes place after three basic ICP steps, we can get nearly quadratic convergence for less than steepest descent cost. This is an interesting accomplishment for a function where derivatives cannot be evaluated. Note that the steepest descent gradient direction is not deliberately computed; we merely observe when a consistent direction is being followed.

Other problems involved with using general-purpose optimization methods are the following: 1) If any angles are used as in [54], angular cycles across 360° must be handled correctly, and 2) if a unit quaternion becomes a nonunit quaternion, as would be expected taking arbitrary direction steps in 4 space, the quaternion must be renormalized somewhere. Unfortunately, if the objective function evaluator changes the values in the state vector during the optimization iteration, this has a bad effect on most nonlinear optimization algorithms.

To summarize, any method that allows one to move from an initial state to its corresponding local minimum could theoretically be used in place of the ICP algorithm. For example, consider Szeliski's [54] work with steepest descent and three rotation angles. However, the arguments above indicate that one would have a hard time trying to find an algorithm that was only ten times slower on average. The key benefit of the ICP algorithm is that the convergence is fast and monotonic. No expensive closest point evaluations are spent on registration vectors that have worse mean square errors than the current state. Because of the ICP convergence theorem, one does not have to "feel around" in the multidimensional space to determine the direction in which to move.

V. The Set of Initial Registrations

Even though the ICP algorithm must converge monotonically to a local minimum from any given rotation and translation of the data point set, it may or may not converge on the desired global minimum. The only way to be sure is to find the minimum of all the local minima. The problem with reaching the desired global minimum with certainty is that it is difficult to precisely characterize in general the partitioning of the registration state space into local minima wells (regions of attraction) because this partitioning is potentially different for every possible different shape encountered.

To be precise, consider a 6-D state space Ω, where the quaternion component q_0 is determined from the other quaternion components: $q_0 = \sqrt{1-(q_1^2+q_2^2+q_3^2)}$. The actual state space Ω is a subset of the space $\Omega' = [-1,1]^3 \times \mathcal{R}^3$, where $\mathcal{R} = (-\infty, +\infty)$ is the real line. The subset Ω is specified by the "inside or on the unit 3 sphere" constraint that $q_1^2 + q_2^2 + q_3^2 \leq 1$. Therefore, Ω may be viewed as a type of hyper-cylinder in 6 space.

For any given nonpathological shape X that represents a real-world surface or object (e.g., pathological shape descriptions based on $\sin(1/x)$ near zero not allowed) and for any given point set P_{reg} already correctly registered with X, consider that any initial state $\vec{q} \in \Omega$ of the point set $P = \vec{q}(P_{\text{reg}})$ will converge to a local minimum as it is matched to X. There are a finite number of local minima $N_m(X,P)$ after one has fixed X and P. (The shape X is considered pathological if this is not true.) Let $\Psi(X,P)$ be the set of all local minima:

$$\Psi(X,P) = \{\psi_n\}_{n=1}^{N_m}. \tag{42}$$

This induces a natural partitioning of Ω into equivalence classes, labeled Ψ_n, where every value of $\vec{q}$ that converges via the ICP algorithm to ψ_n is a member of the class Ψ_n. This allows us to state that

$$\Omega = \bigcup_{n=1}^{N_m} \Psi_n \text{ and } \Psi_n \cap \Psi_m = \phi \text{ if } n \neq m. \tag{43}$$

Let Ψ_1 be the equivalence class that maps to the correct global minimum ψ_1.

To guarantee that the global minimum is found for a given shape X and a given set P not already registered with X, one

must use an appropriate set of initial states so that transforming P by at least one initial registration will place the point set into the correct equivalence class Ψ_1 of registrations. This allows it to converge to the correct global minimum ψ_1. Two fundamental questions are 1) how to construct an initial set of states for any given object that guarantees a correct global minimum and 2) how to construct an initial set of states that guarantees all shapes in a given class of shapes when those shapes converge to the correct respective global minimum.

By using a sufficiently dense uniform sampling of quaternions on the unit sphere combined with a sufficiently dense sampling of translation vectors occupying the total volume about the shape X, it is possible to determine the complete finite set of local minima with a sufficiently small probability of error for one given object. One could construct a set of initial states to include all these local minima solutions along with the halfway states between the k nearest neighbors (e.g., $k = 12$) to attempt to avoid having a set of initial states that lie on or near the boundaries between the equivalence classes.

A method that could be useful for computing guaranteed initial states for a set of model shapes is to use a 6-D occupancy array to compute hypervoxel-based descriptions of the equivalence classes for each shape of interest and then computing an overall partitioning by refining the first shape's partition by intersecting the equivalence classes of each subsequent shape. Such methods can be very memory intensive; a $20 \times 20 \times 20 \times 20 \times 20 \times 20$ hypercubic-hypervoxel grid of the smallest hypercylinder containing all relevant registrations of all shapes of interest requires an 8-Mbyte (64 Mb) array for a single object. Clearly, this is a test of patience for anyone wishing to construct an initial set of states customized to a given set of objects to yield guaranteed results, but indeed, it is possible and relatively straightforward, except for the high dimension of the problem.

A. Initial States for Global Matching

There are simpler methods of dealing with the initial state problem that are very effective on most shapes one comes across. Let us adopt the following definition of the first two moments of the distribution of geometry in P and X: $\vec{\mu}_p = E[\vec{p}|\vec{p} \in P]$, $\vec{\mu}_x = E[\vec{r}|\vec{r} \in X]$, $\Sigma_p = E[(\vec{p}-\vec{\mu}_p)(\vec{p}-\vec{\mu}_p)^t|\vec{p} \in P]$, and $\Sigma_x = E[(\vec{r}-\vec{\mu}_x)(\vec{r}-\vec{\mu}_x)^t|\vec{r} \in X]$, where $E[\cdot]$ represents the sample expectation (averaging) operator. If the point data set P covers a significant portion of the model shape X such that the condition

$$\alpha_1 \sqrt{tr(\Sigma_x)} \leq \sqrt{tr(\Sigma_p)} \leq \sqrt{tr(\Sigma_x)} \tag{44}$$

holds for a sufficiently large factor, say $\alpha_1 = 1/\sqrt{2} \approx 0.71$, then we have found that it is generally not necessary to use multiple initial translation states, as long as enough rotation states are used. This factor α_1 is the allowable occlusion percentage for global matching. The exact value of α_1 could be computed for any given class of object shapes via exhaustive testing if that is desired.

There are two reasonable options for the initial translation state: 1) Apply the ICP algorithm directly to the point set P using multiple rotation states about its center of mass $\vec{\mu}_p$, or 2) transform it first so that the centers of mass of P and X coincide, and then apply ICP. We have found no differences in the final registration results between 1) not translating and 2) pretranslating the point set when using an adequately large set of initial rotations (e.g., 24). In fact, any translation state suffices because the ICP algorithm is very insensitive to the initial translation state when used for global shape matching. We have observed that pretranslating the data point set's center of mass to the model shape's center of mass generally saves a few iterations (e.g., 2 to 4) out of the usual 20 or so total.

A further simplification in the global shape matching algorithm can be accomplished for most generic shapes, where principal moments demonstrate some level of distinctness. Let $p_x \geq p_y \geq p_z$ be the square roots of the eigenvalues of Σ_p, and let $r_x \geq r_y \geq r_z$ be the square roots of the eigenvalues of Σ_x. If the following sets of conditions hold:

$$p_y \leq \alpha_2 p_x \qquad p_z \leq \alpha_2 p_y \tag{45}$$

$$r_y \leq \alpha_2 r_x \qquad r_z \leq \alpha_2 r_y \tag{46}$$

for a specified value of α_2, e.g., $\alpha_2 = 1/\sqrt{2} \approx 0.71$, one can reliably match the basic shape structure of data and model using only the eigenvectors of the matrices Σ_p and Σ_x. Again, the exact value of α_2 could be computed for any given set of objects and any given level of sensor noise via exhaustive testing if needed. In this case of eigenvalue distinctness, the identity transformation and the 180° rotations about the eigenvector axes corresponding to r_x, r_y, and r_z provide a very good set of only four initial rotations that will yield the correct global minimum for a wide class of model shapes.

If two of the three eigenvalues are approximately equal but significantly different from the third for both data and model shapes, the number of initial states need only be expanded for rotations about the nonambiguous axis, thereby reducing the total number of initial rotational states.

If neither of the above cases for global matching hold true (i.e., $p_x \approx p_y \approx p_z$ and $r_x \approx r_y \approx r_z$), then one must use a fine sampling of quaternion states that cover the entire surface of the northern hemisphere of the unit 4 sphere uniformly.

The rotation groups of the regular polyhedra, which have been well known to crystallographers since the 1800's [29], provide a convenient set of uniformly sampled initial rotations: (a) 12 tetrahedral group states, (b) 24 octahedral/hexahedral group states, and (c) 60 icosahedral/dodecahedral group states. The tetrahedral states are a proper subset (subgroup) of the octahedral/hexahedral states and the icosahedral/dodecahedral states. The octahedral/hexahedral states are not properly contained in the icosahedral/dodecahedral states. For a convenient listing of these rotations in quaternion form, see Appendix A of [32].

From an implementation point of view, one has the option of using precomputed lists or nested loops. For the nested loop case, all normalized combinations of $q_0 = \{1, 0\}$, $q_1 = \{+1, 0, -1\}$, $q_2 = \{+1, 0, -1\}$, and $q_3 = \{+1, 0, -1\}$ provide an easy-to-compute set of 40 rotation states that includes the tetrahedral and the octahedral/hexahedral groups. (One must ensure that the first nonzero quaternion component is positive to avoid duplication of states.) For a really complicated set

TESLAIV000023932

of shapes, all normalized combinations of $q_0 = \{1, 0.5, 0\}$, $q_1 = \{+1, 0.5, 0, -0.5, -1\}$, $q_2 = \{+1, 0.5, 0, -0.5, -1\}$, and $q_3 = \{+1, 0.5, 0, -0.5, -1\}$ provides another easy-to-compute set of 312 initial rotation states. Another scheme for a very dense sampling of states is to refine the 60 states of the icosahedral group by subdividing triangles using a 1 to 4 split and using an increased number of rotations about each axis specified by each vertex of the refined icosahedron.

The general rule of thumb is the more complicated the object, the more initial states required. Any method of getting a sufficiently dense, uniform distribution of quaternions over the northern hemisphere of the unit 4 sphere (or over the full interior and surface of the unit 3 sphere) is adequate. In general, every application may want to use a customized set of initial quaternions that will maximize the probability of choosing a good starting point early for the shapes of interest.

B. A Counter Example

Although the above methods for global shape matching will work very well for many shapes with a small probability of error, we can also state categorically that for any given fixed set of initial rotation states, one can construct a shape X that cannot be correctly registered by the algorithm. Begin with a sphere of radius R. Add a thin spike of length S to the surface of the sphere for each specified axis and for each rotation about that axis as indicated by the fixed set of initial rotation states. Next, add one or more spikes of length $S + \epsilon$ anywhere that the extra spikes will fit. If the extras will not fit, make the original spikes thinner. Then, sample points on the surface of this shape with any desired scheme so that there is at least one point per spike, including one at the tip of each spike. The ICP algorithm combined with the given fixed set of initial rotation states will not be able to register a generic repositioning of this point set with the original object in such a way that the longer (or shorter) spikes are correctly registered with each other. It can safely be predicted that the proposed registration algorithm will have difficulty correctly registering "sea urchins" and "planets." These shapes are characterized as having almost exactly equal eigenvalues of the covariance matrix Σ_x and as having small shape features at a fine scale relative to the overall shape. Of course, for any given fixed set of object shapes, the set of initial rotations can be increased to guarantee correct registration.

C. Local Shape Matching

The proposed registration algorithm is definitely *not* useful if only a subset of the data point shape P corresponds to the model shape X or a subset of the model shape X, that is, the data point set includes a significant number of data points that do not correspond to any points on the model shape. Unfortunately, this is a trait of the majority of the shape matching algorithms that have ever been implemented. Moreover, this is a common problem in computer vision since data segmentation algorithms often misgroup one set of data points with another distinct group of points that should not be grouped together.

The ICP algorithm is still useful for local matching problems where the entire set of data points P matches a subset of the model shape X. The drawback is that more than one initial translation must be used, which increases the cost of computing the correct registration. If N_t is the number of initial translation states, N_q is the number of initial rotation states, and each closest point evaluation of P relative to X costs $O(N_p N_x)$, in the worst case, the total cost of local matching is $O(N_t N_q N_x N_p)$ as opposed to the cost of global matching $O(N_q N_x N_p)$. The number of initial translations is also dependent on the relative size of the data point set P compared with the model shape X. By defining a quantity η as the ratio of the "sizes" of the data and model shapes

$$\eta = \frac{m(X)}{m(P)} \geq 1 \tag{47}$$

where $m(\cdot)$ is a general measure that measures approximate 1) arc length if X is a curve, 2) area if X is a surface, and 3) volume if X is a volumetric point configuration, then one can estimate the basic qualitative behavior of the required number of translation states as $N_t \approx c(X)\eta$ for most shapes, where $c(X)$ is an approximate proportionality factor that depends on the complexity of the shape X being matched.

Computing such an estimate $m(X)$ is straightforward for the shape X, but evaluating $m(P)$ is more difficult because P may simply be a point set, and the corresponding length, area, and volume on the shape model X is unknown. One can use a convex hull, surface Delaunay, or closest point connection algorithm to get accurate measures for volumes, areas, and arc lengths, but it is more likely that one would design an algorithm to tolerate up to a particular percentage of occlusion, and a fixed set of translations would be computed to handle all objects in a given class of objects with up to that level of occlusion. Examples of local shape matching are demonstrated in the next section.

VI. Experimental Results

This section is divided into three sections: 1) point set matching, 2) curve matching, and 3) surface matching. All programs were written in C. Any quoted approximate times are given for execution on a single-processor computer rated at 1.6 Mflops on the 100×100 double-precision Linpack benchmark.

A. Point Set Matching

In this section, we demonstrate the ability of the ICP algorithm to perform local point set matching without correspondence. Table I lists a point set with eight points that is to be matched against a set of 11 points. Fig. 4 shows the two point sets prior to registration. Fig. 5 shows the two point sets after registration by the ICP algorithm after six iterations, which took less than 1 s. The computed registration is

```
   Translation: -48.078 6.65685 119.479
 Rotation Axis: (0.0321865 0.998188 -0.0508331)
Rotation Angle: 55.7188 degs
     RMS Error: 0.437608 mm
```

The algorithm does not pay attention to the extra points or to the ordering of the points because it always pairs a given

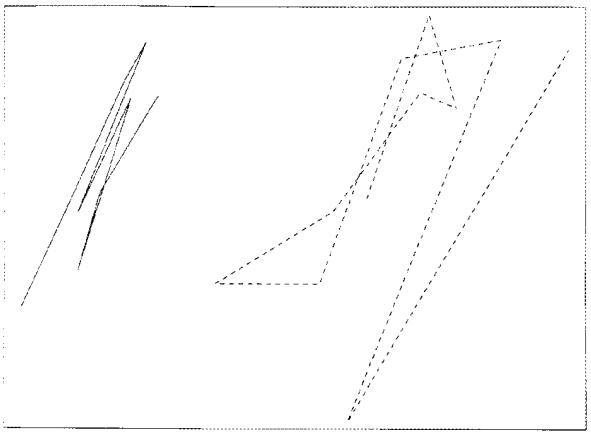


Fig. 4. Set 1 and set 2 prior to registration.

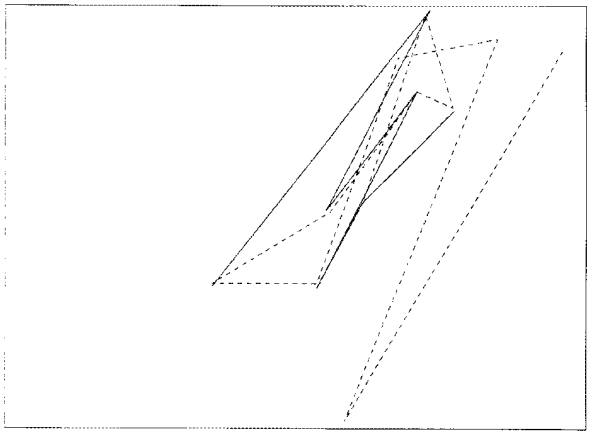


Fig. 5. Set 1 and set 2 after ICP registration.

TABLE I
TWO 3-D POINT SETS: SET 1 IS A SUBSET OF SET 2

$x1$	$y1$	$z1$	$x2$	$y2$	$z2$
43.89	-5.88	106.99	72.78	7.12	146.10
42.02	20.52	112.52	70.19	24.80	148.67
42.01	25.39	113.25	76.21	18.28	147.20
44.95	4.69	112.60	72.71	17.69	148.09
44.12	17.96	115.15	70.67	4.62	145.95
48.26	-1.37	113.59	64.38	-5.40	143.59
46.28	7.03	114.58	72.47	-1.16	143.85
47.00	18.52	117.65	69.82	19.81	148.32
			77.00	25.00	150.00
			80.00	-10.0	140.00
			83.00	30.00	145.00

point on the first set to the closest point on the other set. Only one initial rotation state and one initial translation state were used in this example of local matching. In general, one must use an initial set of rotations combined with an initial set of translations to achieve local matching. Compared with basic point set matching, which requires the same number of points listed in direct correspondence, we are essentially trading off additional CPU time for local matching capability and point order insensitivity.

1) Computational Issues: If one were to use a brute-force tree search (testing every possible correspondence and choosing the best one) in order to find the best match, this type of registration would require $\frac{N_x!}{(N_x-N_p)!}$ operations of the basic least squares quaternion match. For the simple example of local matching above with $N_p = 8$ and $N_x = 11$, a brute force test would require 6 652 800 quaternion registration operations. The above registration required only six operations since the initial state was already a member of the equivalence class of the global minimum. Even with 24 initial translation states and 60 initial rotation states allowing ten iterations for each combined initial state, we would require only 14 400 operations to provide an exhaustive and very capable matching ability. Moreover, it only takes a few minutes for these types of small point sets. The computation reduction ratio of the ICP algorithm compared with brute-force testing for this simple case is 462:1. Of course, other alternatives are possible, but we see that considering 1440 initial states is not unreasonable when the ICP algorithm is used to move from initial state to local minimum.

In the African mask example in the surface section below, we accurately registered a point set with N_p=2500 points to a point set with N_x=4200 points using 60 initial rotation states in less than 8 min. The amount of brute-force enumeration testing required for this case is ridiculously large; more than $1700^{2500} (> 10^{7500})$ operations are required. Even at 1 TeraOp/s (10^{12}) for the age of the universe 10^{18} s, this exact brute-force enumeration of all possible combinations would require well over 10^{250} universe lifetimes!

B. Curve Matching

In this section, the ability of the ICP algorithm to do local free-form curve matching is demonstrated. A 3-D parametric space curve spline was defined as a linear combination of cubic B-splines and control points. A copy of it was translated and rotated to be relatively difficult to match. The rotated and translated curve was converted to a polyline description with 64 points. Each x, y, z component of each point of the polyline was then corrupted by zero-mean Gaussian noise with a standard deviation of $\sigma = 0.1$ (compared with a curve size of $2.3 \times 2 \times 1$ units). The $\pm 3\sigma$ range of 0.6 units is clearly visible compared with the size of the object. We then cut off over half of the noisy polyline leaving a partial noisy curve shape. Fig. 6 shows the two space curves prior to registration. Fig. 7 shows the two space curves after ICP registration using 12 initial rotation states and six initial translation states for a total of 72 initial registration states. If the registration to move the curve away from the original curve is post-multiplied by the registration recovered using the ICP algorithm, we should obtain a matrix close to the identity matrix except for the effects of noise. The match of the partial noisy curve to the original spline curve yielded the following registration matrix,

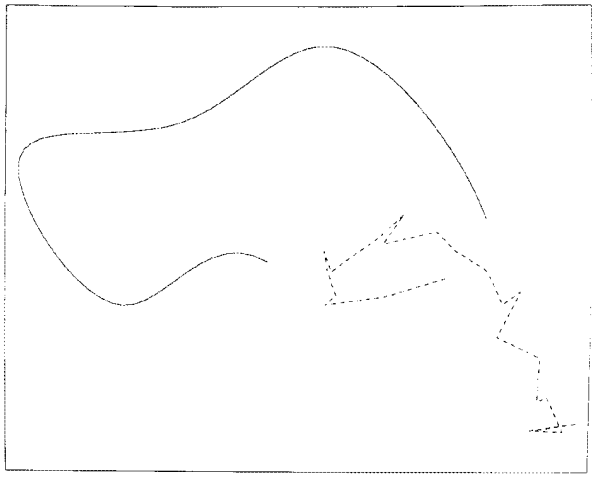


Fig. 6. Ideal and a partial noisy space curve before registration.

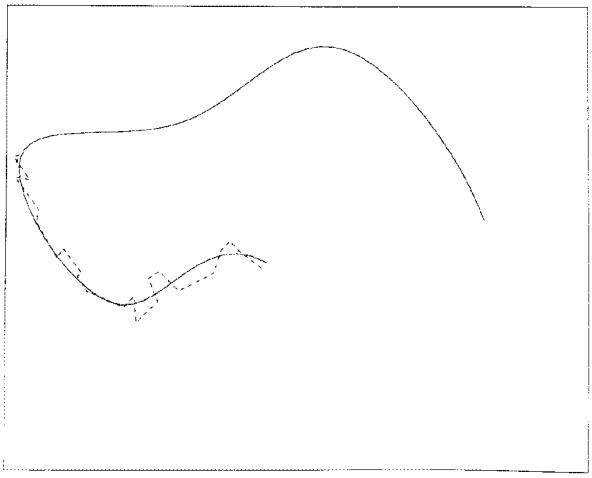


Fig. 7. Ideal and a partial noisy space curve after registration.

which is very close to the identity matrix:

```
Translation: 0.023540  0.006925 -0.015471
   Rotation: 0.999925 -0.012227 -0.001007
     Matrix: 0.012262  0.998647  0.050550
             0.000387 -0.050557  0.998721
```

C. Surface Matching

In this section, the ability of the ICP algorithm to register free-form surface shapes is demonstrated.

1) A Bezier Surface Patch: A simple parametric Bezier surface patch was constructed for quick testing of the free-form surface matching capability of the ICP algorithm. A set of 250 randomly positioned points was evaluated in the interior of the domain of the surface patch and translated and rotated in a random manner. The points of this point set are connected by lines indicating the point list sequence; they *do not* indicate line geometry to be matched. The surface patch is drawn

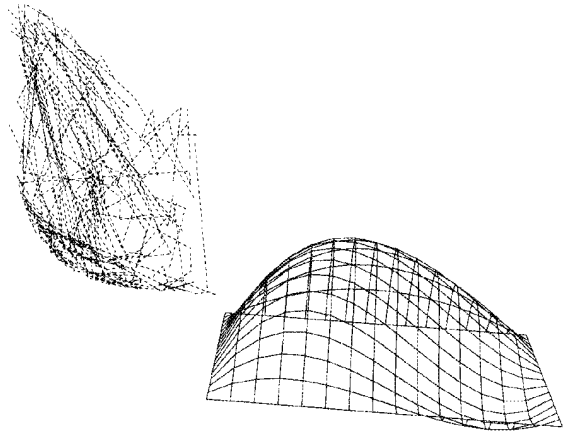


Fig. 8. Noisy point set and surface patch prior to registration: 250 points matched to 450 triangles.

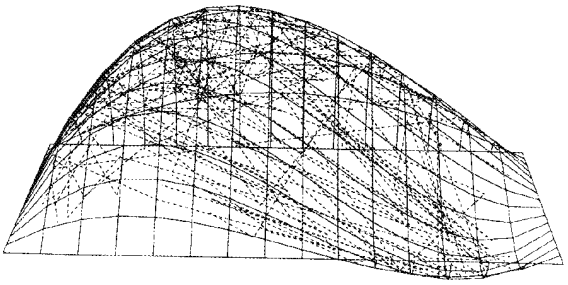


Fig. 9. Noisy point set and surface patch after registration.

via isoparametric lines indicating 450 triangles and fits in a $3 \times 3 \times 1$ units box. Following the space curve example, 3-D vector noise with a standard deviation of 0.1 units in each direction was added to the point set data to create a noisy point set. The surface patch and the noisy point set are shown prior to registration in Fig. 8. After running the ICP algorithm with same 24 initial rotation states for a total of about 3 min, we obtained the result shown in Fig. 9. The initial positioning transformation multiplied by the recovered transformation for the noisy point set yields the following approximate identity transformation:

```
Translation: -0.057329  0.013923  0.018430
   Rotation:  0.999357  0.034041  0.011264
     Matrix: -0.033883  0.999328 -0.013935
             -0.011731  0.013545  0.999840
```

This demonstrates that global matching under noisy conditions works quite well.

A subset of 138 noisy points was used to test the local matching ability. The surface patch and the noisy point subset are shown prior to registration in Fig. 10. After running the ICP algorithm with 24 initial rotation states and six initial translation states for a total of about 6 min, we obtained the result shown in Fig. 11. The initial positioning trans-

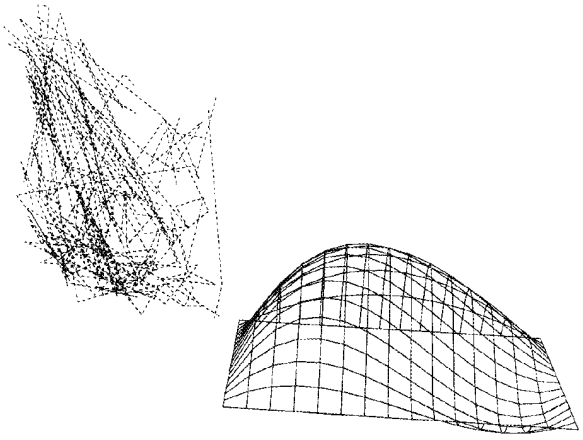


Fig. 10. Noisy point subset and surface patch prior to registration: 138 points and 450 triangles.

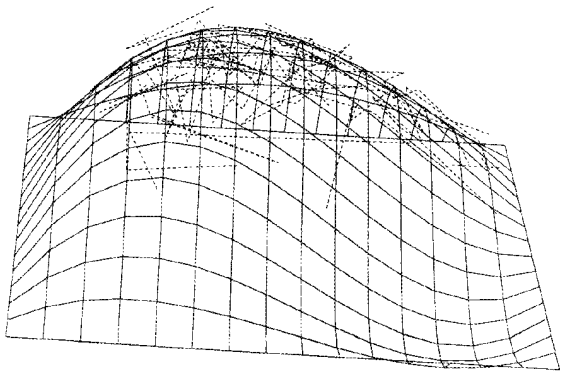


Fig. 11. Noisy point subset and surface patch after registration.

formation multiplied by the recovered transformation for the noisy point subset yields the following approximate identity transformation:

```
Translation: -0.166476  0.159480  0.128289
   Rotation:  0.990806  0.113548  0.073548
     Matrix: -0.113595  0.993521 -0.003545
             -0.073474 -0.004842  0.997285
```

This matrix approximation to the identity is much less precise than the global matching case, but the data is so noisy and the shape is so featureless that we found it surprising that the registration came out as well as it did.

2) The NRCC African Mask: In this experiment, range data from the National Research Council of Canada's African mask example was obtained using the commercially available Hyscan laser triangulation sensor from Hymarc, Ltd. A low-resolution 64 × 68 gridded image was computed from the original data set for use in our experiments and is shown in Fig. 12. This data will serve as the model surface description with 8442 triangles (4221 quadrilateral polygons). A thinned

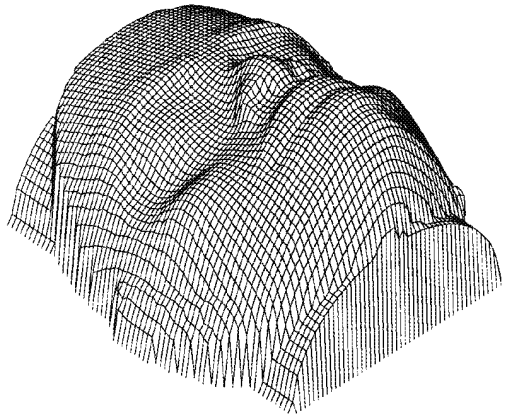


Fig. 12. Model surface: Range image of mask: 8442 triangles.

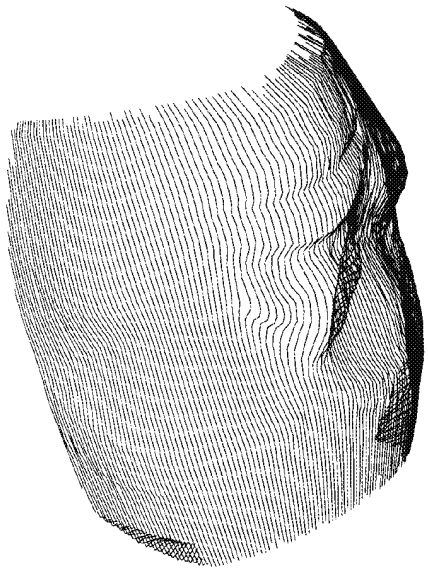


Fig. 13. Data: Rotated and translated point set of mask: 2546 points.

version of the measured data point set containing 2546 points is used as the data point set and is shown as scan lines in a test registration view in Fig. 13.

All trial positionings of the 90-mm object, including the one shown above, converged to the correct solution in about 10 min, and all cases had a 0.59-mm RMS error. This includes six iterations worth of testing on each of 24 initial state vectors and full iteration on the best state of the six initial iterations. A side view of both the digital surface model and the measured point set as registered is shown in Fig. 14. The registration is quite accurate.

A Bezier surface patch model of the mask was created to test the parametric surface matching capability on the given shape. This model is shown in in Fig. 15. The point set in various rotations and translations was then registered to the parametric surface model. We had expected about a 1.2-

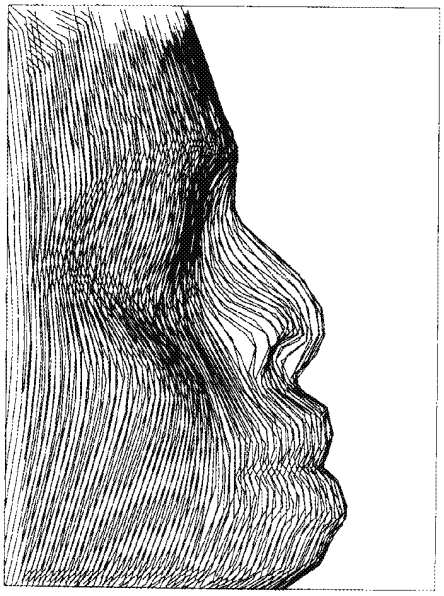


Fig. 14. Side view of range image model and registered point set.

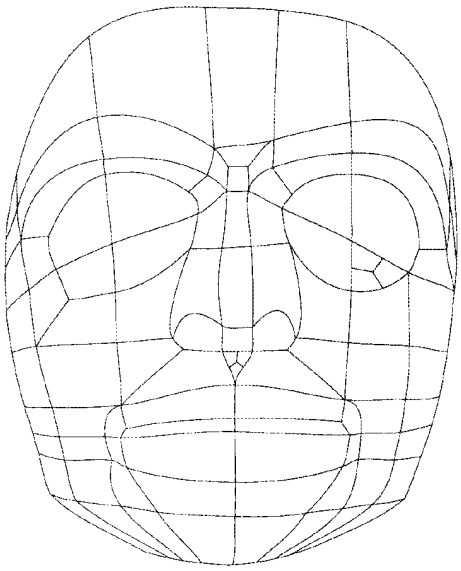


Fig. 15. Surface patch boundaries of a set of parametric surfaces: 97 cubic Bezier patches.

mm rms distance, but the final solution had a 3.4-mm rms distance. After examining the results closely, we discovered that surfaces had not been created in regions where there were measured points. Note the extra data points for which there was no possible surface correspondence and the evidence of misregistration in Fig. 16. Overall, this match was not bad considering the circumstances. After a post-processing step to ignore any measured points whose point-to-surface vectors were not within a few degrees of the surface normal at the corresponding points, the misregistration disappeared, and the data locked in on the surface with the expected rms distance. Although the ICP algorithm is not designed to handle data points that do not correspond to the model shape, one can conclude that a minor misgrouping of nearby points will usually have a minor effect. Another important point to keep in mind is that the matching algorithm does not care about the partitioning of the composite surface model into separate surface patches.

As a final test on the mask, about 30% of the points from the measured data point set were deleted as shown in Fig. 17. The registration algorithm locked in on the solution and gave a slightly improved rms distance in less time than the full data set.

3) Terrain Data: For the final set of experimental results, some terrain data for an area near Tucson was obtained from the University of Arizona. Fig. 18 shows a shaded image of the rugged terrain. The dimensions of the model surface are 6700 × 6840 × 1400 units. A point set was extracted by performing 56 planar section cuts at regular intervals along one axis and then thinning out the data using a chord length deviation check. An interior section of this data set was extracted such that about 60% of the surface area of the original data set was covered. The resulting data set

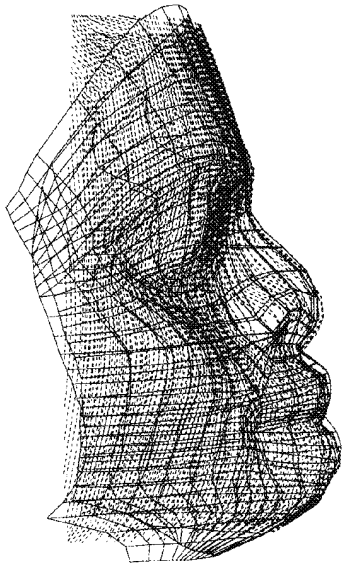


Fig. 16. Side view of parametric surface model and registered point set.

contained 13 655 points and is shown in Fig. 19. The data point set was then lifted above the model surface and rotated to be approximately orthogonal to the model set as shown in Fig. 20. The ICP algorithm performed local matching to the model surface using 24 initial rotations and one initial translation. The registration process for these larger data sets took about 1 hr. The results are shown in Fig. 21. The initial positioning transformation multiplied by the recovered transformation yields the following approximate identity transformation, which demonstrates that surface matching for very

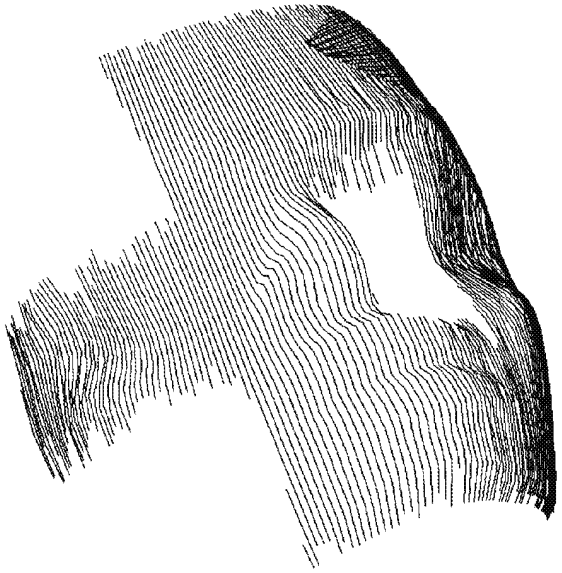


Fig. 17. Subset of the data point set: 1781 points.

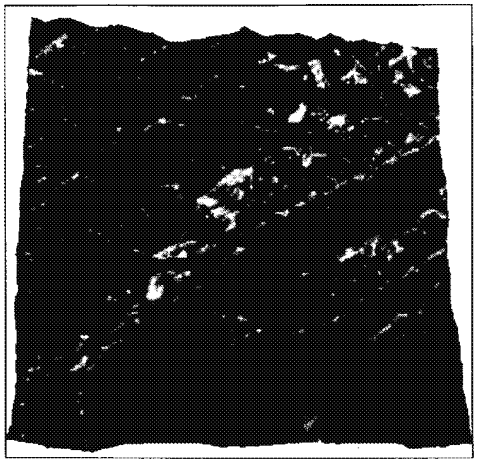


Fig. 18. Model surface: Rugged terrain near Tucson, AZ: 45 900 triangles.

complex surface shapes works quite well:

```
Translation: -11.330336 1.404177  0.934043
   Rotation:   0.999994 0.003308 -0.000231
     Matrix:  -0.003309 0.999994 -0.000505
               0.000230 0.000505  1.000000
```

VII. Conclusions

The iterative closest point (ICP) algorithm is able to register a data shape with N_p points to a model shape with N_x primitives. The model shape can be a point set, a set of polylines, a set of parametric curves, a set of implicit curves, a set of triangles, a set of parametric surfaces, or a set of implicit surfaces. Any other type of shape representation can be incorporated if a procedure for computing the closest point on the model shape to a given point is available. If a data shape were to come in a form other than point set form, a dense set of points on the data shape can serve as the data point set.

The accelerated ICP algorithm converges to a local minimum quickly in comparison with generic nonlinear optimization methods. It is fast enough that global shape matching can be achieved using a sufficiently dense sampling of unit quaternions used as initial rotation states, and local shape matching can be achieved by combining a sufficiently dense

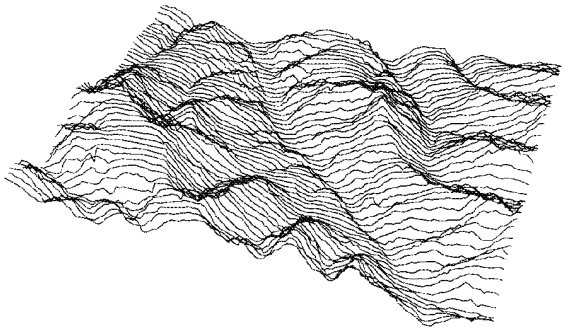


Fig. 19. Data: Rugged terrain near Tucson, AZ: 13 655 points.

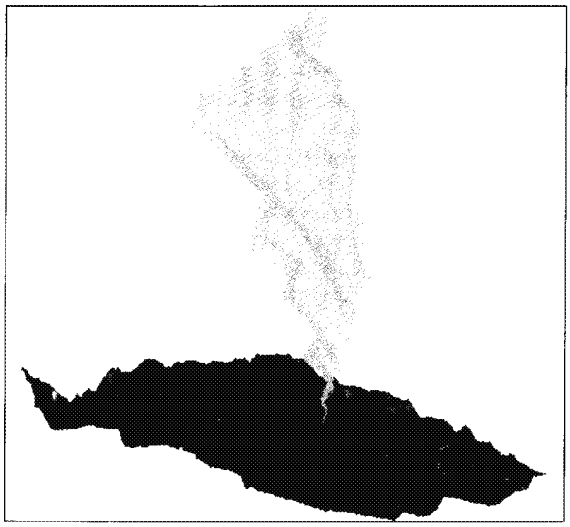


Fig. 20. Data and model of rugged terrain prior to registration.

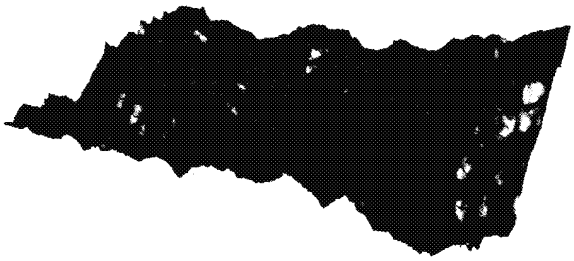


Fig. 21. Data and model of rugged terrain after registration.

sampling of relevant translations. The complexity of a single ICP iteration is $O(N_p N_x)$ in the worst case. For N_t initial translations and N_q initial rotations, the overall worst case complexity of local matching is $O(N_t N_q N_p N_x)$.

The advantages of the ICP shape matching algorithm are as follows:

- It handles the full six degrees of freedom.
- It is independent of shape representation.
- The surface patch or curve segment partitioning of parametric or implicit entities is essentially ignored by the matching procedure. This is important for using CAD data in its native form without elaborate user-guided preprocessing.
- It does not require preprocessing of 3-D point data, such as smoothing, as long as the number of statistical outliers is near zero. This is usually the case with accurate noncontact sensors used for inspection purposes.
- It does not require any derivative estimation or any local feature extraction.
- Almost all aspects of the algorithm are ideally suited for coarse-grain or fine-grain parallel architectures. For large problems, even remote execution procedures and distributed file systems on networks of workstations can provide worthwhile speedup without significant overhead.
- Global matching is achieved at predictable cost based on shape complexity.
- Local matching is achieved at predictable cost based on shape complexity and the percentage of allowable occlusion.
- It can handle a reasonable amount of normally distributed vector noise, with standard deviations of up to 10% of object size demonstrated above.
- The method generalizes to n dimensions by substituting the SVD algorithm [23] for the quaternion algorithm with the added feature that reflections are allowed.
- The method can be made statistically robust by substituting iterations of the SVD-based algorithm by Haralick *et al.* [28] for the quaternion algorithm to identify outliers. The increased computational requirements are significant.
- It can easily be used in conjunction with other algorithms, such as the covariance matrix alignment, which preorient the data so that fewer initial rotation states are necessary.
- Shapes with three sufficiently distinct principal moments (eigenvalues) can be globally matched at a cost of only four initial rotation states.
- It is relatively insensitive to minor data segmentation errors as indicated by the performance of the registration of points with the African mask parametric surface model.
- The results of the last iteration of closest point registration can be used directly as inspection results since the distance to the closest point on a surface is computed as a byproduct.
- The accelerated ICP algorithm can achieve Newton-type quadratic convergence steps at less cost than a numerical steepest descent step. No time is spent evaluating the objective function to find worse mean square errors off the path to the local minimum goal.

The disadvantages of the ICP shape matching algorithm are as follows:

- It is susceptible to gross statistical outliers unless a statistically robust method is substituted at some point (either in preprocessing or registration computation) for outlier detection.
- The fast least squares quaternion and SVD methods are not so easily adapted to weighted least squares extensions, meaning that it is difficult to extend the *fast* algorithm to allow the assignment of unequal uncertainties to points, as was done in [54]. This is not a major inconvenience for inspection applications but would yield noticably suboptimal results for navigation laser radars and long depth of field triangulation systems, where uncertainty increases significantly with range.
- The cost of local matching can get quite large for small allowable occlusion percentages, e.g., 10% or less. We do not advocate our proposed method if feature extraction techniques will successfully solve the problem.
- The generalization to matching deformable models with high order deformations [57] is not straightforward without, e.g., enumerating a dense set of possible deformations.
- As an extension of the outlier rejection issue, the stated algorithm does not solve the segmentation problem, of course. If data points from two shapes are intermixed and matched against the individual shapes, the registrations will be wrong, and the mean square distance metric will be large. This is a problem with almost all of the shape matching algorithms in the literature.
- For any given fixed initial set of rotations, the global shape matching capability can be defeated even without sensor noise by constructing "sea urchin" or "planetoid" shaped objects based on the set of rotations such that correct registration cannot be guaranteed. On the other hand, for a fixed set of objects and no sensor noise, one can determine an initial set of registrations in a finite amount of time such that one can guarantee registration in a finite amount of time with a sufficiently small probability of error.
- In the limit of very complicated sea urchins or perfectly spherical planets with a single 1 μm bump or in the limit of very localized matching (1% of object shape or less) on any object, the ICP algorithm degenerates to brute-force 3-D template matching. Feature extraction techniques, if possible, are preferable in such circumstances.

VIII. Future Directions

Although we have tried to present compelling results, no method with such promise is likely to be widely accepted until more testing has been done. We believe the algorithm for point sets, polyline sets, and triangle sets are simple enough to be widely implemented and tested.

The accelerated ICP algorithm is quite efficient, but there is plenty of room for further computational speedup. For example, we have not discussed the use of $k-d$ trees to ensure expected cost of $O(\log N_x)$ rather than the worst case $O(N_x)$

TESLAIV000023939

performance in the closest point operation. Actual testing on parallel architectures also needs to be done.

Given a large but finite amount of off-line processing, it seems reasonable to make the following statements: For a fixed set of objects, a given level of allowable occlusion, a given maximum possible (Gaussian) noise level, and a set of initial registration states, it is possible to estimate the probability of registration failure by carrying out exhaustive tests. For a 1% (of the smallest shape size) noise level and a 40% maximum allowable occlusion, our tests indicate that very low probabilities of failure should be achievable with minimal extra work. The method needs to be characterized in such detail.

A single complicated object in a set of simpler objects may require a large set of initial registrations to handle the entire set of objects with certainty. Unfortunately, much time is then spent testing initial registrations with simpler objects such that one is continually homing in on the same local minimum over and over again. With some extra bookkeeping, it may be possible to quickly recognize a familiar local minimum well that you have already fallen into a few times and abort that iteration or to use penalty function methods to penalize walking down a path that has already been explored. The appropriate shape of the penalty functions would depend on the shape of the region of attraction in 6-D, which is difficult to quantify and analyze.

It may be possible to extend the basic least squares registration solution to allow deformations (independent axis scaling and bending) of the model shapes when matching to the data shapes. Shears and separate axis scaling can be easily accomodated by allowing a general affine transformation; allowing even quadratic bending about the center of mass significantly complicates matters.

Finally, these free-form shape matching methods are likely to be useful as part of a 3-D object recognition system.

ACKNOWLEDGMENT

The authors wish to thank A. Morgan, R. Khetan, R. Tilove, W. Wiitanen, R. Bartels, D. Field, W. Meyer, C. Wampler, D. Baker, R. Smith, N. Sapidis, and the reviewers for their comments.

REFERENCES

[1] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least square fitting of two 3-D point sets," *IEEE Trans. Patt. Anal. Machine Intell.* vol. PAMI-9, no. 5, pp. 698–700, 1987.

[2] R. Bajcsy and F. Solina, "Three-dimensional object representation revisited," *Proc. 1st Int. Conf. Comput. Vision* (London), June 8–11, 1989, pp. 231–240.

[3] P. J. Besl, "Geometric modeling and computer vision," *Proc. IEEE*, vol. 76, no. 8, pp. 936–958, Aug. 1988.

[4] P. J. Besl, "Active optical range imaging sensors," in *Advances in Machine Vision* (J. Sanz, Ed.). New York: Springer-Verlag, 1989; see also *Machine Vision and Applications*, vol. 1, pp. 127–152, 1989.

[5] ———, "The free-form surface matching problem," *Machine Vision for Three-Dimensional Scenes* (H. Freeman, Ed.). New York: Academic, 1990.

[6] P. J. Besl and R. C. Jain, "Three dimensional object recognition," *ACM Comput. Surveys*, vol. 17, no. 1, 75–145, March 1985,

[7] J. Blumenthal, "Polygonizing implicit surfaces," Xerox Parc Tech. Rep. EDL-88–4, 1988.

[8] B. Bhanu and C. C. Ho, "CAD-based 3-D object representation for robot vision," *IEEE Comput.*, vol. 20, no. 8, pp. 19–36, Aug. 1987.

[9] W. Boehm, G. Farin, and J. Kahmann, "A survey of curve and surface methods in CAGD," *Comput. Aided Geometric Des.*, vol. 1, no. 1, pp. 1–60, July 1984.

[10] R. C. Bolles and P. Horaud, "3DPO: A three-dimensional part orientation system," *Int. J. Robotic Res.*, vol. 5, no. 3, pp. 3–26, Fall 1986.

[11] P. Brou, "Using the Gaussian image to find the orientation of an object," *Int. J. Robotics Res.*, vol. 3, no. 4, pp. 89–125, 1983.

[12] J. Callahan and R. Weiss, "A model for describing surface shape," in *Proc. Conf. Comput. Vision. Patt. Recogn.* (San Francisco, CA), June 1985, pp. 240–247.

[13] C. H. Chen and A. C. Kak, "3DPOLY: A robot vision system for recognizing 3-D objects in low-order polynomial time," Tech. Rep. 88–48, Elect. Eng. Dept., Purdue Univ., West Lafayette, IN, 1988.

[14] R. T. Chin and C. R. Dyer, "Model-based recognition in robot vision," *ACM Comput. Surveys*, vol. 18, no. 1, pp. 67–108, Mar. 1986.

[15] C. DeBoor, *A Practical Guide to Splines*. New York: Springer-Verlag, 1978.

[16] C. DeBoor and K. Hollig, "B-splines without divided differences," *Geometric Modeling: Algorithms and New Trends* (G. Farin, Ed.), *SIAM*, pp. 21–28, 1987.

[17] G. Farin, *Curves and Surfaces in Computer Aided Geometric Design: A Practical Guide*. New York: Academic, 1990.

[18] O. D. Faugeras and M. Hebert, "The representation, recognition, and locating of 3-D objects," *Int. J. Robotic Res.* vol. 5, no. 3, pp. 27–52, Fall 1986.

[19] T. J. Fan, "Describing and recognizing 3-D objects using surface properties," Ph.D. dissertation, IRIS-237, Inst. Robotics Intell. Syst., Univ. of Southern California, Los Angeles, 1988.

[20] P. J. Flynn and A. K. Jain, "CAD-based computer vision: From CAD models to relational graphs," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 13, no. 2, pp. 114–132, 1991; see also Ph.D. Thesis, Comput. Sci. Dept., Michigan State Univ., E. Lansing, MI.

[21] E. G. Gilbert and C. P. Foo, "Computing the distance between smooth objects in 3D space," RSD-TR-13–88, Univ. of Michigan, Ann Arbor, 1988.

[22] E. G. Gilbert, D. W. Johnson, S. S. Keerthi, "A fast procedure for computing the distance between complex objects in 3D space," *IEEE J. Robotics Automat.*, vol. 4, pp. 193–203, 1988.

[23] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ. Press, 1983.

[24] W. E. L. Grimson, "The combinatorics of local constraints in model-based recognition and localization from sparse data," *J. ACM*, vol. 33, no. 4, pp. 658–686, 1986.

[25] W. E. L. Grimson and T. Lozano-Pérez, "Model-based recognition and localization from sparse range or tactile data," *Int. J. Robotics Res.*, vol. 3, no. 3, pp. 3–35, Fall 1984.

[26] K. T. Gunnarsson and F. B. Prinz, "CAD model-based localization of parts in manufacturing," *IEEE Comput.*, vol. 20, no. 8, pp. 66–74, Aug. 1987.

[27] E. Hall, J. Tio, C. McPherson, and F. Sadjadi, "Measuring curved surfaces for robot vision," *Comput.* vol. 15, no. 12, pp. 42–54, Dec. 1982.

[28] R. M. Haralick *et al.*, "Pose estimation from corresponding point data," in*Machine Vision for Inspection and Measurement* (H. Freeman, Ed.). New York: Academic, 1989.

[29] H. Hilton, *Mathematical Crystallography and the Theory of Groups of Movements*. Oxford: Clarendon, 1963; London: Dover, 1963.

[30] B. K. P. Horn, "Extended Gaussian images," *Proc. IEEE*, vol. 72, no. 12, pp. 1656–1678, Dec. 1984

[31] ———, "Closed-form solution of absolute orientation using unit quaternions," *J. Opt. Soc. Amer. A* vol. 4, no. 4, pp. 629–642, Apr. 1987

[32] ———, "Relative orientation," A.I. Memo 994, AI Lab, Mass. Inst. Technol., Cambridge, Sept. 1987.

[33] B. K. P. Horn and J. G. Harris, "Rigid body motion from range image sequences," *Comput. Vision Graphics Image Processing*, 1989.

[34] K. Ikeuchi, "Generating an interpretation tree from a CAD model for 3-D object recognition in bin-picking tasks," *Int. J. Comput. Vision*, vol. 1, no. 2, pp. 145–165, 1987.

[35] A. K. Jain and R. Hoffman, "Evidence-based recognition of 3-D objects," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 10, no. 6, pp. 793–802, 1988.

[36] B. Kamgar-Parsi, J. L. Jones, and A. Rosenfeld, "Registration of multiple overlapping range images: Scenes without distinctive features," *Proc. IEEE Comput. Vision Patt. Recogn. Conf.* (San Diego, CA), June 1989.

[37] Y. Lamdan and H. J. Wolfson, "Geometric hashing: A general and efficient model-based recognition scheme," *Proc. 2nd Int. Conf. Comput. Vision* (Tarpon Springs, FL), Dec. 5–8, 1988, pp. 238–251.

TESLAIV000023940

[38] S. Z. Li, "Inexact matching of 3D surfaces," VSSP-TR-3–90, Univ. of Surrey, England, 1990.
[39] P. Liang, "Measurement, orientation determination, and recognition of surface shapes in range images," Cent. Robotics Syst., Univ. of California, Santa Barbara, 1987.
[40] D. Luenberger, *Linear and Nonlinear Programming*. Reading, MA: Addison-Wesley, 1984.
[41] A. P. Morgan, *Solving Polynomial Systems Using Continuation for Engineering and Scientific Problems*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
[42] M. E. Mortenson, *Geometric Modeling*. New York: Wiley, 1985.
[43] J. Mundy *et al.*, "The PACE system," in *Proc. CVPR '88 Workshop*; see also *DARPA IUW*.
[44] D. W. Murray, "Model-based recognition using 3D shape alone.," *Comput. Vision Graphics Image Processing*, vol. 40, pp. 250–266, 1987.
[45] Press, Flannery, Teukolsky, and Vetterling, *Numerical Recipes in C*. Cambridge, UK: Cambridge University Press, 1988.
[46] J. Rieger, "On the classification of views of piecewise smooth objects," *Image and Vision Computing*, vol. 5, no. 2, pp. 91–97, 1987.
[47] *Proc. IEEE Robust Methods Workshop*. Univ. of Washington, Seattle.
[48] P. Sander, "On reliably inferring differential structure from 3D images," Ph.D. dissertation, Dept. of Elect. Eng., McGill Univ., Montreal, Canada, 1988.
[49] P. H. Schonemann, "A generalized solution to the orthogonal procrustes problem," *Psychometrika*, vol. 31, no. 1, 1966
[50] J. T. Schwartz and M. Sharir, "Identification of partially obscured objects in two and three dimensions by matching noisy characteristic curves," *Int. J. Robotics Res.*, vol. 6, no. 2, pp. 29–44, 1987.
[51] T. W. Sederberg, "Piecewise algebraic surface patches," *Comput. Aided Geometric Des.*, vol. 2, no. 1, pp. 53–59, 1985.
[52] B. M. Smith, "IGES: A key to CAD/CAM systems integration," *IEEE Comput. Graphics Applications*, vol. 3, no. 8, pp. 78–83, 1983.
[53] G. Stockman, "Object recognition and localization via pose clustering," *Comput. Vision Graphics Image Processing*, vol. 40, pp. 361–387, 1987.
[54] R. Szeliski, "Estimating motion from sparse range data without correspondence," *2nd Int. Conf. Comput. Vision* (Tarpon Springs, FL), Dec. 5–8, 1988, pp. 207–216.
[55] G. Taubin, "Algebraic nonplanar curve and surface estimation in 3-space with applications to position estimation," Tech. Rep. LEMS-43, Div. Eng., Brown Univ., Providence, RI, 1988.
[56] ——, "About shape descriptors and shape matching," Tech. Rep. LEMS-57, Div. Eng., Brown Univ., Providence, RI, 1989.
[57] D. Terzopoulos, J. Platt, A. Barr, and K. Fleischer, "Elastically deformable models," *Comput. Graphics*, vol. 21, no. 4, pp. 205–214, July 1987.
[58] B. C. Vemuri, A. Mitiche, and J. K. Aggarwal, "Curvature-based representation of objects from range data," *Image Vision Comput.*, vol. 4, no. 2, pp. 107–114, May 1986.
[59] B. C. Vemuri and J. K. Aggarwal, "Representation and recognition of objects from dense range maps," *IEEE Trans. Circuits Syst.*, vol. CAS-34, no. 11, pp. 1351–1363, Nov. 1987.
[60] A. K. C. Wong, S. W. Lu, and M. Rioux, "Recognition and shape synthesis of 3-D objects based on attributed hypergraphs," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, no. 3, pp. 279–290, Mar. 1989.



Paul J. Besl (M'86) received the B.S. degree *summa cum laude* in physics from Princeton University, Princeton, NJ, in 1978 and the M. S. and Ph. D. degrees in electrical engineering and computer science from the University of Michigan, Ann Arbor, in 1981 and 1986, respectively.

Prior to 1984, he worked for Bendix Aerospace Systems in Ann Arbor and Structural Dynamics Research Corporation, Cincinnati, OH. Since 1986, he has been a research scientist at General Motors Research Laboratories, Warren, MI, where his primary research interest is computer vision, especially the practical applications of range imaging sensors.

Dr. Besl is a member of the Association for Computing Machinery. He received a Rackham Distinguished Dissertation Award for his Ph. D. work on range image understanding, which was later published in book form by Springer Verlag.



Neil D. McKay received the B. S. degree in electrical engineering from the University of Rochester, Rochester, NY, in 1980, the M.S. degree in computer and systems engineering from Rensselaer Polytechnic Institute in 1982, and the Ph. D. degree in electrical engineering from the University of Michigan in 1985.

He has been with the Computer Science Department of General Motors Research Laboratories since 1985. His interests include robot path planning and optimal control.

TESLAIV000023941

EXHIBIT 9

IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF TEXAS
AUSTIN DIVISION

INTELLECTUAL VENTURES II LLC, Plaintiff, v. TESLA, INC., Defendant.	Civil Action No. 1:24-cv-884-ADA

DECLARATION OF MAURICE HERLIHY, PH.D. IN SUPPORT OF TESLA'S OPENING CLAIM CONSTRUCTION BRIEF

I, Maurice Peter Herlihy, hereby declare as follows:

1. I have been retained as an expert witness on behalf of Tesla, Inc. in the above-referenced litigation.

2. I am being compensated for my time in connection with this Declaration at my standard consulting rate of $400 per hour. My compensation is in no way dependent on the outcome of this matter.

3. I have been asked to opine on matters of claim construction for U.S. Patent No. 8,898,395 ("the '395 patent").

4. In preparing this declaration, I have reviewed the '395 patent and its file history, the constructions proposed by the parties, and any other documents referenced in this Declaration.

5. In forming the opinions expressed in this Declaration, I relied upon my education and experience in the relevant field of the art and have considered the viewpoint of a person having ordinary skill in the relevant art as of 2005.

I. BACKGROUND AND QUALIFICATIONS

6. My background and expertise that qualify me as an expert in the technical issues in this case are as follows:

7. As indicated on my *Curriculum* Vitae, which is appended as **Attachment A**, I currently hold the An Wang Chair of Computer Science at Brown University. I have been a professor of computer science at Brown University since 1998, and before that I was an associate professor in the same department (1994-1998).

8. I received an A.B. in Mathematics from Harvard University (1975), an M.S. in Computer Science from M.I.T. (1980), and a Ph.D. in Computer Science from M.I.T. (1984). I served on the faculty of Carnegie Mellon University (1984-1989) and the staff of Digital Equipment Corporation, Cambridge Research Laboratory (1989-1994).

9. I am the recipient of the 2003 Dijkstra Prize in Distributed Computing, the 2004 Gödel Prize in theoretical computer science, the 2008 ISCA influential paper award, the 2012 Edsger W. Dijkstra Prize, and the 2013 Wallace McDowell award. I received a 2012 Fulbright Distinguished Chair in the Natural Sciences and Engineering Lecturing Fellowship, and I am a fellow of the ACM, a fellow of the National Academy of Inventors, the National Academy of Engineering, and the National Academy of Arts and Sciences. In 2022, I won my third Dijkstra Prize.

10. I am the co-inventor of Transactional Memory, a cache-coherence architecture that forms the foundation for the Transactional Synchronization Extensions (TSX) added to Intel's x86 instruction set architecture.

11. I have authored or co-authored two books and more than 200 publications in journals and major international conference proceedings. A list of my publications is included in my CV, which is appended as **Attachment A**.

12. I have previously offered testimony as an expert witness. A list of my prior engagements in which I testified as an expert at trial or by deposition is appended as **Attachment B**.

13. Based on my background and experience, as set forth more fully in my CV, I am familiar with the state of the art in the field of data processing, including memory management, as of 2005, which is the earliest possible priority date for the '395 patent. I am a technical expert in the fields relating to the '395 patent and other related fields, and I remain an active researcher in these fields.

14. Based on my professional experience, I believe I am qualified to testify as an expert on matters related to the '395 patent.

II. UNDERSTANDING OF THE LAW

15. I am not a legal expert and therefore I offer no opinions on the law. However, I have been informed and am aware of legal standards that are relevant to my analysis, as summarized below.

16. I have been informed and understand that claims are generally given their ordinary and customary meaning as understood by a person having ordinary skill in the art ("POSITA") at the time of the alleged invention, in light of the patent specification and the prosecution history. Such an individual is considered to possess ordinary skills and knowledge in a particular technical field and is not an automaton. I have been informed that claim construction is a matter of law and that the final claim constructions for this proceeding will be determined by the Court.

17. It has been explained to me that the patent statute includes a requirement that claims be definite. I have been informed that, to meet this "definiteness" requirement, the claims must particularly point out and distinctly claim the subject matter that the applicant regards as his or her invention. I understand that indefiniteness is evaluated from the perspective of a POSITA at the time of a patent's filing.

18. It has been explained to me that a patent claim is indefinite if the claim, when read in light of the specification and the prosecution history, fails to inform, with reasonable certainty, a POSITA of the scope of the invention.

19. It was further explained to me that absolute or mathematical precision in claim language is not required. However, it is not enough that some meaning can be ascribed to a patent's claims. A claim is indefinite if its language might mean several different things and no informed and confident choice is available among the contending definitions.

III. THE '395 PATENT

A. Summary of the '395 Patent

20. The '395 patent was filed on May 15, 2008, and claims to be a continuation of an application filed on April 7, 2005, which in now U.S. Patent No. 7,376,798.

21. The '395 patent relates to "computer system memory, in particular the management of cache memory." '395 patent at 1:15-17 ("Field of the Invention"). Various cache management schemes were well known in the prior art. For example, the '395 patent identifies the MESI (Modified, Exclusive, Shared, Invalid) protocol as "a popular cache consistency (coherency) protocol." '395 patent at 1:30-32.

22. The '395 patent alleges that "[p]rior attempts to achieve a sequentially consistent, lumped architecture are limited with respect to the number of memory operations that can be included in an instruction group." '395 patent at 2:1-4. According to the '395 patent, "such attempts are limited to a single memory operation per instruction group." '395 patent at 2:4-5. The patent then purports to give a solution to "overcome this disadvantage." '395 patent at 2:5-7.

23. The '395 patent describes a basic computer system **5**, that includes a processor **10**, a main memory **18**, a cache **14** associated with the processor, and one or more external agents **11-13** that can access the main memory and the processor's cache. '395 patent at 3:40-45, Figure 1. The processor **10** and the external agents **11-13** may each be executing a group of instructions including memory operations (*e.g.*, read, load, store, write, lock acquisition, etc.). '395 patent at 4:3-7. An instruction (or instructions) in the group of instructions being executed by the processor **10** may seek access to its cache **14**; likewise, an instruction (or instructions) in the group of instructions being executed by the external agents **11-13** may also seek access to that cache **14**. '395 patent at 4:11-15.

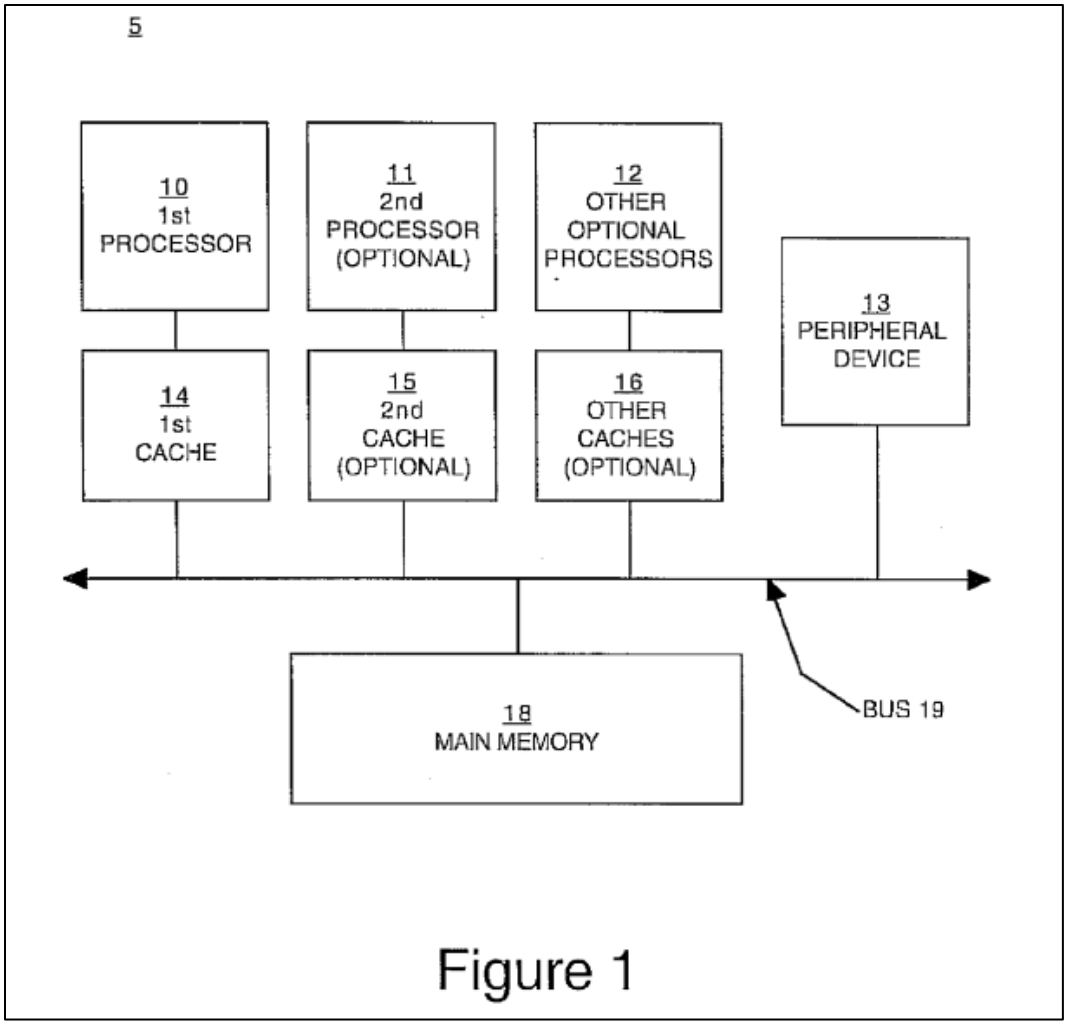


Figure 1

24. An observed bit in cache **14** is set for a cache line if execution of an instruction in a group of instructions executed by the processor **10** causes that cache line to be accessed (*e.g.*, load, store, read, or write). '395 patent at 4:61-65. For example, observed bit **24** is set if an instruction in instruction group A, when executed by processor **10**, causes cache line **21** to be accessed. '395 patent at 4:65-5:2, Figure 2. Similarly, observed bit **25** is set if an instruction in instruction group B, when executed by processor **10**, causes cache line **21** to be accessed. '395 patent at 4:57-60.

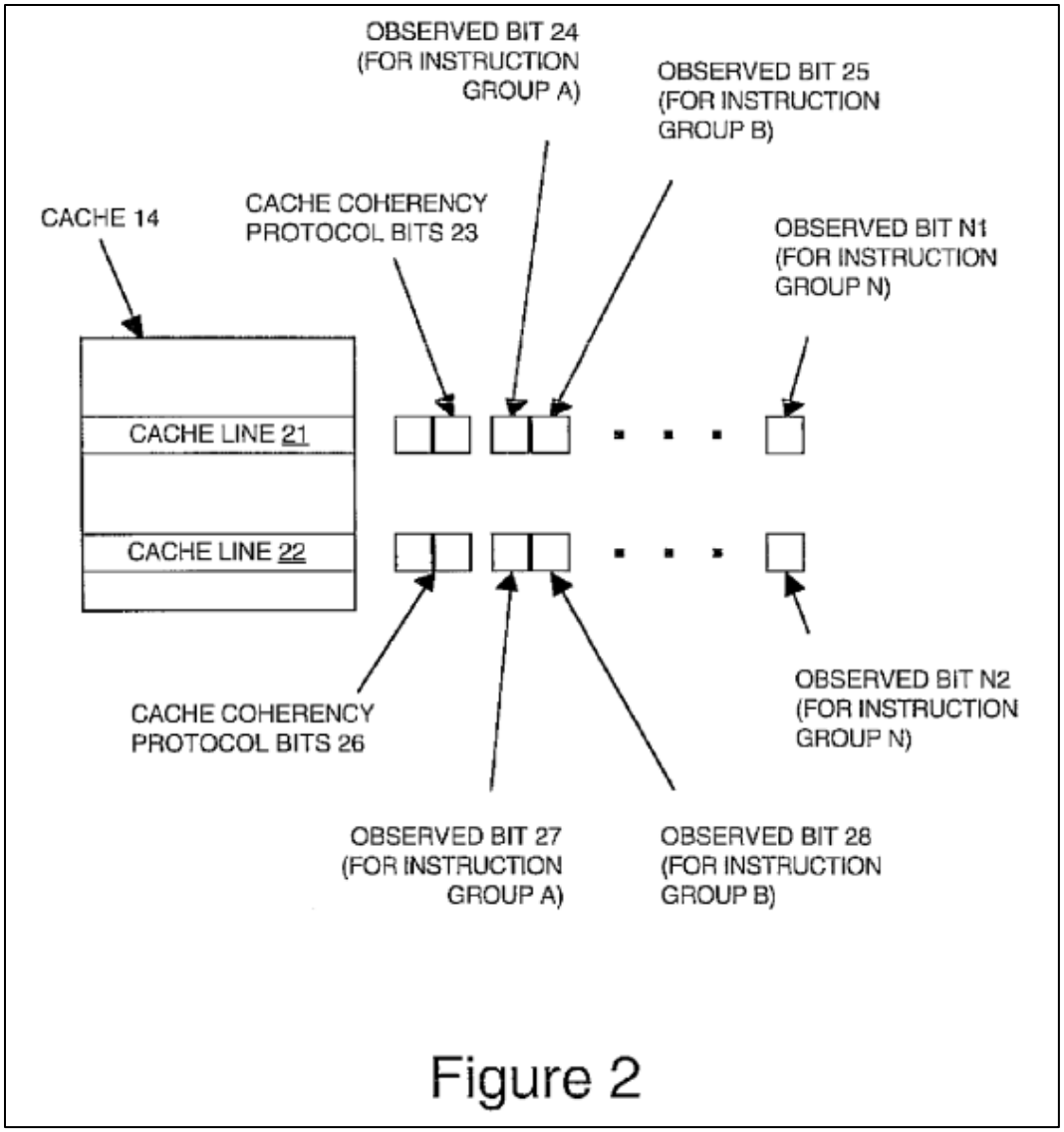


Figure 2

25. The observed bits associated with instruction group A, if set, remain set until the execution of instruction group A is ended. '395 patent at 5:25-27. For example, when instruction group A is either committed, or it is rolled back, or it is aborted, then the associated observed bit **24** in cache line **21** is cleared. Similarly, when instruction group B is either committed, or it is rolled back, or it is aborted, then the associated observed bit **25** in cache line **21** is cleared.

26. Alternatively, if an external agent **11-13** snoops cache line **21** and observed bits **24** and **25** have been set, then instruction group A (corresponding to observed bit **24**) and instruction group B (corresponding to observed bit **25**) may be rolled back and observed bits **24** and **25** may

be cleared. '395 patent at 5:58-61. Further, the patent describes an embodiment wherein if cache line **21** is in the Shared state, then there is no roll back of the instruction groups corresponding to any observed bit(s) in that cache line. '395 patent at 7:53-58.

B. The '395 Prosecution History

27. On May 15, 2008, prior to examination of the original claims, the Applicant filed a Preliminary Amendment. Among other changes, that Amendment cancelled all original claims (1-32) and replaced them with new claims 33-57. The pertinent claims for this Declaration, shown below, are claims 33-35.

> 33. (New) A method of managing memory in a computer system, said method comprising:
>
> executing a first group of instructions that causes a cache line of a cache memory to be read;
>
> setting a first bit of an indicator associated with said cache line to indicate that said cache line has been read;
>
> executing a second group of instructions that causes said cache line to be read; and
>
> setting a second bit of said indicator while said first bit remains set to indicate said cache line has been accessed by both said first group and said second group of instructions.

> 34. (New) The method of Claim 33 further comprising:
>
> executing a third group of instructions that causes said cache line to be accessed;
>
> using a value of said indicator to determine whether said cache line has been read and generating a result thereof; and
>
> handling said first group and said second group of instructions according to said result.

> 35. (New) The method of Claim 34 wherein said handling comprises rolling back said first group and said second group of instructions before granting said access provided said indicator indicates said cache line has been read, and otherwise granting said access without rolling back said first group and said second group of instructions.

28. On March 22, 2010, the Examiner issued a Non-Final Rejection. In pertinent part, claims 33-34 were rejected under 35 U.S.C. § 102(b) as being anticipated by prior art. Claim 35 was indicated to be allowable if the double patenting rejection were overcome and if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

29. On June 17, 2010, the Applicant submitted its Response to the Non-Final Rejection. Among other changes, each of claims 33-35 was amended as shown below:

> 33. (Currently Amended) A method of managing memory in a computer system, said method comprising:
>
> ~~executing a first group of instructions that causes a cache line of a cache memory to be read;~~
>
> setting a first bit of an indicator associated with a ~~said~~ cache line in a cache memory if ~~to indicate that~~ said cache line has been accessed in response to a processor executing an instruction in a first group of instructions ~~read~~;
>
> ~~executing a second group of instructions that causes said cache line to be read;~~ and
>
> setting a second bit of said indicator while said first bit remains set if ~~to indicate~~ said cache line has also been accessed in response to said processor executing an instruction in a ~~by both said first group and said~~ second group of instructions.

> 34. (Currently Amended) The method of Claim 33 further comprising:
>
> executing a third group of instructions that causes said cache line to be accessed;
>
> ~~using a value of said indicator to determine whether said cache line has been read and generating a result thereof;~~ and
>
> processing ~~handling~~ said first group and said second group of instructions according to a value of said indicator ~~said result~~.

> 35. (Currently Amended) The method of Claim 34 wherein said third group of instructions is executed by an agent other than said processor, wherein said processing ~~handling~~ comprises rolling back said first group of instructions provided said first bit is set and rolling back said second group of instructions provided said second bit is set before allowing an instruction in said third group to access said cache line ~~granting said access provided said indicator indicates said cache line has been read~~, and otherwise granting said access ~~without rolling back said first group and said second group of instructions~~.

30. As a result of the claim 34 amendment, the term "processing" was added to the claims for the first time. As reflected in this amendment, the step of "executing a third group of instructions" could be done by "said processor" recited in claim 33 or it could be done by "an agent other than said processor" (e.g., as reflected in the amendment to claim 35). Specifically, I understand that claim 34 must be broader in scope than its dependent claim 35 and therefore, as amended, a POSITA would understand the "executing" step of claim 34 could be performed by either said processor or an external agent.

31. Claim 35, which depended from claim 34, was further amended to recite "said processing comprises . . . ." As reflected in this amendment of claim 35, "said processing" is referring to the "processing said first group and said second group of instructions according to a value of said indicator" in amended claim 34.

32. The arguments made by the Applicant to traverse the rejection focused on the amendment to claim 33 and did not otherwise address the amendments to claims 34 and 35.

33. On August 26, 2010, the Examiner issued a Final Rejection. In pertinent part, claims 33-34 were rejected under 35 U.S.C. § 103(a) as being obvious in view of prior art. Claim 35 was indicated to be allowable if the double patenting rejection were overcome and if rewritten in independent form including all of the limitations of the base claim and any intervening claims. However, the Examiner's statement of reasons for allowability quoted the text of original claim 35 without acknowledging that some of those limitations were no longer part of the claim as amended. Specifically, I highlighted in the below excerpt of this Office Action the language that was ***deleted*** from amended claim 35:

> 19. Claims 35 and 43 disclose "rolling back said first group and said second group of instructions before granting said access provided said indicator indicates said cache line has been read, and otherwise granting said access without rolling back said first group and said second group of instructions." This limitation is taught in the specification, page 11, lines 15-21. This limitation in combination with other recited limitations of claims 35 and 43 are not taught or suggested by the prior art of record.

34. On October 18, 2010, in response to the Final Rejection, the Applicant argued for reconsideration of the rejection without otherwise amending claims 33-35. The arguments made by the Applicant focused on the limitations of claim 33 and did not otherwise address the limitations of claims 34 and 35.

35. The Applicant's arguments filed on October 18, 2010 were determined to be not persuasive. The Applicant then initiated an Appeal procedure. On April 28, 2014, the Patent Trial and Appeal Board issued its Decision on Appeal affirming the Examiner's finding that the prior art teaches or suggests the limitations of claim 33.

36. On June 30, 2014, in Response to the Decision on Appeal, the Applicant amended claim 33 and cancelled claims 34 and 35. However, as shown below, the amendment did not rewrite claim 35 in independent form:

> 33. (Currently Amended) A method of managing memory in a computer system, said method comprising:
>
> setting a first bit of an indicator associated with a cache line in a cache memory if said cache line has been accessed in response to a processor executing an instruction in a first group of instructions; ~~and~~
>
> setting a second bit of said indicator while said first bit remains set if said cache line has also been accessed in response to said processor executing an instruction in a second group of instructions;
>
> executing a third group of instructions that causes said cache line to be accessed, wherein said third group of instructions is executed by an agent other than said processor, wherein said processing comprises rolling back said first group of instructions provided said first bit is set and rolling back said second group of instructions provided said second bit is set before allowing an instruction in said third group to access said cache line, and otherwise granting said access; and
>
> processing said first group and said second group of instructions according to a value of said indicator.
>
> 34-35. (Canceled).

Specifically, the yellow-highlighted text was amended claim 35 and the purple-highlighted text was amended claim 34. As shown, the "said processing" limitation was added to the "executing" step, and now precedes the "processing" step.

37. The argument made by the Applicant was that "claim 33 is amended to incorporate the subject matter of allowable claim 35 and intervening claim 34" and "is therefore in condition

for allowance." I disagree that this version of claim 33 reflects a rewrite of dependent claim 35 in independent form. Further, as set forth in this Declaration, this amendment introduced further uncertainty over the scope of the claim that would not be reasonably ascertainable to a POSITA.

38. The next correspondence from the Patent Office was its July 22, 2014, Notice of Allowance. That Notice merely identifies the full text of the "executing" method step as not being taught or suggested by the prior art of record.

C. Level of Ordinary Skill in the Art

39. I have also been advised that a POSITA is a hypothetical person who is capable of understanding the scientific and engineering principles applicable to the relevant art of the claimed subject matter. I understand that a POSITA is presumed to have known the relevant prior art at the time of the alleged invention.

40. In my opinion, a POSITA as of the April 2005 priority date of the '395 patent would have had a bachelor's degree in electrical or computer engineering, computer science, or a related subject, and two years of professional experience "in the data processing arts" ('395 patent at 3:4-7) and multiprocessor speculative execution or a related field. Less experience may be necessary with additional education, and likewise, less education may be necessary with additional work experience.

41. I understand that the Plaintiff Intellectual Ventures has not yet offered its view of the level of ordinary skill in the art. To the extent that Intellectual Ventures or its expert provides its view of the level of ordinary skill in the art, I reserve the right to respond to the extent that opinion is different than mine. I also note that my opinions provided in this Declaration would not change in view of any insignificant modifications to this level of skill.

D. "Said Processing" (Claim 1)

Tesla's Proposed Construction	IV's Proposed Construction
Indefinite	Plain and ordinary meaning. Not indefinite.

42. I understand from counsel that the Court of Appeals for the Federal Circuit has held that a limitation lacking antecedent basis is indefinite if there is no clear correction for that lack of antecedent basis. Specifically, the lack of an antecedent basis renders a claim indefinite if the claim fails to apprise a POSITA of its scope and, therefore, fails to serve the statutory public notice function required for patent claims.

43. Claim 1 suffers from the fact that, as written, the limitation "***said*** processing" lacks antecedent basis, and the specification does not otherwise aid a POSITA in reasonably ascertaining what is being claimed. Indeed, the meaning of the term "***said*** processing" is objectively uncertain because it is the first reference to "processing," and it appears in the claim:

a. after the introduction of two structures that could perhaps perform "said processing" (i.e., "said processing" *might* refer to an action taken by the claimed "processor" or by the claimed "agent other than said processor)"; and

b. after the introduction of three groups of instructions that are being processed (i.e., "said processing" *might* refer to an action on "an instruction in a first group of instructions," "an instruction in a second group of instructions," or an instruction in "a third group of instructions").

44. With respect to ***which structure performs*** "said processing," that phrase appears within the method step for executing the third group of instructions "by an agent other than said processor," suggesting that "said processing" is done by ***the external agent***. However, the remainder of the limitation concerns the first and second groups of instructions, which are executed

by ***the processor*** in the earlier-recited method steps. Because the claim recites both "a processor" and "an agent other than said processor" before reciting the "said processing" limitation, a POSITA could not reasonably ascertain which of those structures performs "said processing" in claim 1.

45. With respect to ***what instruction is being processed*** by "said processing," that phrase appears within the method step for executing the third group of instructions, suggesting that the action is being taken on this ***third*** group of instructions. However, the remainder of the limitation concerns rolling back the ***first*** and the ***second*** groups of instructions. Because the claimed computer system is executing instructions from each of the first, second, and third groups of instructions, a POSITA could not reasonably ascertain which group of instructions is subject to "said processing" in claim 1.

46. Further, dependent claim 2—which refers to "said processing"—also adds uncertainty. Notably, "determining a state of said cache line" can be determined when processing any of the first, second, or third groups of instructions by either of the processor or the external agent by, *e.g.*, checking the cache coherency protocol bits shown in Figure 2.

47. Finally, there is added objective uncertainty as to the timing of "said processing," which a POSITA would recognize *might* be performed immediately or, alternatively, the cache line *might* be marked for a later rollback. For example, a POSITA would understand that the rollback *might* be suspended for a brief duration to give a conflicting instruction group a chance to complete. Indeed, the timing uncertainty further compounds the uncertainty over which structure—e.g., the processor or the agent other than the processor—performs "said processing."

* * *

I hereby declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct.

Date: January 9, 2025

By: Maurice Herlihy

Maurice Herlihy, Ph.D.

EXHIBIT 10

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	)	
	)	
Rozas	)	Examiner: OTTO, A.
	)	
Serial No.: 12/121,531	)	Art Unit: 2187
	)	
Filed: May 15, 2008	)	Confirmation No.: 7036
	)	
For: MEMORY MANAGEMENT FOR CACHE CONSISTENCY (As Amended)	)	

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

RESPONSE TO AN OFFICE ACTION

Dear Sir:

In response to the Office Action mailed March 22, 2010, the Applicant respectfully requests further examination and reconsideration of the above captioned patent application in view of the following amendments and remarks.

Amendments to the specification are presented on page 2, amendments to the claims are presented beginning on page 3, and remarks begin on page 12.

AMENDMENTS TO THE SPECIFICATION

Please amend the title of the application, which appears on pages 1 and 30 of the substitute specification, as follows:

MEMORY MANAGEMENT FOR ~~METHODS AND SYSTEMS THAT SUPPORT~~ CACHE CONSISTENCY

AMENDMENTS TO THE CLAIMS

1-32. (Canceled).

33. (Currently Amended) A method of managing memory in a computer system, said method comprising:

~~executing a first group of instructions that causes a cache line of a cache memory to be read;~~

setting a first bit of an indicator associated with a ~~said~~ cache line in a cache memory if ~~to indicate that~~ said cache line has been accessed in response to a processor executing an instruction in a first group of instructions ~~read~~;

~~executing a second group of instructions that causes said cache line to be read;~~ and

setting a second bit of said indicator while said first bit remains set if ~~to indicate~~ said cache line has also been accessed in response to said processor executing an instruction in a ~~by both said first group and said~~ second group of instructions.

34. (Currently Amended) The method of Claim 33 further comprising:

executing a third group of instructions that causes said cache line to be accessed;

~~using a value of said indicator to determine whether said cache line has been read and generating a result thereof;~~ and

processing ~~handling~~ said first group and said second group of instructions according to a value of said indicator ~~said result~~.

35. (Currently Amended) The method of Claim 34 wherein said third group of instructions is executed by an agent other than said processor, wherein said processing ~~handling~~ comprises rolling back said first group of instructions provided said first bit is set and rolling back said second group of instructions provided said second bit is set before allowing an instruction in said third group to access said cache line ~~granting said access provided said indicator indicates said cache line has been read~~, and otherwise granting said access ~~without rolling back said first group and said second group of instructions~~.

36. (Currently Amended) The method of Claim 35 wherein said processing ~~handling~~ further comprises:

determining a state of said cache line, wherein said state is specified according to a cache coherency protocol comprising at least a modified state, a shared state, and an invalid state; and

rolling back said first group of instructions provided said first bit is set and rolling back said second group of instructions provided said second bit is set, and then granting said access provided said cache line is in other than said shared state ~~and said indicator indicates said cache line has been read~~, and otherwise granting said access ~~without rolling back said first group and said second group of instructions~~ only when said cache line is in said shared state.

37. (Original) The method of Claim 35 further comprising maintaining a first count of how many times said first group of instructions has been rolled back and a second count of how many times said second group of instructions has been rolled back.

38. (Original) The method of Claim 37 further comprising eliminating one or more instructions from said first group of instructions if said first count exceeds a threshold and eliminating one or more instructions from said second group of instructions if said second count exceeds said threshold.

39. (Original) The method of Claim 33 further comprising clearing said first bit when said execution of said first group of instructions is ended, wherein said second bit remains set until execution of said second group of instructions is ended.

40. (Original) The method of Claim 33 further comprising:

executing a plurality of groups of instructions in parallel;

changing an indicator associated with each cache line in said cache that has been accessed because of an instruction in any of said plurality of groups of instructions; and

provided all cache lines in said cache memory are indicated as having been accessed, rolling back a group of instructions that includes an instruction that when executed attempts to replace information in a cache line of said cache memory, and otherwise continuing said executing of said plurality of groups of instructions.

41. (Currently Amended) A computer system comprising:

a ~~first~~ processor;

a cache memory used by said ~~first~~ processor; and

a memory unit coupled to said ~~first~~ processor and having stored therein instructions, said instructions ~~said memory unit containing instructions that when executed implement a memory management method, said method~~ comprising:

instructions to execute ~~executing~~ a first group of instructions using said processor, wherein said first group of instructions includes an instruction that causes a cache line of said cache memory to be read;

instructions to set ~~setting~~ a first bit of an indicator associated with said cache line to indicate that said cache line has been read;

instructions to execute ~~executing~~ a second group of instructions using said ~~first~~ processor, wherein said second group of instructions includes an instruction that causes said cache line to be read; and

instructions to set ~~setting~~ a second bit of said indicator while said first bit remains set to indicate said cache line has been accessed by both said first group of instructions and said second group of instructions.

42. (Currently Amended) The computer system of Claim 41 wherein said instructions ~~method~~ further comprise ~~comprises~~:

instructions to execute ~~executing~~ a third group of instructions that causes said cache line to be accessed;

~~using a value of said indicator to determine whether said cache line has been read and generating a result thereof;~~ and

instructions to process ~~handling~~ said first group and said second group of instructions according to a value of said indicator ~~said result~~.

43. (Currently Amended) The computer system of Claim 42 wherein said third group of instructions is for execution by an agent other than said processor,

wherein said instructions ~~method~~ further comprise instructions to roll ~~comprises rolling~~ back said first group of instructions provided said first bit is set and to roll back said second group of instructions provided said second bit is set before allowing an instruction in said third group to access said cache line, wherein otherwise said access is granted. ~~granting said access provided said indicator indicates said cache line has been read, and otherwise granting said access without rolling back said first group and said second group of instructions.~~

44. (Currently Amended) The computer system of Claim 43 wherein said instructions ~~method~~ further comprise ~~comprises~~:

instructions to determine ~~determining~~ a state of said cache line, wherein said state is specified according to a cache coherency protocol comprising at least a modified state, a shared state, and an invalid state; and

instructions to roll ~~rolling~~ back said first group of instructions provided said first bit is set and to roll back said second group of instructions provided said second bit is set, wherein ~~and then granting~~ said access is granted provided said cache line is in other than said shared state ~~and said indicator indicates said cache line has been read~~, and wherein otherwise ~~granting~~ said access is granted ~~without rolling back said first group and said second group of instructions~~ only when said cache line is in said shared state.

45. (Currently Amended) The computer system of Claim 44 wherein said instructions ~~method~~ further comprise instructions to maintain ~~comprises maintaining~~ a first count of how many times said first group of instructions has been rolled back and a second count of how many times said second group of instructions has been rolled back.

46. (Currently Amended) The computer system of Claim 45 wherein said instructions ~~method~~ further comprise instructions to eliminate ~~comprises eliminating~~ one or more instructions from said first group of instructions if said first count exceeds a threshold and to eliminate ~~eliminating~~ one or more instructions from said second group of instructions if said second count exceeds said threshold.

47. (Currently Amended) The computer system of Claim 41 wherein said instructions ~~method~~ further comprise instructions to clear ~~comprises clearing~~ said first bit when said execution of said first group of instructions is ended, wherein said second bit remains set until execution of said second group of instructions is ended.

48. (Currently Amended) The computer system of Claim 41 wherein said instructions ~~method~~ further comprise ~~comprises~~:

instructions to execute ~~executing~~ a plurality of groups of instructions in parallel using ~~on~~ said ~~first~~ processor;

instructions to change ~~changing~~ an indicator associated with each cache line in said cache memory that has been accessed because of an instruction in any of said plurality of groups of instructions; and

~~provided all cache lines in said cache memory are indicated as having been accessed,~~ instructions to roll ~~rolling~~ back a group of instructions that includes an instruction that when executed attempts to replace information in a cache line of said cache memory provided all cache lines in said cache memory

are indicated as having been accessed, wherein ~~and~~ otherwise ~~continuing said executing of~~ said plurality of groups of instructions continue to be executed.

49. (Currently Amended) A method of managing shared memory in a computer system, said method comprising:

executing a first group of instructions with ~~on~~ a ~~first~~ processor;

associating a first state with a cache line in a cache comprising a plurality of cache lines, wherein said first state is specified according to a cache coherency protocol comprising at least a modified state, a shared state, and an invalid state;

associating a second state with said cache line during said executing of said first group of instructions, wherein said second state indicates that said cache line has been accessed because of said first group of instructions;

executing a second group of instructions with ~~on~~ said ~~first~~ processor; and

associating a third state with said cache line during said executing of said second group of instructions, wherein said third state indicates that said cache line has been accessed because of said second group of instructions.

50. (Currently Amended) The method of Claim 49 wherein said first group of instructions includes multiple memory operations and said second group of instructions ~~each~~ includes multiple memory operations.

51. (Currently Amended) The method of Claim 49 further comprising:

executing a third group of instructions that causes said cache line to be accessed;

~~using said second state and said second state to indicate said cache line has been accessed;~~ and

rolling back said first group of instructions and said second group of instructions before granting said access.

52. (Original) The method of Claim 51 wherein said rolling back is only performed provided said cache line is in other than said shared state.

53. (Original) The method of Claim 51 further comprising maintaining a first count of how many times said first group of instructions is rolled back and a second count of how many times said second group of instructions is rolled back.

54. (Original) The method of Claim 53 further comprising eliminating one or more instructions from said first group of instructions if said first count exceeds a threshold and eliminating one or more instructions from said second group of instructions if said second count exceeds said threshold.

55. (Original) The method of Claim 49 wherein said second state is indicated by setting a first bit associated with said cache line, wherein said first bit is cleared when said execution of said first group of instructions is ended, and wherein said third state is indicated by setting a second bit associated with said cache line, wherein said second bit is cleared when said execution of said second group of instructions is ended.

56. (Currently Amended) A computer system comprising:

a processor;

a cache memory coupled to said processor and having a plurality of cache lines; and

a memory unit coupled to said processor, wherein said memory unit has a ~~plurality of entries and a~~ plurality of indicator bits per cache line ~~entry~~, wherein a first indicator bit associated with a cache line of said cache memory ~~plurality of indicator bits~~ is set to indicate that said [[a]] cache line ~~of said cache memory~~ has been accessed responsive to said processor executing ~~because of~~ a first group of instructions and wherein said first indicator bit is cleared when execution of said first group of instructions is ended, wherein a second indicator bit associated with said cache line ~~of said plurality of indicator bits~~ is set to indicate that said cache line has been accessed responsive to said processor executing ~~because of~~ a second group of instructions in parallel with said first group and wherein said second indicator bit is cleared when execution of said second group of instructions is ended.

57. (Currently Amended) The computer system of Claim 56 wherein a subset of address bits of said cache line is used to identify an ~~said~~ entry in said memory unit ~~of said plurality of entries~~, wherein ~~in response to said cache line being accessed,~~ an indicator bit of said entry ~~corresponding to said subset of address bits~~ is set in response to said cache line being accessed.

REMARKS

Claims 33-57 are pending. Claims 33-36, 41-51, 56, and 57 are amended herein. No new matter is introduced by the claim amendments.

Claims 35-38, 40, 43-46, and 51-54 are allowable. Applicant thanks the Examiner for indicating the allowability of those claims.

§ 112 Rejections

According to the Office Action, Claims 51-54 are rejected under 35 U.S.C. § 112, second paragraph. Specifically, Claim 51 is rejected, and Claims 52-54 are rejected because of their dependence on Claim 51.

Applicant respectfully submits that, as amended, Claim 51 and its dependent claims traverse the basis for rejection under 35 U.S.C. § 112, second paragraph.

Double Patenting

According to the Office Action, Claims 33-57 are rejected on the ground of non-statutory obviousness-type double patenting as being unpatentable over U.S. Patent No. 7,376,798.

A timely filed terminal disclaimer may be filed to overcome a rejection based on non-statutory double patenting. Applicant respectfully submits that a terminal disclaimer will be considered once the application is otherwise considered to be allowable. The Applicant respectfully requests that the double-patenting rejection be held in abeyance until the claims are otherwise allowable.

§ 102 Rejections

According to the Office Action, Claims 33, 34, 39, 41, 42, 47, 56, and 57 are rejected under 35 U.S.C. § 102(b) as allegedly being anticipated by Gray et al. ("Gray;" U.S. Patent No. 5,086,429). The Applicant has reviewed the cited reference and respectfully submits that the features recited in Claims 33, 34, 39, 41, 42, 47, 56, and 57 are not shown or suggested by Gray.

According to Gray (column 3, lines 53-62), there are two pairs of processors, "with one pair being arbitrarily designated the 'active pair' while the other is designated as a 'hot stand-by pair.' Each processor in each pair … is made to simultaneously execute the same program of instructions …, [and] the two processors in each pair share a common memory." "During a read by the processors, the appropriate code word is read from the memory and the syndrome generators determine whether an error exists" (column 4, lines 8-10, of Gray). "Checkers 28 and 29 provide two redundant output signals for indicating that an error has been detected. These signals are called FAIL 1 and FAIL 2" (column 9, lines 1-4, of Gray). "Checkers 28 and 29 comprise digital comparators for comparing, bit-for-bit, the address bits and the control bits from the dual processors" (column 9, lines 12-14, of Gray). "If there is a miscompare …, then the error flag signals FAIL 1 and FAIL 2 become true" (column 9, lines 15-19, of Gray).

In other words, according to Gray, the error flag signals are set if a pair of redundant processors reads the same information differently.

Therefore, Applicant respectfully submits that Gray does not teach or suggest "setting a first bit of an indicator associated with a cache line in a cache memory if said cache line has been accessed in response to a processor executing an instruction in a first group of instructions; and setting a second bit of said indicator while said first bit remains set if said cache line has also been accessed in response to said processor executing an instruction in a second group of instructions" as recited in independent Claim 33 (emphasis added); "instructions to execute a first group of instructions using said processor, wherein said first group of instructions includes an instruction that causes a cache line of said cache memory to be read; instructions to set a first bit of an indicator associated with said cache line to indicate that said cache line has been read; instructions to execute a second group of instructions using said processor, wherein said second group of instructions includes an instruction that causes said cache line to be read; and instructions to set a second bit of said indicator while said first bit remains set to indicate said cache line has been accessed by both said first group of instructions and said second group of instructions" as recited in independent Claim 41; or "wherein a first indicator bit associated with a cache line of said cache memory is set to indicate that said cache line has been accessed responsive to said processor executing a first group of instructions …, wherein a second indicator bit associated with said cache line is set to indicate that said cache line has been accessed responsive to said processor executing a second group of instructions in parallel with said first group" as recited in independent Claim 56.

Accordingly, Applicant respectfully submits that Gray does not show or suggest the features of independent Claims 33, 41, and 56, and that those claims are in condition for allowance.

Each of the Claims 34 and 39 includes all of the features of Claim 33 plus additional features. Each of the Claims 42 and 47 includes all of the features of independent Claim 41 plus additional features. Claim 57 includes all of the features of independent Claim 56 plus additional features. Applicant respectfully submits that Gray does not show or suggest the features of Claims 34, 39, 42, 47, and 57 in combination with the features of their respective base claim, and that Claims 34, 39, 42, 47, and 57 are in condition for allowance at least because they depend from an allowable claim.

In conclusion, Applicant respectfully submits that Claims 33, 34, 39, 41, 42, 47, 56, and 57 traverse the basis for rejection under 35 U.S.C. § 102(b).

§ 103 Rejections

According to the Office Action, Claims 49, 50, and 55 are rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over Gray in view of Examiner's official notice (in view of what is known in the art).

By rationale similar to that presented above, Applicant respectfully submits that Gray does not teach or suggest "executing a first group of instructions with a processor; … associating a second state with said cache line during said executing of said first group of instructions, wherein said second state indicates that said cache line has been accessed because of said first group of

instructions; executing a second group of instructions with said processor; and associating a third state with said cache line during said executing of said second group of instructions, wherein said third state indicates that said cache line has been accessed because of said second group of instructions" as recited in independent Claim 49.

Accordingly, Applicant respectfully submits that Claim 49 is in condition for allowance.

Each of the Claims 50 and 55 includes all of the features of Claim 49 plus additional features. Applicant respectfully submits that Gray does not show or suggest the features of Claims 50 and 55 in combination with the features of Claim 49, and that Claims 50 and 55 are in condition for allowance at least because they depend from an allowable claim.

In conclusion, Applicant respectfully submits that Claims 49, 50, and 55 traverse the basis for rejection under 35 U.S.C. § 103(a).

Conclusions

In light of the above remarks, Applicant respectfully requests reconsideration of the rejected claims.

Based on the arguments presented above, Applicant respectfully asserts that Claims 33-57 overcome the rejections of record, and therefore Applicant respectfully solicits allowance of those claims.

The Examiner is invited to contact Applicant's undersigned representative if the Examiner believes such action would expedite resolution of the present application.

Please charge any fees or apply any credits to our PTO deposit account number: 50-4160.

Respectfully submitted,

MURABITO HAO & BARNES LLP

Date: 6/17/2010

/William A. Zarbis/

William A. Zarbis
Reg. No. 46,120

Address: Two North Market Street, Third Floor
San Jose, California 95113
Phone: (408) 938-9080 Ext. 113

EXHIBIT 11

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	)	
	)	
Rozas	)	Examiner: OTTO, A.
	)	
Serial No.: 12/121,531	)	Art Unit: 2187
	)	
Filed: May 15, 2008	)	Confirmation No.: 7036
	)	
For: MEMORY MANAGEMENT FOR CACHE CONSISTENCY (As Amended)	)	

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

RESPONSE TO DECISION ON APPEAL

Dear Sir:

In response to the Decision on Appeal mailed April 28, 2014, and per the Request for Continued Examination filed herewith, the Applicant respectfully requests further examination and reconsideration of the above captioned patent application in view of the following amendments and remarks. Amendments to the claims are presented beginning on page 2, and remarks begin on page 10.

AMENDMENTS TO THE CLAIMS

1-32. (Canceled).

33. (Currently Amended) A method of managing memory in a computer system, said method comprising:

setting a first bit of an indicator associated with a cache line in a cache memory if said cache line has been accessed in response to a processor executing an instruction in a first group of instructions; ~~and~~

setting a second bit of said indicator while said first bit remains set if said cache line has also been accessed in response to said processor executing an instruction in a second group of instructions;

executing a third group of instructions that causes said cache line to be accessed, wherein said third group of instructions is executed by an agent other than said processor, wherein said processing comprises rolling back said first group of instructions provided said first bit is set and rolling back said second group of instructions provided said second bit is set before allowing an instruction in said third group to access said cache line, and otherwise granting said access; and

processing said first group and said second group of instructions according to a value of said indicator.

34-35. (Canceled).

36. (Currently Amended) The method of Claim 33 [[35]] wherein said processing further comprises:

determining a state of said cache line, wherein said state is specified according to a cache coherency protocol comprising at least a modified state, a shared state, and an invalid state; and

rolling back said first group of instructions provided said first bit is set and rolling back said second group of instructions provided said second bit is set, and then granting said access provided said cache line is in other than said shared state, and otherwise granting said access only when said cache line is in said shared state.

37. (Currently Amended) The method of Claim 33 [[35]] further comprising maintaining a first count of how many times said first group of instructions has been rolled back and a second count of how many times said second group of instructions has been rolled back.

38. (Original) The method of Claim 37 further comprising eliminating one or more instructions from said first group of instructions if said first count exceeds a threshold and eliminating one or more instructions from said second group of instructions if said second count exceeds said threshold.

39. (Original) The method of Claim 33 further comprising clearing said first bit when said execution of said first group of instructions is ended, wherein said second bit remains set until execution of said second group of instructions is ended.

40. (Original) The method of Claim 33 further comprising:

executing a plurality of groups of instructions in parallel;

changing an indicator associated with each cache line in said cache that has been accessed because of an instruction in any of said plurality of groups of instructions; and

provided all cache lines in said cache memory are indicated as having been accessed, rolling back a group of instructions that includes an instruction that when executed attempts to replace information in a cache line of said cache memory, and otherwise continuing said executing of said plurality of groups of instructions.

41. (Currently Amended) A computer system comprising:

a processor;

a cache memory for use by said processor; and

a memory unit coupled to said processor and having stored therein instructions, said instructions comprising:

instructions to execute a first group of instructions using said processor, wherein said first group of instructions includes an instruction that causes a cache line of said cache memory to be read;

instructions to set a first bit of an indicator associated with said cache line to indicate that said cache line has been read;

instructions to execute a second group of instructions using said processor, wherein said second group of instructions includes an instruction that causes said cache line to be read; ~~and~~

instructions to set a second bit of said indicator while said first bit remains set to indicate said cache line has been accessed by both said first group of instructions and said second group of instructions;

instructions to execute a third group of instructions that causes said cache line to be accessed, wherein said third group of instructions is for execution by an agent other than said processor, wherein said instructions further comprise instructions to roll back said first group of instructions provided said first bit is set and to roll back said second group of instructions provided said second bit is set before allowing an instruction in said third group to access said cache line, wherein otherwise said access is granted; and

instructions to process said first group and said second group of instructions according to a value of said indicator.

42-43. (Canceled).

44. (Currently Amended) The computer system of Claim 41 [[43]] wherein said instructions further comprise:

instructions to determine a state of said cache line, wherein said state is specified according to a cache coherency protocol comprising at least a modified state, a shared state, and an invalid state; and

instructions to roll back said first group of instructions provided said first bit is set and to roll back said second group of instructions provided said second bit is set, wherein said access is granted provided said cache line is in other than said shared state, and wherein otherwise said access is granted only when said cache line is in said shared state.

45. (Previously Presented) The computer system of Claim 44 wherein said instructions further comprise instructions to maintain a first count of how many times said first group of instructions has been rolled back and a second count of how many times said second group of instructions has been rolled back.

46. (Previously Presented) The computer system of Claim 45 wherein said instructions further comprise instructions to eliminate one or more instructions from said first group of instructions if said first count exceeds a threshold and to eliminate one or more instructions from said second group of instructions if said second count exceeds said threshold.

47. (Previously Presented) The computer system of Claim 41 wherein said instructions further comprise instructions to clear said first bit when said execution of said first group of instructions is ended, wherein said second bit remains set until execution of said second group of instructions is ended.

48. (Previously Presented) The computer system of Claim 41 wherein said instructions further comprise:

instructions to execute a plurality of groups of instructions in parallel using said processor;

instructions to change an indicator associated with each cache line in said cache memory that has been accessed because of an instruction in any of said plurality of groups of instructions; and

instructions to roll back a group of instructions that includes an instruction that when executed attempts to replace information in a cache line of said cache

memory provided all cache lines in said cache memory are indicated as having been accessed, wherein otherwise said plurality of groups of instructions continue to be executed.

49. (Currently Amended) A method of managing shared memory in a computer system, said method comprising:

executing a first group of instructions with a processor;

associating a first state with a cache line in a cache comprising a plurality of cache lines, wherein said first state is specified according to a cache coherency protocol comprising at least a modified state, a shared state, and an invalid state;

associating a second state with said cache line during said executing of said first group of instructions, wherein said second state indicates that said cache line has been accessed because of said first group of instructions;

executing a second group of instructions with said processor; ~~and~~

associating a third state with said cache line during said executing of said second group of instructions, wherein said third state indicates that said cache line has been accessed because of said second group of instructions;

executing a third group of instructions that causes said cache line to be accessed; and

rolling back said first group of instructions and said second group of instructions before granting said access.

50. (Previously Presented) The method of Claim 49 wherein said first group of instructions includes multiple memory operations and said second group of instructions includes multiple memory operations.

51. (Canceled).

52. (Currently Amended) The method of Claim 49 [[51]] wherein said rolling back is only performed provided said cache line is in other than said shared state.

53. (Currently Amended) The method of Claim 49 [[51]] further comprising maintaining a first count of how many times said first group of instructions is rolled back and a second count of how many times said second group of instructions is rolled back.

54. (Original) The method of Claim 53 further comprising eliminating one or more instructions from said first group of instructions if said first count exceeds a threshold and eliminating one or more instructions from said second group of instructions if said second count exceeds said threshold.

55. (Original) The method of Claim 49 wherein said second state is indicated by setting a first bit associated with said cache line, wherein said first bit is cleared when said execution of said first group of instructions is ended, and wherein said third state is indicated by setting a second bit associated with said cache line, wherein said second bit is cleared when said execution of said second group of instructions is ended.

56. (Currently Amended) A computer system comprising:

a processor;

a cache memory coupled to said processor and having a plurality of cache lines; and

a memory unit coupled to said processor, wherein said memory unit has a plurality of indicator bits per cache line, wherein a first indicator bit associated with a cache line of said cache memory is set to indicate that said cache line has been accessed responsive to said processor executing a first group of instructions and wherein said first indicator bit is cleared when execution of said first group of instructions is ended, wherein a second indicator bit associated with said cache line is set to indicate that said cache line has been accessed responsive to said processor executing a second group of instructions in parallel with said first group and wherein said second indicator bit is cleared when execution of said second group of instructions is ended;

wherein further, in response to an agent other than said processor executing a third group of instructions that causes said cache line to be accessed, said first group of instructions is rolled back provided said first bit is set and said second group of instructions is rolled back provided said second bit is set before allowing an instruction in said third group to access said cache line, wherein otherwise said access is granted.

57. (Previously Presented) The computer system of Claim 56 wherein a subset of address bits of said cache line is used to identify an entry in said memory unit, wherein an indicator bit of said entry is set in response to said cache line being accessed.

REMARKS

The Office Action mailed August 26, 2010, and the Advisory Action mailed October 29, 2010, were the last Office Actions issued prior to the Decision on Appeal. According to those Office Actions, Claims 33-57 stand rejected on the ground of non-statutory obviousness-type double patenting; Claims 33, 34, 39, 41, 42, 47, 49, 50, 55, 56, and 57 stand rejected under 35 U.S.C. § 103(a), as more fully described below; and Claims 35-38, 40, 43-46, and 51-54 are allowable if written in independent form once the double patenting rejection is overcome.

Applicant thanks the Examiner for indicating the allowability of Claims 35-38, 40, 43-46, and 51-54.

Please cancel Claims 34, 35, 42, 43, and 51 without prejudice. Claims 33, 36-41, 44-50, and 52-57 are pending. Claims 33, 36, 37, 41, 44, 49, 52, 53, and 56 are amended.

Double Patenting

Remaining Claims 33, 36-41, 44-50, and 52-57 stand rejected on the ground of non-statutory obviousness-type double patenting as being unpatentable over U.S. Patent No. 7,376,798. A terminal disclaimer is filed along with this response, thereby overcoming the double patenting rejection.

§ 103 Rejections

Claims 33, 39, 41, 47, 56, and 57

Claims 33, 39, 41, 47, 56, and 57 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over Gray et al. ("Gray;" U.S. Patent No. 5,086,429) in view of Hady et al. ("Hady;" U.S. Patent Application Publication No. 2006/0112226).

Claim 33 is amended to incorporate the subject matter of allowable Claim 35 and intervening Claim 34. Claim 33 is therefore in condition for allowance, and Claim 39 is also in condition for allowance because of its dependence on Claim 33. Therefore, a discussion of the rejection of Claims 33 and 39 under 35 U.S.C. § 103(a) is rendered moot at this time.

Claim 41 is amended to incorporate the subject matter of allowable Claim 43 and intervening Claim 42. Claim 41 is therefore in condition for allowance, and Claim 47 is also in condition for allowance because of its dependence on Claim 41. Therefore, a discussion of the rejection of Claims 41 and 47 under 35 U.S.C. § 103(a) is rendered moot at this time.

With regard to Claim 56, Applicant respectfully submits that neither Gray nor Hady nor the combination thereof teaches or suggests the features included in the amendment to that claim, namely "in response to an agent other than said processor executing a third group of instructions that causes said cache line to be accessed, said first group of instructions is rolled back provided said first bit is set and said second group of instructions is rolled back provided said second bit is set before allowing an instruction in said third group to access said cache line,

wherein otherwise said access is granted." Both of those references are silent with regard to the recited features. Thus, Claim 56 overcomes the basis for rejection under 35 U.S.C. § 103(a) and is in condition for allowance, and Claim 57 is in condition for allowance at least because of its dependence on an allowable claim.

Claims 49, 50, and 55

Claims 49, 50, and 55 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over Gray and Hady in view of Examiner's official notice.

Claim 49 is amended to incorporate the subject matter of allowable Claim 51. Claim 49 is therefore in condition for allowance, and Claims 50 and 55 are also in condition for allowance because of their dependence on Claim 49. Therefore, a discussion of the rejection of Claims 49, 50, and 55 under 35 U.S.C. § 103(a) is rendered moot at this time.

Conclusions

In light of the above amendments and remarks, Applicant respectfully requests reconsideration of the rejected claims.

Based on the amendments and arguments presented above, Applicant respectfully asserts that pending Claims 33, 36-41, 44-50, and 52-57 overcome the rejections of record, and therefore Applicant respectfully solicits allowance of those claims.

The Examiner is invited to contact Applicant's undersigned representative if the Examiner believes such action would expedite resolution of the present application.

Please charge any fees or apply any credits to our PTO deposit account number: 50-4160.

Respectfully submitted,

MURABITO HAO & BARNES LLP

Date: June 30, 2014

/William A. Zarbis/

William A. Zarbis
Reg. No. 46,120
Address: Two North Market Street, Third Floor
San Jose, California 95113
Phone: (408) 938-9080 Ext. 113

EXHIBIT 12

US008493909B2

(12) United States Patent
Choi

(10) Patent No.: US 8,493,909 B2
(45) Date of Patent: *Jul. 23, 2013

(54) **METHOD FOR THE TRANSMISSION OF DATA FIELD OF TECHNOLOGY**

(75) Inventor: **Hyung-Nam Choi**, Hamburg (DE)

(73) Assignee: **Intel Mobile Communications GmbH**, Neubiberg (DE)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **13/273,256**

(22) Filed: **Oct. 14, 2011**

(65) **Prior Publication Data**

US 2012/0026921 A1 Feb. 2, 2012

Related U.S. Application Data

(63) Continuation of application No. 10/595,245, filed as application No. PCT/EP2004/051958 on Aug. 30, 2004, now Pat. No. 8,040,833.

(30) **Foreign Application Priority Data**

Sep. 29, 2003 (DE) 103 45 220

(51) **Int. Cl.**
H04Q 7/00 (2006.01)

(52) **U.S. Cl.**
USPC **370/322**; 370/329; 370/342; 370/338; 370/335; 370/441; 370/350; 370/328; 370/479

(58) **Field of Classification Search**
None
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

7,359,345	B2 *	4/2008	Chang et al.	370/322
8,089,879	B2 *	1/2012	Hu et al.	370/235
2004/0228313	A1	11/2004	Cheng et al.	
2009/0274086	A1 *	11/2009	Petrovic et al.	370/312

FOREIGN PATENT DOCUMENTS

CN	1402566	A	3/2003
EP	0993137	A1	4/2000
EP	993137	A1 *	4/2000
WO	2004073273	A1	8/2004

OTHER PUBLICATIONS

3GPP TS 25.321 Version 5.6.0.*
3GPP TS 25.321 V5.6.0 (Sep. 2003) 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Medium Access Control (MAC) protocol specification Release 5; pp. 1-55.
3GPP TS 25.321 V5.5.0 (Jun. 2003) 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Medium Access Control (MAC) protocol specification Release 5; pp. 1-54.
English abstract of CN 1402566 A (Mar. 12, 2003).
3GPP TS 36.321 V9.3.0, 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification (Release 9), Jun. 2010, pp. 1-48.

* cited by examiner

Primary Examiner — Olumide T Ajibade Akonai
Assistant Examiner — Randy Peaches

(57) **ABSTRACT**

A method, system and computer program product for the transmission of data from a transmitter to receiver. The present invention is directed to improving data transmission. To this end, the transmission of data is accelerated, and an inband-signaling of information is carried out on an MAC-layer plane, wherein the information is particularly relevant to the base station (BS).

20 Claims, 10 Drawing Sheets

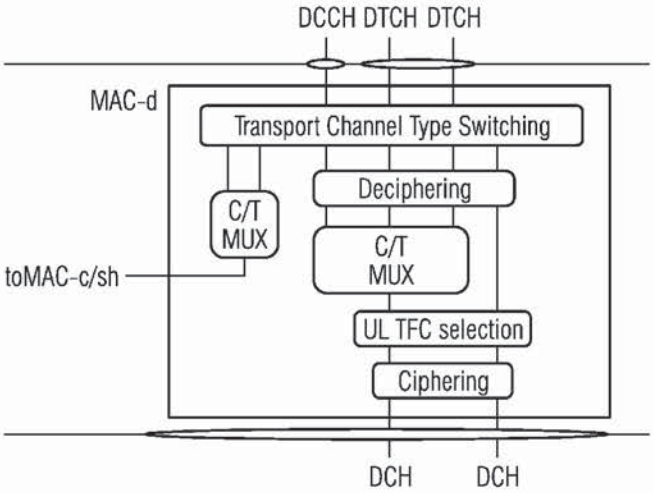


FIG 1

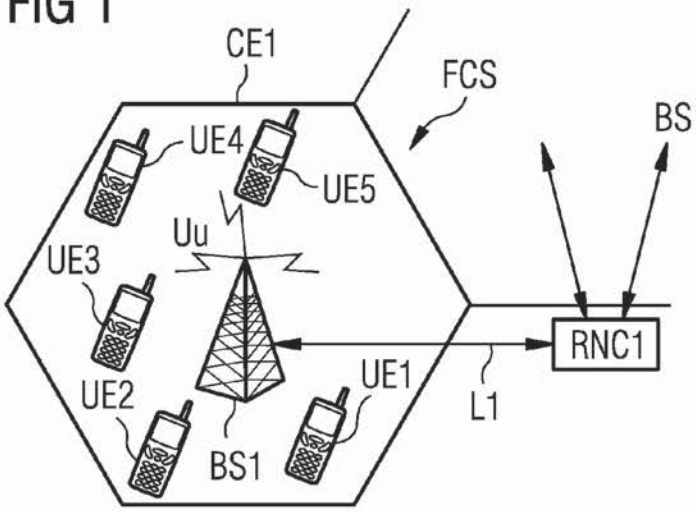

CE1
FCS
BS
UE4
UE5
Uu
UE3
RNC1
UE1
UE2
L1
BS1

FIG 2

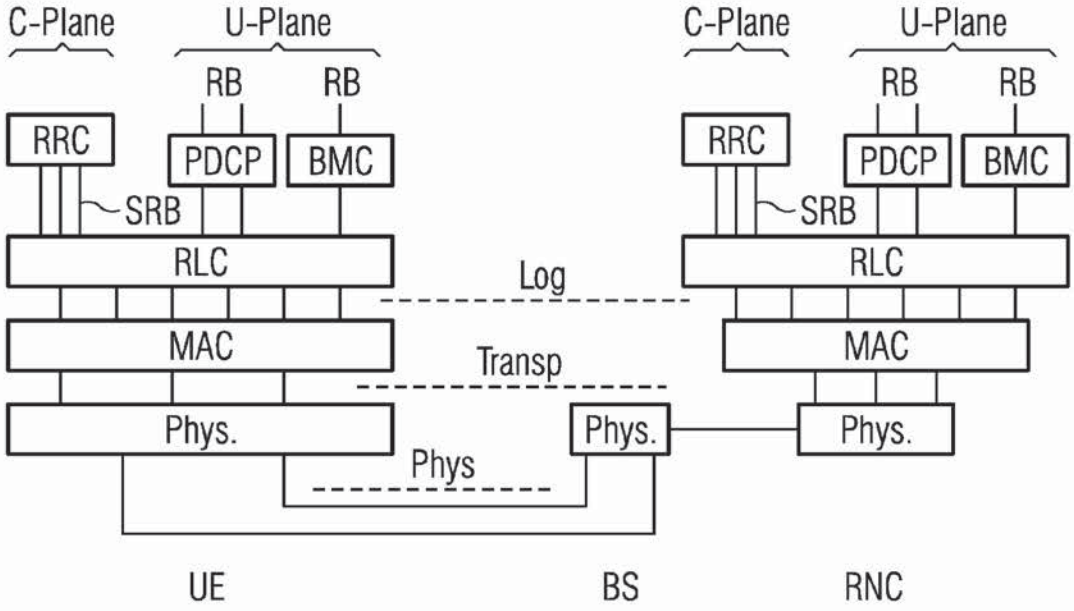

C-Plane
U-Plane
C-Plane
U-Plane
RB
RB
RB
RB
RRC
PDCP
BMC
RRC
PDCP
BMC
SRB
SRB
RLC
RLC
Log
MAC
MAC
Transp
Phys.
Phys.
Phys.
Phys
UE
BS
RNC

FIG 3

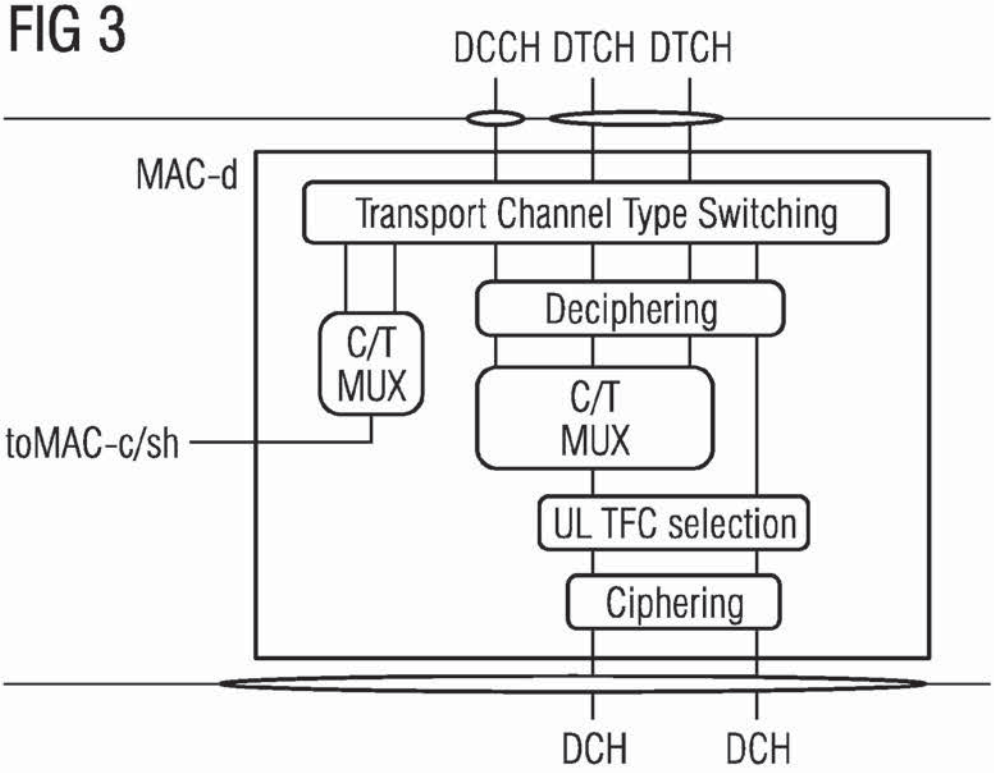

DCCH DTCH DTCH
MAC-d
Transport Channel Type Switching
Deciphering
C/T MUX
C/T MUX
toMAC-c/sh
UL TFC selection
Ciphering
DCH
DCH

FIG 4

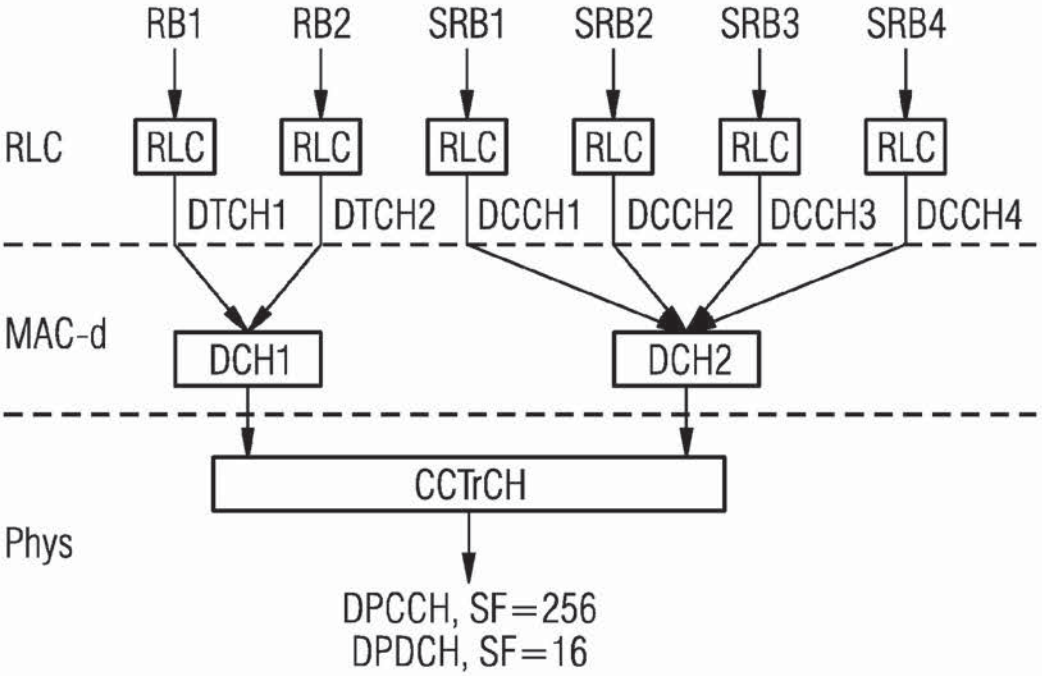

RB1
RB2
SRB1
SRB2
SRB3
SRB4
RLC
RLC
RLC
RLC
RLC
RLC
RLC
DTCH1
DTCH2
DCCH1
DCCH2
DCCH3
DCCH4
MAC-d
DCH1
DCH2
CCTrCH
Phys
DPCCH, SF=256
DPDCH, SF=16

FIG 5

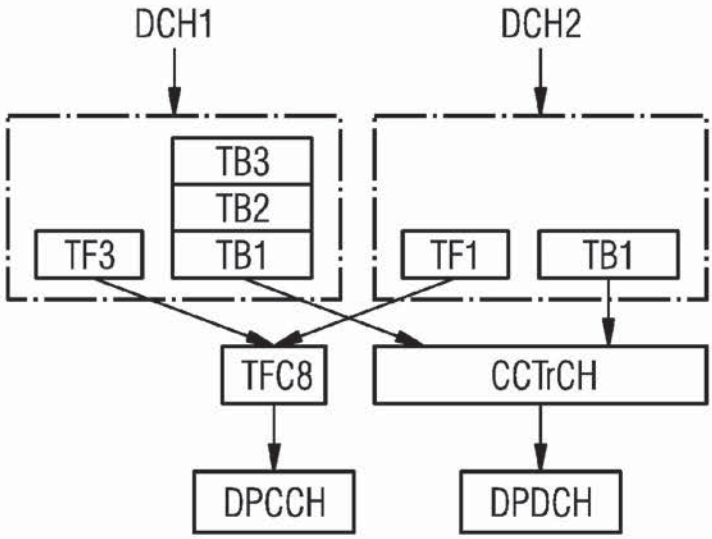

DCH1
DCH2
TB3
TB2
TB1
TF3
TF1
TB1
TFC8
CCTrCH
DPCCH
DPDCH

FIG 6

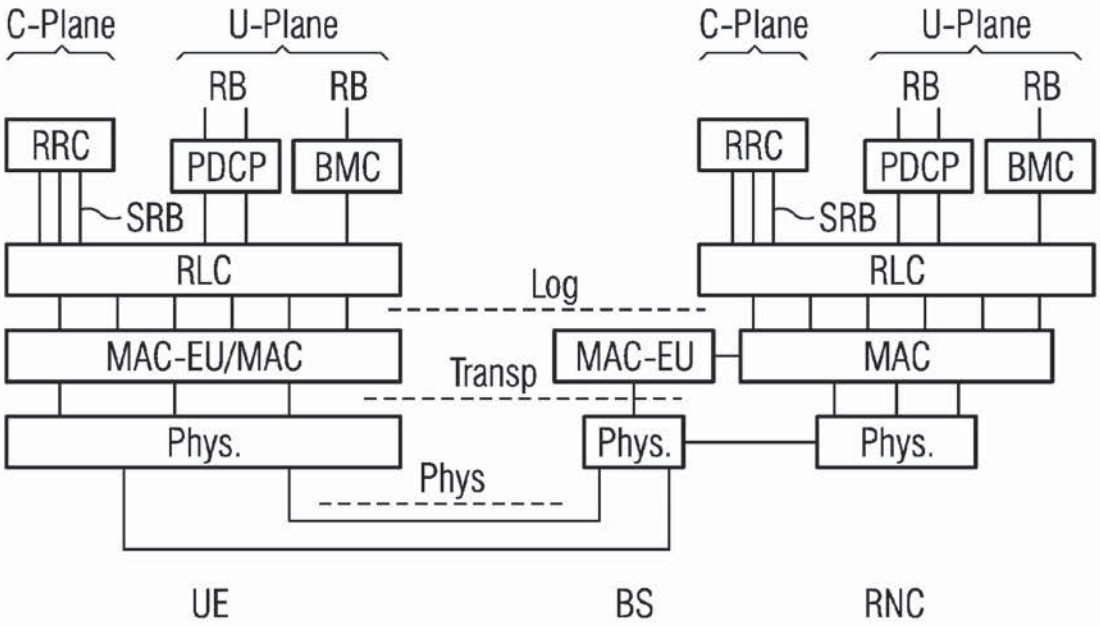

C-Plane
U-Plane
C-Plane
U-Plane
RB
RB
RB
RB
RRC
PDCP
BMC
RRC
PDCP
BMC
SRB
SRB
RLC
RLC
Log
MAC-EU/MAC
Transp
MAC-EU
MAC
Phys.
Phys.
Phys.
Phys
UE
BS
RNC

FIG 7A

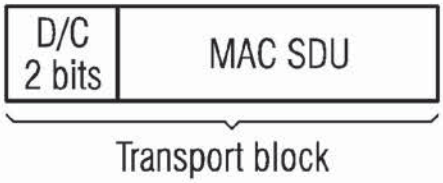


FIG 7B

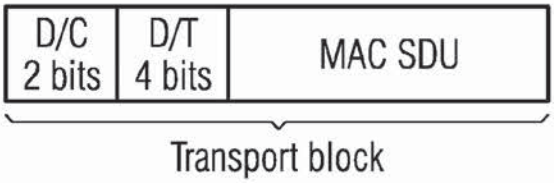


FIG 8



FIG 9

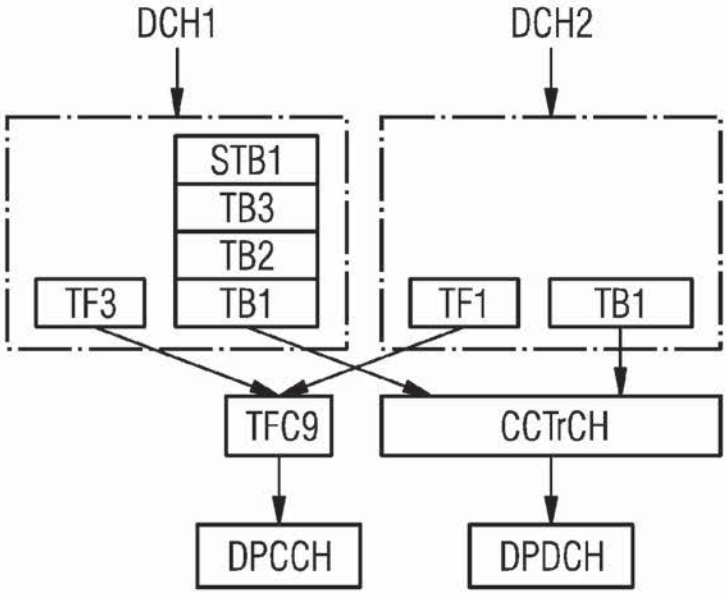


FIG 10

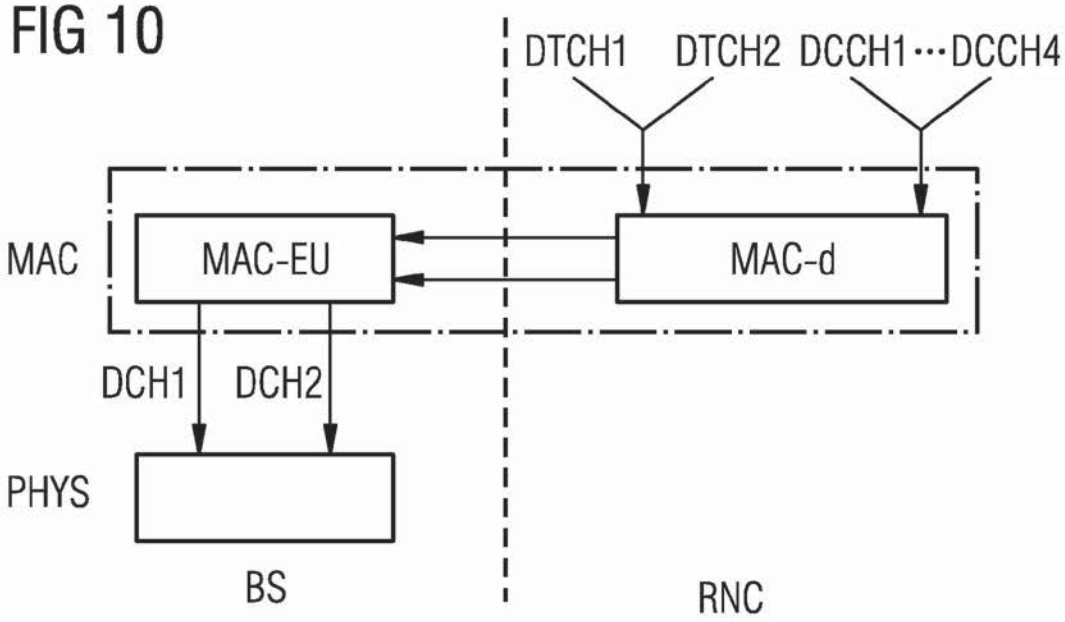

DTCH1
DTCH2
DCCH1 ··· DCCH4
MAC
MAC-EU
MAC-d
DCH1
DCH2
PHYS
BS
RNC

FIG 11

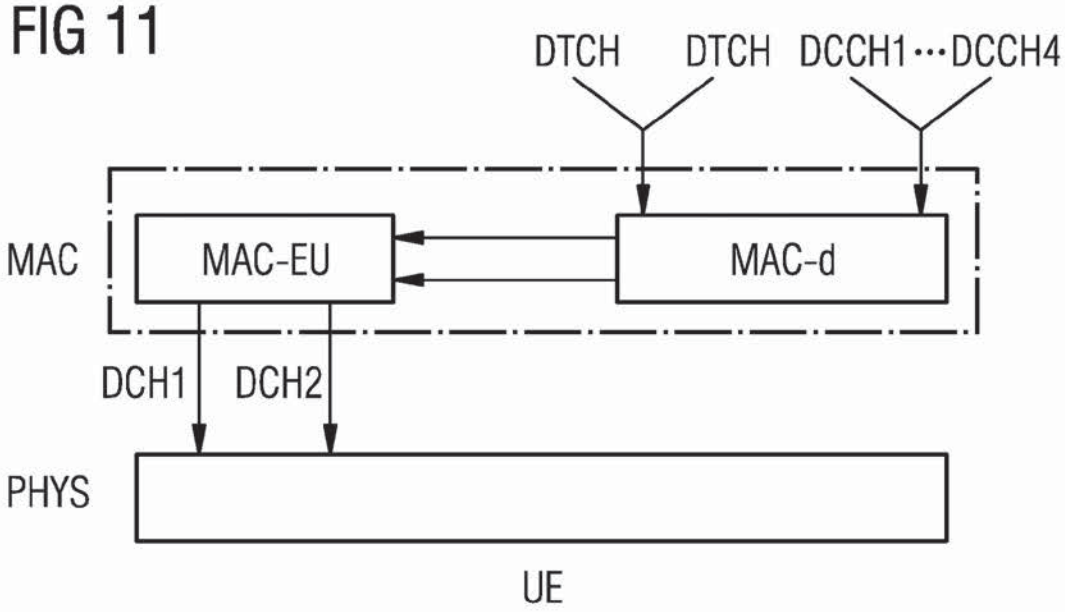

DTCH
DTCH
DCCH1 ··· DCCH4
MAC
MAC-EU
MAC-d
DCH1
DCH2
PHYS
UE

FIG 12

Table 1

RRC	Signaling Radio Bearer	SRB1	SRB2	SRB3	SRB4
RLC	Log. channel type	DCCH	DCCH	DCCH	DCCH
	RLC mode	UM	AM	AM	AM
	Packet size	136 bits	128 bits	128 bits	128 bits
	Priority of logical channel	1	2	3	4
MAC	Multiplexing	Mulitplexing of 4 log. channels			
Phys. layer	Transport channel type	DCH			
	Transport block size	148 bits			
	Transport format set (TFS)	TF0=0x148 bits TF1=1x148 bits			
	TTI	40 ms			
	Coding type	Rate 1/3 convolutional coding			
	CRC length	16 bits			

FIG 13

Table 2

RRC	Radio Bearer	RB1	RB2
RLC	Log. channel type	DTCH	DTCH
	RLC mode	AM	AM
	Packet size	320 bits	320 bits
	Priority of logical channel	1	1
MAC	Multiplexing	Mulitplexing of 2 log. channels	
Phys	Transport channel type	DCH	
	Transport block size	340 bits	
	Transport format set (TFS)	TF0=0x340 bits TF1=1x340 bits TF2=2x340 bits TF3=3x340 bits TF4=4x340 bits	
	TTI	20 ms	
	Coding type	Rate 1/3 turbo coding	
	CRC length	16 bits	

FIG 14

Table 3

TFCS SIZE	10
TFC0	(TF0, TF0)
TFC1	(TF1, TF0)
TFC2	(TF2, TF0)
TFC3	(TF3, TF0)
TFC4	(TF4, TF0)
TFC5	(TF0, TF1)
TFC6	(TF1, TF1)
TFC7	(TF2, TF1)
TFC8	(TF3, TF1)
TFC9	(TF4, TF1)

FIG 15

Table 4

RRC	Radio Bearer	RB1	RB2
RLC	Log. channel type	DTCH	DTCH
	RLC mode	AM	AM
	Packet size	160 bits	160 bits
	Priority of logical channel	1	1
MAC	Multiplexing	Mulitplexing of 2 log. channels	
Phys	Transport channel type	DCH	
	Transport block size	182 bits	
	Transport format set (TFS)	TF0=0x182 bits TF1=1x182 bits TF2=2x182 bits TF3=4x182 bits TF4=6x182 bits TF5=8x182 bits	
	TTI	20 ms	
	Coding type	Rate 1/3 turbo coding	
	CRC length	16 bits	

FIG 16

Table 5

TFCS SIZE	12
TFC0	(TF0, TF0)
TFC1	(TF1, TF0)
TFC2	(TF2, TF0)
TFC3	(TF3, TF0)
TFC4	(TF4, TF0)
TFC5	(TF5, TF0)
TFC6	(TF0, TF1)
TFC7	(TF1, TF1)
TFC8	(TF2, TF1)
TFC9	(TF3, TF1)
TFC10	(TF4, TF1)
TFC11	(TF5, TF1)

METHOD FOR THE TRANSMISSION OF DATA FIELD OF TECHNOLOGY

PRIORITY AND CROSS REFERENCE TO RELATED APPLICATION

This application is a continuation of U.S. Ser. No. 10/595,245, PCT filed on Aug. 30, 2004, which is a national stage entry under 35 U.S.C. §371 of International Application No. PCT/EP04/51958, filed Mar. 29, 2006, which claims priority to German application 103 45 220.6 filed Sep. 29, 2003, all of which are incorporated herein in their entirety by reference.

FIELD OF THE INVENTION

The present invention relates generally to transmitting data from a transmitter to a receiver in accordance with the Universal Mobile Telecommunications System (UMTS) standard.

BACKGROUND

The UMTS radio interface is subdivided into three protocol layers: the physical layer as Layer 1, the data link layer including MAC, RLC, BMC, PDCP as Layer 2 and the network layer with RRC as Layer 3. Within the protocol structure of the UMTS air interface, a radio resource control RRC in the radio network control entity RNC is responsible for the control and allocation of radio resources for all user equipment located in a radio cell. Resource management is currently performed on a relatively slow time basis because the corresponding signaling between the user equipment and the RNC is effected via the RRC messages.

An essential task of the MAC-d entity in the MAC layer in the transmitter is, in the transmit case, to map the data applied via the dedicated logical channels above the MAC layer to the dedicated transport channels of the physical layer. In the receiver, the task of the Mac-d is to distribute the data received on the dedicated transport channels to the dedicated logical channels. In the receive case, the MAC-d entity again distributes the data received via the dedicated transport channels to the relevant dedicated logical channels. On the transport channels, the data is transmitted in the form of fixed-length packet units, the so-called transport blocks. With regard to the further standardization and evolution of UMTS within the Third Generation Partnership Project or 3GPP bodies, improvements for fast and efficient data transmission via the dedicated transport channel are needed.

SUMMARY

An object of the present invention is to improve a data transmission method according to the UMTS standard so as to speed up data transmission. According to the invention, a method for inband signaling (at the MAC layer level) of information relevant to the UMTS base station is disclosed. The present invention implements fast and efficient signaling between the user terminal equipment UE and a particular UMTS base station at the MAC layer level. In the MAC layer level it is, therefore, possible according to the invention to differentiate between data transport blocks and signaling transport blocks and to handle them differently, thereby enabling user data and RRC signaling data to be exchanged between user terminal equipment and base station in the normal way. Signaling data additionally relevant to the base Station is only exchanged between user terminal equipment and base station. This speeds up data transmission particularly in the uplink direction, i.e., from user terminal equipment to a network or to the base station as part of a following network.

According to the invention, in an architecture of a corresponding communication system, RRC functionalities in the form of at least one control and/or data processing means are, therefore, moved from the RNC to the base station. Structural conformity with the existing a UMTS standard is also implemented. To this end, in particular, suitable signaling is created by providing appropriate signaling means in the base station and user terminal equipment. In addition, special signaling transport blocks and two different transport block formats are created. Additional features and advantages of the present invention are described in, and will be apparent from, the following Detailed Description, Figures and Tables that follow.

BRIEF DESCRIPTION OF THE FIGURES

FIG. **1** illustrates components of a radio communication system according to the UMTS standard.

FIG. **2** illustrates a protocol structure of the UMTS air interface.

FIG. **3** illustrates an architecture of the MAC-d entity on the UE side.

FIG. **4** illustrates an uplink transmission scenario.

FIG. **5** illustrates mapping of data from transport channels to the physical channels.

FIG. **6** illustrates an extended protocol structure of the UMTS air interface.

FIG. **7** illustrates transport block formats.

FIG. **8** illustrates an MAC SDU format in the signaling-transport block case.

FIG. **9** illustrates inband signaling.

FIG. **10** illustrates a signaling procedure in the MAC layer according to Example 1 in the downlink direction.

FIG. **11** illustrates a signaling procedure in the MAC layer according to Example 2 in the uplink direction.

FIG. **12** Table 1 lists parameters for the signaling radio bearers.

FIG. **13** Table 2 lists parameters for the radio bearers.

FIG. **14** Table 3 lists the configurations permitted for transport format combinations.

FIG. **15** Table 4 lists the parameters for the radio bearers according to an embodiment of the invention.

FIG. **16** Table 5 lists the configurations permitted for transport format combinations according to an embodiment of the invention.

DETAILED DESCRIPTION

In the Figures and further explanations, the same reference numerals and abbreviations are consistently used for identical parts, functional blocks, layers, and the like. Because of progressive standardization at international level, the technical terms and designations are predominantly borrowed from the Anglo-Saxon language area and should be understood by one of ordinary skill in the art.

A solution is proposed according to the invention in which only base-station-relevant information can be exchanged at MAC layer level between a base station and a user equipment via special signaling transport blocks. A base station according to the invention, therefore, likewise possesses resource management functions, e.g., for reconfiguring the physical channels or for switching the transport channel type. In this regard, new signaling mechanisms between a base station and a user equipment, enabling the base station to perform radio

resource management as quickly and efficiently as possible, will also be defined and described below. To facilitate orientation, the basic principles of the network and protocol architecture according to the UMTS standard will first be explained.

FIG. **1** illustrates an example radio communication system FCS operated in accordance with the Universal Mobile Telecommunications System or UMTS standard. In FIG. **1**, a radio cell CE**1**, a base station BS**1** and a higher-order radio network control entity RNC**1** are disclosed. The base station BS**1** is controlled via an associated data line L**1** by the higher-order radio network control entity RNC**1**, which monitors the assignment of radio resources in the radio cell CE**1** served by the base station BS**1**. The base station BS**1** is representative of a plurality of other base stations BS (not shown in detail in FIG. **1**) of the radio communication system FCS that possess and cover the corresponding radio cells CE. Between the base station BS**1** and a radio communication equipment (in this case, one of the mobile telephones UE**1**-UE**5** in the radio cell CE**1**) message and/or data signals are transmitted over at least one predefined air interface Uu according to a multiple access transmission method.

In UMTS frequency division duplex or FDD mode, for example, separate signal transmission in the uplink and downlink direction is achieved by means of corresponding separate assignment of frequencies or frequency ranges, uplink signifying signal transmission from the user equipment to the relevant base station, and downlink denoting signal transmission from the assigned base station to the user equipment. A plurality of users or more precisely mobile telephones UE**1**-UE**5** in the same radio cell CE**1** are preferably separated via orthogonal codes, in particular, according to the so-called code division multiple access or CDMA method. In this example, a number of user equipments UE**1**, UE**2**, UE**3**, UE**4** and UE**5** are present in the radio cell CE**1** of the base station BS**1**. The UMTS air interface Uu is subdivided into three protocol layers. FIG. **2** illustrates the protocol structure in terms of the dedicated transport channel DCH. The lowest layer is the physical layer, Layer 1. The next higher layer is the data link layer, Layer 2, comprising MAC, RLC, BMC and PDCP. The highest layer is the network layer with the RRC, Layer 3. This architecture is present both in the user equipment UE and in the UMTS network, also known as UMTS Terrestrial Radio Access Network or UTRAN, including the base stations BS and the radio network control entities RNC. Each protocol layer provides its services to the next higher layer via defined service access points.

To make the architecture easier to understand, these service access points are provided with commonly used and unique names such as logical channels, transport channels, radio bearers RB, and signaling radio bearers SRB. The protocol architecture shown in FIG. **2** is not only subdivided horizontally into the abovementioned layers and units, but also vertically into the control plane (C-plane), including the physical layer, MAC, RLC and RRC, and the user plane (U-plane), including the physical layer, MAC, RLC, PDCP and BMC. Only control data required for setting up and maintaining a connection is transmitted via the C-plane, whereas the actual user data is transported via the U-plane. Each protocol layer or protocol entity has particular functions. The protocol architecture is described in detail in [1]. On the network side, the physical layer is in the relevant base station and radio network control entity, whereas the MAC, RLC, PDCP, BMC and RRC are in the radio network control entity only. On the transmitter side, the task of the physical layer Phys is to ensure reliable transmission of the data coming from the MAC layer via the air interface Uu, the data being mapped to the relevant physical channels Phy. The physical layer Phys provides its services to the MAC layer via transport channels which specify how and with which characteristic the data is to be transported over the air interface Uu. The essential functions of the physical layer Phys include channel coding, modulation and CDMA code spreading. Correspondingly, on the receiver side, the physical layer Phys performs CDMA code de-spreading, demodulation and decoding of the received data and then forwards it to the MAC layer for further processing.

The MAC layer provides its services to the RLC layer via logical channels Log that characterize the data type of the transported data. The task of the MAC layer in the transmitter is to map the data present on a logical channel Log above the MAC layer to the transport channels Transp of the physical layer Phys. For this purpose, the physical layer Phys provides the transport channels with different transmission rates. One of the essential functions of the MAC layer in the transmitter is, therefore, to select a suitable transport format TF for each configured transport channel Transp depending on the instantaneous transmission rate, the transmitted power and the data priority of the logical channels Log mapped to that transport channel Transp. For example, a transport format TF specifies how many MAC packet units, known as a transport block, are sent to the physical layer Phys via the transport channel Transp per transmission time interval TTI. In the receiver, the MAC layer distributes the transport blocks received on the transport channels Transp to the logical channels Log. The MAC layer consists of three logical units.

The MAC-d entity handles the user and control data that are mapped to the dedicated transport channels DCH via the corresponding dedicated logical channels DTCH (Dedicated Traffic Channel) and DCCH (Dedicated Control Channel). The MAC-c/sh (MAC-control/shared) entity handles the user and control data of logical channels that are mapped to the common transport channels such as RACH in the uplink or FACH in the downlink. The MAC-b (MAC-broadcast) entity handles only the radio-cell-relevant system information which is broadcast to all the UEs in the relevant radio cell by mapping the logical broadcast control channel BCCH to the transport broadcast channel BCH. The RLC layer provides its services, in the case of RRC, via the signaling radio bearers SRB. In the case of PDCP and BMC, this takes place via the radio bearers RB. The SRB or RB characterize how the RLC layer has to deal with the data packets.

For this purpose, for example, the transmission mode for each configured SRB or RB is defined by the RRC layer: transparent mode TM, unacknowledged mode UM or acknowledged mode AM, the RLC layer being modeled in such a way that there is one independent RLC entity per RB or SRB. In addition, the task of the RLC protocol in the transmitter is to divide or combine the user and signaling data of RBs or SRBs into packets. In the case of the UM and AM transmission modes, the relevant RLC entity stores the copies of the data packets present at an RB or SRB in an RLC transmit buffer until they can be successfully transported over the air interface Uu by the layers below RLC. The RLC layer transfers the data packets resulting from the dividing or combining to the MAC layer for further transportation. The RRC protocol is responsible for setup and cleardown, the reconfiguration of physical channels Phy, transport channels Transp, logical channels Log, signaling radio bearers and radio bearers, as well as negotiation of all the parameters of the Layer 1 and 2 protocols. For this purpose, the RRC entities in the RNC and UE exchange corresponding RRC messages via the SRBs. For RRC layer details, see [2]. The PDCP protocol is only responsible for the transmission of packet-

switched (PS) domain data. Its main function is to compress or decompress EP header information. The BMC protocol is used on the network side to transmit so-called cell broadcast messages over the air interface Uu.

The functional basics of the MAC-d entity will now be described. The MAC-d entity in the MAC layer handles the user and control data which are mapped via the corresponding dedicated logical channels DTCH (Dedicated Traffic Channel) and DCCH (Dedicated Control Channel) to the dedicated transport channels DCH. The relevant details are described in [3]. By way of example, FIG. **3** shows the architecture of the MAC-d entity on the UE side:

- If configured by RRC, the user and control data are mapped by block transport channel type switching from DTCH and DCCH to common transport channels such as RACH, for example, and forwarded to the MAC-c/sh entity for further processing.
- C/T MUX is used if multiplexing of a plurality of dedicated logical channels to the same transport channel is performed. In this case, in order to ensure unambiguous identification, the data packets from the relevant logical channels have a 4-bit C/T field added as the MAC header in which the identity of the logical channel is entered. This enables the MAC-d entity on the receiver side to clearly identify the logical channel from which the received data originates.
- In the case of RLC transparent mode, TM, the data packets are encrypted (ciphering) in the transmit case or decrypted (deciphering) in the receive case.
- The task of the UL TFC selection block is uplink scheduling, i.e. selecting a suitable transport format combination TFC for all the configured DCHs depending on the instantaneous transmission rate, the transmitted power and the data priority of the dedicated logical channels which are mapped to the transport channels.

To facilitate understanding of how the protocols relate to one another, an example will now be explained. For this purpose, a scenario is assumed in which the user equipment UE**1** in the radio cell CE**1** is using two uplink packet services of 64 kbps each in parallel, e.g., for Internet browsing and streaming of data. The UE**1** has been allocated dedicated radio resources by the RRC layer in the RNC**1** on the basis of the current traffic situation in the radio cell CE**1** and the requested quality of service QoS. In detail, the individual protocol layers or protocol entities have been configured by the RRC layer in the RNC**1** for the downlink and uplink in such a way that a particular QoS, such as a certain guaranteed or maximum data rate and/or a defined transmission delay shall be ensured by the Layer 1 and 2 protocols for the duration of the mobile connection. The configuration specified by the RNC**1** is then signaled to the RRC layer in the user equipment UE**1**.

FIG. **4** illustrates a typical configuration for the uplink transmission scenario illustrated. In the U-plane, two RBs are specified, i.e., RB**1** and RB**2**, via which the user data of the relevant packet service is transmitted. In the RLC layer, each RB is mapped to an RLC entity and logical traffic channel DTCH. In the C-plane, four 3.4-kbps SRBs SRB**1** to SRB**4** specified on the basis of the different types of control messages are mapped to an RLC entity and logical control channel DCCH in the RLC layer. In the MAC-d entity, two transport channels DCH**1** and DCH**2** are configured. The two logical traffic channels DTCH**1** and DTCH**2** are multiplexed onto the transport channel DCH**1** in the U-plane and the four logical control channels DCCH**1** to DCCH**4** are multiplexed onto the transport channel DCH**2** in the C-plane. In the physical layer, the two transport channels are channel-coded and multiplexed onto a coded composite transport channel CCTrCH of 10 ms frame length. Based on FDD radio transmission technology, the data is transmitted over the air interface Uu to UTRAN on the CCTrCH after spreading and modulation via the dedicated physical data channel DPDCH with SF=16. Specific physical layer control information is transmitted on the dedicated physical control Channel DPCCH with spreading factor SF=256 to enable the physical layer in the base station BS**1** to also correctly decode the data on the DPDCH after decoding the control information on the DPCCH.

FIGS. **12** to **14**, Tables 1 to 3 summarize the configured parameters for the signaling radio bearers, radio bearers and the permitted transport format combinations. For the processing of the data packets in the transmit buffers of the relevant RLC-entities, the logical channels are assigned different priorities from 1 to 8. Priority 1 constitutes the highest and priority 8 constitutes the lowest priority. On the basis of these priorities, the data packets are preferred by the logical channels having a higher priority. In the event of a stalemate situation, i.e., both or a plurality of logical channels on the same transport channel have the same priority, the buffer occupancy BO is taken into account as a further criterion. If, in the case of equal priority of, e.g., two logical channels on the same transport channel, the buffer status of logical channel **1** is higher than that of logical channel **2**, the data is first processed by channel **1**.

For the transport channel DCH**1**, five transport formats TF**0** to TF**4** in the transport format set TFS are configured. For example, the transport format TF**2** specifies that, in each transmission time interval TTI of 20 ms, two transport blocks TB of size 340 bits are transmitted via the DCH**1** to the physical layer where 16 CRC checksum bits are added to each transport block for error detection. The two transport blocks are then jointly channel-coded by a rate 1/3 turbo coder in order to protect them from transmission errors that may be caused by the transmission channel. For the transport channel DCH**2**, on the other hand, only two transport formats TF**0** and TF**1** in the transport format set TFS are configured. Thus, the transport format TF**1** specifies that, in each transmission time interval TTI=40 ms, one transport block of size 148 bits is transmitted via the DCH**2** to the physical layer where 16 CRC checksum bits are added to the transport block for error detection. The transport block is then channel-coded by a rate 1/3 convolutional coder.

The coded data of the two transport channels are then jointly multiplexed onto a CCTrCH frame depending on their respective TTIs. On the basis of TTI=20 ms, the data from DCH**1** is transmitted over the air interface to UTRAN in two consecutive frames, whereas the data from DCH**2**, on the basis of TTI=40 ms, is transmitted in four consecutive frames. The permissible combination of transport formats of the two transport channels DCH**1** and DCH**2** on the CCTrCH is specified by the transport format combination set TFCS. The maximum number of possible transport format combinations TFC is the product of the number of transport formats configured for each transport channel. It lies within the responsibility and control of UTRAN to correctly specify the size of the TFCS, i.e., the number and type of the permitted combinations of transport formats of different transport channels. In practice, the permitted number of TFCs in a TFCS is less than the theoretically possible maximum. In this embodiment, the permitted size of the TFCS=10 is also the actual maximum number, also 5 TFs from DCH**1** and 2 TFs from DCH**2**. These 10 permitted transport format combinations are listed in FIG. **14** Table 3. The notation of the TFCs is defined with i=0 . . . 4 and j=0.1 for TF#i from DCH**1**, TF#j from DCH**2**.

FIG. **5** illustrates an example of uplink scheduling in which the MAC-d entity has selected the transport format combination TFC**8** for data transmission as a function of the instantaneous transmission situation, the combination TFC**8**=(TF**3**, TF**1**) specifying that the relevant portions of the coded data of three transport blocks TB**1**, TB**2**, TB**3** from DCH**1** (=TF**3**) and of one transport block (TB**1**) from DCH**2** (=TF**1**) are transmitted on the CCTrCH. To ensure that the physical layer in the base station BS**1** can correctly decode the data on the DPDCH, the transport format combination TFC**8** used on the CCTrCH is signaled on the DPCCH as control information. An essential aspect of the present invention is the definition of special signaling transport blocks STB for inband signaling of base-station-relevant information at the MAC layer level. This allows fast and efficient control of radio resources. Without limiting generality, it will now be assumed that the base station BS possesses the following RRC functions:

- Reconfiguration of physical channels in the uplink and downlink
- Reconfiguration of the transport formats and transport format combinations in the uplink and downlink
- Switching of the transport channel type, i.e. from common transport channels to dedicated transport channels and vice versa
- Setting of the uplink SIR_{target} for fast performance control of dedicated physical channels

Specifically, a solution according to the invention looks as follows:

1. Extended UTRAN Protocol Architecture:

Within the UTRAN protocol architecture, a new entity with the designation Medium Access Control Enhanced Uplink, abbreviated to MAC-EU, is defined in the MAClayer. A correspondingly extended UTRAN protocol architecture is shown in FIG. **6** analogously to the depiction in FIG. **2**. On the network side the MAC-EU entity is only in the base station, the MAC-EU performing all the functions required for inband signaling of base-station-relevant information concerning the signaling transport blocks for radio resource control.

These functions include:

- Generating one or more signaling transport blocks for inband signaling;
- Selecting a transport channel for transmitting the signaling transport blocks;
- Multiplexing signaling transport blocks within the transport blocks of a transport channel that are to be transmitted;
- Demultiplexing signaling transport blocks within the received transport blocks of a transport channel;
- Forwarding the information transmitted in the signaling transport blocks to the RRC entity in the base station or UE for further processing and
- Checking for reliable transmission or reception of messages in signaling transport blocks.

Depending on the RRC functionality of the base station BS, various types of messages are exchanged between the base station BS and a user equipment UE via an STB. This new transport block typically contains, according to the following non-exhaustive list:

- Physical Channel Reconfiguration Control: message from the base station to the UE to reconfigure the physical channels in the uplink and downlink.
- TF Reconfiguration Control: message from the base station to the UE to reconfigure the transport format and transport format combinations in the uplink and downlink.
- Buffer Status Control: message from the base station to the UE to transmit the current data volume of a particular transport channel, i.e. the RLC buffer status of all the RBs or logical channels which are multiplexed into the transport channel.
- Buffer Status Report: reply from the UE to the base station in response to the buffer status control message with signaling of the data volume of the transport channel.

2. Definition of Transport Block Types:

Two new transport block formats are defined for the dedicated logical channels DTCH and DCCH depending on the MAC-multiplexing, see FIG. **7**. Without limiting generality, a DCH transport channel is considered in FIG. **7**, i.e., in principle the new formats can also be used for the common transport channels such as RACH in the uplink and FACH in the downlink:

- Case a): DTCH or DCCH are mapped to a DCH transport channel without multiplexing. In this case only a 2-bit D/C field is added as a MAC header.
- Case b): DTCH or DCCH are mapped to a DCH transport channel with multiplexing. In this case the MAC header consists of the 2-bit D/C field and the 4-bit C/T field in which the relevant identity of the logical channel is transmitted.

The field is designated D/C as an abbreviation for Data/Control. The D/C field indicates the transport block type:

- With D/C=00 a signaling transport block STB is signaled. The MAC SDU then constitutes the packet unit via which only base-station-relevant information for radio resource control is exchanged between user equipment and base station.
- Correspondingly, D/C=11 signals a normal transport block by means of which user or RRC signaling data is transmitted, as hitherto. The MAC SDU then constitutes the packet unit which receives the MAC layer via DTCH or DCCH.

3. Structure of the Signaling Transport Block:

FIG. **8** shows the general structure of the MAC SDU part of a signaling transport block STB via which n messages can be transmitted:

- TN UL: this status field or field transmits an uplink transmission number and is used for tracking the transmission status in the uplink. The field is k bits long.
- TN DL: this field transmits a downlink transmission number and is used for tracking the transmission status in the downlink. The field is k bits long.
- Poll: this field is used to request from the receiver an acknowledgment of successful transmission of a signaling transport block within a specified time. The field is 1 bit long.
- MT: in this field the message type transmitted in the following message part is specified. The field is 1 bit coded.
- MP: in this field the message specified by the MT part is transmitted. The field has a variable length of m bits depending on the type of message to be transmitted.
- Flag: this field is used to indicate whether or not the field MT, i.e., another message, is transmitted in the following field. The field is 1 bit long.
- Pad: this field is used to pad out the unused part in the MAC SDU with so-called dummy bits.

The status fields TN UL, TN DL and Poll are used to check for reliable transmission of messages in a signaling transport block. This is implemented by the following mechanism:

- The MAC-EU entity in the UE has an uplink transmission counter Z**1** with an integer value range of 0 to N−1, coded with k bits. For each STB transmitted in the uplink direction, the current value of this uplink counter is transmitted in the field TN UL and then incremented by one. The value of the last DL-STB received is addition-

ally transmitted in the field TN DL. When the maximum count has been reached, Z1 is reset to 0 and incremented once again.

Similarly, the MAC-EU entity in the base station has a downlink transmission counter Z2 with an integer value range of 0 to N−1, coded with k bits. For each STB transmitted in the downlink direction, the current value of this downlink counter is transmitted in the field TN DL and then incremented by one. The value of the last UL-STB received is additionally transmitted in the field TN UL. When the maximum count has been reached, Z2 is reset to 0 and incremented once again.

Using the status field Poll, the relevant MAC-EU entities can if necessary request acknowledgment of successful reception of the STB within a specified time from the relevant receiver entities, i.e., by a set poll bit=1.

In the case of error-free transmission conditions, the relevant MAC-EU entities in the receiver receive a sequential numerical order of signaling transport blocks, i.e., any transmission errors are detected by breaks in the sequential numerical order.

To summarize, according to the invention there are defined, within the MAC_protocol layer, special signaling transport blocks via which in future fast and efficient inband signaling for radio resource control between a base station and a user equipment can be implemented, thereby providing the following advantages. The invention supports an extended UTRAN protocol architecture with RRC functionality in the base station, so that radio resources can in future be managed closer to the air interface. In this way, reconfigurations of radio resources in the uplink and downlink can be carried out much more quickly and efficiently for a user equipment as a function of the traffic load in a radio cell. Data transmission in the downlink and particularly in the uplink can be significantly improved in terms of transmission delay and data throughput.

In the examples which follow, the extended UTRAN protocol architecture shown in FIG. **6** with the new MAC-EU entity in the MAC layer will be considered. Without limiting generality, it will be assumed that the base station possesses the following RRC functions, as already stated above:

Reconfiguration of physical channels in the uplink and downlink

Reconfiguration of the transport formats and transport format combinations in the uplink and downlink

Switching of the transport channel type, i.e., from common transport channels to dedicated transport channels and vice versa

Setting of the uplink SIR_{target} for fast performance control of dedicated physical channels

Data transmission between a UE and UTRAN via a dedicated connection with the following configuration is considered:

For uplink and downlink a transmission scenario according to FIG. **4** is considered.

In the U-plane the user data is transmitted on two RBs, i.e., RB#**1** and RB#**2**. The configuration of the two RBs is summarized in FIG. **15**_Table 4.

In the C-plane, 4 SRBs (SRB#**1** to SRB#**4**) are configured. Their parameters are summarized in FIG. **12**_Table 1.

FIG. **16** Table 5 lists the permitted transport format combinations, a total of 12 combinations now being defined.

A transport block format according to case b) in FIG. **7** is considered, i.e., the MAC header consists of the fields D/C and C/T.

In respect of the format of a signaling transport block according to FIG. **8**, the following configuration is considered: fields TN UL, TN DL and MT 3 bits long in each case.

Example 1 of an embodiment: inband signaling in the downlink. The downlink transmission counter Z**2** is in the initial state with the value 0 and MAC-EU in the base station has yet to receive a UL-STB from the UE. Because of the current traffic Situation in the radio cell, the base station wants to send two radio resource control messages to the UE via a DL-STB within the existing dedicated data transmission between UE and UTRAN:

Physical Channel Reconfiguration Control for reconfiguring the dedicated physical channels in the uplink and downlink, e.g., new parameters for SF, channelization code and scrambling code.

Buffer Status Control for transmitting the current data volume of the dedicated UL transport channel DCH**1**.

Based on the downlink scheduling, the MAC-d entity in the RNC has selected TFC**9** for data transmission on the CCTrCH, i.e., the relevant portions of the coded data of four transport blocks (TB**1**, TB**2**, TB**3**, TB**4**) from DCH**1** (=TF**3**) and of one transport block (TB**1**) from DCH**2** (=TF**1**) are to be transmitted in the physical layer every 10 ms.

Because of the available transmission capacity on DCH**1**, the MAC-EU in the base station BS selects this transport channel for transmitting its 182-bit signaling transport block. MAC-EU signals the requirement to the MAC-d entity so that only three normal transport blocks are actually transmitted over the DCH**1**. The MAC-EU now generates an STB with the following configuration:

D/C=00

C/T=dummy bits, as this field has no significance in the case of an STB

TN UL=0

TN DL=0

Poll=0

MT=Physical Channel Reconfiguration Control

MP**1**=content of the Physical Channel Reconfiguration Control message

Flag-1=1, in Order to indicate that another message follows

MT=Buffer Status Control

MP**2**=content of the Buffer Status Control message

Flag-2=0, in order to indicate that another message follows

Pad=dummy bits, if required

This STB is then multiplexed by the MAC-EU within the normal transport blocks of DCH**1** to be transmitted, as shown in FIG. **9**, and passed on to the physical layer for further processing. The basic signal flow is shown in FIG. **10**, the dash-dotted line clearly indicating the physical separation of the MAC layer as a logical entity. In order that the physical layer in the UE can correctly decode the data on the DPDCH, the transport format combination TFC**9** used on the CCTrCH is signaled on the DPCCH as control information. In the MAC-EU entity in the UE, the received transport blocks on the DCH**1** are evaluated on the basis of the D/C field in the MAC header, and, if D/C=00, the DL-STB**1** is demultiplexed accordingly. The three other transport blocks TB **1**, TB**2** and TB**3** are passed on to MAC-d entity for further processing.

Example 2 of an embodiment: inband signaling in the uplink. The uplink transmission counter Z**1** is in the initial state with the value 0 and the MAC-EU in the UE has received the DL-STB from the base station. On the basis of the received messages, on the one hand the physical channels in the uplink and downlink are reconfigured while, on the other, measurement of the data volume on the UL-DCH**1** is per-

formed. As a response thereto, the Buffer Status Report is now to be sent to the base station via a UL-STB. Based on the uplink scheduling, the MAC-d entity in the UE has selected TFC**9** for data transmission on the CCTrCH, i.e., the relevant portions of the coded data of four transport blocks (TB**1**, TB**2**, TB**3**, TB**4**) from DCH**1** (=TF**3**) and of one transport block (TB**1**) from DCH**2** (=TF**1**) are to be transmitted in the physical layer every 10 ms.

Because of the available transmission capacity on DCH**1**, the MAC-EU selects this transport channel for transmitting its 182-bit signaling transport block. MAC-EU signals the requirement to the MAC-d entity so that only three normal transport blocks are actually transmitted over the DCH**1**. The MAC-EU now generates an STB with the following configuration:

D/C=00

C/T=dummy bits, as this field has no significance in the case of an STB

TN UL=0

TN DL=0

Poll=0

MT=Buffer Status Report

MP **1**=content of the Buffer Status Report message

Flag-1=0, in order to indicate that no other message follows

Pad=dummy bits, if required

This STB is then multiplexed by the MAC-EU within the normal transport blocks of DCH**1** to be transmitted, as shown in FIG. **9**, and passed on to the physical layer for further processing. The basic signal flow in the uplink direction is shown in FIG. **11** where, in contrast to the situation in FIG. **10**, there is no physical Separation of the MAC layer. In order that the physical layer in the base station can correctly decode the data on the DPDCH, the transport format combination TFC**9** used on the CCTrCH is signaled on the DPCCH as control information. In the MAC-EU entity in the base station, the received transport blocks on the DCH**1** are evaluated on the basis of the D/C field in the MAC header, and, if D/C=00, the UL-STB**1** is demultiplexed accordingly. The three other transport blocks TB**1**, TB**2** and TB**3** are passed on to MAC-d entity in the RNC for further processing.

Within the scope of the description of the present invention, reference is made particularly to the following literature:

3GPP TS 25.301: Radio Interface Protocol Architecture

3GPP TS 25.331: Radio Resource Control (RRC) protocol specification

3GPP TS 25.321: Medium Access Control (MAC) protocol specification

In addition, the following abbreviations are used:

3GPP Third Generation Partnership Project
AM Acknowledged Mode
BCCH Broadcast Control Channel
BCH Broadcast Channel
BMC Broadcast Multicast Control
BO Buffer Occupancy
BS Base Station
CCTrCH Coded Composite Transport Channel
CDMA Code Division Multiple Access
CE Radio cell
CRC Cyclic Redundancy Check
D/C Data/Control
DCCH Dedicated Control Channel
DCH Dedicated Channel
DL Downlink
DPCCH Dedicated Physical Control Channel
DPDCH Dedicated Physical Data Channel
DTCH Dedicated Traffic Channel
FACH Forward Access Channel
FCS Radio communication system
FDD Frequency Division Duplex
IP Internet Protocol
kbps Kilobits per second
Log Logical Channel
MAC Medium Access Control
MAC-b MAC broadcast
MAC-c/sh MAC control/shared
MAC-d MAC dedicated
MAC-EU MAC Enhanced Uplink
MP Message Part
MT Message Type
PDCP Packet Data Convergence Protocol
PS Packet-switched
Phy Physical Channel
Phys Physical layer
QOS Quality of Service
RACH Random Access Channel
RB Radio Bearer
RLC Radio Link Control
RNC Radio Network Controller, radio network control entity
RRC Radio Resource Control
SDU Service Data Unit
SF Spreading Factor
SIR Signal to Interference Ratio
SRB Signaling Radio Bearer
STB Signaling Transport Block
TB Transport Block
TF Transport Format
TFC Transport Format Combination
TFCS Transport Format Combination Set
TFS Transport Format Set
TM Transparent Mode
TN Transmission Number
Transp Transport Channel
TTI Transmission Time Interval
UE User Equipment
UL Uplink
UM Unacknowledged Mode
UMTS Universal Mobile Telecommunications System
UTRAN UMTS Terrestrial Radio Access Network

It should be understood that various changes and modifications to the presently preferred embodiments described herein will be apparent to those skilled in the art. Such changes and modifications can be made without departing from the spirit and scope of the present invention and without diminishing its intended advantages. It is therefore intended that such changes and modifications be covered by the appended claims.

The invention claimed is:

1. A method for transmitting data from user terminal equipment to a base station, the base station configured to process at least one resource management function, comprising:

performing inband signaling of information relevant to the base station at the MAC layer level between a user terminal equipment and the base station by:

multiplexing the inband signaling information with data of a transport channel that is to be transmitted, wherein the inband signaling information comprises Buffer Status Report information signaling volume of the data of the transport channel from the user terminal equipment to the base station based on at least one radio resource control message received from the base station.

2. The method according to claim **1**, wherein the transport channel is a dedicated or common transport channel.

3. The method according to claim 1, further comprising providing an uplink transmission number in the inband signaling information for tracking the transmission status in uplink.

4. The method according to claim 1, further comprising providing a downlink transmission number in the inband signaling information for tracking the transmission status in downlink.

5. The method according to claim 1, further comprising providing a message type in the inband signaling information indicating a type of message to be transmitted.

6. The method according to claim 5, further comprising providing a message part in the inband signaling information having a variable length based on the message type to be transmitted.

7. The method according to claim 1, further comprising providing a Flag in the inband signaling information indicating whether another message is transmitted in the subsequent field.

8. The method according to claim 1, further comprising: providing a Pad in the inband signaling information for padding out unused part in MAC Service Data Unit (MAC SDU) with dummy bits.

9. The method according to claim 1, further comprising exchanging, in the inband signaling information, various radio resource control messages between the base Station (BS) and the user terminal equipment (UE).

10. The method according to claim 1, further comprising transmitting the inband signaling information using CDMA via an air interface.

11. The method according to claim 1, further comprising transmitting the inband signaling information in FDD mode.

12. The method according to claim 1, further comprising transferring the inband signaling information to the physical layer.

13. The method according to claim 1, wherein performing inband signaling of information comprises:

reconfiguring physical channels in uplink and downlink;

reconfiguring transport formats and transport format combinations in uplink and downlink;

switching transport channel type between common transport channels and dedicated transport channels; or

setting uplink SIR_{target} for fast performance control of dedicated physical channels.

14. The method according to claim 1, further comprising selecting a transport channel for transmitting the inband signaling information.

15. The method according to claim 1, further comprising increasing a transmission counter after transmitting the inband signaling information.

16. A user terminal device configured to transmit data to a base station, wherein the base station is configured to process at least one resource management function, the user terminal device configured to:

perform inband signaling of information relevant to the base station at the MAC layer level between a user terminal device and the base station by:

multiplexing the inband signaling information with data of a transport channel that is to be transmitted, wherein the inband signaling information comprises Buffer Status Report information signaling volume of the data of the transport channel from the user terminal device to the base station based on at least one radio resource control message received from the base station.

17. The user terminal device according to claim 16, further configured to transmit the inband signaling information using CDMA via an air interface.

18. The user terminal device according to claim 16, further configured to transmit the inband signaling information in FDD mode.

19. The user terminal device according to claim 16, further configured to transfer the inband signaling information to the physical layer.

20. A base station configured to receive data from a user terminal device, wherein the base station is configured to process at least one resource management function, the base station configured to:

receive inband signaling information relevant to the base station at the MAC layer level from a user terminal device by:

de-multiplexing the inband signaling information from data of a transport channel that is transmitted by the user terminal device, wherein the inband signaling information comprises Buffer Status Report information signaling volume of the data of the transport channel from the user terminal device to the base station based on at least one radio resource control message transmitted by the base station.

* * * * *

EXHIBIT 13

US006111865A

United States Patent [19]

Butler et al.

[11] **Patent Number: 6,111,865**

[45] **Date of Patent: Aug. 29, 2000**

[54] **DUAL CHANNEL SLOTTED PAGING**

[75] Inventors: **Brian K. Butler**, San Diego, Calif.; **Klein S. Gilhousen**, Bozeman, Mont.

[73] Assignee: **Qualcomm Incorporated**, San Diego, Calif.

[21] Appl. No.: **08/865,650**

[22] Filed: **May 30, 1997**

[51] **Int. Cl.**[7] **H04J 13/00**

[52] **U.S. Cl.** **370/335**; 370/365; 455/31.3

[58] **Field of Search** 435/31.2, 31.3, 435/38.1, 38.3, 575, 574, 343, 458; 370/311, 312, 320, 321, 328, 335, 336, 342, 365, 368; 340/825.44, 825.53

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,845,491	7/1989	Fascenda et al.	340/825.44
5,301,225	4/1994	Susuki et al.	379/59
5,377,183	12/1994	Dent	370/335
5,382,949	1/1995	Mock et al.	340/825.44
5,392,287	2/1995	Tiedemann et al.	370/311
5,392,452	2/1995	Davis	455/38.1
5,398,021	3/1995	Moore	340/825.27
5,450,071	9/1995	DeLuca et al.	340/825.44
5,481,254	1/1996	Gaskill et al.	340/825.52
5,485,463	1/1996	Godoroja	340/825.44
5,491,718	2/1996	Gould et al.	370/335
5,511,067	4/1996	Miller	370/335
5,511,110	4/1996	Drucker	455/458
5,537,097	7/1996	Ericksson et al.	340/825.52
5,544,223	8/1996	Robbins et al.	370/335
5,629,940	5/1997	Gaskill	370/311
5,659,573	8/1997	Bruckert et al.	375/200
5,710,784	1/1998	Kindred et al.	375/262

FOREIGN PATENT DOCUMENTS

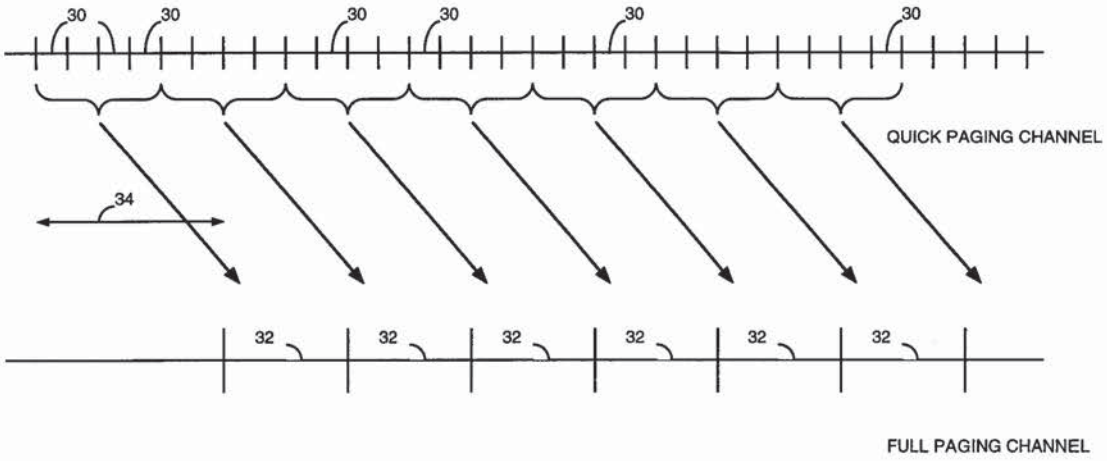
0427126	11/1990	European Pat. Off.	H04Q 7/04
2244409	10/1989	United Kingdom	H04Q 7/04
9010987	9/1990	WIPO .	
9318596	9/1993	WIPO	H04B 3/16
9409599	4/1994	WIPO	H04Q 7/04

Primary Examiner—Chau Nguyen
Assistant Examiner—Kenneth Vanderpuye
Attorney, Agent, or Firm—Russell B. Miller; Charles D. Brown; Brian S. Edmonston

[57] **ABSTRACT**

A novel and improved method and system for paging a cellular telephone or other wireless terminal using two paging channels that reduces standby mode power consumption described. In one embodiment of the invention, a minimally encoded quick paging channel is established over which short, quick page messages are transmitted during one of a set of quick paging slots. The quick page message indicates that a communications request has been received, and that the receiving communication terminals should process a highly encoded full paging channel over which more detailed, full page messages are transmitted during the next full paging slot. A terminal monitors full paging channel only after a quick page message has been received on the quick paging channel.

17 Claims, 6 Drawing Sheets



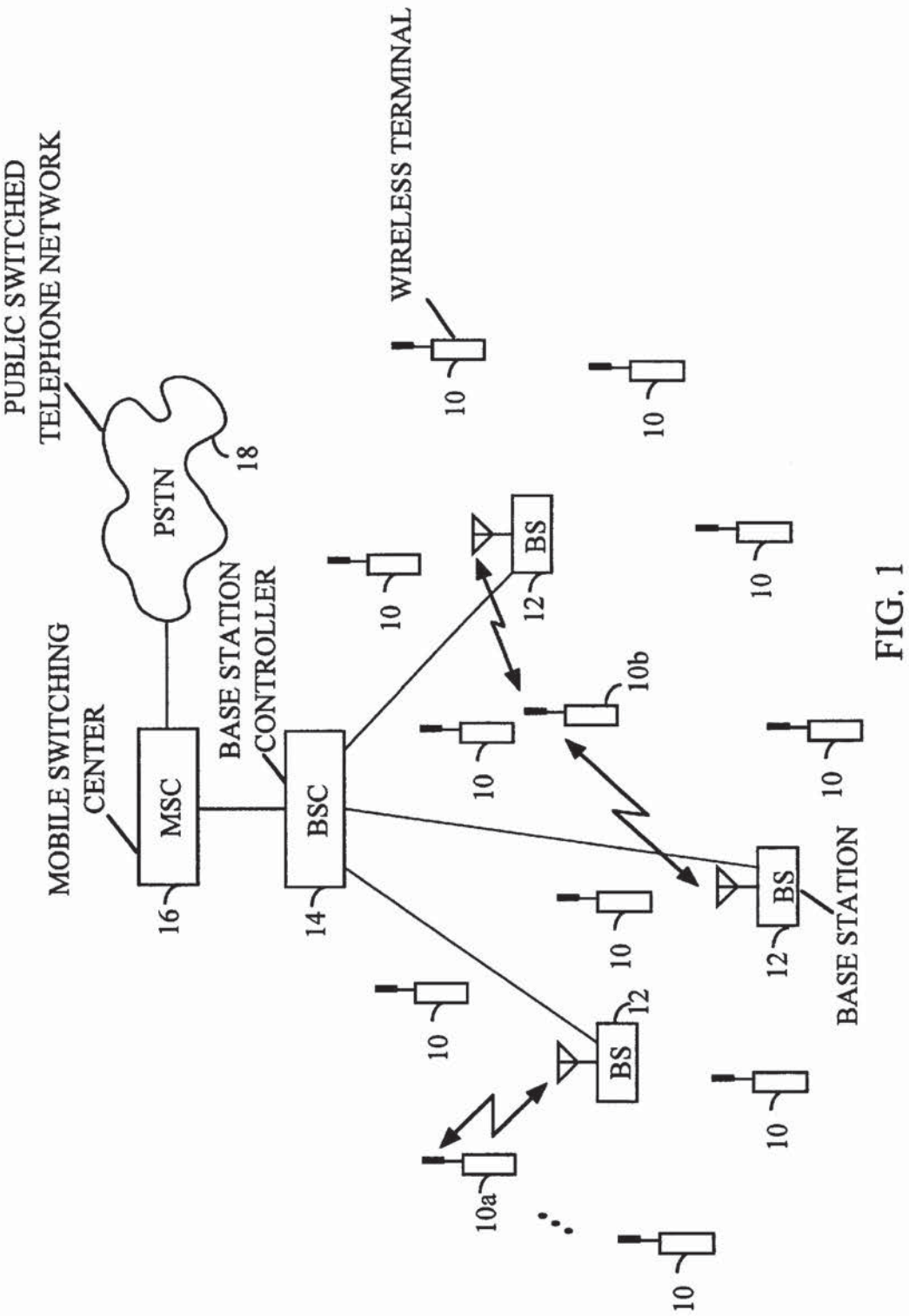

PUBLIC SWITCHED TELEPHONE NETWORK
PSTN
18
WIRELESS TERMINAL
10
MOBILE SWITCHING CENTER
MSC
16
BASE STATION CONTROLLER
BSC
14
BS
12
BASE STATION
10b
10a

FIG. 1

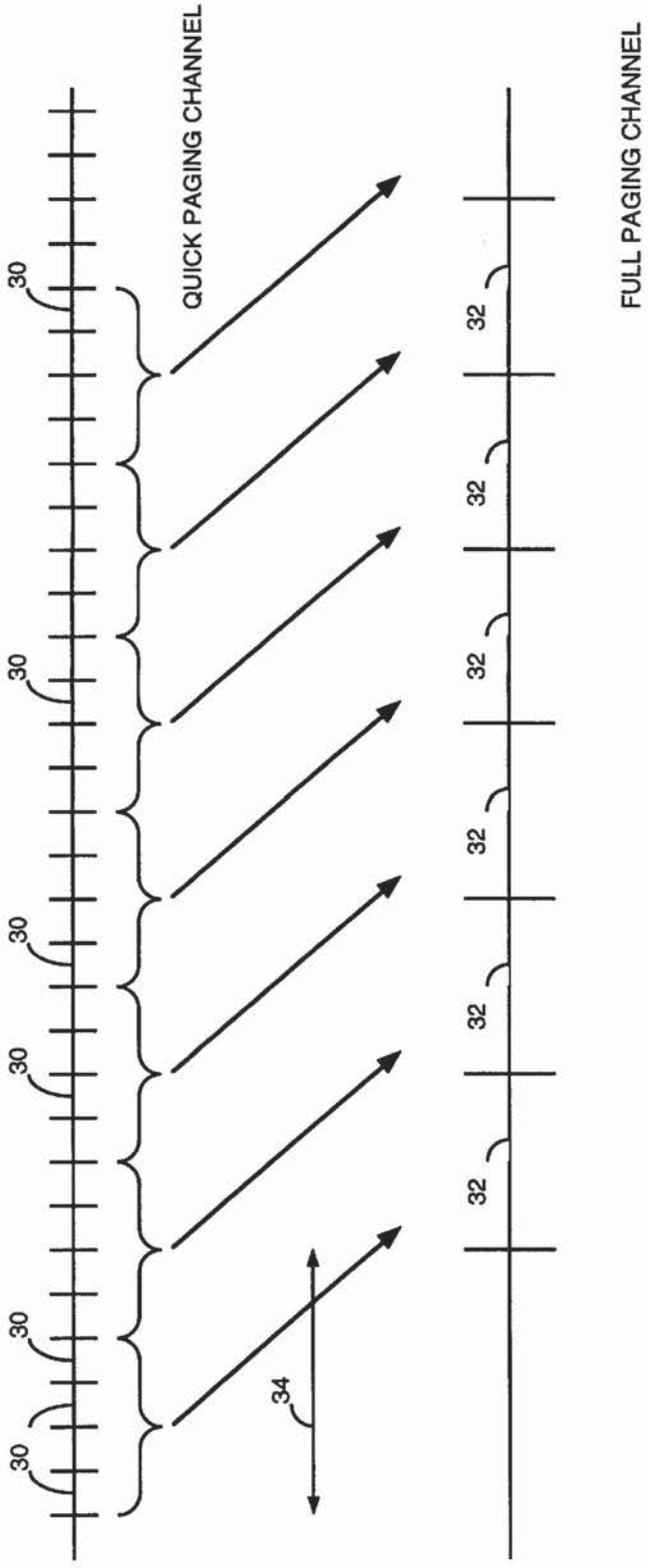


FIG. 2

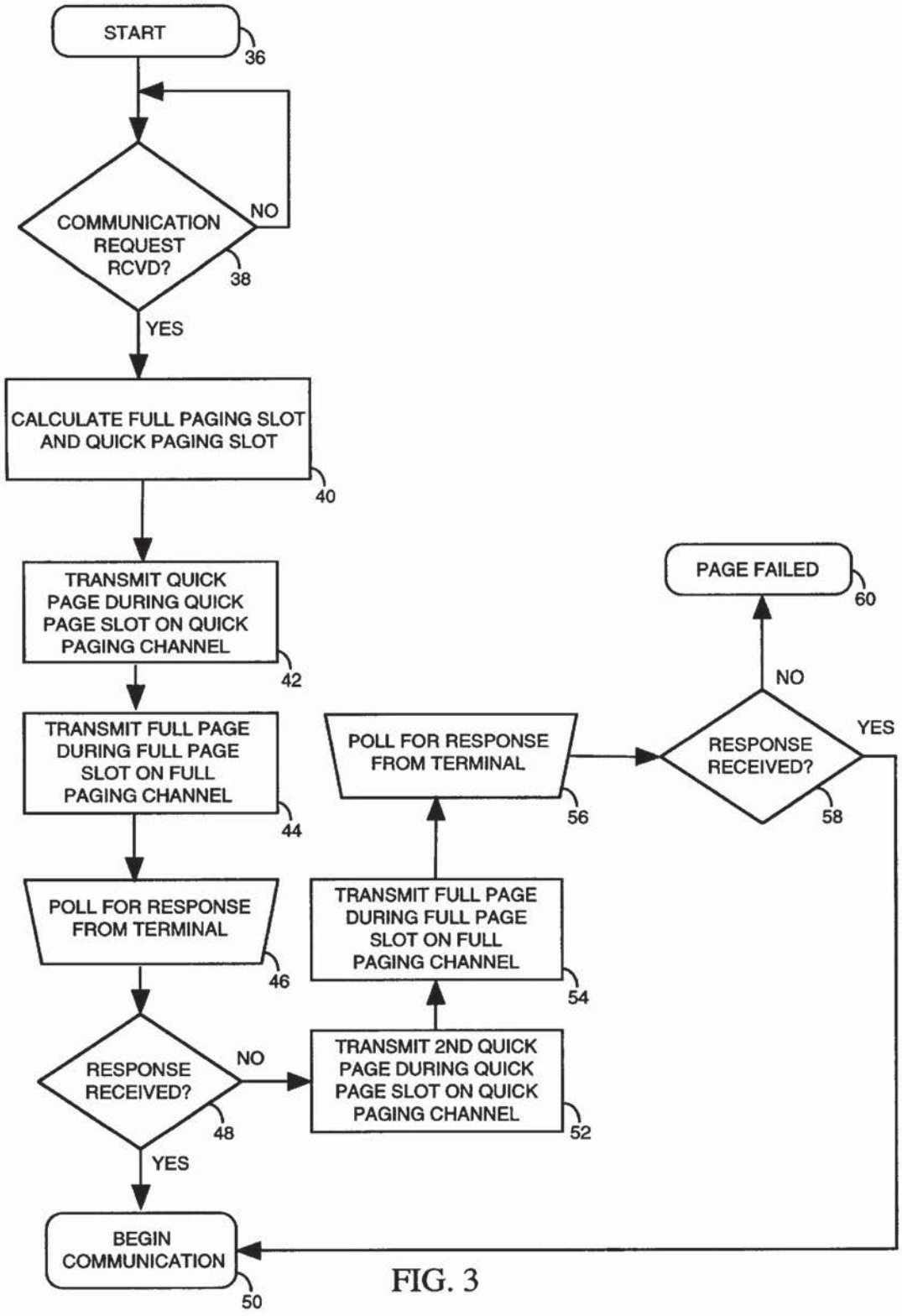

START
36
COMMUNICATION REQUEST RCVD?
38
NO
YES
CALCULATE FULL PAGING SLOT AND QUICK PAGING SLOT
40
TRANSMIT QUICK PAGE DURING QUICK PAGE SLOT ON QUICK PAGING CHANNEL
42
TRANSMIT FULL PAGE DURING FULL PAGE SLOT ON FULL PAGING CHANNEL
44
POLL FOR RESPONSE FROM TERMINAL
46
RESPONSE RECEIVED?
48
NO
YES
TRANSMIT 2ND QUICK PAGE DURING QUICK PAGE SLOT ON QUICK PAGING CHANNEL
52
TRANSMIT FULL PAGE DURING FULL PAGE SLOT ON FULL PAGING CHANNEL
54
POLL FOR RESPONSE FROM TERMINAL
56
RESPONSE RECEIVED?
58
NO
YES
PAGE FAILED
60
BEGIN COMMUNICATION
50

FIG. 3

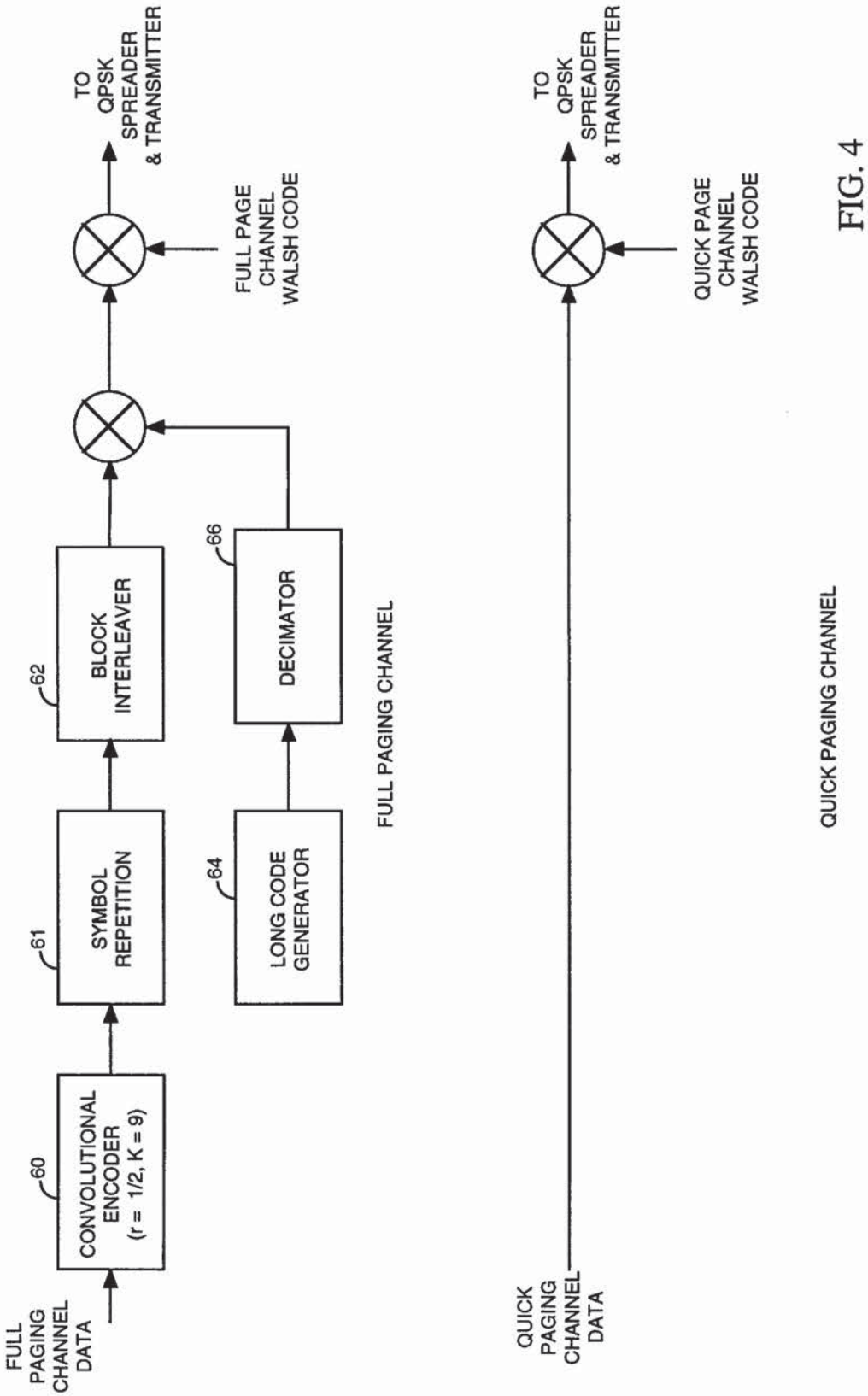

FULL PAGING CHANNEL DATA
60
CONVOLUTIONAL ENCODER (r = 1/2, K = 9)
61
SYMBOL REPETITION
62
BLOCK INTERLEAVER
64
LONG CODE GENERATOR
66
DECIMATOR
FULL PAGING CHANNEL
FULL PAGE CHANNEL WALSH CODE
TO QPSK SPREADER & TRANSMITTER
QUICK PAGING CHANNEL DATA
QUICK PAGE CHANNEL WALSH CODE
TO QPSK SPREADER & TRANSMITTER
QUICK PAGING CHANNEL

FIG. 4

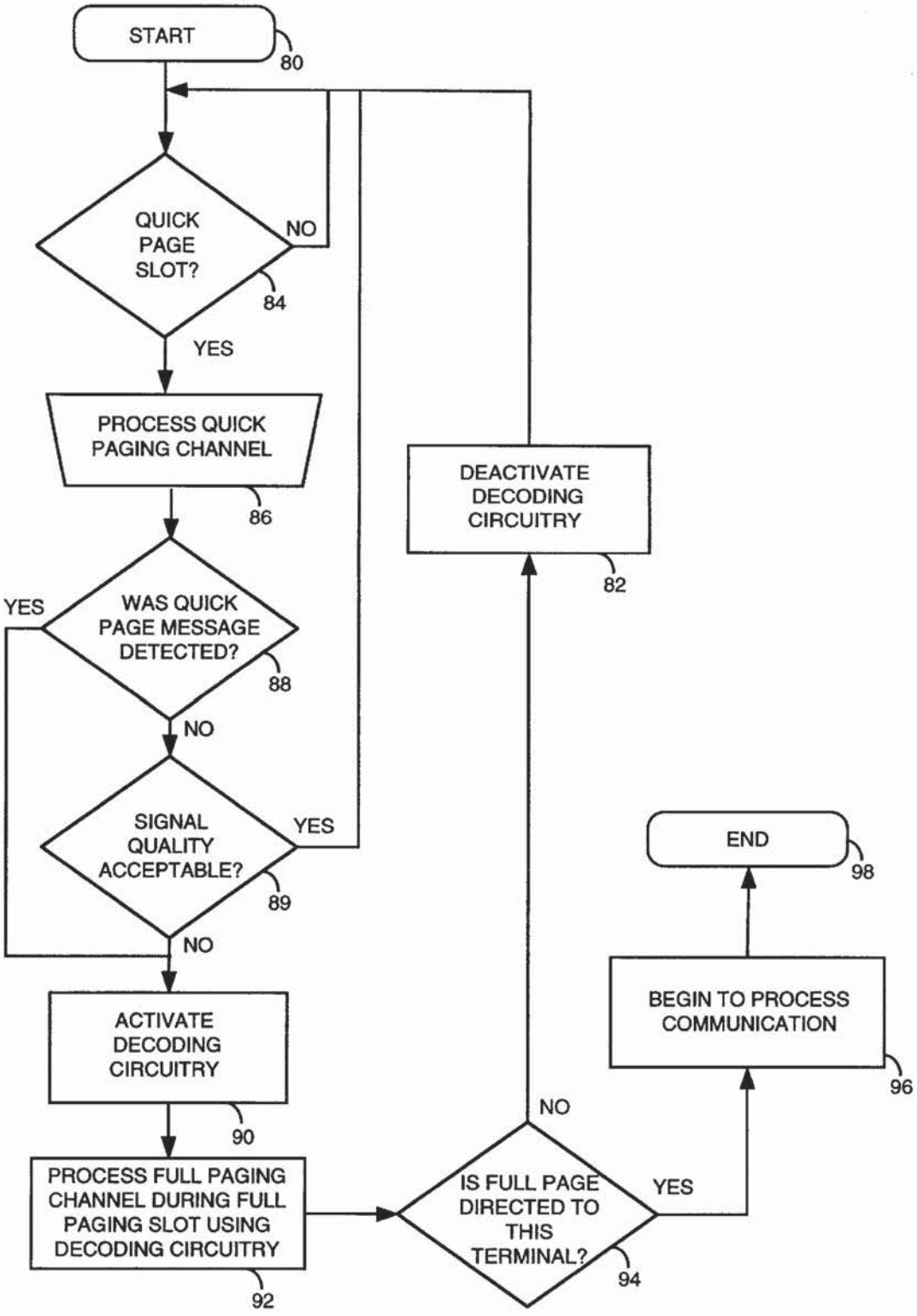

START
80
QUICK PAGE SLOT?
84
NO
YES
PROCESS QUICK PAGING CHANNEL
86
WAS QUICK PAGE MESSAGE DETECTED?
88
YES
NO
SIGNAL QUALITY ACCEPTABLE?
89
YES
NO
ACTIVATE DECODING CIRCUITRY
90
PROCESS FULL PAGING CHANNEL DURING FULL PAGING SLOT USING DECODING CIRCUITRY
92
IS FULL PAGE DIRECTED TO THIS TERMINAL?
94
NO
YES
DEACTIVATE DECODING CIRCUITRY
82
BEGIN TO PROCESS COMMUNICATION
96
END
98

FIG. 5

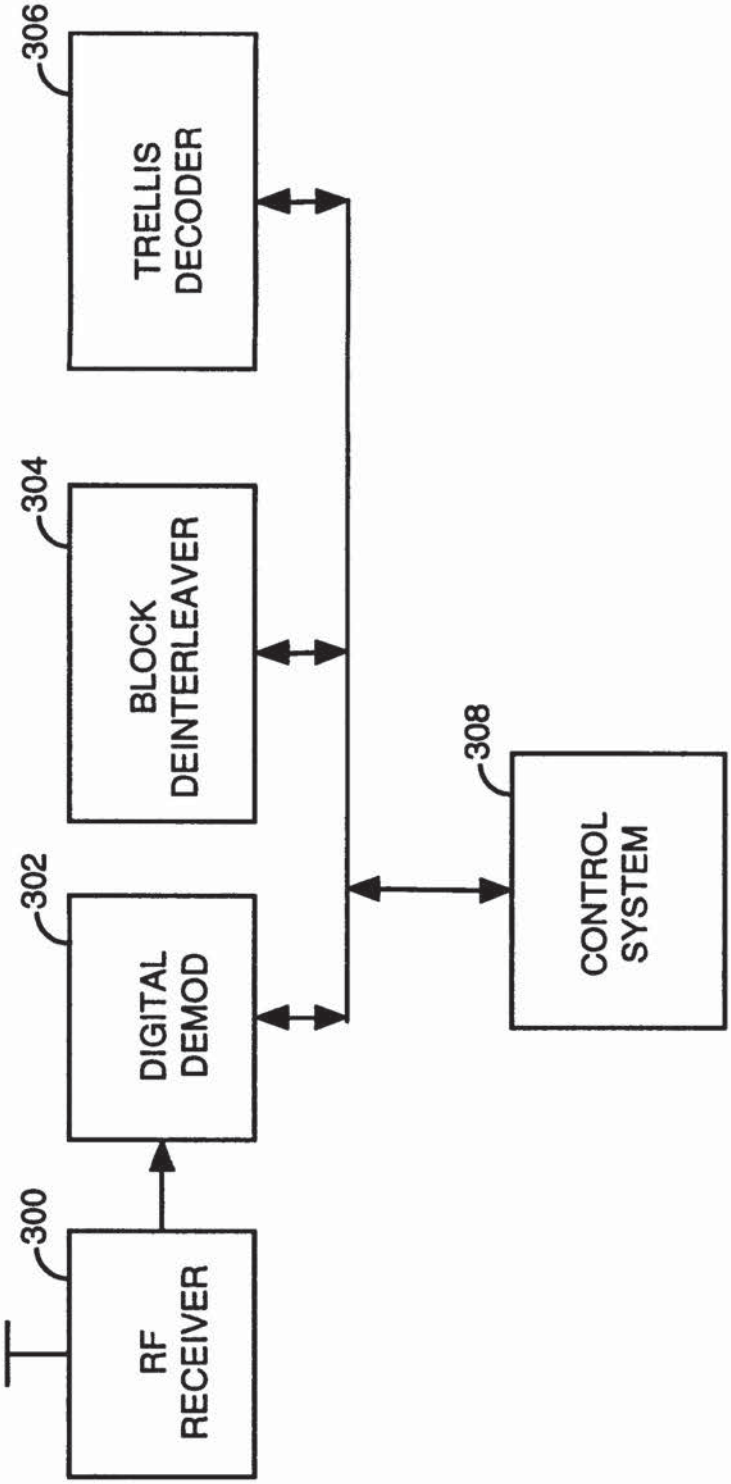

306
TRELLIS DECODER
304
BLOCK DEINTERLEAVER
308
CONTROL SYSTEM
302
DIGITAL DEMOD
300
RF RECEIVER

FIG. 6

DUAL CHANNEL SLOTTED PAGING

BACKGROUND OF THE INVENTION

I. Field of the Invention

The present invention relates to wireless telecommunications. More particularly, the present invention relates to a novel and improved method and apparatus for paging a cellular telephone or other wireless communications device using a quick paging channel and a full paging channel.

II. Description of the Related Art

The IS-95 cellular telephone standard (and its derivatives such as IS-95A and ANSI J-STD-008 referred to herein collectively as IS-95) uses advanced signal processing techniques to provide efficient and high quality cellular telephone service. For example, an IS-95 compliant cellular telephone system utilizes vocoding, error detection, forward error correction (FER), interleaving and spread spectrum modulation in order to make more efficient use of the available RF bandwidth, and to provide more robust connections. In general, the benefits provided by IS-95 include longer talk time and fewer dropped calls when compared to other types of cellular telephone systems.

To conduct communications in an orderly manner, IS-95 provides a set of highly encoded channels over which data having different functions is transmitted. These highly encoded channels include a paging channel over which paging messages are transmitted notifying cellular telephones or other types of wireless terminals that an incoming request to communicate is pending. In accordance with the IS-95 standard, paging messages are transmitted at low to medium data rates during time slots that are preassigned to groups of cellular telephones. Table I provides the data included in a General Page Message as an example of a typical paging message generated substantially in accordance with the IS-95A standard.

TABLE I

Message Field	Length (Bits)
MSG_TYPE (Message Type)	8
CONFIG_MSG_SEQ	6
ACC_MSG_SEQ	6
CLASS_0_DONE	1
CLASS_1_DONE	1
RESERVED	2
BROADCAST_DONE	1
RESERVED	4
ADD_LENGTH	3
ADD_PFIELD	8 × ADD_LENGTH
And zero or more occurrences of the following page record:	
PAGE_CLASS	2
PAGE_SUBCLASS	2
Page class specific fields	Typically 2–12 bytes.

Table I is provided simply to illustrate the length of a typical paging message, so a detailed description of the function of each field is not included herein. Such a detailed description may be obtained, however, by referring to the well known, and publicly available, IS-95 standard (in particular the IS-95A standard).

To monitor for paging messages, a cellular telephone periodically monitors the paging channel during the assigned paging slot. In particular, the cellular telephone periodically activates complex RF and digital signal processing circuitry for as long as is necessary to successfully process the paging message. Since the typical paging message is relatively long, and transmitted via a highly encoded low to medium rate channel, the associated processing during each paging slot requires a significant amount of time and signal processing resources, and therefore requires a significant amount of power to complete. This reduces the amount of time an IS-95 cellular telephone can remain in standby mode using a battery of given capacity, and therefore is highly undesirable.

SUMMARY OF THE INVENTION

The present invention is a novel and improved method and system for paging a cellular telephone or other wireless terminal using two paging channels that reduces standby mode power consumption. In accordance with one embodiment of the invention, a minimally encoded quick paging channel is established over which short, quick page messages are transmitted during one of a set of quick paging slots. The quick page message indicates that a request to communicate has been received and that the receiving communication terminals should monitor a highly encoded full paging channel during the next full paging slot for a more detailed, full page message. A communications terminal monitors the full paging channel after a quick page message has been received on the quick paging channel.

To page the communications terminal, a base station controller (BSC) first generates the quick page message during a quick page slot assigned to a set of communications terminal that includes the particular communications terminals being paged. This is followed by a full page message identifying the particular communications terminal. The communications terminal periodically monitors the quick paging slot and, upon detecting the quick page, activates decoding circuitry for processing the full paging channel. Upon processing the full paging channel, the communications terminal determines if the full page message is directed to it, and if not, deactivates the decoding circuitry and returns to processing the quick paging channel.

BRIEF DESCRIPTION OF THE DRAWINGS

The features, objects, and advantages of the present invention will become more apparent from the detailed description set forth below when taken in conjunction with the drawings in which like reference characters identify corresponding elements throughout and wherein:

FIG. **1** is a block diagram of a cellular telephone system;

FIG. **2** is a timing diagram illustrating the timing slots in the quick paging channel and the full paging channel;

FIG. **3** is a flow diagram illustrating the steps performed during the paging of a wireless terminal;

FIG. **4** is a block diagram illustrating the coding performed on the full paging channel and the quick paging channel;

FIG. **5** is a flow diagram illustrating the steps performed by a wireless terminal during standby mode; and

FIG. **6** is a block diagram of a receiver configured in accordance with one embodiment of the invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

A method and system for paging a cellular telephone or other wireless terminal that uses two paging channels to reduce standby mode power consumption is described. In the following description the invention is set forth in the context of a cellular telephone system operating substantially in accordance with the IS-95 standard. While the invention is particularly suited for operation in such an

environment, many other digital communication systems can benefit from the use of the present invention, including TDMA based wireless communication systems, satellite based communication system, and wireline systems over which encoded signaling is transmitted.

FIG. **1** is a block diagram of a highly simplified cellular telephone system configured in accordance with the use of the present invention. Wireless terminals **10** (typically cellular telephones) are located among base stations **12**. Wireless terminals **10***a* and **10***b* are in active mode and therefore are interfacing with one or more base stations **12** using radio frequency (RF) signals modulated in accordance with the CDMA signal processing techniques of the IS-95 standard. A system and method for processing RF signals substantially in accordance with the use of the IS-95 standard is described in U.S. Pat. No. 5,103,459 entitled "System and Method for Generating Signal Waveforms in a CDMA Cellular Telephone System" assigned to the assignee of the present invention and incorporated herein by reference ('459 patent). The other wireless terminals **10** are in standby mode and therefore monitoring for page messages indicating a request to communicate.

In the preferred embodiment of the invention, each base station generates a forward link signals comprised of a set of forward link channels. The channels are established by a set of orthogonal 64 chip Walsh codes, each of which is used to modulate the data associated with a particular channel. The channels are categorized by function and include a pilot channel over which a phase offset pattern is repeatedly transmitted, a synch channel over which synchronization data is transmitted including the absolute system time and the phase offset of the associated pilot channel, and traffic channels over which the data directed to the terminals are transmitted. The traffic channels are normally assigned to transmit data to a particular wireless terminal **10** for the duration of the interface with that particular base station.

Additionally, in accordance with one embodiment of the invention, one or more of the Walsh channels are designated as quick paging channels, and one or more of the Walsh channels are designated as full paging channels. The designation and operation of the full paging channels is preferably performed in accordance with the paging channel specified by the IS-95 standard. Some methods and apparatus for performing paging substantially in accordance with the IS-95 standard are described in U.S. Pat. Nos. 5,392,287 ('287 patent) entitled "APPARATUS AND METHOD FOR REDUCING POWER CONSUMPTION IN A MOBILE COMMUNICATIONS RECEIVER" and 5,509,015 ('015 patent) entitled "METHOD AND APPARATUS FOR SCHEDULING COMMUNICATIONS BETWEEN TRANSCEIVERS" both assigned to the assignee of the present invention and incorporated herein by reference.

As described in the '287 and '015 patents, and as specified by the IS-95 standard, the full paging channel is time divided into time "slots". The slots are, in turn, assigned to groups of wireless terminals, where the assignment is performed based on the International Mobile Subscriber ID (IMSI) which is unique for each wireless terminal **10**, or other terminal identification information such as one or more Mobile Identification Numbers (MIN). In alternative embodiments of the invention other identification information may also be used including the electronic serial number (ESN) of the wireless terminal or the temporary mobile subscriber ID (TMSI). The various possible types of identification information that may be used will be referred to hereinafter collectively as the MOBILE ID. The quick paging channels are also divided into time slots.

FIG. **2** is a timing diagram illustrating the time slots of a full paging channel and a quick paging channel when configured in accordance with one embodiment of the invention. The quick paging channel is divided into quick paging slots **30** and the full paging channel is divided into full paging slots **32** which are preferably longer in duration than the quick paging slots. Sets, or groups, of quick paging slots **30** are assigned to single full paging slots **32** as illustrated by the diagonal arrows, although the use of a one-to-one correspondence between quick paging slots and full paging slots, or other ratios, is consistent with the use of the invention. The assignment of the quick paging slots **30** to a particular set of wireless terminals is preferably performed via application of a hashing function to the MOBILE ID of the wireless terminal **10**.

To page a particular wireless terminal **10**, a quick page message is transmitted during the quick paging slot and a full page message is transmitted during the full page slot assigned to that wireless terminal. The quick paging slot and full paging slots occur in a periodically repeating fashion, which ensures that a slot associated with a particular terminal occurs after a sufficient period of time. As illustrated in FIG. **2**, the full page slot **32** occurs a delay **34** after the associated quick paging slots **30** to allow the wireless terminal to process the quick page message and activate additional decoding circuitry before the next full page slot.

FIG. **3** is a block diagram of the steps performed by BSC **14** during the paging process. The paging process begins at step **36** and at step **38** it is determined whether a request to communicate has been received. If not step **38** is performed again.

If a request to communicate has been received, the full paging slot and quick paging slot associated with the wireless terminal to which the communication request is directed are calculated at step **40** based on the MOBILE ID or other identification information of that wireless terminal **10**. BSC **14** preferably performs the necessary processing using one or more microprocessors running software stored in memory (not shown).

The full paging slot is preferably determined in accordance with the above referenced '287 and '015 patents, and the quick paging slot is determined by application of a hashing function to the MOBILE ID, although other methods for assigning paging slots to wireless terminals are consistent with the use of the present invention. If the MOBILE ID is not contained directly within the request to communicate, it can be obtained via look-up in a database using any other identification information contained in the request such as the telephone number or mobile identification number (MIN) of the wireless terminal **10**.

Once the quick paging slot and the full paging slot are determined, BSC **14** transmits a quick page message at step **42** on a quick paging channel, and transmits a full page message at step **44** on a full paging channel, via one or more base stations **12**. The base stations **12** encode and modulate the paging channels as described in greater detail below, and transmission of the two page messages occurs during the corresponding quick paging slot and full paging slot.

After transmission of the quick page message and the full page message, the BSC **14** polls for a response at step **46** indicating that the page was received. If a response was received the communication begins at step **50**.

If no response was received after a timeout period, a second quick page message is transmitted at step **52**, and a second full page message is transmitted at step **54**. At step **56** the BSC **14** polls for a response from the wireless

terminal **10** and determines at step **58** if a response was received. If a response was received, the communication begins at step **50**. If it is determined at step **58** that no response was received, the page fails at step **60**. In an alternative embodiment of the invention, two quick paging messages and two corresponding full page messages are generated for each page. The second quick page message and full page message increase the likelihood of the page being received, without introducing the delay necessary to determine whether an acknowledgment message has been received from the wireless terminal **10**.

In the preferred embodiment of the invention, the quick page message is comprised of an INCOMMING_PAGE bit. An INCOMMING_PAGE bit in a first state (such as logic high) indicates that a communication request has been received for one of the wireless terminals **10** associated with that quick paging slot, and therefore that those wireless terminals should process the full paging channel during the next designated full paging slot. An INCOMMING_PAGE bit in a second state (such as logic low) indicates that no communication request has been received for those wireless terminals **10**, and therefore the full paging channel should not be processed during next full paging slot. Hence, the quick page message is more highly encoded than the full page message, since the page is represented by a single bit, rather than a substantially higher number of bits, and therefore can be processed with fewer resources. Such "message" encoding should not be confused with the "channel" coding described below, where a greater amount of encoding requires more data processing resources.

In the preferred embodiment of the invention, the full page message contains the information specified in the IS-95 standard for a normal page message that allows each wireless terminal **10** to determine whether the page is directed to it. An example of a page generated in accordance with the IS-95A standard is provided in Table I listed above. As illustrated by Table I, the full page message contains significantly more information than the quick page message, which is preferably comprised of a single bit. Therefore, the quick page message can be processed more easily by each wireless terminal **10**, and with less power, than a full page message.

In alternative embodiments of the invention, multi-bit quick page messages are used. These multi-bit quick paging messages are used to encode and convey additional information beyond simply indicating that the wireless terminal **10** should monitor the full paging channel during the next assigned full paging slot **32**. For example, the multi-bit quick page message could be used to indicate more particularly which wireless terminal **10** is being paged from the subset of wireless terminal assigned to the corresponding quick page slot **30**. The multi-bit quick page message could also be used to indicate that the full paging channel should be performed for a longer duration so that system parameter changes can be broadcast to all the wireless terminals **10**. Those skilled in the art will recognize various useful types of information that can be transmitted using a multi-bit quick page message. Also, in another alternative embodiment of the invention, reduced forward error correction encoding is performed on the quick page message.

In addition to transmitting less information in a quick page message than in a full page message, the preferred embodiment of the invention incorporates a minimal coding scheme for the quick paging channel when compared to the full paging channel. FIG. **4** provides an illustration of the coding schemes employed for the full paging channel and the quick paging channel in accordance with one embodiment of the invention.

As shown in FIG. **4**, data transmitted via the full paging channel is convolutionally encoded by convolutional encoder **60** and the resulting code symbols repeated by symbol repeater **61** in order to generate symbols at a predetermined rate. The repeated code symbols are then block interleaved by block interleaver **62**. The data from the block interleaver is scrambled via EXCLUSIVE-OR (EX-OR) with a decimated long code generated by long code generator **64** and decimator **66**. The long code is binary code generated in a predetermined manner as a function of a seed number, and is known to all wireless terminals **10**. The scrambled data is then modulated with a Walsh channel code designated for a full paging channel, and the Walsh channel code modulated data is QPSK spread using a pseudorandom noise code (PN code), summed with the data from the other channels, and upconverted for transmission, preferably in accordance with the IS-95 standard (spreading, summing and upconversion not shown).

Still referring to FIG. **4**, data transmitted via the quick paging channel is applied directly to a Walsh channel code designated for a quick paging channel, and then spread, summed and upconverted as described above. Preferably, a single bit of data transmitted via the quick channel is modulated multiple times by the same Walsh code, effectively transmitting the bit multiple times. One could also transmit the data bit repeatedly using a symbol repeater like symbol repeater **61** used for the full paging channel. In still another embodiment of the invention, the quick paging channel could be scrambled using the long code as performed for the full paging channel.

As should be apparent from FIG. **4**, the processing associated with transmitting information over the quick paging channel is substantially less than that associated with the full paging channel. Therefore, the amount of processing necessary to perform receive processing of the quick paging channel is also substantially less, and therefore requires less energy than that necessary for the full paging channel. While the reduced amount of processing performed for the quick paging channel does increase the likelihood of error during the processing of any particular bit, transmitting the same bit multiple times reduces the total overall likelihood of error, while preserving low processing complexity.

FIG. **5** is a flow diagram of the processing performed by a wireless terminal **10** in standby mode when performed in accordance with one embodiment of the invention. The processing is preferably performed using a microprocessor controlled by software instructions stored in memory coupled to other integrated circuits and systems (not shown) that are well known in the art. The processing begins at step **80** and at step **84** it is determined whether the assigned quick paging slot has arrived, and if not step **84** is performed again.

If the assigned quick paging slot has arrived, the wireless terminal **10** processes the quick paging channel at step **86**. Preferably, the processing is performed using a significantly smaller subset of the signal processing circuitry contained in the wireless terminal than used to process full page messages. In accordance with the transmit processing performed for the quick paging channel shown in FIG. **4**, the receive processing is preferably comprised of downconversion of the RF energy received, despreading with the PN spreading code, and demodulation with the designated Walsh code. The resulting soft decision data is processed directly to determine the logic level transmitted.

Referring again to FIG. **5**, at step **88** it is determined whether a quick page message was received at step **86** based on the logic level of the data detected. If a quick page

message was detected processing continues at step **90** as described below. If a quick page message was not detected, it is further determined at step **89** whether the signal quality during the processing of the quick paging channel was acceptable. If so, the wireless terminal **10** returns to step **82**. If the signal quality was not acceptable, processing continues at step **90** as described below.

The received signal quality can be determined by various well known methods including determining when the receive power of the signal transmitted from transmitter **50** falls below a threshold, or by determining when the signal-to-noise ratio of the pilot channel falls below a predetermined threshold. By monitoring for a full page message when the received signal quality is unacceptable, the number of missed full page messages due to undetected quick page messages resulting from unacceptable signal quality is minimized.

If a quick page message was detected, or the received signal quality was not acceptable, the wireless terminal **10** activates additional decoding circuitry at step **90** and, at step **92** processes the full paging channel during the assigned full paging slot using the activated circuitry. The time between the quick paging slot and the full paging slot assigned to a particular terminal must be sufficient to allow activation of the additional decoding circuitry within the wireless terminal **10** after detection of the quick page message before the full page slot occurs.

At step **94**, the wireless terminal **10** determines whether the full page message processed at step **92** was directed to it based on the address contained in that message, and if not decoding circuitry within wireless terminal **10** is deactivated at step **82** and step **84** is performed again. If the full page message was directed to the wireless terminal **10**, processing for the corresponding communication begins within the wireless terminal at step **96**, and the wireless terminal enters active mode at step **98**.

FIG. **6** is a block diagram providing a highly simplified illustration of wireless terminal **10** when configured in accordance with one embodiment of the invention. Digital demodulator **302**, block interleaver **304**, trellis decoder **306** and control system **308** are coupled via a digital bus, and RF receiver **300** is coupled to digital demodulator **302**.

During standby mode, the control system periodically activates RF receiver **300** and digital demodulator **302** to process the pilot and quick paging channels. RF receiver **300** downconverts and digitizes RF signals and digital demodulator **302** performs digital demodulation for a first duration generating soft decision data for the channels being processed. Control system **308** examines the pilot channel soft decision data to determine the quality of the signal and examines the quick paging channel to determine if a quick page message has been received.

If a quick page message has been received, or the signal has been received with poor quality, control system **308** activates block deinterleaver **304** and trellis decoder **306** and configures digital demodulator to begin processing the full paging channel for a second duration that is longer than the first duration. Control system **308** then monitors the data received over the full paging channel for a full paging message directed to it, and if none is detected, deactivates block deinterleaver **304** and trellis decoder **306** and continues in standby mode. If a full page message is detected, control system **308** places the wireless terminal in active mode during which the associated communication is conducted.

As should be apparent from the description provided above, by paging using a quick page message with a minimal number of bits, and which is transmitted over a minimally encoded channel, the present invention allows a wireless terminal to consume less power when monitoring for page messages during standby mode. Consuming less power in standby mode allows a wireless terminal to operate longer on a given battery, and therefore extends the standby time of that wireless terminal. Since wireless terminals are typically used in mobile telecommunications, it is often necessary to go extended periods of time without recharging or replacing the battery of the wireless terminal. Thus, in order to provide increased convenience, and to reduce the likelihood of missing page messages due to battery depletion, extending the standby time for a given battery size is highly desirable.

Thus, a dual channel method and system for paging cellular telephones and other wireless terminals that reduces standby power consumption has been described. The previous description of the preferred embodiments is provided to enable any person skilled in the art to make or use the present invention. The various modifications to these embodiments will be readily apparent to those skilled in the art, and the generic principles defined herein may be applied to other embodiments without the use of the inventive faculty. Thus, the present invention is not intended to be limited to the embodiments shown herein but is to be accorded the widest scope consistent with the principles and novel features disclosed herein.

I claim:

1. A method for paging a wireless terminal in a wireless telecommunications system comprising:

transmitting a quick page message via a less encoded channel, said quick page message notifying the wireless terminal to begin monitoring a highly encoded channel for a full page message; and

transmitting a full page message via said highly encoded channel;

said highly encoded channel comprising a series connected convolutional encoder, an interleaver, and a modulator for performing modulation with a Walsh code; and

said less encoded channel comprising a modulator for modulating data with a Walsh code.

2. The method as set forth in claim **1** wherein said quick page message contains less information than said full page message.

3. The method as set forth in claim **1** wherein said quick page message is comprised of a single bit of data.

4. A method for paging a particular wireless terminal included within a set of wireless terminals comprising:

a) transmitting a quick page message directed to a subset of said set of wireless terminals said subset including the particular wireless terminal, said quick page message for directing said subset to monitor a highly encoded channel for a full page message; and

b) transmitting a full page message identifying said particular wireless terminal;

said highly encoded channel comprising a series connected convolutional encoder, an interleaver, and a modulator for performing modulation with a Walsh code.

5. The method as set forth in claim **4** wherein step a) is comprised of the step of transmitting said quick page message via a less encoded channel during a first time slot assigned to the subset of wireless terminals, and said less encoded channel comprises a modulator for modulating data with a Walsh code.

6. The method as set forth in claim **5** wherein step b) is comprised of the step of transmitting said full page message during a second time slot assigned to the subset of wireless terminals that occurs after said first time slot.

7. The method as set forth in claim **4** wherein said quick page message is transmitted over a less encoded channel and said full page message is transmitted over said highly encoded channel.

8. The method as set forth in claim **7** wherein said less encoded channel is established via direct sequence modulation with a first channel code, and said highly encoded channel is established via direct sequence modulation with a second channel code.

9. The method as set forth in claim **4** wherein said quick page message has less information than said full page message.

10. A method for paging a wireless terminal comprising:

a) generating a quick page message;

b) transmitting said quick page message via a less encoded channel;

c) said quick page message causing said wireless terminal to monitor a highly encoded channel for a full page message;

d) generating a full page message; and

e) transmitting said full page message via said highly encoded channel;

said highly encoded channel comprising a series connected convolutional encoder, an interleaver, and a modulator for performing modulation with a Walsh code; and

said less encoded channel comprising a modulator for modulating data with a Walsh code.

11. The method as set forth in claim **10** further comprising the steps of:

direct sequence spreading said full page message; and

direct sequence spreading said quick page message.

12. The method as set forth in claim **10** further comprising the steps of:

adding error detection information to said full page message;

direct sequence spreading said full page message; and

direct sequence spreading said quick page message.

13. The method as set forth in claim **10** further comprising the steps of adding a time delay between when said quick page message is transmitted and when said full page message is transmitted.

14. The method as set forth in claim **10** further comprising the steps of:

receiving said quick page message;

activating signal processing circuitry; and

processing a full page message using said signal processing circuitry.

15. The method as set forth in claim **14** wherein said signal processing circuitry is comprised of a trellis decoding system.

16. The method as set forth in claim **14** wherein said signal processing circuitry is comprised of a deinterleaver.

17. The method as set forth in claim **14** wherein said signal processing circuitry is comprised of a cyclical redundancy check circuit.

* * * * *

EXHIBIT 14

US008446849B2

(12) **United States Patent**
Damnjanovic

(10) **Patent No.: US 8,446,849 B2**
(45) **Date of Patent: May 21, 2013**

(54) **METHODS AND APPARATUSES FOR POWER CONTROL**

(75) Inventor: **Aleksandar Damnjanovic**, Del Mar, CA (US)

(73) Assignee: **QUALCOMM Incorporated**, San Diego, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 687 days.

(21) Appl. No.: **12/142,121**

(22) Filed: **Jun. 19, 2008**

(65) **Prior Publication Data**

US 2008/0316950 A1 Dec. 25, 2008

Related U.S. Application Data

(60) Provisional application No. 60/945,325, filed on Jun. 20, 2007.

(51) **Int. Cl.**
G08C 17/00 (2006.01)
H04W 4/00 (2009.01)
H04B 7/00 (2006.01)

(52) **U.S. Cl.**
USPC **370/311**; 370/329; 455/522

(58) **Field of Classification Search** 370/311, 370/328, 329; 455/522
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,604,730	A	2/1997	Tiedemann, Jr.	
6,567,459	B1 *	5/2003	Hakkinen et al.	375/132
2001/0017848	A1 *	8/2001	Tiedemann, Jr.	370/318
2002/0018453	A1 *	2/2002	Yu et al.	370/333
2002/0123337	A1 *	9/2002	Dharia et al.	455/423
2003/0026219	A1	2/2003	Moon et al.	
2004/0233870	A1 *	11/2004	Willenegger et al.	370/329
2006/0003696	A1 *	1/2006	Diaz Cervera et al.	455/11.1
2006/0262840	A1 *	11/2006	Wang et al.	375/221
2007/0178930	A1 *	8/2007	Xiao et al.	455/522
2008/0045259	A1 *	2/2008	Shen et al.	455/522
2008/0108379	A1 *	5/2008	Cho et al.	455/522
2008/0171520	A1 *	7/2008	Steer	455/77
2009/0185577	A1	7/2009	Kishiyama et al.	
2011/0010705	A1 *	1/2011	Adams et al.	717/178

FOREIGN PATENT DOCUMENTS

CN	1320307		10/2001
EP	1871028	*	12/2007
EP	2033338	A1	3/2009

(Continued)

OTHER PUBLICATIONS

International Search Report and Written Opinion—PCT/US2008/067754, International Search Authority—European Patent Office—Feb. 27, 2009.

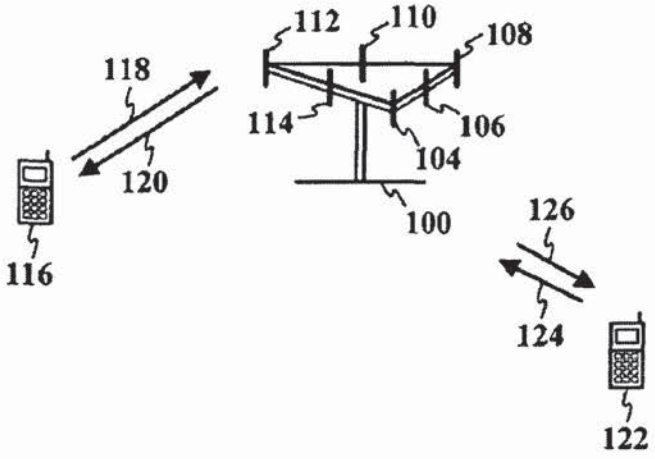
(Continued)

Primary Examiner — Ronald Abelson
(74) *Attorney, Agent, or Firm* — Liem T. Do

(57) **ABSTRACT**

Downlink power control commands are mapped to resources used for uplink communications within a wireless communication system. An eNode B receives communications from UEs and determines the resources used by the UEs for those transmissions on the uplink which are transmitted at non-optimum power levels. Power control messages are formulated wherein the location of the power control commands is mapped to particular resources used by the UEs for their uplink transmissions. This facilitates the eNode B to dynamically assign resources for the power control commands while permitting the UEs to decode the power control messages to adjust their power accordingly.

53 Claims, 14 Drawing Sheets



FOREIGN PATENT DOCUMENTS

JP	10503337		3/1998
JP	2002508620	A	3/2002
JP	2002520943	A	7/2002
JP	PCT/JP2006/305499	*	10/2006
TW	453058	B	9/2001
TW	453059	B	9/2001
WO	WO9603813	A1	2/1996
WO	WO9949595	A1	9/1999
WO	WO0003495		1/2000

OTHER PUBLICATIONS

NTT DoCoMo: "Intra-Node BTPC for Non-Scheduled Data Transmission in E-UTRA Uplink" 3GPP Draft; R1-072419 UL TPC for Non-Scheduled Data, 3rd Generation Partnership Project (SGPP), Mobile Competence Centre; 650,Route Des Lucioles; F-06921 Sophia-Antipolis Cedex ; France, vol. tsg-ran\WGI-RLI\TSGRI_49\Docs, no. Kobe, Japan; May 7, 2007, May 2, 2007, XP050106145.

Taiwan Search Report—TW097123263—TIPO—Oct. 26, 2011.

* cited by examiner

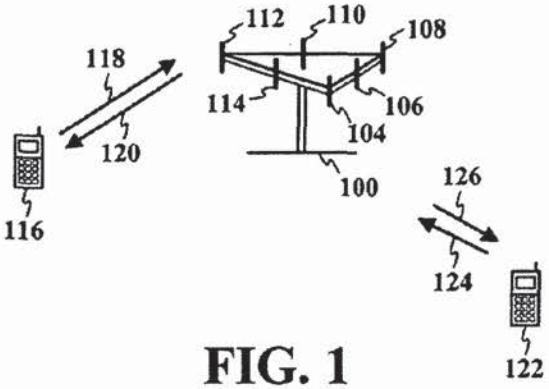

112
110
108
118
114
106
104
120
100
126
116
124
122

FIG. 1

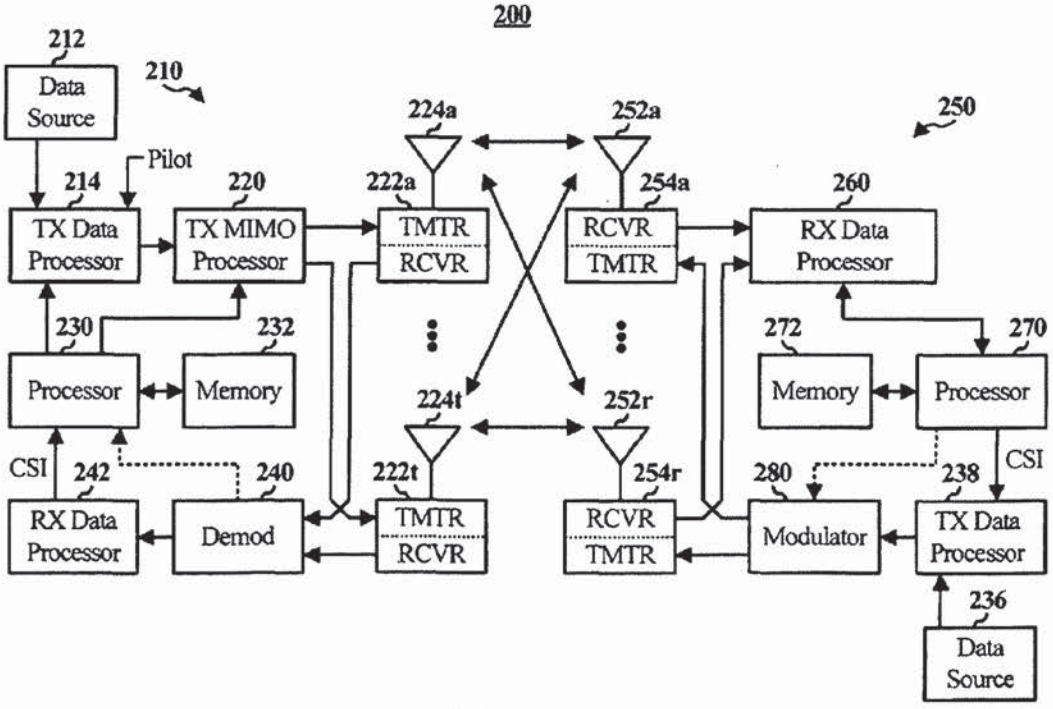

200
212
Data Source
210
224a
252a
250
Pilot
214
220
222a
254a
260
TX Data Processor
TX MIMO Processor
TMTR RCVR
RCVR TMTR
RX Data Processor
230
232
272
270
Processor
Memory
Memory
Processor
224t
252r
CSI
242
240
222t
254r
280
238
CSI
RX Data Processor
Demod
TMTR RCVR
RCVR TMTR
Modulator
TX Data Processor
236
Data Source

FIG. 2

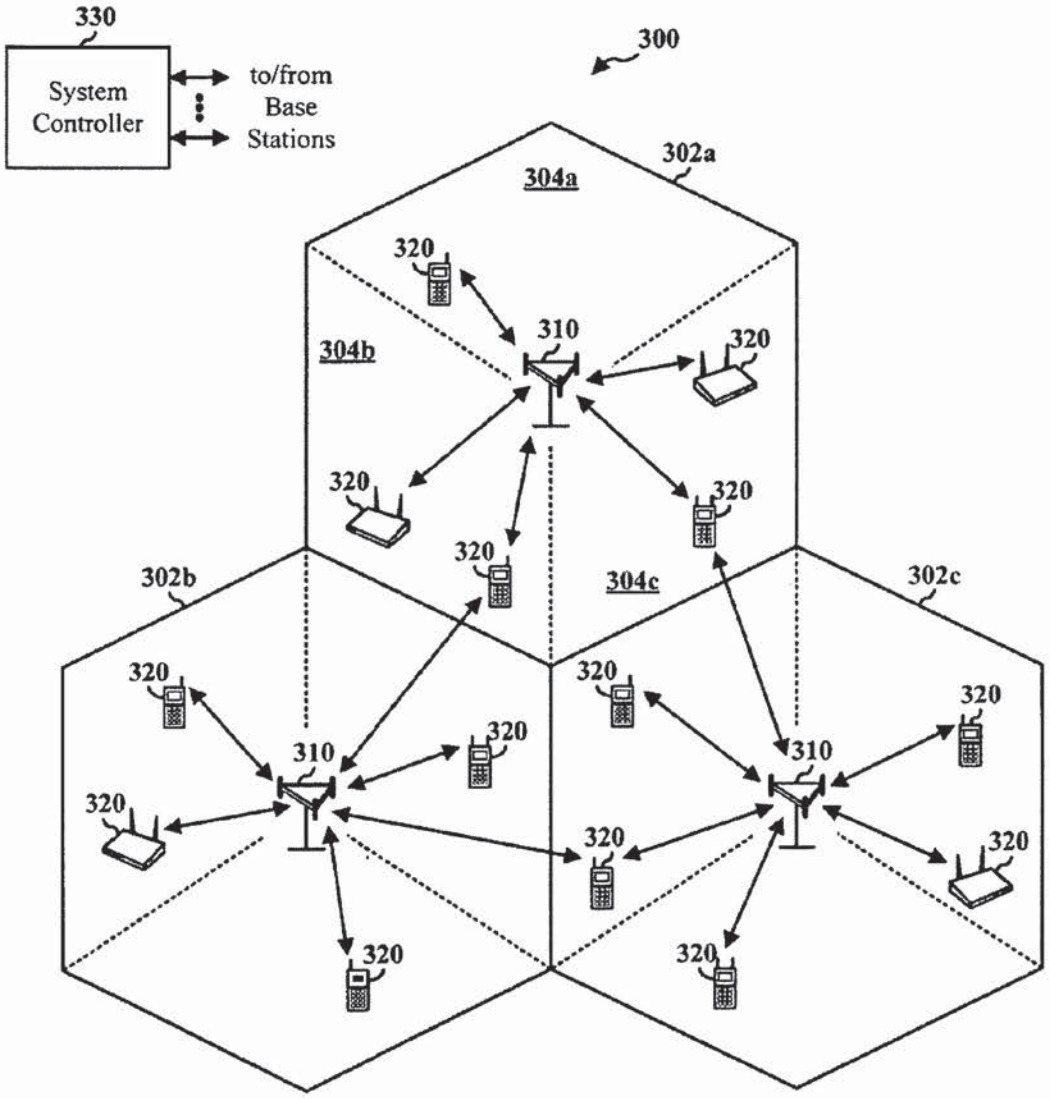

330
System Controller
to/from Base Stations
300
302a
304a
304b
304c
302b
302c
310
320

FIG. 3

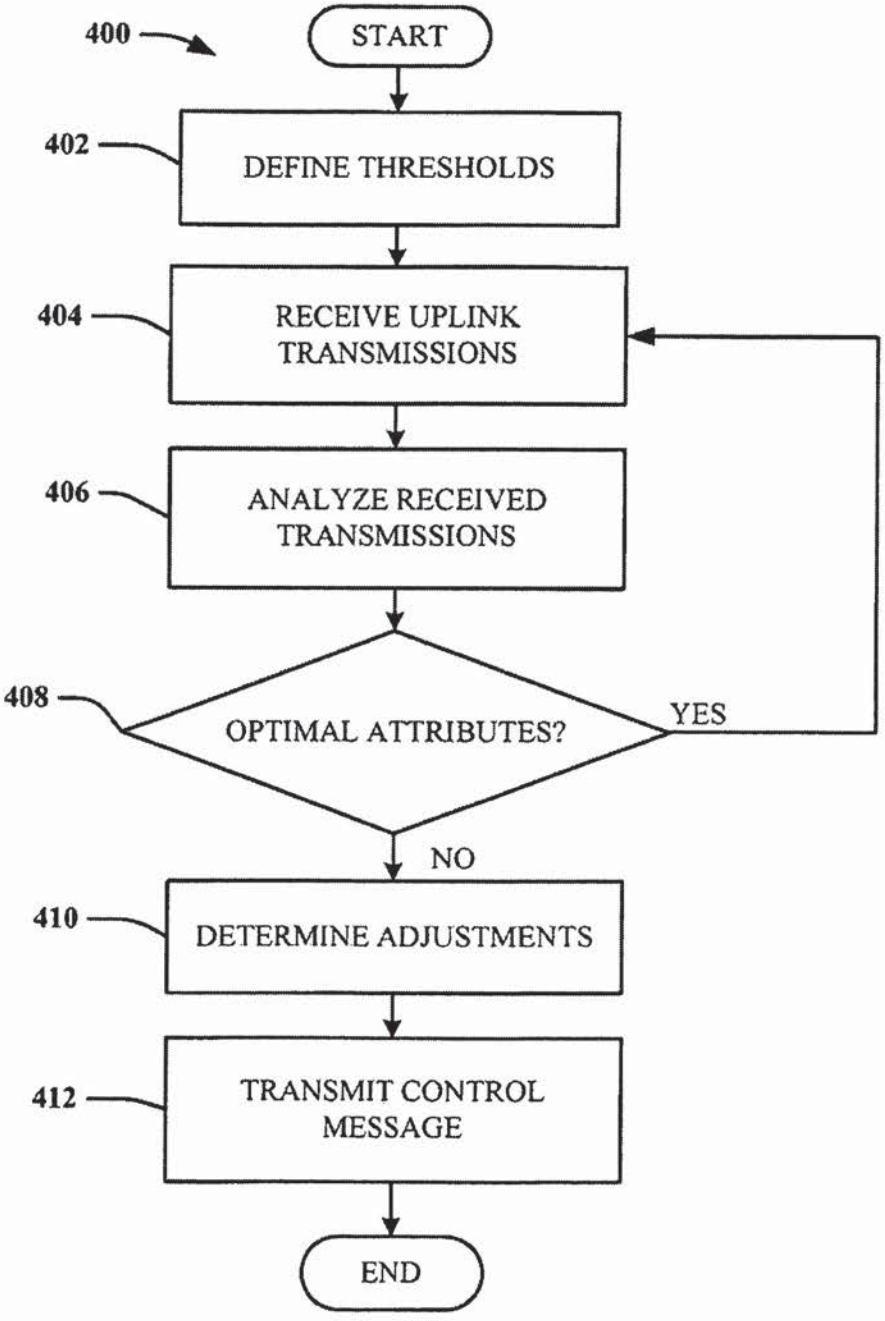

400
START
402
DEFINE THRESHOLDS
404
RECEIVE UPLINK TRANSMISSIONS
406
ANALYZE RECEIVED TRANSMISSIONS
408
OPTIMAL ATTRIBUTES?
YES
NO
410
DETERMINE ADJUSTMENTS
412
TRANSMIT CONTROL MESSAGE
END

FIG. 4

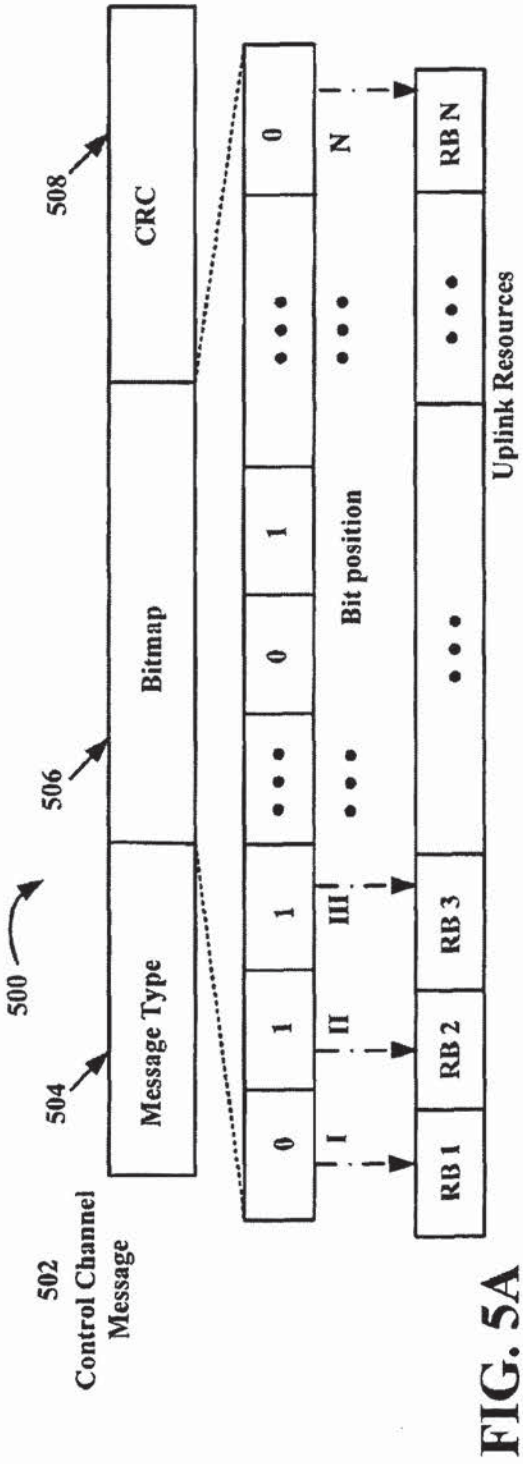

500
502
Control Channel Message
504
Message Type
506
Bitmap
508
CRC
0
1
1
0
1
0
N
RB 1
RB 2
RB 3
RB N
I
II
III
Bit position
Uplink Resources

FIG. 5A

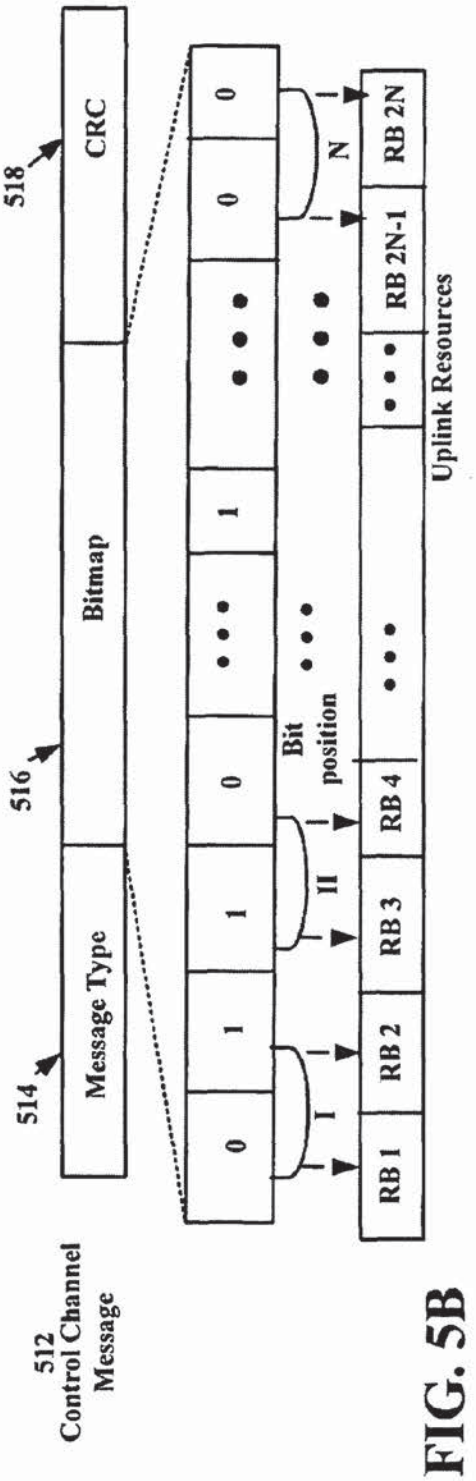

512
Control Channel Message
514
Message Type
516
Bitmap
518
CRC
0
1
1
0
1
0
0
N
RB 1
RB 2
RB 3
RB 4
RB 2N-1
RB 2N
I
II
Bit position
Uplink Resources

FIG. 5B

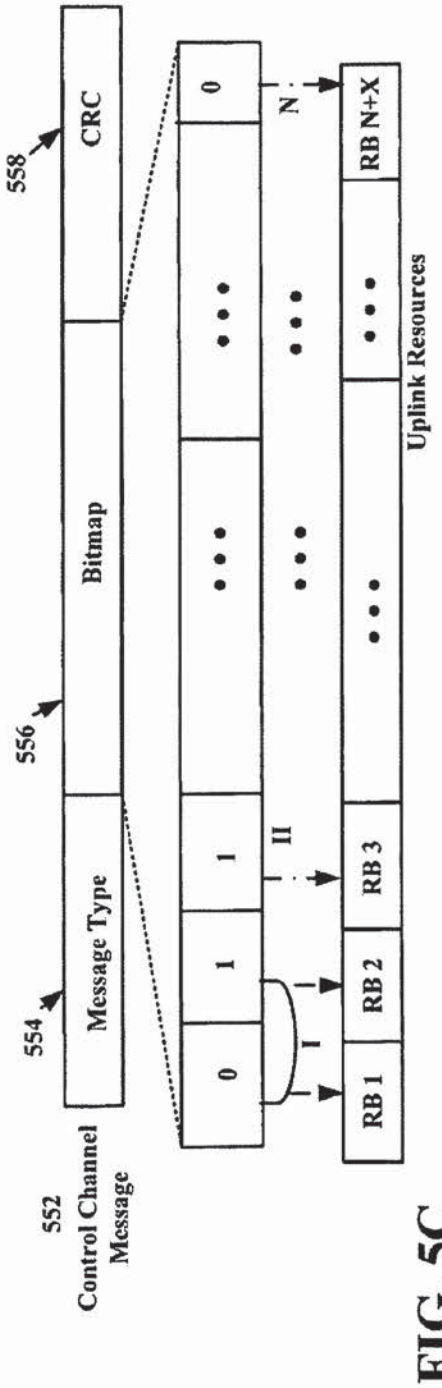

552
Control Channel Message
554
Message Type
556
Bitmap
558
CRC
0
1
1
I
II
N
RB 1
RB 2
RB 3
RB N+X
Uplink Resources

FIG. 5C

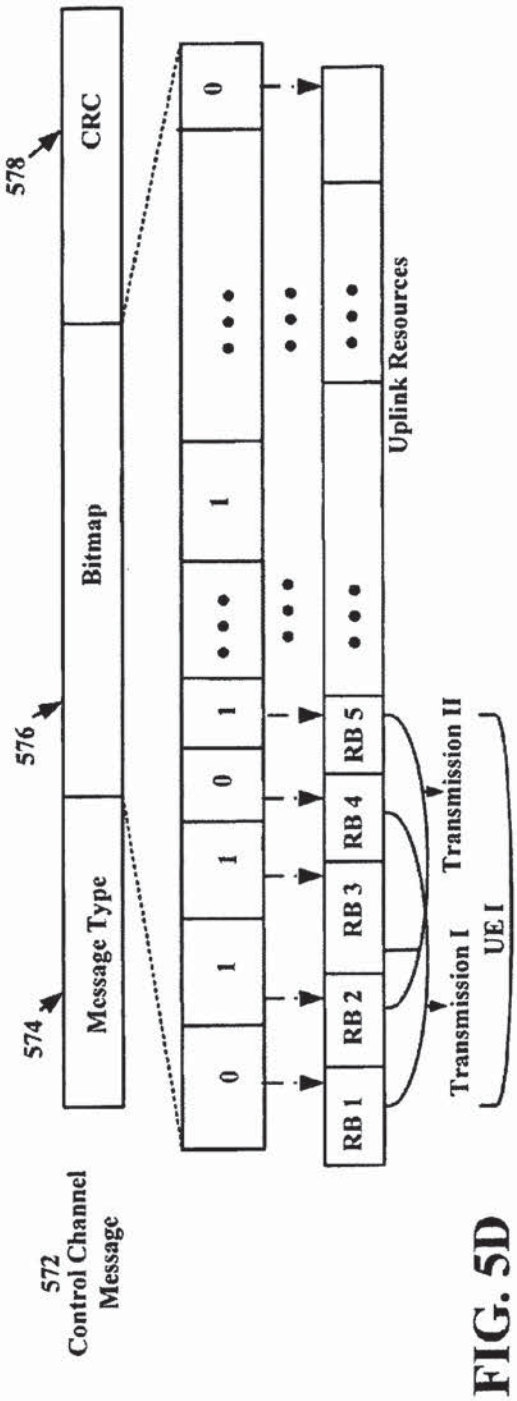

572
Control Channel Message
574
Message Type
576
Bitmap
578
CRC
0
1
1
0
1
1
0
RB 1
RB 2
RB 3
RB 4
RB 5
Uplink Resources
Transmission I
Transmission II
UE I

FIG. 5D

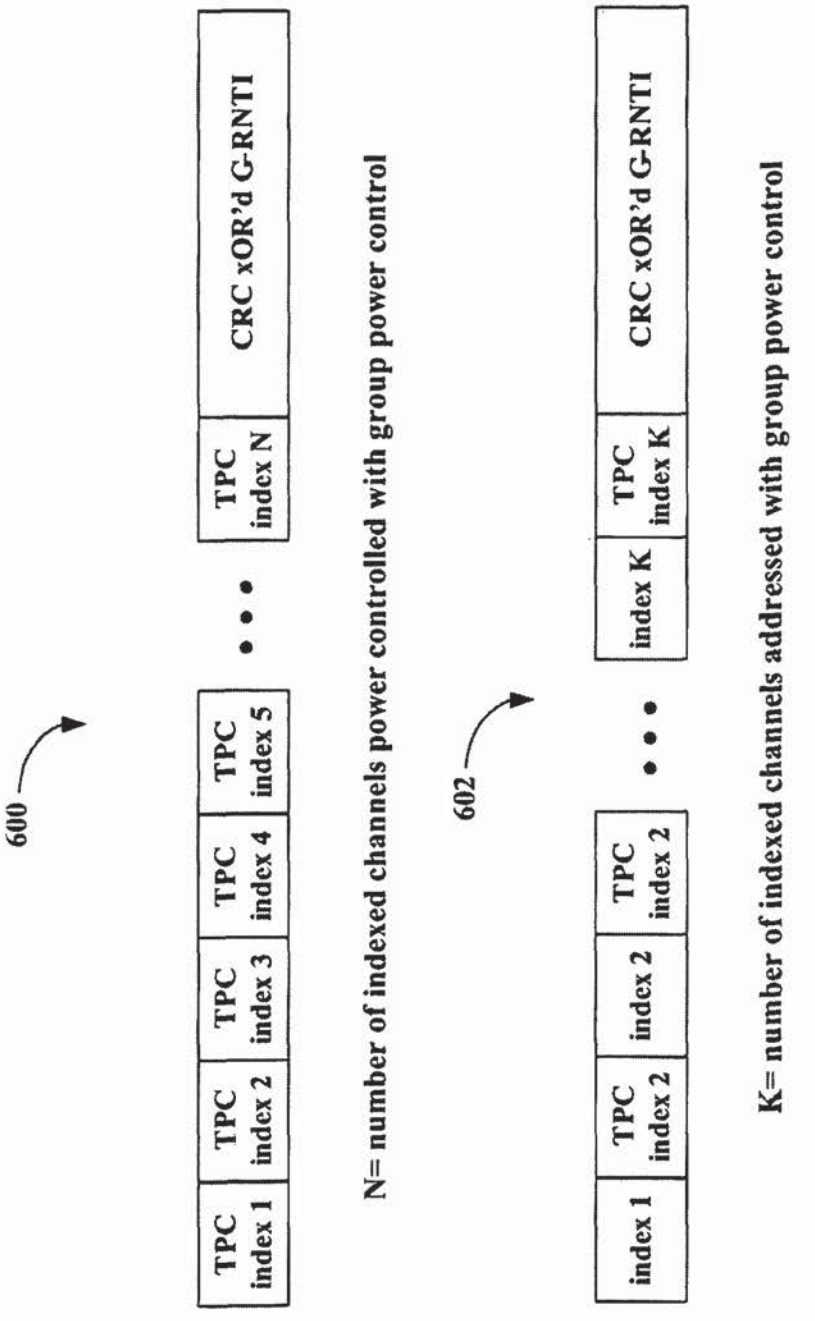

600
TPC index 1
TPC index 2
TPC index 3
TPC index 4
TPC index 5
• • •
TPC index N
CRC xOR'd G-RNTI
N= number of indexed channels power controlled with group power control
602
index 1
TPC index 2
index 2
TPC index 2
• • •
index K
TPC index K
CRC xOR'd G-RNTI
K= number of indexed channels addressed with group power control

FIG. 6

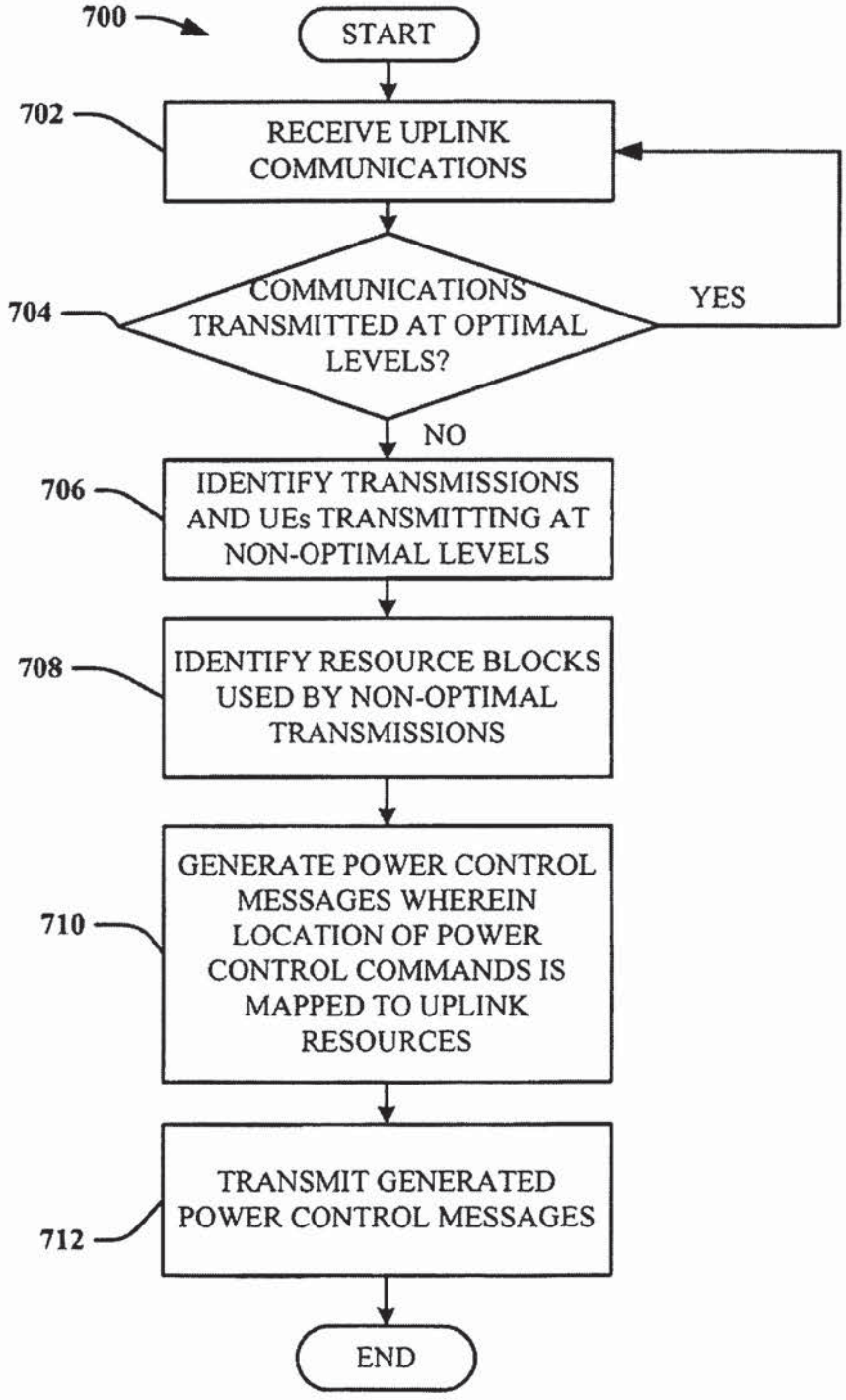

700
START
702
RECEIVE UPLINK COMMUNICATIONS
704
COMMUNICATIONS TRANSMITTED AT OPTIMAL LEVELS?
YES
NO
706
IDENTIFY TRANSMISSIONS AND UEs TRANSMITTING AT NON-OPTIMAL LEVELS
708
IDENTIFY RESOURCE BLOCKS USED BY NON-OPTIMAL TRANSMISSIONS
710
GENERATE POWER CONTROL MESSAGES WHEREIN LOCATION OF POWER CONTROL COMMANDS IS MAPPED TO UPLINK RESOURCES
712
TRANSMIT GENERATED POWER CONTROL MESSAGES
END

FIG. 7

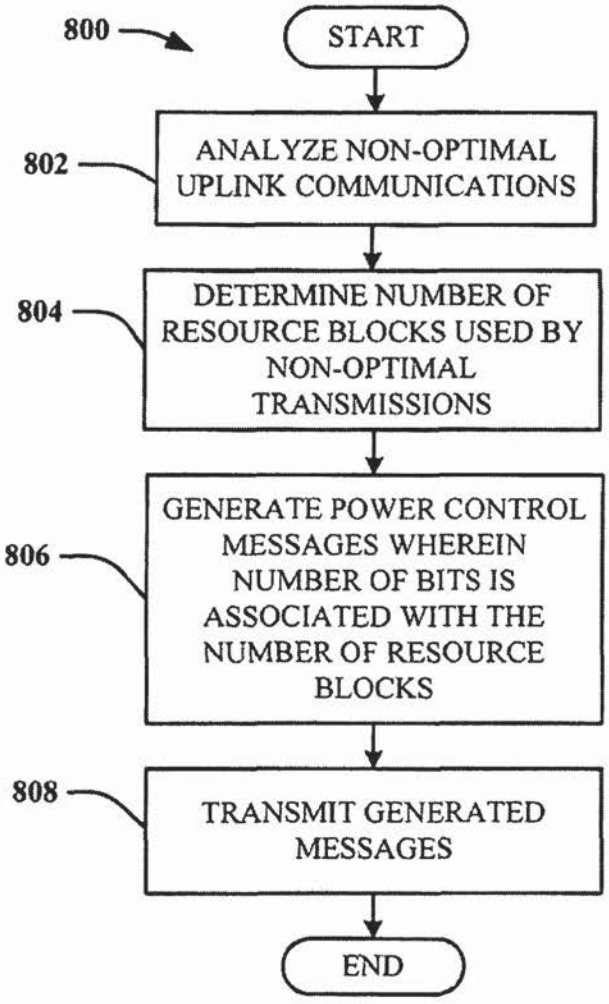

800
START
802
ANALYZE NON-OPTIMAL UPLINK COMMUNICATIONS
804
DETERMINE NUMBER OF RESOURCE BLOCKS USED BY NON-OPTIMAL TRANSMISSIONS
806
GENERATE POWER CONTROL MESSAGES WHEREIN NUMBER OF BITS IS ASSOCIATED WITH THE NUMBER OF RESOURCE BLOCKS
808
TRANSMIT GENERATED MESSAGES
END

FIG. 8

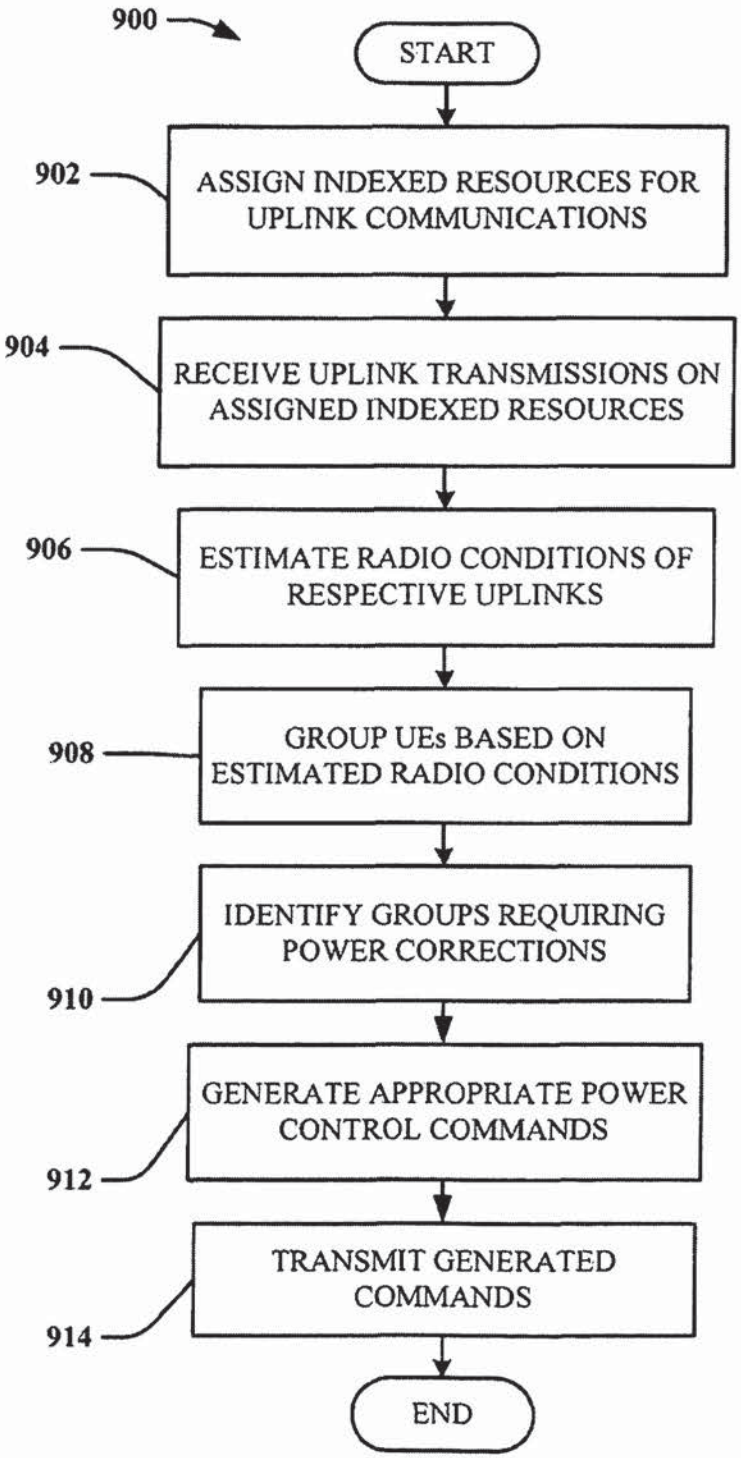

900
START
902
ASSIGN INDEXED RESOURCES FOR UPLINK COMMUNICATIONS
904
RECEIVE UPLINK TRANSMISSIONS ON ASSIGNED INDEXED RESOURCES
906
ESTIMATE RADIO CONDITIONS OF RESPECTIVE UPLINKS
908
GROUP UEs BASED ON ESTIMATED RADIO CONDITIONS
910
IDENTIFY GROUPS REQUIRING POWER CORRECTIONS
912
GENERATE APPROPRIATE POWER CONTROL COMMANDS
914
TRANSMIT GENERATED COMMANDS
END

FIG. 9

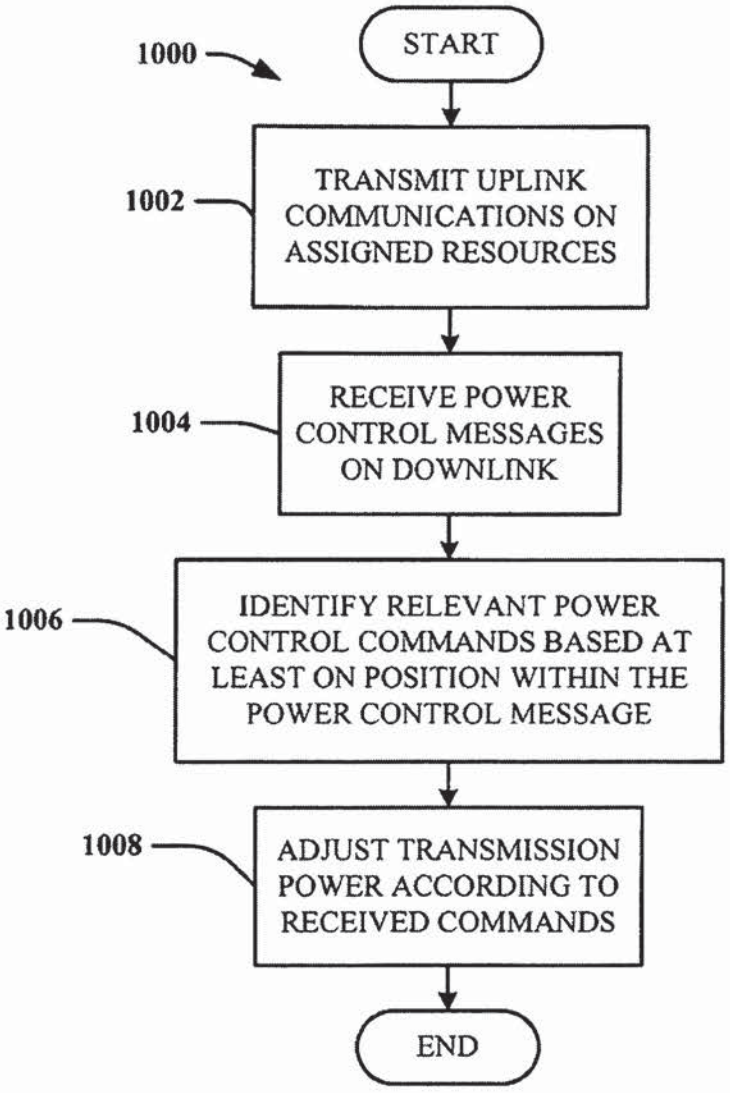

1000
START
1002
TRANSMIT UPLINK COMMUNICATIONS ON ASSIGNED RESOURCES
1004
RECEIVE POWER CONTROL MESSAGES ON DOWNLINK
1006
IDENTIFY RELEVANT POWER CONTROL COMMANDS BASED AT LEAST ON POSITION WITHIN THE POWER CONTROL MESSAGE
1008
ADJUST TRANSMISSION POWER ACCORDING TO RECEIVED COMMANDS
END

FIG. 10

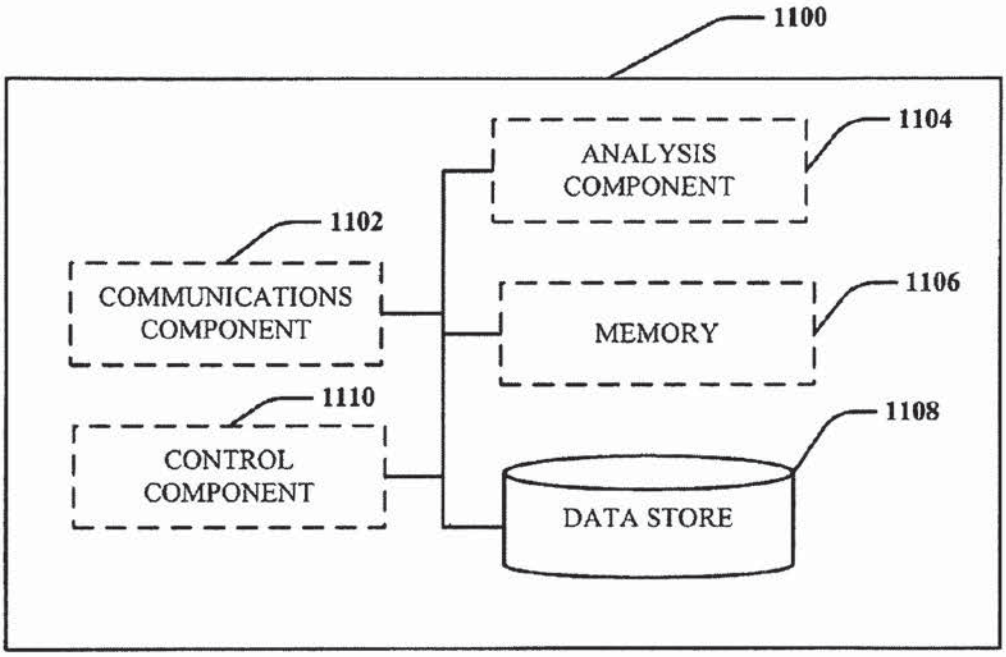

1100
1104
ANALYSIS COMPONENT
1102
COMMUNICATIONS COMPONENT
1106
MEMORY
1110
CONTROL COMPONENT
1108
DATA STORE

FIG. 11

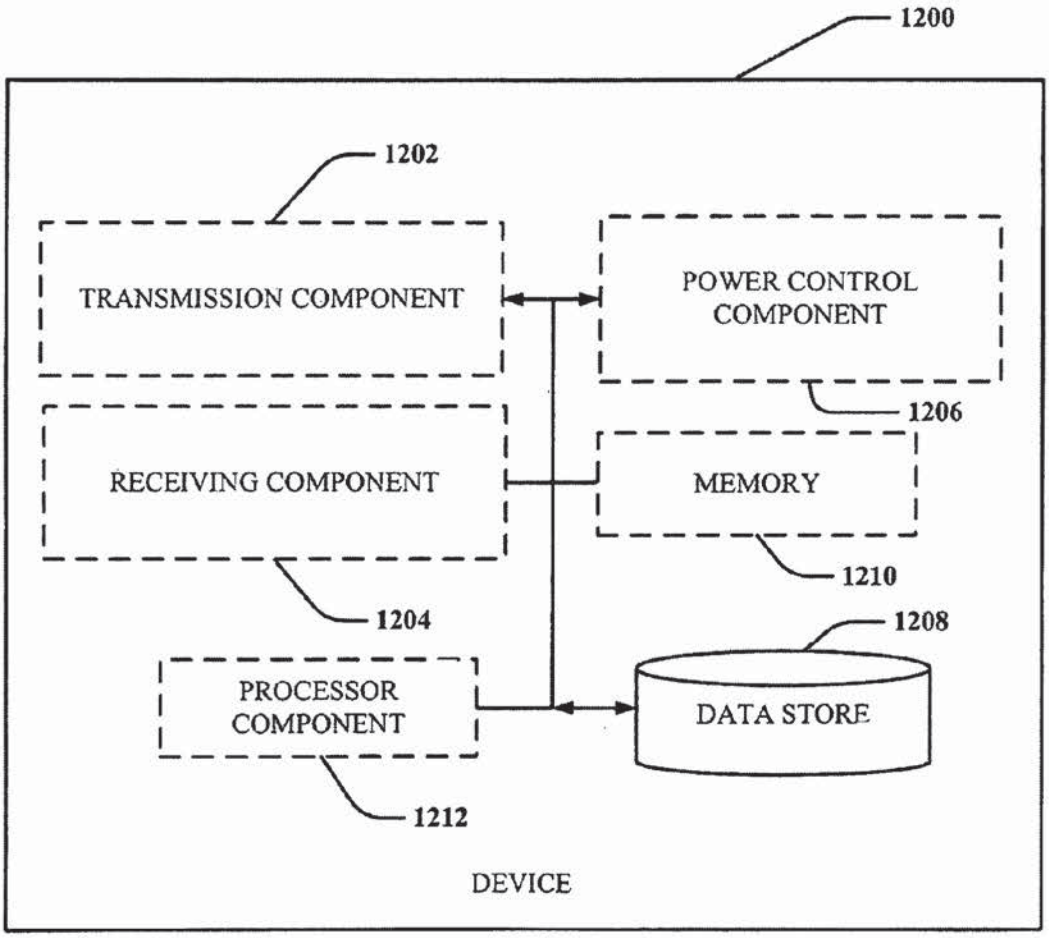

1200
1202
TRANSMISSION COMPONENT
POWER CONTROL COMPONENT
1206
RECEIVING COMPONENT
MEMORY
1210
1204
1208
PROCESSOR COMPONENT
DATA STORE
1212
DEVICE

FIG. 12

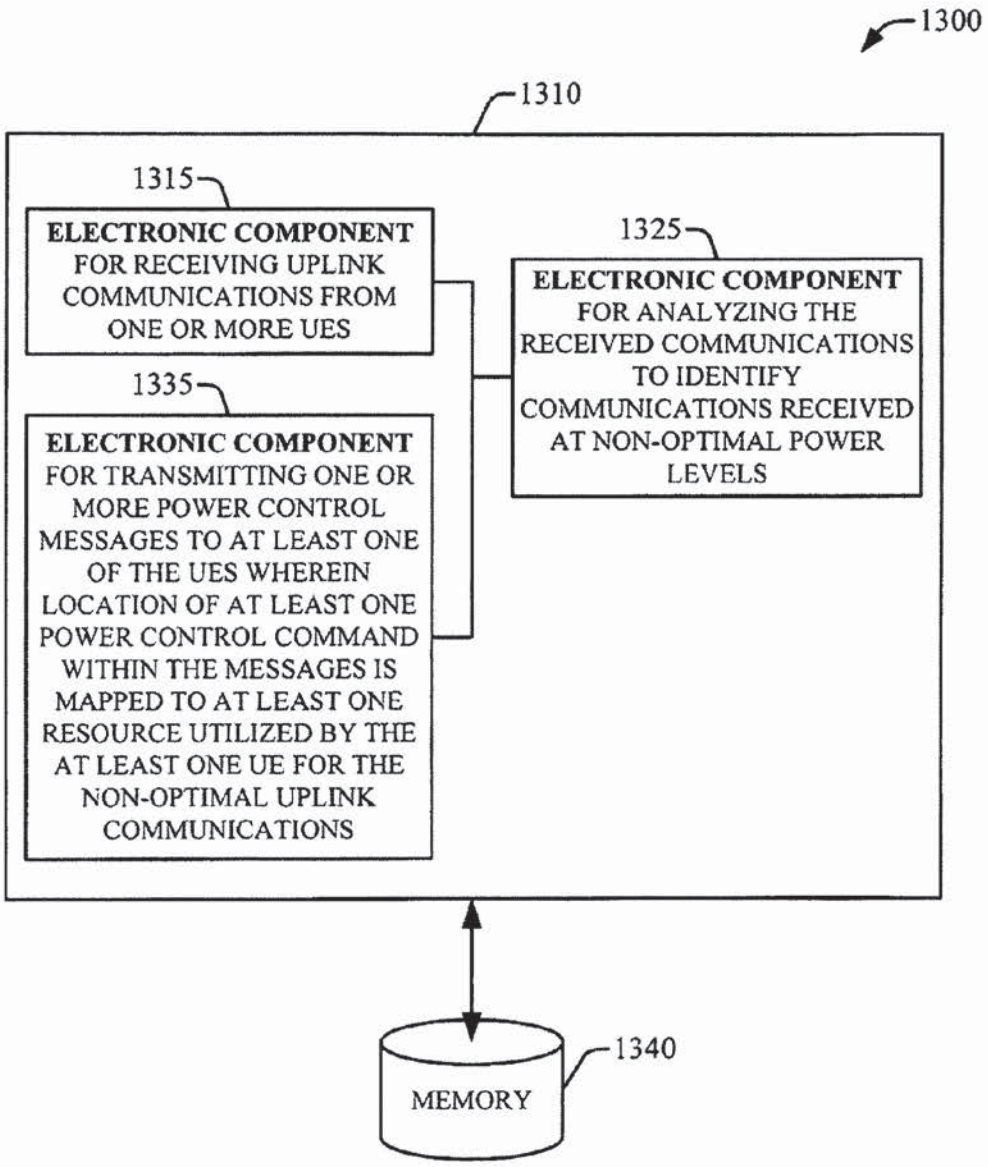

1300
1310
1315
ELECTRONIC COMPONENT FOR RECEIVING UPLINK COMMUNICATIONS FROM ONE OR MORE UES
1325
ELECTRONIC COMPONENT FOR ANALYZING THE RECEIVED COMMUNICATIONS TO IDENTIFY COMMUNICATIONS RECEIVED AT NON-OPTIMAL POWER LEVELS
1335
ELECTRONIC COMPONENT FOR TRANSMITTING ONE OR MORE POWER CONTROL MESSAGES TO AT LEAST ONE OF THE UES WHEREIN LOCATION OF AT LEAST ONE POWER CONTROL COMMAND WITHIN THE MESSAGES IS MAPPED TO AT LEAST ONE RESOURCE UTILIZED BY THE AT LEAST ONE UE FOR THE NON-OPTIMAL UPLINK COMMUNICATIONS
1340
MEMORY

FIG. 13

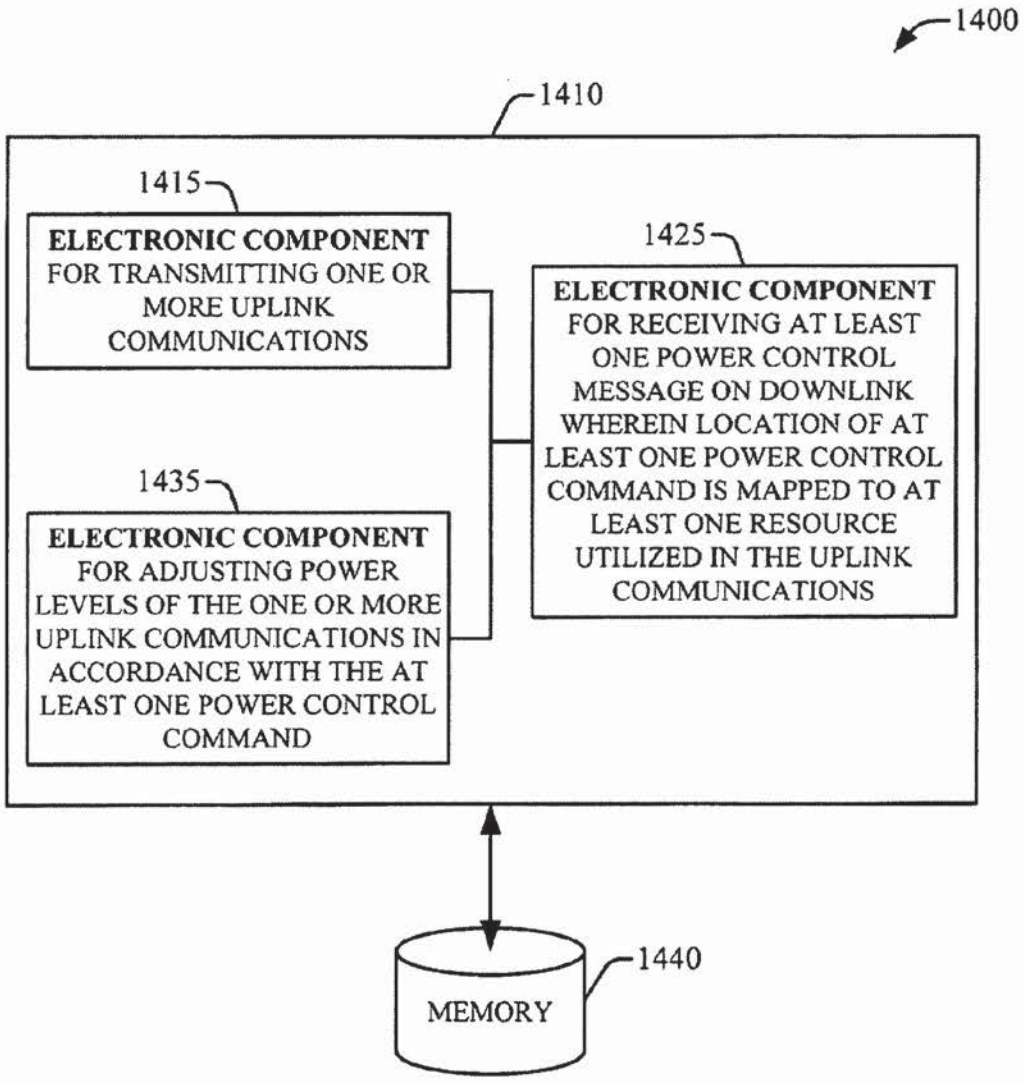

1400
1410
1415
ELECTRONIC COMPONENT FOR TRANSMITTING ONE OR MORE UPLINK COMMUNICATIONS
1425
ELECTRONIC COMPONENT FOR RECEIVING AT LEAST ONE POWER CONTROL MESSAGE ON DOWNLINK WHEREIN LOCATION OF AT LEAST ONE POWER CONTROL COMMAND IS MAPPED TO AT LEAST ONE RESOURCE UTILIZED IN THE UPLINK COMMUNICATIONS
1435
ELECTRONIC COMPONENT FOR ADJUSTING POWER LEVELS OF THE ONE OR MORE UPLINK COMMUNICATIONS IN ACCORDANCE WITH THE AT LEAST ONE POWER CONTROL COMMAND
1440
MEMORY

FIG. 14

METHODS AND APPARATUSES FOR POWER CONTROL

CROSS-REFERENCE TO RELATED APPLICATION(S)

This application claims the benefit of U.S. Provisional Patent Application Ser. No. 60/945,325 filed on Jun. 20, 2007 and entitled "METHODS AND APPARATUSES FOR POWER CONTROL," the entirety of which is incorporated herein by reference.

BACKGROUND OF THE INVENTION

Wireless communication systems are widely deployed to provide various types of communications such as voice, data, videos etc. These systems may be multiple-access systems capable of supporting communication with multiple access terminals by sharing available system resources (e.g., bandwidth and transmit power). Examples of such multiple-access systems include code division multiple access (CDMA) systems, time division multiple access (TDMA) systems, frequency division multiple access (FDMA) systems, 3GPP Long Term Evolution (LTE) systems, and orthogonal frequency division multiple access (OFDMA) systems. Typically, a wireless communication system comprises several base stations, wherein each base station communicates with a mobile station using a forward link and each mobile station (or access terminal) communicates with base station(s) using a reverse link.

Generally, a wireless multiple-access communication system can simultaneously support communication for multiple wireless terminals. Each terminal communicates with one or more base stations via transmissions on the forward and reverse links. The forward link (or downlink) refers to the communication link from the base stations to the terminals, and the reverse link (or uplink) refers to the communication link from the terminals to the base stations. This communication link may be established via a single-in-single-out (SISO), multiple-in-signal-out (MISO) or a multiple-in-multiple-out (MIMO) system.

A MIMO system employs multiple (NT) transmit antennas and multiple (NR) receive antennas for data transmission. A MIMO channel formed by the NT transmit and NR receive antennas may be decomposed into NS independent channels, which are also referred to as spatial channels, where $N_S \leqq \min\{N_T, N_R\}$. Each of the NS independent channels corresponds to a dimension. The MIMO system can provide improved performance (e.g., higher throughput and/or greater reliability) if the additional dimensionalities created by the multiple transmit and receive antennas are utilized.

A MIMO system supports a time division duplex (TDD) and frequency division duplex (FDD) systems. In a TDD system, the forward and reverse link transmissions are on the same frequency region so that the reciprocity principle allows the estimation of the forward link channel from the reverse link channel. This enables the eNode B to extract transmit beamforming gain on the forward link when multiple antennas are available at the eNode B.

When a user generates data for transmission, a service request is conveyed to the base station or an eNode B via established communication channels and in turn the base station assigns resources contingent upon time, bandwidth requirements/availability or service options of the user. The access terminal of the user or user equipment (UE) accordingly utilizes the assigned resources for communicating with the eNode B via the uplink. However, various factors such as

power constraints, signal attenuation, obstacles, etc. can lead to fluctuations in transmission from the UE. As a result, the eNode B transmits power control signals to the UE based on the transmissions received from the UE thereby facilitating the UE to transmit with optimal power. However, the resources assigned for transmitting power control commands are assigned statically so that even if a UE is not transmitting or transmitting at optimum power levels, resources for power control commands, remain assigned thereby leading to a non-optimal use of resources within the wireless communication system.

SUMMARY OF THE INVENTION

The following presents a simplified summary of the claimed subject matter in order to provide a basic understanding of some aspects of the claimed subject matter. This summary is not an extensive overview of the claimed subject matter. It is intended to neither identify key or critical elements of the claimed subject matter nor delineate, the scope of the claimed subject matter. Its sole purpose is to present some concepts of the claimed subject matter in a simplified form as a prelude to the more detailed description that is presented later.

In accordance with an aspect, a method for controlling power within a wireless communication system is disclosed. Upon receiving one or more uplink communications it is determined if any of the received communications have been transmitted at non-optimum power levels. If yes, then the UEs transmitting the communications at non-optimum power levels and the resources employed by the UEs for the non-optimal uplink communications are identified. Accordingly, power control messages for such UEs are generated. In a more detailed aspect, position of commands for controlling power within the messages is mapped to the resources employed by the respective UEs. The power control messages thus generated are subsequently transmitted to the UEs.

In a further aspect, power control messages can comprise a bitmap that communicates the power control commands wherein the bits within the bitmap represent the power control commands and location of the bits within the bitmap is mapped to the resources employed by the respective UEs. The number of bits for controlling power of an UE is associated with number of resources utilized by the UE for its uplink communications. In accordance with different aspects, a common power control message can be transmitted to a plurality of UEs or alternatively individual power control messages comprising the power control bits can be sent to each of the UEs separately. If at least one of the UEs is transmitting a plurality of data streams at different frequencies then the power control messages can individually adjust power of one or more resource blocks utilized by the data streams being transmitted at non-optimum power levels.

In another aspect, the resources assigned to the UEs for their uplink communications can be indexed. The power control messages transmitted to the UEs can comprise the indexing information in addition to the power control commands. This facilitates grouping UEs based at least on their radio conditions so that a group power control message can be transmitted to UEs having similar radio conditions. This facilitates dynamic signaling wherein only those UEs that require power correction receive power control commands while no resources are wasted in sending redundant communications to UEs already transmitting at optimum power levels.

Another aspect relates to an apparatus for conveying power control information within a communication system. The

apparatus comprises a receiver that receives one or more communications from one or more UEs. A processor within the apparatus generates power control messages for those communications transmitted at non-optimum power levels. The messages are created such that the location power control commands within them are mapped to the resources utilized by the uplink communications. A transmitter subsequently transmits the power control messages to the UEs. The apparatus also comprises a memory that stores information associated with the power control commands or the resources.

In different aspects, a power control message can comprise a plurality of power control commands associated with a plurality of UEs or a plurality of power control messages can be transmitted separately to each of the plurality of UEs. Location of the power control commands within the power control message is mapped to resources utilized by the UEs for their uplink communications. Another aspect relates to UEs utilizing a plurality of resources for transmitting on the uplink. In this aspect, the power control message can comprise a plurality of power control commands whose position within the power control message is mapped to the plurality of resources utilized by a UE. Yet another aspect is associated with UEs transmitting a plurality of the communications at a plurality of frequencies. In accordance with this aspect, the processor generates a power control message for the plurality of communications received from a UE wherein the power control message comprises power control commands for one or more of the communications being transmitted at different frequencies.

The apparatus also facilitates a more dynamic signaling wherein the resources utilized by the UEs are indexed and the power control messages can comprise indexing information associated with the resources along with the power control commands. This facilitates grouping UEs based at least on their respective radio conditions and transmitting a common group power control messages to only those UEs requiring power corrections.

In another aspect, the subject innovation discloses a computer program product comprising a computer-readable medium comprising: code for causing at least a computer to receive one or more uplink communications; code for causing at least a computer to determine if the communications have been transmitted at optimum power levels; code for causing at least a computer to determine one or more resources employed by one or more UEs for the non-optimal uplink communications for uplink communications transmitted at non-optimal power levels; code for causing at least a computer to generate one or more power control messages for the UEs wherein position of control commands for controlling power is mapped to the one or more resources employed by the respective UEs; and code for causing at least a computer to transmit the power, control messages to the one or more UEs.

A apparatus for communicating power control information is disclosed in accordance with yet another aspect. The apparatus comprises a means for receiving uplink communications from UEs. Means for analyzing the received communications, also comprised within the system, identify communications received at non-optimal power levels. Accordingly, means for transmitting transmit power control messages to the UEs wherein location of at least one power control command within the messages is mapped to at least one resource utilized by the at least one UE for the non-optimal uplink communications.

In yet another aspect, a method of receiving power control commands is disclosed. The method comprises transmitting uplink communications, and receiving at least one power control message on the downlink. The power control message comprises at least one power control command wherein location of the power control command is mapped to the resources utilized in the uplink communications. Based on its location within the message, the power control command is identified and power levels of the one or more uplink communications are corrected in accordance with the received power control command.

An apparatus for adjusting transmission power levels within a communication system is disclosed in accordance with this aspect. The apparatus comprises a transmitter that transmits one or more uplink communications. Some of the uplink communication may be transmitted at non-optimal power levels. Accordingly, a receiver receives at least one power control message wherein location of at least one power control command within the at least one power control message is mapped to one or more resources utilized by the non-optimal uplink communications. A processor also comprised within the apparatus controls power levels of the transmitter based on the received power control command.

Another aspect is associated with a computer program product comprising a computer-readable medium comprising: code for causing at least a computer to transmit one or more uplink communications; code for causing at least a computer to receive at least one power control message on downlink wherein location of at least one power control command is mapped to at least one resource utilized in the uplink communications; and code for causing at least a computer to adjust power levels of the one or more uplink communications in accordance with the at least one power control command.

An apparatus for adjusting power within a wireless communication system is disclosed in accordance with this aspect. The apparatus comprises means for transmitting, means for receiving and means for adjusting power levels. The means for transmitting transmits one or more uplink communications. Means for receiving facilitates receiving at least one power control message on downlink wherein location of at least one power control command is mapped to at least one resource utilized in the uplink communications. Accordingly, means for adjusting corrects power levels of the one or more uplink communications in accordance with the at least one power control command.

The following description and the annexed drawings set forth in detail certain illustrative aspects of the claimed subject matter. These aspects are indicative, however, of but a few of the various ways in which the principles of the claimed subject matter may be employed and the claimed subject matter is intended to include all such aspects and their equivalents. Other advantages and distinguishing features of the claimed subject matter will become apparent from the following detailed description of the claimed subject matter when considered in conjunction with the drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** illustrates a multiple access wireless communication system according to one embodiment.

FIG. **2** is a block diagram of an embodiment of an eNode B and an access terminal (or a UE) in a MIMO system.

FIG. **3** is an illustration of a wireless multiple-access communication system in accordance with various aspects described herein.

FIG. **4** is a flow chart detailing a closed loop power control of UE uplink transmissions.

FIG. **5**A illustrates a schematic diagram of a power control message transmitted to a plurality of access terminals that receive power control commands in accordance with an aspect.

FIG. **5**B relates to another aspect wherein more than one resource block is employed by the access terminal for transmission of its uplink communications.

FIG. **5**C relates to yet another aspect wherein different access terminals employ varying number of resource blocks for their uplink communications.

FIG. **5**D illustrates another aspect related to transmitting power control message when an UE is transmitting a plurality of data streams in accordance with a MC-FDMA scheme.

FIG. **6** relates to a methodology of transmitting power control message to one or more access terminals in accordance with an aspect.

FIG. **7** relates to a methodology of transmitting power control message to one or more UEs in accordance with an aspect.

FIG. **8** relates to a more detailed aspect of generating power control messages for various UEs.

FIG. **9** is associated with a methodology of transmitting power control commands using indexing as described herein.

FIG. **10** is a flow chart detailing a method of adjusting power at the UEs based on the received commands.

FIG. **11** illustrates a high-level system diagram of various components of a device in accordance with various aspects.

FIG. **12** is another high level diagram illustrating various components of a device in accordance with different aspects described herein.

FIG. **13** illustrates a block diagram of an example system that enables communication of power control information in accordance with aspects disclosed in the subject specification.

FIG. **14** illustrates a block diagram of an example system that enables adjustment of power within a wireless communication system in accordance with aspect described in the subject specification.

DESCRIPTION OF THE INVENTION

The claimed subject matter is now described with reference to the drawings, wherein like reference numerals are used to refer to like elements throughout. In the following description, for purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the claimed subject matter. It may be evident, however, that the claimed subject matter may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form in order to facilitate describing the claimed subject matter.

Various embodiments are now described with reference to the drawings, wherein like reference numerals are used to refer to like elements throughout. In the following description, for purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of one or more aspects. It may be evident however, that such embodiment(s) may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form in order to facilitate describing one or more embodiments. As used in this application, the terms "component," "module," "system," and the like are intended to refer to a computer-related entity, either hardware, firmware, a combination of hardware and software, software, or software in execution. For example, a component may be, but is not limited to being, a process running on a processor, a processor, an integrated circuit, an object, an executable, a thread of execution, a program, and/or a computer. By way of illustration, both an application running on a computing device and the computing device can be a component. One or more components can reside within a process and/or thread of execution and a component may be localized on one computer and/or distributed between two or more computers. In addition, these components can execute from various computer readable media having various data structures stored thereon. The components may communicate by way of local and/or remote processes such as in accordance with a signal having one or more data packets (e.g., data from one component interacting with another component in a local system, distributed system, and/or across a network such as the Internet with other systems by way of the signal).

Various embodiments will be presented in terms of systems that may include a number of devices, components, modules, and the like. It is to be understood and appreciated that the various systems may include additional devices, components, modules, etc. and/or may not include all of the devices, components, modules etc. discussed in connection with the figures. A combination of these approaches may also be used.

The word "exemplary" is used herein to mean "serving as an example, instance, or illustration." Any embodiment or design described herein as "exemplary" is not necessarily to be construed as preferred or advantageous over other embodiments or designs. The word "listening" is used herein to mean that a recipient device (eNode B or UE) is receiving and processing data received on a given channel.

Various aspects can incorporate inference schemes and/or techniques in connection with transitioning communication resources. As used herein, the term "inference" refers generally to the process of reasoning about or inferring states of the system, environment, and/or user from a set of observations as captured via events and/or data. Inference can be employed to identify a specific context or action, or can generate a probability distribution over states, for example. The inference can be probabilistic—that is, the computation of a probability distribution over states of interest based on a consideration of data and events, or decision theoretic, building upon probabilistic inference, and considering display actions of highest expected utility, in the context of uncertainty in user goals and intentions. Inference can also refer to techniques employed for composing higher-level events from a set of events and/or data. Such inference results in the construction of new events or actions from a set of observed events and/or stored event data, whether or not the events are correlated in close temporal proximity, and whether the events and data come from one or several event and data sources.

Furthermore, various aspects are described herein in connection with a subscriber station. A subscriber station can also be called a system, a subscriber unit, mobile station, mobile, remote station, access point, eNode B, remote terminal, access terminal, user terminal, user agent, a user device, mobile device, portable communications device, or user equipment. A subscriber station may be a cellular telephone, a cordless telephone, a Session Initiation Protocol (SIP) phone, a wireless local loop (WLL) station, a personal digital assistant (PDA), a handheld device having wireless connection capability, or other processing device connected to a wireless modem.

Moreover, various aspects or features described herein may be implemented as a method, apparatus, or article of manufacture using standard programming and/or engineering techniques. The term "article of manufacture" as used herein is intended to encompass a computer program accessible from any computer-readable device, carrier, or media.

For example, computer-readable media can include but are not limited to magnetic storage devices (e.g., hard disk, floppy disk, magnetic strips . . .), optical disks (e.g., compact disk (CD), digital versatile disk (DVD) . . .), smart cards, and flash memory devices (e.g., card, stick, key drive . . .). Additionally, various storage media described herein can represent one or more devices and/or other machine-readable media for storing information. The term "machine-readable medium" can include, without being limited to, wireless channels and various other media capable of storing, containing, and/or carrying instruction(s) and/or data.

The techniques described herein may be used for various wireless communication networks such as Code Division Multiple Access (CDMA) networks, Time Division Multiple Access (TDMA) networks, Frequency Division Multiple Access (FDMA) networks, Orthogonal FDMA (OFDMA) networks, Single-Carrier FDMA (SC-FDMA) networks, etc. The terms "networks" and "systems" are often used interchangeably. A CDMA network may implement a radio technology such as Universal Terrestrial Radio Access (UTRA), cdma2000, etc. UTRA includes Wideband-CDMA (W-CDMA) and Low Chip Rate (LCR). cdma2000 covers IS-2000, IS-95 and IS-856 standards. A TDMA network may implement a radio technology such as Global System for Mobile Communications (GSM). An OFDMA network may implement a radio technology such as Evolved UTRA (E-UTRA), IEEE 802.11, IEEE 802.16, IEEE 802.20, Flash-OFDM®, etc. UTRA, E-UTRA, and GSM are part of Universal Mobile Telecommunication System (UMTS). Long Term Evolution (LTE) is an upcoming release of UMTS, that uses E-UTRA. UTRA, E-UTRA, GSM, UMTS and LTE are described in documents from an organization named "3rd Generation Partnership Project" (3GPP). cdma2000 is described in documents from an organization named "3rd Generation Partnership Project 2" (3GPP2). These various radio technologies and standards are known in the art. For clarity, certain aspects of the techniques are described below for LTE, and LTE terminology is used in much of the description below.

Single carrier frequency division multiple access (SC-FDMA), which utilizes single carrier modulation and frequency domain equalization is a technique. SC-FDMA has similar performance and essentially the same overall complexity as those of OFDMA system. SC-FDMA signal has lower peak-to-average power ratio (PAPR) because of its inherent single carrier structure. SC-FDMA has drawn great attention, especially in the uplink communications where lower PAPR greatly benefits the mobile terminal in terms of transmit power efficiency. It is currently a working assumption for uplink multiple access scheme in 3GPP Long Term Evolution (LTE), or Evolved UTRA.

Referring to FIG. **1**, a multiple access wireless communication system according to one embodiment is illustrated. An eNode B **100** includes multiple antenna groups, wherein a first group includes antennas **104** and **106**, another includes **108** and **110**, and an additional group includes **112** and **114**. In FIG. **1**, only two antennas are shown for each antenna group, however, more or fewer antennas may be utilized for each antenna group. UE (user equipment) or AT (access terminal) **116** is in communication with antennas **112** and **114**, where antennas **112** and **114** transmit information to UE **116** over forward link **120** and receive information from UE **116** over reverse link **118**. UE **122** is in communication with antennas **106** and **108**, where antennas **106** and **108** transmit information to UE **122** over forward link **126** and receive information from UE **122** over reverse link **124**. In a FDD system, communication links **118**, **120**, **124** and **126** may use different frequencies for communication. For example, forward link **120** may use a different frequency than that used by reverse link **118**. Each group of antennas and/or the area in which they are designed to communicate is often referred to as a sector of the access point or eNode B. In the embodiment, antenna groups are each designed to communicate to UEs in a sector within the areas covered by eNode B **100**.

In communication over forward links **120** and **126**, the transmitting antennas of eNode B **100** utilize beamforming in order to improve the signal-to-noise ratio of forward links for the different UEs **116** and **124**. Also, an eNode B using beamforming to transmit to UEs scattered randomly through its coverage area causes less interference to UEs in neighboring cells than an eNode B transmitting through a single antenna to all its UEs.

An eNode B may be a fixed station used for communicating with the terminals and may also be referred to as an access point, a Node B, an enhanced Node B (eNode B) or some other terminology. An access terminal (AT) may also be called a user equipment (UE), a wireless communication device, terminal, or some other terminology.

FIG. **2** is a block diagram of an embodiment of an eNode B **210** and a access terminal (AT) or user equipment (UE) **250** in a MIMO system **200**. At the eNode B **210**, traffic data for a number of data streams is provided from a data source **212** to a transmit (TX) data processor **214**.

In an embodiment, each data stream is transmitted over a respective transmit antenna. TX data processor **214** formats, codes, and interleaves the traffic data for each data stream based on a particular coding scheme selected for that data stream to provide coded data.

The coded data for each data stream may be multiplexed with pilot data using OFDM techniques. The pilot data is typically a known data pattern that is processed in a known manner and may be used at the receiver system to estimate the channel response. The multiplexed pilot and coded data for each data stream is then modulated (i.e., symbol mapped) based on a particular modulation scheme (e.g., BPSK, QSPK, M-PSK, or M-QAM) selected for that data stream to provide modulation symbols. The data rate, coding, and modulation for each data stream may be determined by instructions performed by processor **230**.

The modulation symbols for all data streams are then provided to a TX MIMO processor **220**, which may further process the modulation symbols (e.g., for OFDM). TX MIMO processor **220** then provides NT modulation symbol streams to NT transmitters (TMTR) **222***a* through **222***t*. In certain embodiments, TX MIMO processor **220** applies beamforming weights to the symbols of the data streams and to the antenna from which the symbol is being transmitted.

Each transmitter **222** receives and processes a respective symbol stream to provide one or more analog signals, and further conditions (e.g., amplifies, filters, and upconverts) the analog signals to provide a modulated signal suitable for transmission over the MIMO channel. NT modulated signals from transmitters **222***a* through **222***t* are then transmitted from NT antennas **224***a* through **224***t*, respectively.

At the UE **250**, the transmitted modulated signals are received by NR antennas **252***a* through **252***r* and the received signal from each antenna **252** is provided to a respective receiver (RCVR) **254***a* through **254***r*. Each receiver **254** conditions (e.g., filters, amplifies, and downconverts) a respective received signal, digitizes the conditioned signal to provide samples, and further processes the samples to provide a corresponding "received" symbol stream.

An RX data processor **260** then receives and processes the NR received symbol streams from NR receivers **254** based on

a particular receiver processing technique to provide NT "detected" symbol streams. The received symbols or other information can be stored in an associated memory **272**. The RX data processor **260** then demodulates, deinterleaves, and decodes each detected symbol stream to recover the traffic data for the data stream. The processing by RX data processor **260** is complementary to that performed by TX MIMO processor **220** and TX data processor **214** at the eNode B **210**.

A processor **270** periodically determines which pre-coding matrix to use (discussed below). Processor **270** formulates a reverse link message comprising a matrix index portion and a rank value portion.

The reverse link message may comprise various types of information regarding the communication link and/or the received data stream. Information received on the reverse link can be stored in an associated memory **232**. The reverse link message is then processed by a TX data processor **238**, which also receives traffic data for a number of data streams from a data source **236**, modulated by a modulator **280**, conditioned by transmitters **254***a* through **254***r*, and transmitted back to transmitter system **210**.

At the eNode B **210**, the modulated signals from receiver system **250** are received by antennas **224**, conditioned by receivers **222**, demodulated by a demodulator **240**, and processed by a RX data, processor **242** to extract the reserve link message transmitted by the receiver system **250**. Processor **230** then determines which pre-coding matrix to use for determining the beamforming weights then processes the extracted message.

In an aspect, logical channels are classified into Control Channels and Traffic Channels. Logical Control Channels comprises Broadcast Control Channel (BCCH) which is DL channel for broadcasting system control information. Paging Control Channel (PCCH) which is DL channel that transfers paging information. Multicast Control Channel (MCCH) which is Point-to-multipoint DL channel used for transmitting Multimedia Broadcast and Multicast Service (MBMS) scheduling and control information for one or several MTCHs. Generally, after establishing RRC connection this channel is only used by UEs that receive MBMS. Dedicated Control Channel (DCCH) is Point-to-point bi-directional channel that transmits dedicated control information and used by UEs having an RRC connection. In aspect, Logical Traffic Channels comprises a Dedicated Traffic Channel (DTCH) which is Point-to-point bi-directional channel, dedicated to one UE, for the transfer of user information. Also, a Multicast Traffic Channel (MTCH) for Point-to-multipoint DL channel for transmitting traffic data.

In an aspect, Transport Channels are classified into DL and UL. DL Transport Channels comprises a Broadcast Channel (BCH), Downlink Shared Channel (DL-SCH) and a Paging Channel (PCH), the PCH for support, of UE power saving (DRX cycle is indicated by the network to the UE), broadcasted over entire cell and mapped to PHY resources which can be used for other control/traffic channels. DL transport channel associated with MBMS is Multicast Channel (MCH) The UL Transport Channels comprises a Random Access Channel (RACH), Uplink Shared Data Channel (UL-SDCH) and pluarlity of PHY channels. The PHY channels comprises a set of DL channels and UL channels.

The DL PHY channels and signals comprises:
Reference signal (RS)
Primary and Secondary Synchronization Signals (PSS/SSS)
Physical Downlink Shared Channel (PDSCH)
Physical Downlink Control Channel (PDCCH)
Physical Multicast Channel (PMCH)
Physical HARQ Indicator Channel (PHICH)
Physical Control Format Indicator Channel (PCFICH)

The UL PHY Channels comprises:
Physical Random Access Channel (PRACH)
Physical Uplink Control Channel (PUCCH)
Channel Quality Indicator (CQI)
Precoding Matrix Indicator (PMI)
Rank Indicator (RI)
Scheduling request (SR)
Uplink ACK/NAK
Physical Uplink Shared Channel (PUSCH)
Sounding Reference Signal (SRS)

In an aspect, a channel structure is provided that preserves low PAR (at any given time, the channel is contiguous or uniformly spaced in frequency) properties of a single carrier waveform.

For the purposes of the present document, the following abbreviations apply:

AM	Acknowledged Mode
AMD	Acknowledged Mode Data
ARQ	Automatic Repeat Request
BCCH	Broadcast Control CHannel
BCH	Broadcast CHannel
C-	Control-
CCCH	Common Control CHannel
CCH	Control CHannel
CCTrCH	Coded Composite Transport Channel
CP	Cyclic Prefix
CRC	Cyclic Redundancy Check
CTCH	Common Traffic CHannel
DCCH	Dedicated Control CHannel
DCH	Dedicated CHannel
DL	DownLink
DSCH	Downlink Shared CHannel
DTCH	Dedicated Traffic CHannel
FACH	Forward link Access CHannel
FDD	Frequency Division Duplex
L1	Layer 1 (physical layer)
L2	Layer 2 (data link layer)
L3	Layer 3 (network layer)
LI	Length Indicator
LSB	Least Significant Bit
MAC	Medium Access Control
MBMS	Multimedia Broadcast Multicast Service
MCCH	MBMS point-to-multipoint Control CHannel
MRW	Move Receiving Window
MSB	Most Significant Bit
MSCH	MBMS point-to-multipoint Scheduling CHannel
MTCH	MBMS point-to-multipoint Traffic CHannel
PCCH	Paging Control CHannel
PCH	Paging CHannel
PDU	Protocol Data Unit
PHY	PHYsical layer
PhyCH	Physical CHannels
RACH	Random Access CHannel
RLC	Radio Link Control
RRC	Radio Resource Control
SAP	Service Access Point
SDU	Service Data Unit
SN	Sequence Number
SUFI	SUper FIeld
TCH	Traffic CHannel
TDD	Time Division Duplex
TFI	Transport Format Indicator
TM	Transparent Mode
TMD	Transparent Mode Data
TTI	Transmission Time Interval
U-	User-
UE	User Equipment
UL	UpLink
UM	Unacknowledged Mode
UMD	Unacknowledged Mode Data
UMTS	Universal Mobile Telecommunications System
UTRA	UMTS Terrestrial Radio Access
UTRAN	UMTS Terrestrial Radio Access Network
MBSFN	multicast broadcast single frequency network
MCE	MBMS coordinating entity

-continued

MCH	multicast channel
DL-SCH	downlink shared channel
MSCH	MBMS control channel
PDCCH	physical downlink control channel
PDSCH	physical downlink shared channel
MBSFN	multicast broadcast single frequency network
MCE	MBMS coordinating entity
MCH	multicast channel
DL-SCH	downlink shared channel
MSCH	MBMS control channel
PDCCH	physical downlink control channel
PDSCH	physical downlink shared channel
PUCCH	physical uplink control channel
PUSCH	physical uplink shared channel

FIG. **3** is an illustration of a wireless multiple-access communication system **300** in accordance with various aspects. In one example, the wireless multiple-access communication system **300** includes multiple eNode Bs **310** and multiple UEs **320**. Each eNode B **310** provides communication coverage for a particular geographic area **302** (e.g., **302***a*, **302***b*, **302***c*). The term "cell" can refer to an eNode B and/or its coverage area depending on the context in which the term is used. To improve system capacity, an access terminal coverage area may be partitioned into multiple smaller areas, e.g., three smaller areas **304***a*, **304***b*, and **304***c*. Each smaller area is served by a respective eNode B. The term "sector" can refer to an eNode B and/or its coverage, area depending on the context in which the term is used. For a sectorized, cell, the eNode Bs for all sectors of that cell are typically co-located within the base station for the cell. The signaling transmission techniques described herein may be used for a system with sectorized cells as well as a system with un-sectorized cells. For simplicity, in the following description, the term "base station" or eNode B is used generically for a station that serves a sector as well as a station that serves a cell.

Terminals or UEs **320** are typically dispersed throughout the system, and each UE may be fixed or mobile. A terminal may also be called, and may contain some or all of the functionality of, a mobile station, a user equipment, and/or some other device. A terminal may be a wireless device, a cellular phone, a personal digital assistant (PDA), a wireless modem card, and so on. A terminal may communicate with zero, one, or multiple base stations on the forward and reverse links at any given moment.

For a centralized architecture, a system controller **330** couples to APs **310** and provides coordination and control for these base stations. System controller **330** may be a single network entity or a collection of network entities. For a distributed architecture, the APs **310** may communicate with one another as needed.

One or more aspects of a wireless communication system design are described that support full & half duplex FDD (Frequency Division Duplex) and TDD (Time Division Duplex) modes of operation, with support for scalable bandwidth. However, this need not be the case, and other modes may also be supported, in addition to, or in lieu, of the previous, modes. Further, it should be noted that the concepts and approaches herein, need not be used in conjunction with any other of the concepts or approaches described herein.

As detailed supra, uplink power control can have considerable impact on the operation of a RN (radio network) as signals are affected by various factors such as location of the AT from an eNode B or atmospheric effects or other factors. The attenuation experienced by uplink signals due to these various factors can be compensated by constantly transmitting them at a high power level. However, this may result in undesirable consequences such as high level of interference between users or wastage of resources if at least some of the factors causing such attenuation are not present. As a result, it is desirable to control the power at which uplink signals are transmitted by an AT. This is generally achieved by transmitting power control commands to the AT by a serving base station through its control channel transmissions. The process of controlling power of uplink transmissions can either be an aperiodic message based procedure or it may occur periodically.

FIG. **4** is a flow chart **400** detailing a closed loop power control of uplink transmissions. The procedure begins at **402** wherein thresholds can be defined for one or more attributes associated with uplink transmissions received from an UE at a base station. For example, SNR (Signal-to-Noise Ratio) can be an attribute detected from the UE uplink communications which has a threshold defined for optimal transmissions. At **404**, uplink transmissions are received from an UE at the base station. These transmissions may relate to resource requests, data/voice transmissions or other communications. The received transmissions are analyzed at **406** in order to detect requisite attributes such as SNR at **406**. At **408**, it is determined if the uplink communications have optimal attributes. If yes, the base station continues to receive transmissions without affecting the power levels of the UE. If at **408** it is concluded that the transmissions received from the UE are non-optimal, at **410** the attributes are compared with the predefined threshold in order to determine an appropriate adjustment that should be made to the attributes to set them at optimal levels. Based at least on the conclusions derived at **410** appropriate power control messages are transmitted at **412**. For example, if one of the attributes measured is the SNR of the uplink signal, upon comparison with a predefined SNR threshold an appropriate adjustment of power level is determined in order to facilitate transmissions at the optimal SNR. Accordingly, if at **410** it is determined that the SNR is less than the predefined threshold it can be concluded that the signal should be transmitted at a higher power level. Therefore a power control command is sent to the UE to transmit the signals at higher power levels. If at **410**, it is determined that the SNR of the received signal is not less than the predefined threshold, it is concluded that the signal has an SNR greater than the threshold and accordingly an power control command is sent to the UE at **412** to decrease the power level of the uplink transmissions. It can be appreciated that the procedure detailed supra is just one methodology of controlling power at the UE. Other methods are possible for example, an UE can be programmed to adjust its own power level via measuring the attributes of a downlink signal etc.

Accordingly, resources are assigned by a serving base station to transmit power control commands to an UE in order to facilitate the terminal to transmit uplink communications at an optimal level. These resources are generally assigned statically per user (or per UE) and therefore remain assigned irrespective of whether or not the UE is transmitting. This leads to a non-optimal use of communication system resources. Various aspects discussed herein mitigate such wastage of resources by associating downlink power control commands to resources used for uplink transmissions thereby leading to a more dynamic assignment of resources for power control signaling.

FIG. **5**A illustrates a schematic diagram **500** of a power control message transmitted to a plurality of UEs I, II, . . . N that receive power control commands mapped in a manner that is associated with the resources on which the respective UEs transmit their uplink communications. In the figure, **502** is a control channel message carrying power control bits for

the terminals I, II, . . . N. In this embodiment, a control channel message **502** can comprise a message type field **504**, a bitmap **506**, and a cyclic redundancy check (CRC) field **508**. The message type field **504** carries a specific value to identify this message as a control channel message. Different messages may be assigned different message type values. The bitmap field **506** comprises a bitmap that carries power control bits for each of the associated UEs. Accordingly, the control channel message **502** is transmitted with N bits each bit carrying a power control message for each of the I, II . . . N UEs.

In a further aspect, the location of these power control bits in the bitmap is associated with the resources used by each of the I, II, . . . N UEs for their uplink transmissions. Accordingly, if UE I is transmitting uplink communications on resource block **1** (RB **1**), the power control bit meant for UE I would be the first bit **0** in the bitmap. Similarly, if UE II is transmitting uplink communications on resource block (RB) **2**, the second bit in the bitmap **504** would be the power control bit for the terminal II. In accordance with an aspect, a resource block can be made up of resource elements. For example, a resource block can be comprised of predetermined resource elements such as a specified number of consecutive SC FDMA symbols in time domain and a specific number of consecutive subcarriers in the frequency domain. Thus, if the UE N is transmitting on the Nth resource block, the Nth bit in the bitmap field would be the power control bit for the UE N. Hence, the location of the power control bits in the bitmap transmitted to the UEs on the downlink is mapped to the resources employed by the UEs for their uplink communications. These resources can be persistently assigned to respective UEs in accordance with a further aspect. If the resources are not persistently assigned, the power control bits can be included in uplink grants wherein **502** would accordingly be an assignment message. In addition to persistently scheduled resources, power control bitmap could also be utilized for implicitly scheduled retransmissions, where instead of UL grants, PHICH is used to negatively acknowledge PUSCH transmission and implicitly schedule retransmission on the same virtual resources. Although it has been illustrated that a single message comprising a single bit for controlling power of each terminal is transmitted simultaneously to all the terminals, it can be appreciated that this is not necessary. Separate control messages can be transmitted to each of the UEs wherein each message carries power control bits for individual UEs as seen in further detail infra.

FIG. **5**B relates to another aspect wherein more than one resource block is employed by a UE for transmission of its uplink communications. A UE I can employ two resource blocks (**1** and **2**) while UE II can employ resource blocks **3** and **4** (RB **3**, RB **4**) for their respective uplink transmissions. Accordingly, two bits are used for controlling transmission power of the UEs. In this figure, **512** is the control channel message including a message type field **514**, a bitmap **516**, and a cyclic redundancy check (CRC) field **518**. Employing more than 1 bit in the message permits the base station better control of power in terms of granularity that can be achieved as two bits can correspond to 2^2=4 levels of power. Similarly, depending on the number of resource blocks used by a terminal for its uplink transmission, a base station can use as many bits to manipulate power of a UE thereby achieving greater control.

FIG. **5**C relates to yet another aspect wherein different UEs employ varying number of resource blocks for their uplink communications. In particular, it is a schematic diagram of a power control message transmitted to a plurality of UEs I, II, . . . N wherein the location of the power control bits for each of the UEs is mapped to the resources on which the respective terminals transmit their uplink communications. In the figure, **552** is a control channel message carrying power control bits for the terminals I, II, . . . N. In this embodiment, control message **552** includes a message type field **554**, a bitmap **556**, and a cyclic redundancy check (CRC) field **558**. The message type field **554** carries a specific value to identify this message as a control channel message, the bitmap field **554** comprises a bitmap that carries power control bits for each of the associated UEs while the CRC field is used by the UEs to verify if the control channel message has been correctly decoded. As seen in the figure, UE I has two resource blocks assigned to it and therefore, two bits are used for power control of UE I. UE II has only one resource block assigned to it and has only one bit for controlling its transmission power. Accordingly, the control channel message **552** is transmitted with a number of bits N+X>N wherein each of the I, II, . . . N UEs can have different number of bits associated with its power control. This permits the serving base station to have different power correction steps for different UEs based on the number of resources they employ for their uplink transmissions. Accordingly, the base station can have better control over transmission power levels of UEs using more resources. As mentioned supra, the location of the power control bits in the bitmap transmitted to the UEs on the downlink is mapped to the resources employed by the UEs for their uplink communications. Therefore, various UEs can decode the number of bits and the location of the bits within the bitmap and adjust their transmission power accordingly.

FIG. **5**D illustrates another aspect related to transmitting power control message when an UE is transmitting a plurality of data streams in accordance with a MC-OFDMA scheme. A data stream can comprise audio/video/user data generated at the UE. Accordingly, it may use continuous or discontinuous resource blocks for transmitting on the uplink and the power control bits can be located in contiguous positions or at discontinuous positions within the bitmap conveying the power control message to the UE. In addition, each data stream can be using a different number of resource blocks for its transmissions. Therefore, a different number of bits can be associated with power control of resource blocks utilized by each data stream. However, as discussed supra, since the position of these bits in the downlink communication from a serving base station are mapped to the respective resource blocks used for uplink transmissions, the UE can decode the downlink control message and adjust its transmission power associated with the resource block(s) for each data stream accordingly.

In particular, FIG. **5**D illustrates a downlink control message transmitted to a UE I which communicates on the uplink in accordance with a MC-FDMA mode utilizing a plurality of discontinuous resource blocks. As an example, it can employ two resource blocks (**1** and **4**) for one of its uplink transmissions (transmission I) while resource blocks (**2**, **3** and **5**) are used for another uplink communication (transmission II) being transmitted on a difference frequency. Accordingly, two bits at the first and the fourth positions are used for controlling power of the UE for the communication utilizing the first and the fourth resource blocks while three bits at the second, third and fifth positions are utilized for controlling power of the transmissions utilizing the second, third and fifth resource blocks. In the figure a control message **572** is sent. UE I wherein, the message includes a message type field **574**, a bitmap **576**, and a cyclic redundancy check (CRC) field **578**. As discontinuous resource blocks are used for transmitting on the uplink, bits whose positions are non-contiguous are used for controlling power of respective transmissions. However,

as the position of these bits in the bitmap are associated with the resource blocks used for uplink communications, UE I can decode and appropriately adjust power of its various data streams based on their positions. It can be appreciated that a single control message may be transmitted to a plurality of UEs wherein one or more terminals can be transmitting a plurality of data streams. Therefore, the remaining bits within the message can be associated with power control of other UEs. The transmission power of each of these data streams can be controlled by the power control message by its bits whose position in the message is mapped to the resources utilized by the various uplink data streams originating from the one or more UEs. In addition, the aforementioned aspects can be applicable to both persistent and non-persistent assignments. For non-persistent assignments, the power control commands can be embedded in uplink grants.

Although the embodiments described herein have illustrated that the power control commands are carried to the UEs via bitmaps, it can be appreciated that this is not necessary. Other forms of power control communications wherein the location of the power control messages for respective UEs are mapped to the resources used by them for their uplink communications are possible as further detailed infra. In accordance with the aspects described supra, while a bitmap can be used to transmit power control commands to a plurality of UEs, more efficient means of transmitting these commands individually can be explored. A bitmap can convey individual power control commands separately to a single UE, however, as it is statically configured with predetermined number of bits it utilizes same amount of resources regardless of the number of UEs being affected by it. Such a static process tends to waste resources especially if other UEs do not require any adjustment of their transmission power.

The disadvantages arising from use of bitmaps can be overcome by a more dynamic method of indexing resource blocks that need to be power controlled. Such a process also consumes resources for indexing. For example, information associated with sixteen UEs can be indexed using 4 bits. Therefore, for each UE a total of 6 bits are required wherein 4 bits are used for indexing while 2 bits can be used for power control commands. As such a method requires 6 bits in order to convey 2 bits of power control information, it can be less efficient or 3-4 times less indexed channels can be multiplexed. Similarly if the power control command is conveyed within a single bit, 5-7 times less indexed channels can be multiplexed. However, it provides the flexibility of transmitting power control commands to a subset of UEs as it permits statistical multiplexing whereas communication of power control information via bitmaps is more static. In a more detailed aspect, UEs can be divided into various groups based on their CQI (channel quality indicator) or radio conditions. In addition to indexing individual resource blocks, another aspect relates to indexing assignment of resource blocks. For non-persistent assignments as discussed supra, such indexing can be included in the uplink grants.

Accordingly, FIG. **6** relates to another aspect of transmitting power control bits wherein uplink resources are indexed. A schematic diagram of a group power control command **600** is illustrated wherein the command is transmitted to a group of UEs which require adjustments to their transmission power wherein location of the TPC (transmit power control) command is mapped to indexed uplink resources used by the UEs. For example, PUCCH (Physical Uplink Control Channel) and persistent PUSCH (Physical Uplink Shared Channel) resource blocks are indexed when assigned with RRC (radio resource control). In the figure, each of the blocks indicates an indexed channel whose power is being adjusted via the group, power control command. Therefore, N is the total number of indexed channels power controlled via the group power control command. This aspect relates to a static mapping of power control commands for a given group of UEs as all the UEs in the group have their power varied in accordance with the TPC command associated therewith.

Another aspect relates to a more dynamic mapping of power control commands for a group of access terminals wherein power of specific access terminals within the group can be controlled. A schematic diagram **602** illustrates the group power control command transmitted in accordance with this aspect. As mentioned supra, it can comprise 4-6 bits for indexing thereby facilitating statistical multiplexing and 2 bits for TPC commands. In the figure, index 1 indicates the 4-6 bit index which specifies the resource block to be controlled while TPC index 1 indicates the power control command that adjusts the power of the UE for transmissions utilizing the resource block associated with index 1. Therefore, K is the total number of indexed channels addressed by the group power control command **602**. Thus, indexing persistently assigned resource blocks used for uplink communications and using them to map downlink power control commands can lead to a more dynamic method of signaling. It facilitates grouping UEs according to their radio conditions and transmitting a group power control command only to those UEs that require power correction. If a UE is assigned multiple indexed resources, a single indexed resource (for example the one with lowest index) is signaled in power control message for the UE to interpret the power control command. This leads to significant reduction in power overhead whereas a more static mapping leads to the group power control command being dimensioned with enough power to be transmitted to all UEs within a cell.

FIG. **7** relates to a methodology **700** of transmitting power control message to one or more UEs in accordance with an aspect. The procedure begins at **702** wherein a serving base station receives communications from a plurality of UEs. At **704**, the various communications are analyzed to determine if they have been transmitted with optimum power. As mentioned supra, SNR of the received communications can be one of the attributes of the received communications that indicates the transmission power of the messages. Various other factors can also indicate the power with which the messages are transmitted by the UEs. For example, if the SNR of the received communications is very low then it can be concluded that the message was transmitted with low power and that the UE transmitting it should increase the power of its uplink communications. Similarly location of the UE can be another attribute to indicate a change in transmission power of the message. If it is detected that the UE is proximate to the base station then the transmission power of the UE may be reduced to conserve power. If the messages are all transmitted with optimum power then the procedure returns to receiving uplink communications which are being transmitted at optimum levels, else, the messages which have been transmitted with non-optimum power and the UEs transmitting them are identified at **706**. At **708**, the resource blocks used by these UEs for transmitting the messages are also identified. As detailed supra, depending on the mode of transmission, continuous or discontinuous resource blocks can be employed by the UEs for their uplink communications. At **710**, one or more control messages to adjust power of those UEs transmitting at non-optimum power levels are formulated. In accordance with different aspects, a single power control message may be formulated for a plurality of UEs or individual messages may be composed for each of the UEs. Regardless of the number of messages, the location of the power control bits within these

messages is mapped to the resources utilized by the UEs for their uplink communications. The messages are then communicated to the UEs on the downlink as shown at **712** in order to adjust their transmission power to more optimum levels. This method aids in assigning resources for power control messages dynamically based on the uplink communications. This is in contrast to more a static method of assigning these resources wherein the resources allocated for transmitting power control messages may be under utilized if the UEs are either not transmitting or are transmitting at optimum power levels.

FIG. **8** relates to a more detailed aspect of generating power control messages for various UEs. The method **800** begins at **802** by analyzing communications received from each UE transmitting at non-optimum power levels. At **804** the number of resource blocks utilized by each UE for its uplink communications is identified. At **806**, based on the analysis of step **802**, power control commands to increase or decrease power are generated for each UE statically as bitmaps wherein the number of bits for controlling power of a given UE in the bitmap would depend on the number of resource blocks used by the UE for its uplink communications. In a different aspect, the power control messages can be generated more dynamically wherein indexed resources are employed for uplink communications and the control messages comprise indexing information in addition to the power control commands for respective resources. The power control commands thus generated are communicated to the UEs either in a common message for a plurality of UEs or as individual messages for each of the UEs as shown at **808**.

In accordance with different aspects, an UE can be transmitting a plurality of data streams at different frequencies via adopting a MC-FDMA mode. In this case, the different data streams being transmitted at non-optimal power levels for each UE transmitting in a multi-carrier mode are identified. Additionally, the resource blocks utilized by the UEs for such data streams are also identified. In accordance with this aspect, it can be appreciated that some of the plurality of data streams may be transmitted at optimum power levels and hence do not need adjustments of power levels. Based upon the results obtained on analyzing different transmissions, appropriate power control commands are generated for data streams being transmitted at non-optimal power levels. In accordance with different aspects, a common message with the power control commands incorporated therein may be transmitted to a plurality of UEs or each UE may receive an exclusive message comprising its individual power control commands for adjusting power of its one or more data streams.

FIG. **9** is associated with a methodology **900** of transmitting power control commands using indexing as described herein. The method begins at **902** wherein UEs are assigned resources for uplink communications. These resources can be assigned persistently or non-persistently. Regardless of the type of assignment, the resources can be indexed either individually or an assignment of a group of resource blocks can be indexed in accordance with different aspects. At **904**, uplink communications are received from the different UEs that were assigned indexed resources as described at **902**. At **906**, the radio conditions associated with the received communications are estimated. For example, as detailed supra, SNR can be used as one of the attributes to determine the radio conditions. Based at least on the estimated radio conditions, the UEs are grouped into different groups as shown at **908**. For example, UEs with good conditions that do not need any changes in their transmission power can be classified under one group whereas other UEs that may require adjustments to their power can be placed in another group. Among such groups requiring power corrections, different groups can be formed based on the kind of power adjustments that are required. Within such groups, those groups of UEs requiring power corrections are identified at **910** and appropriate power control commands are generated as shown at **912** such that their position within the power control messages is mapped to the resource blocks used by respective communications for their uplink transmissions. The generated power control messages for different communications are transmitted on the downlink as shown at **914**. However, as the resource blocks are indexed, the power control commands transmitted on the downlink comprise information (or bits) associated with the indexing of the resource blocks in addition to those conveying the power control commands.

Conventionally, transmitting minimum number of bits is encouraged in order to reduce usage of resources. However, the aforementioned aspects offset the cost of greater resources required for transmitting higher number of bits by the unexpected benefits associated with dynamic aspects of such signaling. As mentioned supra, indexing resource blocks facilitates grouping UEs based on their radio conditions thereby permitting transmission of power control commands aperiodically to only for those UEs requiring power corrections. Thus, no resources are wasted in transmitting any communications to UEs already operating at optimal power levels.

FIG. **10** is a flow chart **1000** detailing a method of adjusting power at the UEs based on the received commands. The procedure begins at **1002** wherein an UE transmits uplink communications on assigned resources to a serving base, station. At **1004**, it can receive communications from the base station on the downlink wherein the communications can relate to controlling power at the UE. At **1006**, the received messages are decoded and if they are associated with controlling transmission power of the UEs, the relevant commands are identified. In accordance with an aspect, the location of power control commands in the messages transmitted on the downlink is mapped to the resources utilized by the UE for its uplink communications. If the UE is transmitting a plurality of data streams at different frequencies, then the power control commands for one or more of the data streams are transmitted in a manner that is mapped to the resource blocks used by the one or more data streams for their uplink communications. This facilitates the UE to recognize the appropriate power control commands and adjust its transmission power on the uplink accordingly as shown at **1008**.

FIG. **11** illustrates a high-level system diagram of various components of a device in accordance with various aspects. It is to be appreciated that the device **1100** may be an eNode B, a UE, or a combination thereof. It comprises a communications component **1102** that facilitates receiving and transmitting communications to various entities utilizing hardware, software, and services as described herein. Although the communications component **1102** is depicted as a single entity, it may be appreciated separate transmission and reception components can be employed for sending and receiving communications. In accordance with an aspect the device **1100** can act as an eNode B and the communications component **1102** receives communications from various UEs relating to one or more of resource requests, data transmissions etc. An analysis component **1104** analyzes the communications received from the various entities including a plurality of UEs to identify the entities. In accordance with an aspect, communications received from UEs are analyzed to determine the resource blocks used by the terminals for transmitting their respective uplink communications. Additionally, the analysis compo-

nent can also analyze the communications to determine if the communications are being transmitted at optimum power levels. Although it has been illustrated that a single entity executes the analysis process to determine various aspects of the communications, it may be appreciated that more than one entity can be employed for different analysis procedures. The analysis component **1104** can include a single or multiple set of processors or multi-core processors wherein the processors can carry out other operations such as grouping UEs based on their similar radio conditions or generating indexing information for the resources assigned to the UEs or other functions. Moreover, the analysis component **1104** can be implemented as an integrated processing system and/or a distributed processing system. The information gathered by the analysis component **1104** can be stored in the memory **1106**/data store **1108** for further processing. Memory **1106** can include random access memory (RAM), read only memory (ROM), or a combination thereof. Data store **1108** can be any suitable combination of hardware and/or software that provides for mass storage of information, databases, and programs employed in connection with aspects described herein. The control component **1110** utilizes the information regarding the resource blocks employed by respective UEs on their uplink communications to form control messages for UEs that may be transmitting at non-optimal power levels. As detailed supra, the control component can generate a single message for a plurality of terminals or it can generate individual messages for each terminal. The messages can be stored in the memory **1106** or data store **1108** for transmitting to the UEs by the communications component **1102**.

FIG. **12** is another high level diagram illustrating various components of a device **1200** in accordance with different aspects described herein. The device **1200** can be an eNode B, an UE or a combination thereof. The device comprises a transmission component **1202** for transmitting various communications. If the device is acting as an UE then the transmission component **1202** can transmit various communications on the uplink to a serving eNode B/base station. The communications can include resource requests, data transmission on assigned resources etc. The device also comprises a receiving component **1204** for receiving communications from various entities including eNode B, other UEs etc. Upon transmission of resource requests, the receiving component can receive control messages relating to assignment of resources for uplink communications or data transmissions. In accordance with an aspect, the receiving component **1204** can receive control messages related to adjusting power of the UE based on the attributes of its uplink communications. These messages can be stored in the data store **1206**. Data store **1208** can be any suitable combination of hardware and/or software that provides for mass storage of information, databases, and programs employed in connection with aspects described herein. The device **1200** may optionally comprise a volatile/non-volatile memory **1210** including random access memory (RAM), read only memory (ROM), or a combination thereof. These messages are decoded and processed by a processing component **1212**. In accordance with an aspect, the power control messages received from a serving base station/eNode B can be decoded and processed in order to identify the power control information. As detailed supra, these messages can have power control bits at locations within the message mapped to specific resource blocks employed by the device **1200** for its uplink communications. The information decoded from such power control messages can be stored in the memory **1210** and/or data store **1208** and employed by the power control module **1206** to adjust the power of the device **1200** in accordance with the appropriate power control bits transmitted in these messages.

Next, systems that can enable aspects of the disclosed subject matter are described in connection with FIGS. **13** and **14**. Such systems can include functional blocks, which can be functional blocks that represent functions implemented by a processor or an electronic machine, software, or combination thereof (e.g., firmware).

FIG. **13** illustrates a block diagram of an example system **1300** that enables communication of power control information in accordance with aspects disclosed in the subject specification. System **1300** can reside at least partially within a base station, for example. System **1300** includes a logical grouping **1310** of electronic components that can act in conjunction. In an aspect of the subject innovation, logical grouping **1310** includes an electronic component **1315** for receiving uplink communications from one or more UEs; an electronic component **1325** for analyzing the received communications to identify communications received at non-optimal power levels; and an electronic component **1335** for transmitting one or more power control messages to at least one of the UEs wherein location of at least one power control command within the messages is mapped to at least one resource utilized by the at least one UE for the non-optimal uplink communications.

System **1300** can also include a memory **1340** that retains instructions for executing functions associated with electrical components **1315**, **1325**, and **1235**, as well as measured or computed data that may be generated during executing such functions. While shown as being external to memory **1340**, it is to be understood that one or more of electronic components **1315**, **1325**, and **1335**, and can exist within memory **1340**.

FIG. **14** illustrates a block diagram of an example system **1400** that enables adjustment of power within a wireless communication system in accordance with aspect described in the subject specification. System **1400** can reside at least partially within a mobile, for example. System **1400** includes a logical grouping **1410** of electronic components that can act in conjunction. In an aspect of the subject innovation, logical grouping **1410** includes an electronic component **1415** for transmitting one or more uplink communications; an electronic component **1425** for receiving at least one power control message on downlink wherein location of at least one power control command is mapped to at least one resource utilized in the uplink communications; and an electronic component **1435** for adjusting power levels of the one or more uplink communications in accordance with the at least one power control command.

System **1400** can also include a memory **1440** that retains instructions for executing functions associated with electrical components **1415**, **1425** and **1435**, as well as measured or computed data that may be generated during executing such functions. While shown as being external to memory **1440**, it is to be understood that one or more of electronic components **1415**, **1425** and **1435**, and can exist within

The data transmission techniques described herein may be implemented by various means. For example, these techniques may be implemented in hardware, firmware, software, or a combination thereof. For a hardware implementation, the processing units used for data transmission at a transmitter or data reception at a receiver may be implemented within one or more application specific integrated circuits (ASICs), digital signal processors (DSPs), digital signal processing devices (DSPDs), programmable logic devices (PLDs), field programmable gate arrays (FPGAs), processors, controllers, micro-controllers, microprocessors, electronic devices, other

electronic units designed to perform the functions described herein, or a combination thereof.

For a firmware and/or software implementation, the techniques may be implemented with modules (e.g., procedures, functions, and so on) that perform the functions described herein. The firmware and/or software codes may be stored in a memory and executed by a processor. The memory may be implemented within the processor or external to the processor.

The previous description of the disclosed embodiments is provided to enable any person, skilled in the art to make or use the disclosure. Various modifications to these embodiments will be readily apparent to those skilled in the art, and the generic principles defined herein may be applied to other embodiments without departing from the spirit or scope of the disclosure. Thus, the disclosure is not intended to be limited to the embodiments shown herein but is to be accorded the widest scope consistent with the principles and novel features disclosed herein.

What has been described above includes examples of the various embodiments. It is, of course, not possible to describe every conceivable combination of components or methodologies for purposes of describing the embodiments, but one of ordinary skill in the art may recognize that many further combinations and permutations are possible. Accordingly, the detailed description is intended to embrace all such alterations, modifications, and variations that fall within the spirit and scope of the appended claims.

In particular and in regard to the various functions performed by the above described components, devices, circuits, systems and the like, the terms (including a reference to a "means") used to describe such components are intended to correspond, unless otherwise indicated, to any component which performs the specified function of the described component (e.g., a functional equivalent), even though not structurally equivalent to the disclosed structure, which performs the function in the herein illustrated exemplary aspects of the embodiments. In this regard, it will also be recognized that the embodiments includes a system as well as a computer-readable medium having computer-executable instructions for performing the acts and/or events of the various methods.

In addition, while a particular feature may have been disclosed with respect to only one of several implementations, such feature may be combined with one or more other features of the other implementations as may be desired and advantageous for any given or particular application. Furthermore, to the extent that the terms "includes," and "including" and variants thereof are used in either the detailed description or the claims, these terms are intended to be inclusive in a manner similar to the term "comprising."

What is claimed is:

1. A method for controlling power within a wireless communication system, the method comprising:

receiving one or more uplink communications;

determining if the one or more uplink communications have been transmitted at an optimum power level;

determining one or more resource blocks employed by one or more user equipments (UEs) for transmitting one or more non-optimal uplink communications at a non-optimal power level;

generating one or more power control messages for the one or more UEs that transmitted the one or more non-optimal uplink communications at the non-optimal power level, wherein one or more positions of one or more power control commands within the one or more power control messages are mapped respectively to the one or more resource blocks employed by the one or more UEs for transmitting the one or more non-optimal uplink communications at the non-optimal power level; and

transmitting the one or more power control messages to the one or more UEs that transmitted the one or more non-optimal uplink communications at the non-optimal power level;

wherein the one or more uplink communications include a plurality of data streams transmitted concurrently by at least one of the UEs at different frequencies and the power control messages individually adjust power of the one or more resource blocks utilized for the data streams being transmitted at the non-optimal power level.

2. The method of claim **1**, wherein the power control messages comprise a bitmap that communicates the power control commands.

3. The method of claim **2**, wherein the power control commands are bits within the bitmap and a location of the bits within the bitmap is mapped to the one or more resource blocks employed by the respective UEs.

4. The method of claim **3**, wherein a number of bits for controlling power of a UE is associated with a number of resource blocks utilized by the UE for its uplink communications.

5. The method of claim **1**, wherein the transmitting the power control messages comprises transmitting at least one common power control message to the plurality of UEs.

6. The method of claim **1**, wherein the transmitting the power control messages comprises transmitting individual power control messages comprising power control bits to each of the UEs.

7. The method of claim **1**, further comprising indexing the resource blocks utilized by the respective UEs.

8. The method of claim **7**, wherein the power control messages comprise bits associated with the indexed resource blocks in addition to bits associated with the power control commands.

9. The method of claim **7**, further comprising grouping the UEs based at least on their estimated radio conditions.

10. The method of claim **9**, wherein the transmitting the power control messages comprises transmitting group power control commands to one or more of the UE groups.

11. An apparatus for conveying power control information within a communication system, the apparatus comprising:

a receiver that receives one or more communications from one or more user equipments (UEs);

a processor that generates at least one power control message for the one or more UEs that transmitted at least one of the one or more communications at a non-optimum power level, wherein at least one location of at least one power control command within the at least one power control message is mapped to one or more resource blocks utilized by the one or more UEs for transmitting at least one non-optimum communication at the non-optimum power level; and

a transmitter that transmits the at least one power control message to the one or more UEs that transmitted the at least one non-optimum communication;

wherein the one or more uplink communications include a plurality of data streams transmitted concurrently by at least one of the UEs at different frequencies and the at least one power control message individually adjusts power of the one or more resource blocks utilized for the data streams being transmitted at the non-optimal power level.

12. The apparatus of claim 11, further comprising a memory that stores information associated with one or more of the power control commands or the resource blocks.

13. The apparatus of claim 11, wherein the at least one power control message comprises a plurality of power control commands associated with the plurality of UEs wherein a location of the plurality of power control commands within the power control message is mapped to a plurality of resource blocks utilized by the plurality of UEs for the one or more communications.

14. The apparatus of claim 11, wherein the transmitter transmits a plurality of power control messages separately to each of the plurality of UEs wherein the plurality of power control messages transmit individual power control commands to each of the plurality of UEs.

15. The apparatus of claim 11, wherein at least one of the UEs utilizes a plurality of resource blocks for transmitting the at least one non-optimum communication.

16. The apparatus of claim 15, wherein the at least one power control message comprises a plurality of power control commands whose position within the power control message is mapped to the plurality of resource blocks utilized by the at least one UE.

17. The apparatus of claim 11, wherein the at least one power control message comprises indexing information associated with the resource blocks.

18. The apparatus of claim 17, wherein the power control message comprises power control commands for a plurality of the one or more UEs grouped together based at least on respective radio conditions.

19. A computer program product comprising:

a computer-readable medium comprising:

code for causing at least a computer to receive one or more uplink communications;

code for causing at least a computer to determine if the one or more uplink communications have been transmitted at an optimum power level;

code for causing at least a computer to determine one or more resource blocks employed by one or more user equipments (UEs) for transmitting one or more non-optimal uplink communications at a non-optimal power level;

code for causing at least a computer to generate one or more power control messages for the one or more UEs that transmitted the one or more non-optimal uplink communications at the non-optimal power level, wherein one or more positions of one or more power control commands within the one or more power control messages are mapped respectively to the one or more resource blocks employed by the one or more UEs to transmit the one or more non-optimal uplink communications at the non-optimal power level; and

code for causing at least a computer to transmit the one or more power control messages to the one or more UEs that transmitted the one or more non-optimal uplink communications at the non-optimal power level;

wherein the one or more uplink communications include a plurality of data streams transmitted concurrently by at least one of the UEs at different frequencies and the power control messages individually adjust power of the one or more resource blocks utilized for the data streams being transmitted at the non-optimal power level.

20. The computer program product of claim 19, wherein the computer-readable medium further comprises code for causing at least a computer to include a bitmap into the power control messages in order to communicate the power control commands.

21. The computer program product of claim 20, wherein the computer-readable medium further comprises code for causing at least a computer to generate the power control commands as bits within the bitmap.

22. The computer program product of claim 21, wherein the computer-readable medium further comprises code for causing at least a computer to map a location of the bits within the bitmap to the one or more resource blocks employed by the respective UEs.

23. The computer program product of claim 22, wherein a number of bits for controlling power of at least one of the one or more UEs is associated with a number of resource blocks utilized by the at least one UE for its uplink communications.

24. The computer program product of claim 19, wherein the code for causing at least a computer to transmit the power control messages cause the computer to transmit individual power control messages comprising power control bits to each of the UEs.

25. The computer program product of claim 19, wherein the code for causing at least a computer to transmit the power control messages cause the computer to transmit a common power control message to a plurality of UEs.

26. The computer program product of claim 19, wherein the computer-readable medium further comprises code for causing at least a computer to index the resource blocks employed by the UEs.

27. The computer program product of claim 26, wherein the computer-readable medium further comprises code for causing at least a computer to group the UEs into one or more groups based at least on respective radio conditions.

28. The computer program product of claim 27, wherein the code for causing at least a computer to transmit the power control messages cause the computer to transmit a common power control message to at least one of the UE groups wherein the common power control message comprises indexing information in addition to the control commands.

29. A wireless communications apparatus for communicating power control information, the system comprising:

means for receiving one or more uplink communications from one or more user equipments (UEs);

means for analyzing the received one or more uplink communications to identify one or more non-optimal uplink communications received at a non-optimal power level; and

means for transmitting one or more power control messages to at least one of the one or more UEs that transmitted the one or more non-optimal uplink communications at the non-optimal power level, wherein at least one location of at least one power control command within the one or more power control messages is mapped to at least one resource block utilized by the at least one UE for transmitting the one or more non-optimal uplink communications at the non-optimal power level;

wherein the one or more uplink communications include a plurality of data streams transmitted concurrently by at least one of the UEs at different frequencies and the power control messages individually adjust power of the one or more resource blocks utilized for the data streams being transmitted at the non-optimal power level.

30. The apparatus of claim 29, further comprising means for storing one or more of the received uplink communications or attributes associated with the received uplink communications.

31. The apparatus of claim 29, wherein the means for transmitting the power control messages comprises means for transmitting a bitmap that comprises the power control commands.

32. The apparatus of claim 31, wherein the power control commands are bits within the bitmap and a location of the bits within the bitmap is mapped to the at least one resource block employed by the at least one UE for transmitting the non-optimal uplink communications.

33. The apparatus of claim 32, wherein a number of bits for controlling power of the at least one UE is associated with a number of resource blocks utilized by the UE for the non-optimal uplink communications.

34. The apparatus of claim 29, further comprising means for indexing resource blocks utilized by the UEs for uplink communications.

35. The apparatus of claim 34, wherein the means for analyzing comprises means for grouping the UEs into one or more groups in accordance with respective radio conditions associated therewith.

36. The apparatus of claim 35, wherein the means for transmitting the power control messages comprises means for transmitting a common power control message to at least one of the UE groups wherein the common power control message comprises the resource indexing information in addition to power control commands.

37. A method of receiving power control commands, the method comprising:

transmitting one or more uplink communications;

receiving at least one power control message on a downlink when the one or more uplink communications is determined to be transmitted at a non-optimal power level, wherein at least one location of at least one power control command within the at least one power control message is mapped to one or more resource blocks utilized for transmitting the one or more uplink communications at the non-optimal power level;

identifying the at least one power control command based at least on a location of the at least one power control command within the at least one power control message; and

adjusting power levels of the one or more uplink communications in accordance with the at least one power control command;

wherein the one or more uplink communications include a plurality of data streams transmitted concurrently at different frequencies and the at least one power control message individually adjusts power of the one or more resource blocks utilized for the data streams being transmitted at the non-optimal power level.

38. The method of claim 37, wherein the resource blocks utilized for the uplink communications are indexed.

39. The method of claim 38, further comprising associating with a group of UEs based at least upon radio conditions.

40. The method of claim 39, wherein the receiving the power control message comprises receiving a group power control message that includes power control commands in the group in addition to indexing information associated with the resource blocks utilized by the UEs for the uplink communications.

41. An apparatus for adjusting transmission power levels within a communication system, the apparatus comprising:

a transmitter that transmits one or more uplink communications at a non-optimal power level;

a receiver that receives at least one power control message when the one or more uplink communications is determined to be transmitted at the non-optimal power level, wherein at least one location of at least one power control command within the at least one power control message is mapped to one or more resource blocks utilized for transmitting the one or more non-optimal uplink communications at the non-optimal power level; and

a processor that controls power levels of the transmitter based on the at least one power control command;

wherein the one or more uplink communications include a plurality of data streams transmitted concurrently at different frequencies and the at least one power control message individually adjusts power of the one or more resource blocks utilized for the data streams being transmitted at the non-optimal power level.

42. The apparatus of claim 41, wherein the received power control message comprises a bitmap wherein at least one bit within the bitmap functions as the power control command and a location of the bit within the bitmap is mapped to the resource blocks utilized for the non-optimal uplink communications.

43. The apparatus of claim 42, wherein a number of bits within the bitmap corresponds to a number of resource blocks utilized for the non-optimal uplink communications.

44. The apparatus of claim 41, wherein the resource blocks are persistently assigned.

45. The apparatus of claim 41, wherein the resource blocks utilized by the non-optimal communications are indexed.

46. The apparatus of claim 41, wherein the power control message comprises indexing information associated with the resource blocks in addition to the power control command.

47. A computer program product comprising:

a computer-readable medium comprising:

code for causing at least a computer to transmit one or more uplink communications;

code for causing at least a computer to receive at least one power control message on a downlink when the one or more uplink communications is determined to be transmitted at a non-optimal power level, wherein at least one location of at least one power control command within the at least one power control message is mapped to at least one resource block utilized for transmitting the one or more uplink communications at the non-optimal power level; and

code for causing at least a computer to adjust power levels of the one or more uplink communications in accordance with the at least one power control command;

wherein the one or more uplink communications include a plurality of data streams transmitted concurrently at different frequencies and the at least one power control message individually adjusts power of the one or more resource blocks utilized for the data streams being transmitted at the non-optimal power level.

48. The computer program product of claim 47, wherein the received power control message comprises a plurality of power control commands, wherein a location of the power control commands within the received message is mapped to the plurality of resource blocks used for the uplink communications.

49. The computer program product of claim 47, wherein the resource blocks utilized for the uplink communications are indexed.

50. The computer program product of claim 47, wherein the computer-readable medium further comprises code for causing at least a computer to associate with a group of UEs based at least upon radio conditions.

51. The computer program product of claim 50, wherein the received power control message is a group power control message comprising power control commands in the group and indexing information associated with the resource blocks utilized by the UEs for the uplink communications.

52. An apparatus for adjusting power within a wireless communication system, the apparatus comprising:

means for transmitting one or more uplink communications;

means for receiving at least one power control message on a downlink when the one or more uplink communications is determined to be transmitted at a non-optimal power level, wherein at least one location of at least one power control command within the at least one power control message is mapped to at least one resource block utilized for transmitting the one or more uplink communications at the non-optimal power level; and

means for adjusting power levels of the one or more uplink communications in accordance with the at least one power control command;

wherein the one or more uplink communications include a plurality of data streams transmitted concurrently by at least one of the UEs at different frequencies and the at least one power control message individually adjusts power of the one or more resource blocks utilized for the data streams being transmitted at the non-optimal power level.

53. The apparatus of claim **52**, wherein the resource blocks utilized in the uplink communications are indexed and the power control message comprises the indexing information along with the power control command.

* * * * *

EXHIBIT 15

3GPP TR 21.905 V10.3.0 (2011-03)

Technical Report

3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Vocabulary for 3GPP Specifications (Release 10)





The present document has been developed within the 3rd Generation Partnership Project (3GPP ™) and may be further elaborated for the purposes of 3GPP.
The present document has not been subject to any approval process by the 3GPP Organisational Partners and shall not be implemented.
This Specification is provided for future development work within 3GPP only. The Organisational Partners accept no liability for any use of this Specification.
Specifications and reports for implementation of the 3GPP ™ system should be obtained via the 3GPP Organisational Partners' Publications Offices.

Keywords
GSM, UMTS, LTE, Vocabulary

3GPP

Postal address

3GPP support office address
650 Route des Lucioles - Sophia Antipolis
Valbonne - FRANCE
Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Internet
http://www.3gpp.org

Copyright Notification

No part may be reproduced except as authorized by written permission.
The copyright and the foregoing restriction extend to reproduction in all media.

© 2011, 3GPP Organizational Partners (ARIB, ATIS, CCSA, ETSI, TTA, TTC).
All rights reserved.

UMTS™ is a Trade Mark of ETSI registered for the benefit of its members
3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners
LTE™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners
GSM® and the GSM logo are registered and owned by the GSM Association

Contents

Foreword 5
1 Scope 5
2 References 5
3 Terms and definitions 6
0-9 6
A 6
B 8
C 9
D 11
E 12
F 13
G 14
H 15
I 15
J 17
K 17
L 17
M 18
N 20
O 21
P 22
Q 24
R 24
S 27
T 30
U 32
V 33
W 34
X 34
Y 34
Z 34
4 Abbreviations 34
0-9 34
A 34
B 36
C 37
D 39
E 40
F 41
G 42
H 42
I 43
J 44
K 45
L 45
M 46
N 47
O 48
P 49
Q 51
R 51
S 53
T 55
U 57
V 58

W 58
X 58
Y 59
Z 59

5 Equations 60

Annex A: Change history 62

Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the document.

1 Scope

The purpose of this report is to identify specialist technical terms used within the 3GPP project for the purposes of specifying service requirements. The motivations for this are:

- To ensure that editors use terminology that is consistent across specifications.
- To provide a reader with convenient reference for technical terms that are used across multiple documents.
- To prevent inconsistent use of terminology across documents.

This document is a collection of terms, definitions and abbreviations related to the baseline documents defining 3GPP objectives and systems framework. This document provides a tool for further work on 3GPP technical documentation and facilitates their understanding.

The terms, definitions and abbreviations as given in this document are either imported from existing documentation (ETSI, ITU or elsewhere) or newly created by 3GPP experts whenever the need for precise vocabulary was identified.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] Void

[2] 3GPP TS 25.990: "Technical Specification Group (TSG) RAN; Vocabulary ".

[3] "The Path towards UMTS - Technologies for the Information Society" – Report #2, UMTS Forum.

[4] 3GPP TS 23.122: "Non-Access-Statum (NAS) functions related to Mobile Station (MS) in idle mode".

[5] ETSI TR 180 000: "NGN terminology".

[6] IEC 60050-161: "International Electrotechnical Vocabulary - Chapter 161: Electromagnetic compatibility".

[7] 3GPP TS 23.401: "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access".

[8] 3GPP TS 23.101: "General Universal Mobile Telecommunications System (UMTS) architecture".

3 Terms and definitions

0-9

1.8V technology Smart Card: A Smart Card operating at 1.8V ± 10% and 3V ± 10%.

1.8V technology Terminal: A terminal operating the Smart Card - Terminal interface at 1.8V ± 10% and 3V ± 10%.

3GPP Generic User Profile (GUP): The 3GPP Generic User Profile is the collection of user related data which affects the way in which an individual user experiences services and which may be accessed in a standardised manner.

3GPP system: the telecommunication system standardised by the 3GPP consisting of a core network and a radio access network that may be either GERAN or UTRAN, or both.

3GPP System core network: refers in this specification to an evolved GSM core network infrastructure.

3GPP System coverage: see coverage area.

3GPP System IC Card: An IC card (or 'smartcard') of defined electromechanical specification which contains at least one USIM.

3GPP System mobile termination: part of the 3GPP System Mobile Station which provides functions specific to the management of the radio interface (Um).

3GPP-WLAN Interworking: Used to generically refer to interworking between the 3GPP system and the WLAN family of standards.

3V technology Smart Card: A Smart Card operating at 3V± 10% and 5V ± 10%.

3V technology Terminal: A terminal operating the Smart Card - Terminal interface at 3V-± 10% and 5V ± 10%.

A

A/Gb mode: mode of operation of the MS when connected to the Core Network via GERAN and the A and/or Gb interfaces.

Acceptable Cell: A cell that the UE may camp on to make emergency calls. It must satisfy certain conditions.

Access conditions: A set of security attributes associated with a file.

Access delay: The value of elapsed time between an access request and a successful access (source: ITU-T X.140).

Access Stratum: functional grouping consisting of the parts in the infrastructure and in the user equipment and the protocols between these parts being specific to the access technique (i.e. the way the specific physical media between the User Equipment and the Infrastructure is used to carry information).

Note: For full definition, see 23.101 [8], clause 6.2.

Access Stratum SDU (Service Data Unit): Unit of data transferred over the access stratum SAP (Service Access Point) in the Core Network or in the User Equipment.

Access protocol: A defined set of procedures that is adopted at an interface at a specified reference point between a user and a network to enable the user to employ the services and/or facilities of that network (source: ITU-T I.112).

Accounting: The process of apportioning charges between the Home Environment, Serving Network and User.

Accuracy: A performance criterion that describes the degree of correctness with which a function is performed. (The function may or may not be performed with the desired speed.) (source: ITU-T I.350).

Active communication: a UE is in active communication when it has a CS connection established. For PS active communication is defined by the existence of one or more Activated PDP contexts. Either one or both of the mentioned active communications may occur in the UE.

Active Set: Set of radio links simultaneously involved in a specific communication service between an UE and a UTRAN.

Adjacent Channel Leakage power Ratio (ACLR): The ratio of the average power centered on the assigned channel frequency to the average power centered on an adjacent channel frequency. In both cases the average power is measured with a filter that has Root Raised Cosine (RRC) filter response with roll-off $\alpha = 0.22$ and a bandwidth equal to the chip rate.

Air Interface User Rate: The user rate between Mobile Termination and IWF. For T services it is the maximum possible AIUR not including padding. For NT services it is the maximum possible AIUR.

ALCAP: Generic name for the transport signalling protocols used to set-up and tear-down transport bearers.

Allowable PLMN: A PLMN which is not in the list of forbidden PLMN in the UE.

Allowed CSG list: A list stored in the UE containing the CSG identities and associated PLMN identities of the CSGs to which the subscriber belongs.

Ancillary equipment: Equipment (apparatus), used in connection with a receiver, transmitter or transceiver is considered as an ancillary equipment (apparatus) if:

- the equipment is intended for use in conjunction with a receiver, transmitter or transceiver to provide additional operational and/or control features to the radio equipment, (e.g. to extend control to another position or location); and
- the equipment cannot be used on a stand alone basis to provide user functions independently of a receiver, transmitter or transceiver; and
- the receiver, transmitter or transceiver to which it is connected, is capable of providing some intended operation such as transmitting and/or receiving without the ancillary equipment (i.e. it is not a sub unit of the main equipment essential to the main equipment basic functions).

Applet: A small program that is intended not to be run on its own, but rather to be embedded inside another application

Application: an application is a service enabler deployed by service providers, manufacturers or users. Individual applications will often be enablers for a wide range of services. (UMTS Forum report #2) [3]

Applications / Clients: These are services, which are designed using service capability features.

Application Dedicated File (ADF): an application DF is the entry point to an application on the UICC.

Application Interface: Standardised Interface used by application/clients to access service capability features.

Application protocol: The set of procedures required by the application.

ASCI Generic name to identify the services VGCS, VBS and eMLPP.

Authentication: A property by which the correct identity of an entity or party is established with a required assurance. The party being authenticated could be a user, subscriber, home environment or serving network.

Available PLMN: A PLMN where the UE has found a cell that satisfies certain conditions.

Average power: The thermal power as measured through a root raised cosine filter with roll-off $\alpha = 0.22$ and a bandwidth equal to the chip rate of the radio access mode. The period of measurement shall be one power control group (timeslot) unless otherwise stated.

B

Band category: A group of operating bands for which the same MSR scenarios apply

Base Station: A base station is a network element in radio access network responsible for radio transmission and reception in one or more cells to or from the user equipment. A base station can have an integrated antenna or be connected to an antenna by feeder cables. In UTRAN it terminates the I_{ub} interface towards the RNC. In GERAN it terminates the Abis interface towards the BSC.

Baseline capabilities: Capabilities that are required for a service-less UE to operate within a network. The baseline capabilities for a UE include the capabilities to search for, synchronise with and register (with authentication) to a network. The negotiation of the UE and the network capabilities, as well as the maintenance and termination of the registration are also part of the required baseline capabilities.

Base Station Controller: This equipment in the BSS is in charge of controlling the use and the integrity of the radio resources.

Base station receive period: The time during which the base station is receiving data subframes or UpPTS.

Base Station RF bandwidth: The bandwidth in which a Base Station transmits and receives multiple carriers and/or RATs simultaneously

Base Station RF bandwidth edge: The frequency of one of the edges of the Base Station RF bandwidth

Base Station Subsystem: Either a full network or only the access part of a GERAN offering the allocation, release and management of specific radio resources to establish means of connection between an MS and the GERAN.
A Base Station Subsystem is responsible for the resources and transmission/reception in a set of cells.

Baseline Implementation Capabilities: Set of Implementation capabilities, in each technical domain, required to enable a UE to support the required Baseline capabilities.

Basic OR Basic Optimal Routeing

Basic telecommunication service: This term is used as a common reference to both bearer services and teleservices.

Bearer: A information transmission path of defined capacity, delay and bit error rate, etc.

Bearer capability: A transmission function which the UE requests to the network.

Bearer independent protocol: (UICC) Mechanism by which the ME provides the (U)SIM applications on the UICC with access to the data bearers supported by the ME and the network.

Bearer service: A type of telecommunication service that provides the capability of transmission of signals between access points.

Best effort QoS: The lowest of all QoS traffic classes. If the guaranteed QoS cannot be delivered, the bearer network delivers the QoS which can also be called best effort QoS.

Best effort service: A service model which provides minimal performance guarantees, allowing an unspecified variance in the measured performance criteria.

Billing: A function whereby CDRs generated by the charging function are transformed into bills requiring payment.

Broadcast: A value of the service attribute "communication configuration", which denotes unidirectional distribution to all users (source: ITU-T I.113).

Byte code: (UICC) A hardware machine independent representation of a primitive computer operation that serves as an instruction to a software program called an interpreter or a virtual machine that simulates the hypothetical computer's central processing unit. code generated by a Java compiler and executed by the Java interpreter.