

Filed: June 6, 2025

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GOOGLE LLC,
Petitioner,

v.

POINTWISE VENTURES LLC,
Patent Owner.

Case No. IPR2025-01102
U.S. Patent No. 8,471,812

PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 8,471,812

TABLE OF CONTENTS

I.	PRELIMINARY STATEMENT.....	1
II.	THE '812 PATENT.....	2
	A. Disclosure.....	2
	B. File History	5
	C. Claim Construction	6
III.	THE PRIOR ART	7
	A. Oami.....	7
	B. Du.....	9
	C. Darrell	12
IV.	STATEMENT OF PRECISE RELIEF REQUESTED.....	15
V.	LEVEL OF ORDINARY SKILL.....	16
VI.	CLAIMS 1-12 ARE UNPATENTABLE OVER THE PRIOR ART 17	
	A. Ground 1: <i>Oami</i> alone or in combination with <i>Darrell</i> and further in view of <i>Du</i> renders obvious claims 1-12	17
	1. Claim 1.....	17
	1[pre]: A method for identifying an object, the method comprising:	17
	1[a]: (a) providing a pointing and identification device for pointing at the object, the pointing and identification device comprising:	17
	1[b]: at least one actuation means for actuation by the user when the user points the pointing and identification device at the object;.....	21

1[c]: a digital camera for forming a digital image of the object or of a portion of the object when the user points the pointing and identification device at the object and actuates the at least one actuation means; and.....	23
1[d]: a communication device for communicating the digital image to a different location when the user actuates the at least one actuation means;.....	24
1[e]: (b) communicating the digital image to the different location;.....	25
1[f]: (c) automatically identifying a list of likely pointed-to objects from the digital image at the different location to return the list of likely pointed-to objects; and	26
1[g]: (d) returning the list of likely pointed-to objects to the user to select one of the likely pointed-to objects;.....	31
1[h]: wherein the object is at least one of a spot on a displayed image on a display, a subarea of a space on the displayed image on the display, one of a plurality of objects in the displayed image on the display, an object in space, or near an object in space, a subarea of a surface in space, or one of a plurality of objects in space.....	33
2. Claim 2.....	38
2[a]: The method of claim 1, wherein step (c) comprises: (i) receiving a designation of the object as designated by the user;	38
2[b]: (ii) recording a frame image comprising the designated object;	38
2[c]: (iii) determining a context type of the frame image from the frame image; and.....	39

2[d]: (iv) identifying the object.	40
3. Claim 3: The method of claim 1, further comprising receiving a selection by the user of said one of the likely pointed-to objects as the object and an instruction from the user to obtain information on the object.	40
4. Claim 4: The method of claim 1, further comprising receiving a selection by the user of said one of the likely pointed-to objects as the object and an instruction from the user to purchase the object.	44
5. Claim 5: The method of claim 1, wherein the object is at least one of the spot on the displayed image on the display, the sub-area of the space on the displayed image on the display, and one of the plurality of objects in the displayed image on the display.	48
6. Claim 6: The method of claim 5, wherein the object is the sub-area of the space on the displayed image on the display.	48
7. Claim 7: The method of claim 1, wherein the object is a spot on a displayed image on a display.	48
8. Claim 8: The method of claim 1, wherein the object is a subarea of a space on the displayed image on the display.	48
9. Claim 9: The method of claim 1, wherein the object is one of a plurality of objects in the displayed image on the display.	48
10. Claim 10: The method of claim 1, wherein the object is an object in space, or near an object in space.	48
11. Claim 11: The method of claim 1, wherein the object is a subarea of a surface in space.	48
12. Claim 12: The method of claim 1, wherein the object is one of a plurality of objects in space.	49

B. Ground 2: *Du* in view of *Darrell* renders obvious claims
1-12 49

1.	Claim 1.....	49
	Element 1[pre]:	49
	Element 1[a]:	49
	Element 1[b]:	53
	Element 1[c]:	56
	Element 1[d]:	57
	Element 1[e]:	59
	Element 1[f]:	60
	Element 1[g]:	65
	Element 1[h]:	68
2.	Claim 2.....	72
	Element 2[a]:	72
	Element 2[b]:	72
	Element 2[c]:	73
	Element 2[d]:	74
3.	Claim 3:	74
4.	Claim 4:	77
5.	Claim 5:	78
6.	Claim 6:	78
7.	Claim 7:	78
8.	Claim 8:	78

9.	Claim 9:	78
10.	Claim 10:	78
11.	Claim 11:	78
12.	Claim 12:	78
VII.	DISCRETIONARY DENIAL IS IMPROPER.....	78
VIII.	MANDATORY NOTICES.....	79
A.	Real Party-in-Interest Under 37 C.F.R. § 42.8(b)(1).....	79
B.	Related Matters Under 37 C.F.R. § 42.8(b)(2).....	79
C.	Lead and Back-Up Counsel Under 37 C.F.R. § 42.8(b)(3).....	81
D.	Service Information Under 37 C.F.R. § 42.8(b)(4)	81
IX.	GROUND FOR STANDING	81
X.	CONCLUSION	81

LIST OF EXHIBITS

Exhibit	Description
Ex. 1001	U.S. Patent No. 8,471,812 to <i>Bunch</i> (“’812 patent”)
Ex. 1002	Prosecution File History for U.S. Patent No. 8,471,812
Ex. 1003	Curriculum Vitae of Dr. Daniel P. Lopresti
Ex. 1004	Declaration of Dr. Daniel P. Lopresti (“Lopresti”)
Ex. 1005	Mobile Multimedia Library: an MPEG-7 application with camera-equipped mobile phones by <i>Oami</i> et al. (“ <i>Oami</i> ”)
Ex. 1006	U.S. Patent Pub. 2005/0198095 to <i>Du</i> et al. (“ <i>Du</i> ”)
Ex. 1007	U.S. Patent Pub. 2005/0162523 to <i>Darrell</i> et al. (“ <i>Darrell</i> ”)
Ex. 1008	Seeing-Is-Believing: Using Camera Phones for Human-Verifiable Authentication by <i>McCune</i> et al. (“ <i>McCune</i> ”)
Ex. 1009	IDEixis—Image-Based Deixis for Recognizing Locations by <i>Yeh</i> (“ <i>Yeh</i> ”)
Ex. 1010	A System for Distributed Image Acquisition, Content Analysis and Similarity Retrieval by <i>Noorzaie</i> et al. (“ <i>Noorzaie</i> ”)
Ex. 1011	Adding Hyperlinks to Digital Television by <i>Bove</i> et al. (“ <i>Bove</i> ”)
Ex. 1012	Content-Based Image Retrieval on Mobile Device by <i>Ahmad</i> et al. (“ <i>Ahmad</i> ”)
Ex. 1013	U.S. Patent Pub. No. 2003/0174868 (“ <i>Adachi</i> ”)
Ex. 1014	U.S. Patent Publication No. 2004/0208372 (<i>Boncyk</i>)
Ex. 1015	Declaration of Lauren Gluckman (“Gluckman”)

I. PRELIMINARY STATEMENT

Petitioner Google LLC requests IPR of claims 1-12 of U.S. Patent No. 8,471,812 purportedly assigned to Pointwise Ventures LLC.

The '812 patent claims a method of using a digital camera to form a digital image of an object when the user points a pointing and identification device at the object, communicating the digital image to a different location, identifying a list of likely pointed-to objects from the digital image, and returning a list of likely pointed-to objects to the user.

The Notice of Allowance stated the patent overcame rejection due to two features: (1) “automatically identifying a list of likely pointed-to objects from the digital image”; and (2) “returning the list of likely pointed-to objects to the user to select one of the likely pointed-to objects as the identified objects.” Ex. 1002, 310. But *both* of those features were well known before the earliest possible priority date for the '812 patent, September 23, 2005. For example, *Oami* disclosed identifying a list of candidate images, returning the top four candidates to the user, and allowing the user to select one of the candidate images. *See also, e.g., Ahmad*, 5-6, 9, Figs. 3, 6; *Yeh*, 7, 8, 19, Fig. 3-1; *Noorzaie*, 2, 5. Additionally, *Darrell* disclosed finding images similar to a captured image and presenting the similar images as thumbnails for the user to select. Accordingly, the Board should institute IPR and cancel the challenged claims.

II. THE '812 PATENT

A. Disclosure

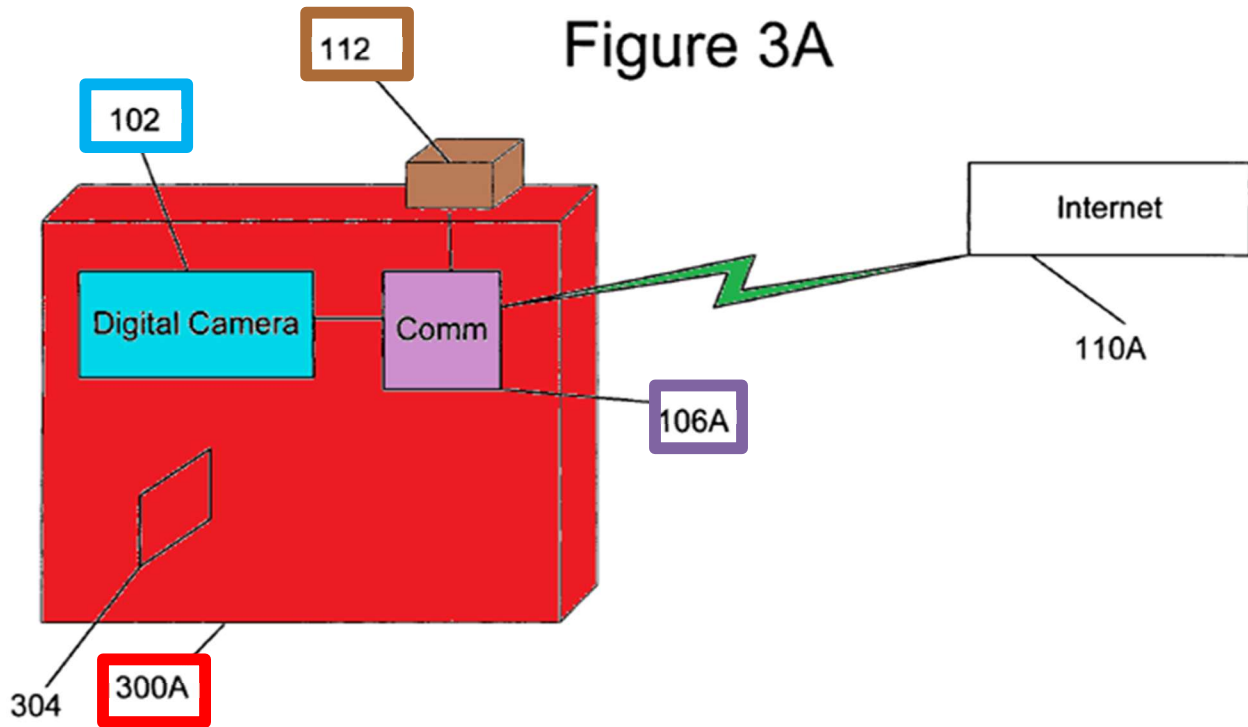
The '812 patent claims a method of taking a picture of an object, sending that image elsewhere to be used in identifying a list of likely matches for the object, and returning the list of likely matches to the user.

The specification describes “a pointing and identification device (PID) 100A.” '812 patent, 5:8-10. According to the patent, a “PID includes a digital camera and one or both of a laser and a reticle for aiming the digital camera.” '812 patent, Abstract. The laser or reticle “allows the user to aim the device.” '812 patent, 2:45-52. The user aims the “viewfinder” at an object to “designate” it. '812 patent, 9:9-19.

Once a user has aimed the camera (pointing and identification device) at an object to “designate” it, “the user clicks to select the item.” '812 patent, 9:9-19. The user “clicks” the “at least one mouse button 112,” thus a “frame” (or image) is “captured” (or “recorded”). '812 patent, 5:8-23; 9:9-19; 10:57-63.

The specification further discloses using “a wireless communication component 106A” to communicate information regarding the captured image to “a computer or Internet 110A.” '812 patent, 5:8-23. The specification further describes this as “data” (including a “Frame”) being “transmitted to a remote system.” '812 patent, 11:46-53. The “computer or internet 110A” then

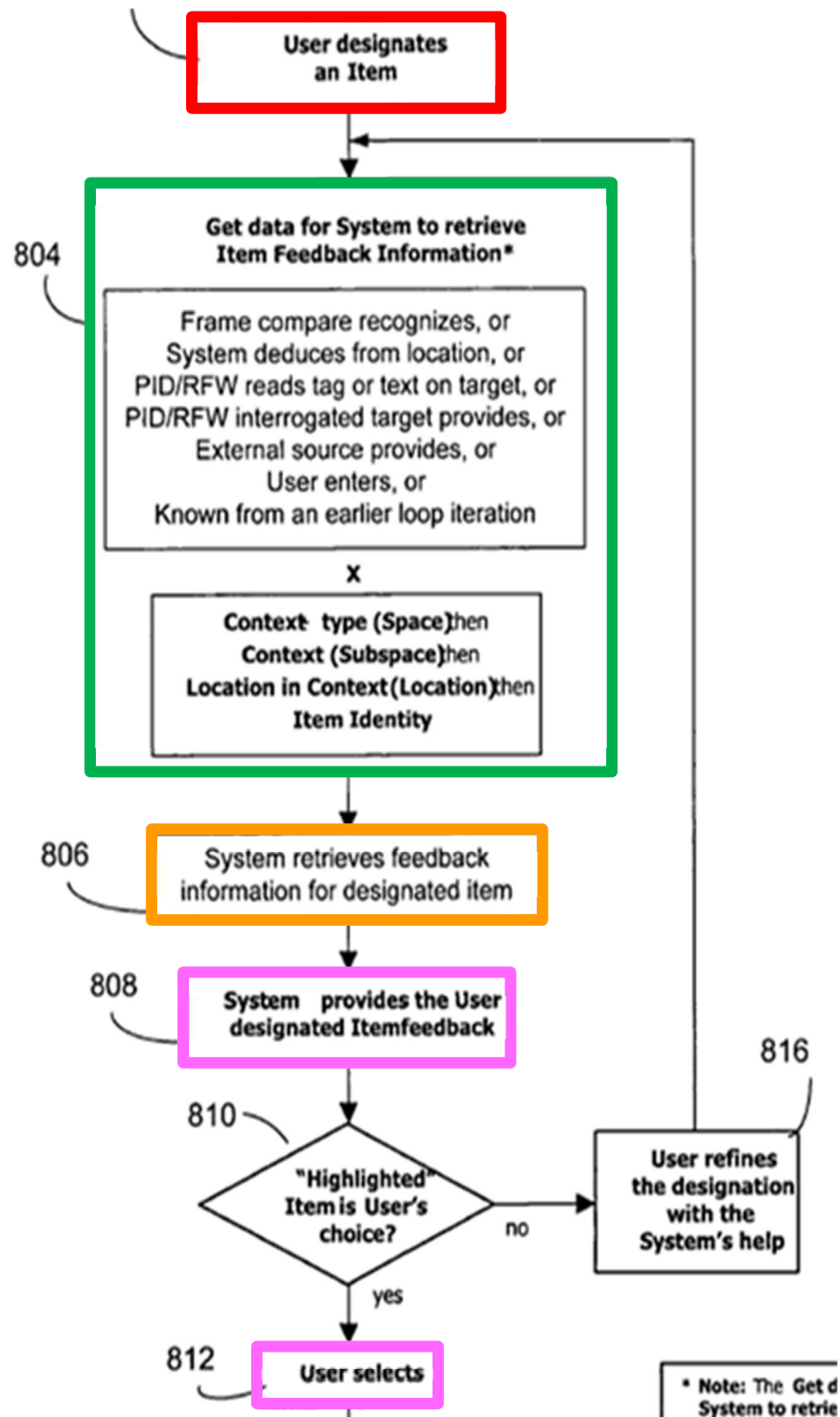
“recogniz[es] a Frame captured” by the camera. ’812 patent, 5:8-23. Figure 3A, as annotated below, shows the components of the pointing and identification device working together to transmit a captured image to a remote system (“Internet”):



’812 patent, Fig. 3A (annotated).

The ’812 specification discloses that the pointing and identification device “communicates with a remote system to obtain a list of likely pointed-to objects from which the user can select.” ’812 patent, 6:30-33. The specification further discloses generating the “list of likely pointed to objects” by analyzing data extracted from the captured “Frame.” ’812 patent, 14:35-59. The “list” is then “transmitted to the appropriate device(s) to provide feedback to the user.” ’812 patent, 15:65-16:10. The user may then “select [one] or more of these objects[.]”

'812 patent, 16:11-17. Figure 8 illustrates the steps of obtaining a list of likely pointed-to objects and the user selecting one of those objects:



'812 patent, Fig. 8 (annotated).

The '812 patent explains that the “Frame” (or “image”) may be, for example, an image from a show displayed on a screen (e.g., “TV Shows, DVDs, and video tapes”); printed material (including a barcode or a logo); or an image of the “Real World.” '812 patent, 12:44-13:15.

B. File History

During prosecution, the examiner rejected the then-pending claims as anticipated in view of U.S. Patent Publication No. 2004/0208372 (*Boncyk*). Ex. 1002, 273. The applicant argued that *Boncyk* did not disclose steps (c) or (d) of the claims:

- “(c) automatically identifying a list of likely pointed-to objects from the digital image at the different location to return the list of likely pointed-to objects”
- “(d) returning the list of likely pointed-to objects to the user to select one of the likely pointed-to objects”

Ex. 1002, 282-84.

The examiner subsequently issued a Notice of Allowance, noting that the “improvement comprises, while identifying an object, automatically identifying a list of likely pointed-to objects from the digital image and returning the list of likely pointed-to objects to the user to select one of the likely pointed-to objects as the identified objects.” Ex. 1002, 310.

C. Claim Construction

Claim 1 includes a limitation that invokes “means for” language consistent with 35 U.S.C. §112(6): “at least one actuation means for actuation by the user when the user points the pointing and identification device at the object.” Lopresti, ¶¶83-85. The function of the “actuation means” is actuation by the user when the user points the pointing and identification device at the object, as claimed. As claim 1 further recites, “when the user actuates the at least one actuation means,” “a digital camera [forms] a digital image of the object or a portion of the object” and “a communication device [communicates] the digital image to a different location.”

The corresponding structure described in the ’812 patent is a button or its equivalent, such as “mouse button 112” or “laser button.” ’812 patent, 5:22-23, 5:41-42, 5:64-66. As the specification explains, “the user clicks to select [an] item,” then “the PID communicates with a remote system to obtain a list of likely pointed-to objects from which the user can select.” ’812 patent, 6:30-33; *see also id.*, 3:52-56.

For the purposes of this proceeding only, no other terms require construction.

III. THE PRIOR ART

A. Oami

Oami was published and publicly accessible as of at least September 16, 2005, and is, therefore, prior art under pre-AIA 35 U.S.C. § 102(a). Gluckman, ¶¶33-55. Moreover, documents from the SPIE Digital Library, the United States Copyright Office, and the Internet Archive demonstrate that *Oami* was widely available to a POSITA using available search tools. Gluckman, ¶¶33-55. *Oami* was not considered during prosecution.

Oami discloses capturing an image of an object with a camera, transmitting that image to a remote server for identification of match candidates for the object, receiving from the server a list of candidates, and selecting a candidate.

Oami discloses that “the user takes a picture of the unknown object by a camera-equipped mobile phone” (with a keypad that can be used to take the photo). *Oami*, § 2.2. *Oami* further discloses using the “internet connection functionalities of mobile phones” to transmit the taken picture “to a remote MML [(Mobile Multimedia Library)] server.” *Oami*, §§ 1, 2.2.

The “picture is analyzed on the MML server to identify the object in the picture” and “several object categories are derived as candidates based on the identification results.” *Oami*, § 2.2.

Oami discloses that “a web page containing the identification results is presented on the screen of the mobile phone.” *Oami*, § 2.2 *Oami* further discloses that “one of the candidates is selected by clicking one of the representative images[.]” *Oami*, § 2.2.

Figure 1, annotated below, illustrates the “camera-equipped mobile phone” (with its keypad) transmitting the taken photo to the “MML server” where it is analyzed and “candidate categories are derived” and returned to the user in the form of an array of thumbnail images:

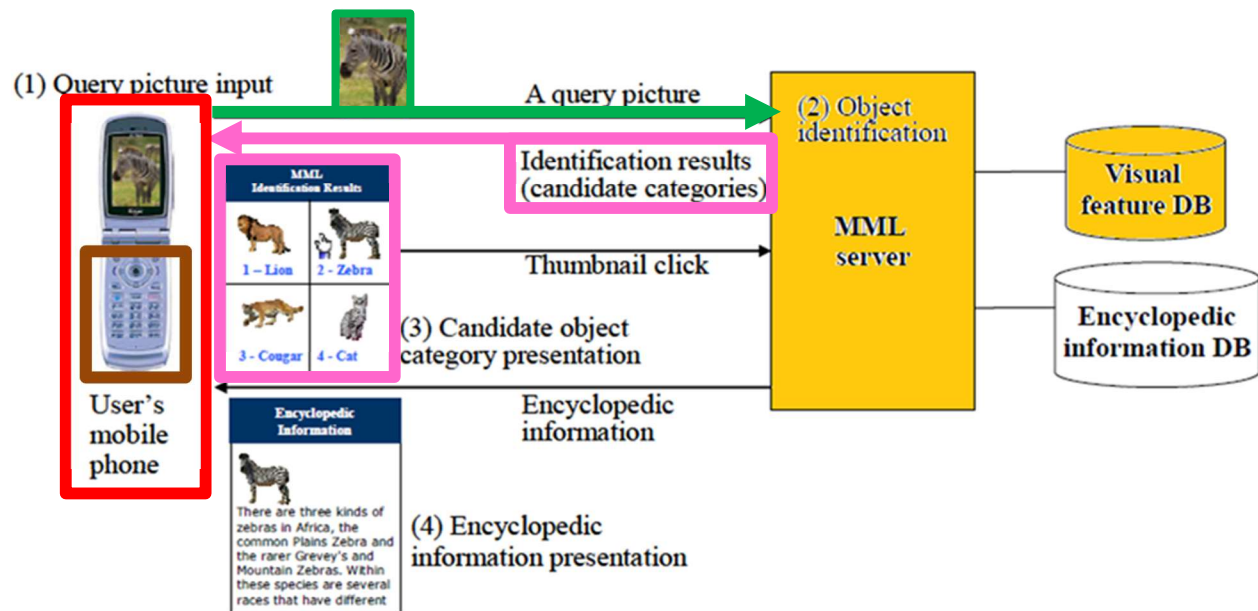


Figure 1: System overview of MML.

Oami, Fig. 1 (annotated), §2.2.

Oami discloses using its object identification system to identify, among other things, animals, flowers, buildings, product information, and price information for products in a store. *Oami*, §§ 2.1, 4.2, 6.

Oami is analogous art at least because it is in the same field of endeavor as the '812 patent—object recognition. Lopresti, ¶67; '812 patent, 4:30-37; *Oami*, Abstract, §§ 2.2, 3. *Oami* is also reasonably pertinent to the particular problem the '812 patent is concerned with because it deals with technologies relating to recognizing and obtaining information about a selected object. Lopresti, ¶67; *Oami*, Abstract; §§ 2.2, 3.

B. Du

Du was filed as a U.S. Patent Application on December 31, 2003, and published on September 8, 2004, and is, therefore, prior art under pre-AIA § 102(a). *Du* was not considered during prosecution.

Du discloses capturing an image of an object using a camera and sending the captured image to a remote server for analysis and identification of the item. Notably, *Du* further discloses presenting a user with likely matches from which to select when different possible identifications exist.

Du discloses that “a consumer may use the camera of a mobile telephone to capture an image of a desired item.” *Du*, [0006]. *Du* discloses that its “mobile telephone” has a keypad, which can be used to capture the image. *Du*, Fig. 4A. Figure 4A, as annotated below, illustrates the “mobile phone 400 having a camera 413” as well as a keypad:

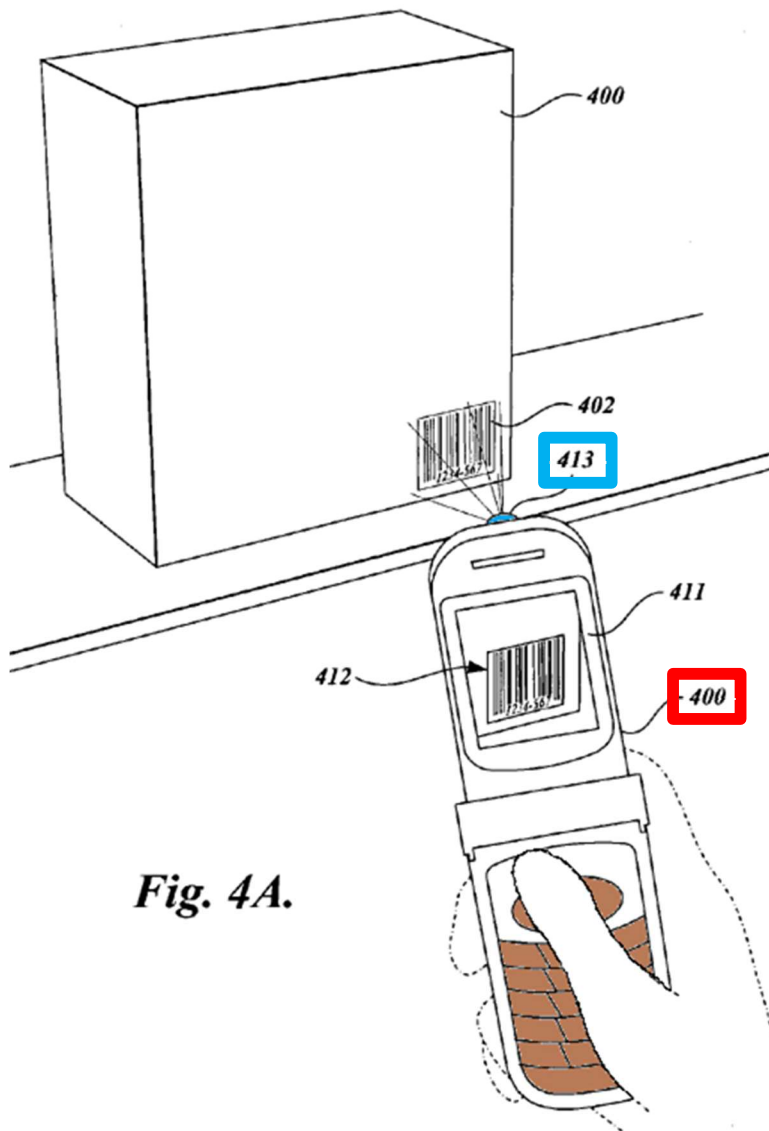
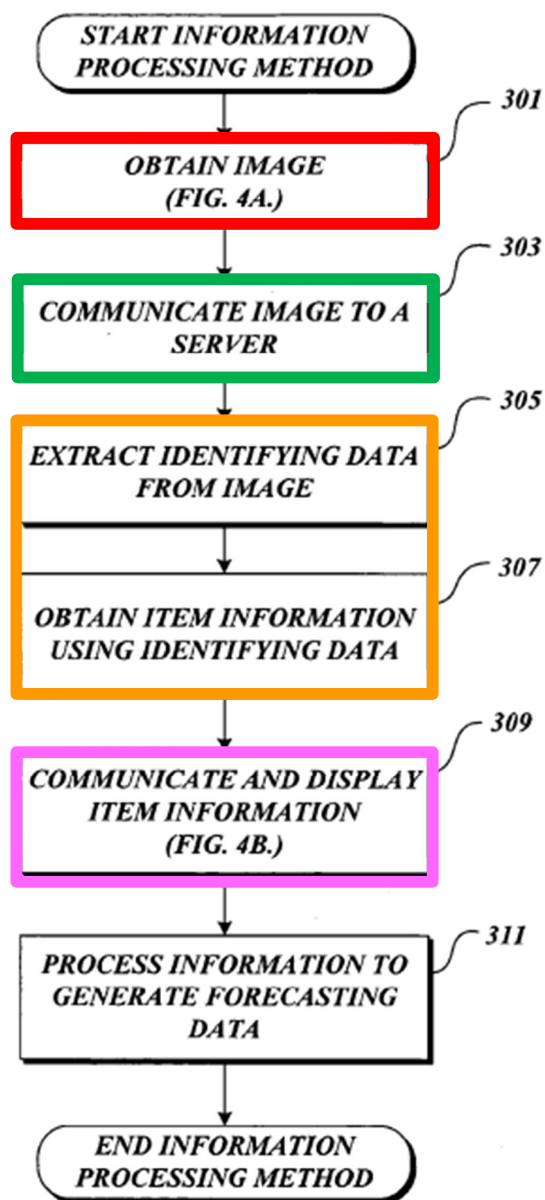


Fig. 4A.

Du, Fig. 4A (annotated); *Du*, [0029].

Du further discloses that the “imaging device . . . communicates the captured image to the server 103.” *Du*, [0019]. The image may include, for example, a barcode, “alphanumeric data,” text, or “a graphic design, such as a symbol or trademark.” *Du*, [0034]. Once at the server, “data identifying the desired item is extracted from the image.” *Du*, [0006]. At times, “conflicting identifying data is

obtained.” *Du*, [0034]. In such cases, *Du* discloses that “the user of the **imaging device may be polled to determine the correct data**[.]” *Du*, [0034]. Figure 3, as annotated below, illustrates these disclosures:



Du, Fig. 3 (annotated).

Du discloses an embodiment in which the user is “provided with a control that allows the user to immediately purchase the selected item from the online source.” *Du*, [0044].

Du is analogous art at least because it is in the same field of endeavor as the ’812 patent—object recognition. Lopresti, ¶74; ’812 patent, 4:30-37; *Du*, [0006]. *Du* is also reasonably pertinent to the particular problem the ’812 patent is concerned with because it deals with technologies relating to recognizing and obtaining information about a selected object. Lopresti, ¶74; *Du*, [0006].

C. Darrell

Darrell was filed as a U.S. Patent Application on January 22, 2004, and published on July 28, 2005, and is, therefore, prior art under pre-AIA § 102(a). *Darrell* was not considered during prosecution.

Darrell discloses that its “invention relates . . . particularly to systems and techniques to identify a location or object in a person’s field of view.” *Darrell*, [0001]. *Darrell* discloses that by using a “hand held device 10 with a camera 12, an image-based query can be formed simply by pointing with the camera 12 and snapping a photo.” *Darrell*, [0024], *see Darrell*, [0031]. According to *Darrell*, “the handheld device 10 includes the camera 12 to capture an image and a wireless communication device 14, . . . to communicate the image with existing data base

25 to find similar images. *Darrell*, [0026]; see *Darrell*, [0021], claims 1-3 and 5-6.

Figure 1, annotated below, illustrates these disclosures:

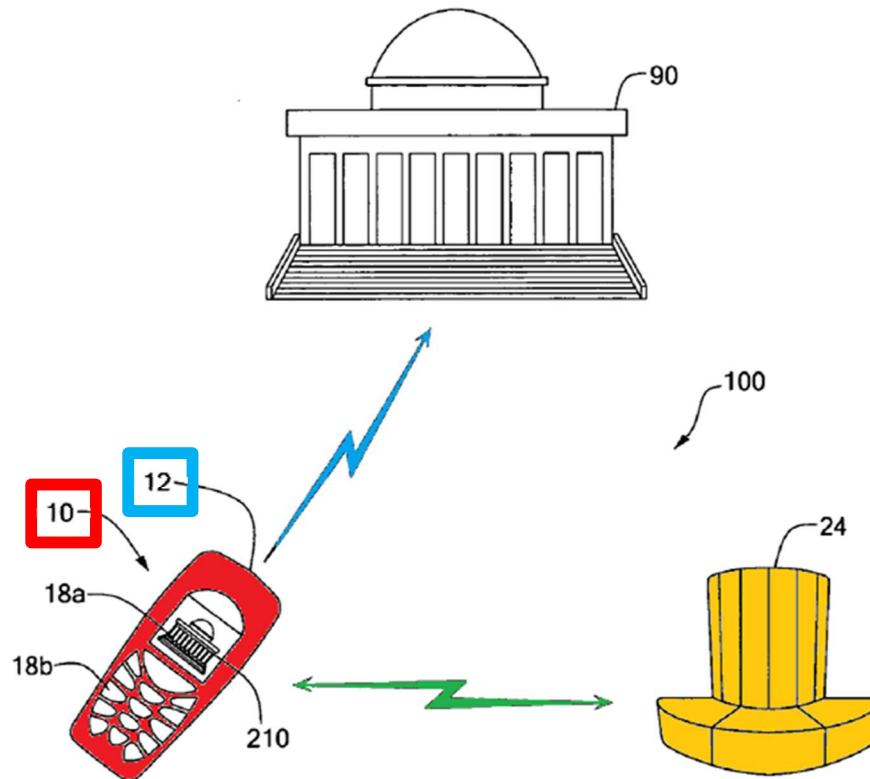


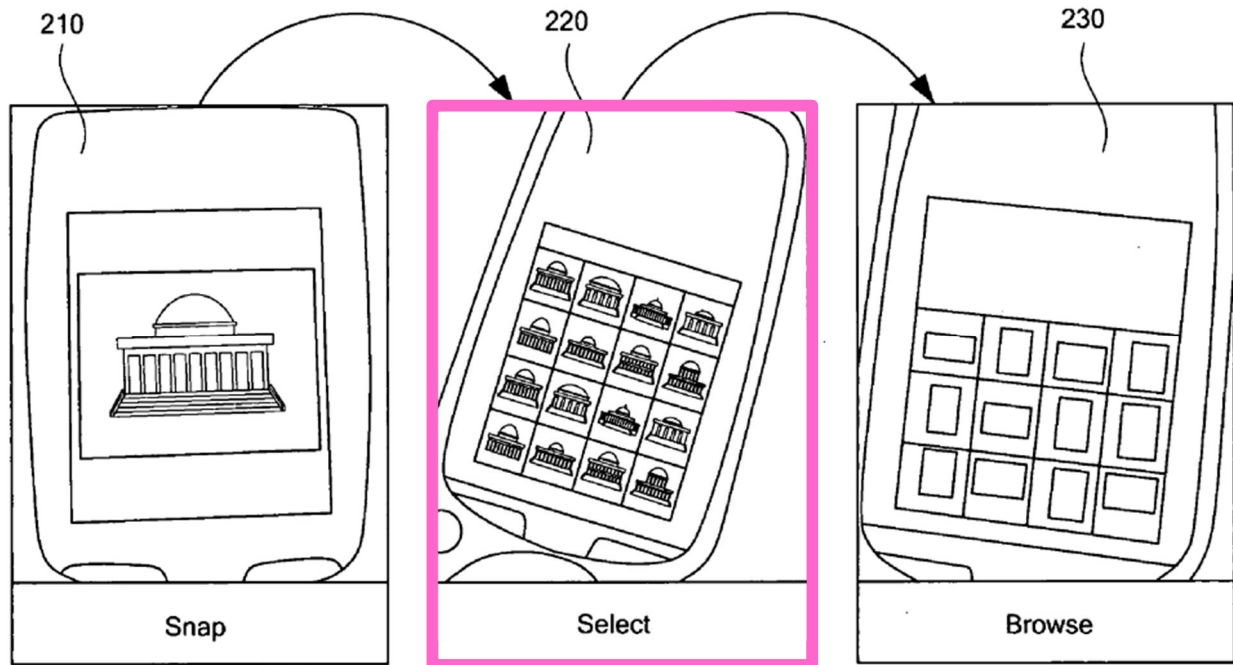
FIG. 1

Darrell, Fig. 1 (annotated).

Darrell further discloses that “the handheld device 10 communicates the captured image to a web server 24 to find images similar to the captured image.”

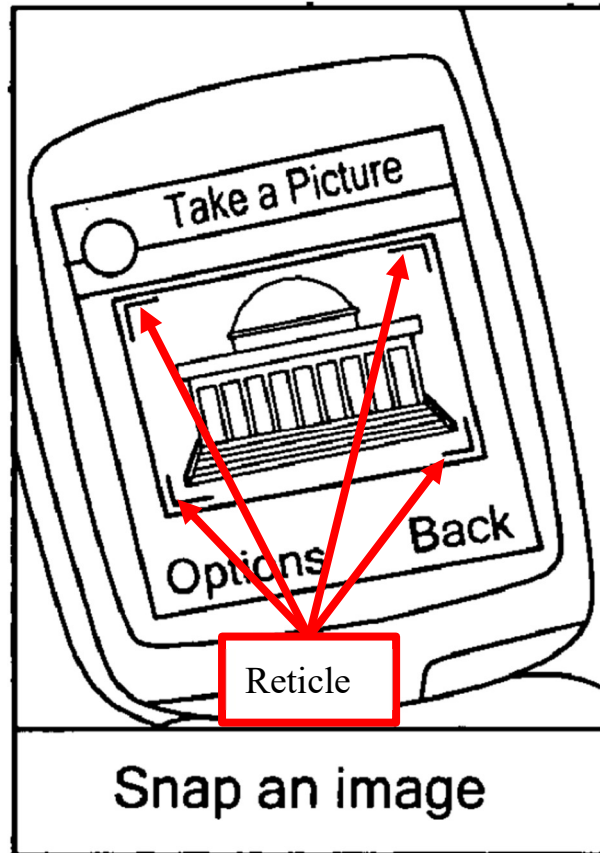
Darrell, [0028]; see, e.g., *Darrell*, [0021], [0024], [0026]. *Darrell* further discloses that the “mobile device 10 displays the result by arranging candidate images into a thumbnail mosaic 220 (FIG. 2).” *Darrell*, [0025]. The “user, as the final judge, can easily identify what sub-images are ‘really relevant.’” *Darrell*, [0025]. Figure 2,

annotated below, shows the “thumbnail mosaic 22” from which the user can
“easily identify what sub-images are ‘really relevant’”:



Darrell, Fig. 2 (annotated), [0025].

Darrell further discloses using a reticle to aid in capturing an image, as shown in the following figure:



Darrell, Fig. 3 (annotated).

Darrell is analogous art at least because it is in the same field of endeavor as the '812 patent—object recognition. Lopresti, ¶79; '812 patent, 4:30-37; *Darrell*, Abstract, [0001], [0006]. *Darrell* is also reasonably pertinent to the particular problem the '812 patent is concerned with because it deals with technologies relating to recognizing and obtaining information about a selected object. Lopresti, ¶79; *Darrell*, Abstract, [0006]-[0007].

IV. STATEMENT OF PRECISE RELIEF REQUESTED

Petitioner requests review and cancellation of the challenged claims based on the following prior art and grounds.

Prior Art References	
	<i>Oami</i> (Ex. 1005), <i>Mobile Multimedia Library: an MPEG-7 application with camera-equipped mobile phones</i> , published and publicly accessible as of at least September 16, 2005 (pre-AIA § 102(a)).
	<i>Du</i> (Ex. 1006), U.S. Patent Pub. 2005/0198095, filed December 31, 2003, and published September 8, 2005 (pre-AIA § 102(a)).
	<i>Darrell</i> (Ex. 1007), U.S. Patent Pub. 2005/0162523, filed January 22, 2004, published July 28, 2005 (pre-AIA § 102(a)).

Grounds of Unpatentability	
1	<i>Oami</i> alone or in view of <i>Darrell</i> and/or <i>Du</i> renders obvious claims 1-12 under Section 103.
2	<i>Du</i> in view of <i>Darrell</i> renders claims 1-12 obvious under Section 103.

V. LEVEL OF ORDINARY SKILL

A POSITA would have had at least two years of education or experience in the field of computer science or computer vision. Lopresti, ¶¶82. More education can supplement practical experience and vice versa. Lopresti, ¶¶82. Petitioner's expert met this standard by 2005. Lopresti, ¶¶4-35.

VI. CLAIMS 1-12 ARE UNPATENTABLE OVER THE PRIOR ART

A. Ground 1: *Oami* alone or in combination with *Darrell* and further in view of *Du* renders obvious claims 1-12

1. Claim 1

1[pre]: A method for identifying an object, the method comprising:

Oami discloses 1[pre]. Lopresti, ¶¶117-19. For example, *Oami* describes a method of identifying an object using a mobile phone and a “Mobile Multimedia Library (MML)” that “allows users to get information of an unknown object anywhere and anytime.” *Oami*, Abstract. In *Oami*’s method, a “user simply takes a picture of the object by a mobile phone and sends it directly to an MML server on which the picture is analyzed to identify the object.” *Oami*, Abstract.

1[a]: (a) providing a pointing and identification device for pointing at the object, the pointing and identification device comprising:

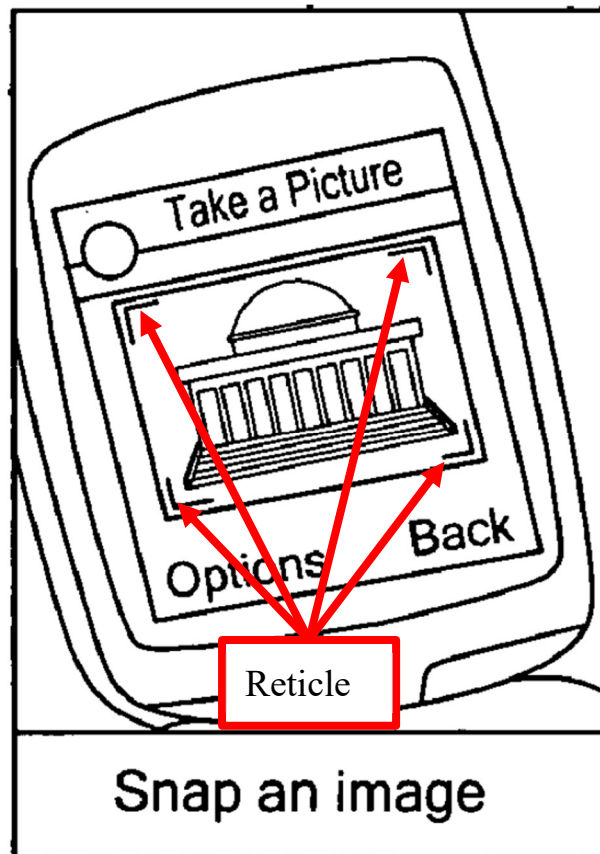
Oami discloses or in view of *Darrell* renders obvious 1[a]. Lopresti, ¶¶120-25. In *Oami*, “camera-equipped mobile phones” are used for pointing at target objects and capturing images. *Oami*, Abstract.

The ’812 patent states that the “pointing and identification device . . . includes a digital camera and one or both of a laser and a reticle for aiming the digital camera.” ’812 patent, Abstract; *see* ’812 patent, 5:8-23.

To the extent the claimed “pointing and identification device” requires “one or both of a laser and a reticle for aiming the digital camera,” *Oami* in view of

Darrell renders obvious 1[a]. For example, it would have been obvious to combine *Oami*'s object recognition system with a reticle, as taught by *Darrell*. Lopresti, ¶¶ 123, 108-116.

Darrell discloses a camera reticle in the form of four reticle corners, as shown in the following image:



Darrell, Fig. 3 (annotated). *Darrell*'s reticle is, like the reticle described in the '812 patent, "any means that the user sees in a viewfinder to designate a spot in the view or a subarea of the view." '812 patent, 9:9-19. A POSITA would have recognized that a reticle was a common aiming aid in digital cameras at the time of the alleged invention. Lopresti, ¶110; *see, e.g., Darrell*, Fig. 3; *McCune*, §§ 8.2-9.2 (disclosing

a mobile phone that uses a crosshair reticle to aid the user in pointing at an object).

Additionally, a POSITA would have known that a motivation existed to use aiming aids—such as reticles or laser pointers—to assist digital cameras in identifying pointed-at objects. Lopresti, ¶110; *see, e.g., Darrell*, Fig. 3; *McCune*, §§ 8.2-9.2; *Bove*, 2, 6 (disclosing a system that uses “a small digital camera” in cooperation with “a laser pointer” to identify an object the laser is pointed at).

A POSITA would have been motivated to incorporate *Darrell*’s reticle with *Oami*’s object recognition system. Lopresti, ¶111. First, a POSITA would have recognized that a reticle allows a user to more precisely identify an object in the camera’s frame. Lopresti, ¶111. For example, a POSITA would have recognized that an image might contain more than one object. Lopresti, ¶111. In *Oami* itself, the illustrative image—Figure 1—contains two animals and foliage.



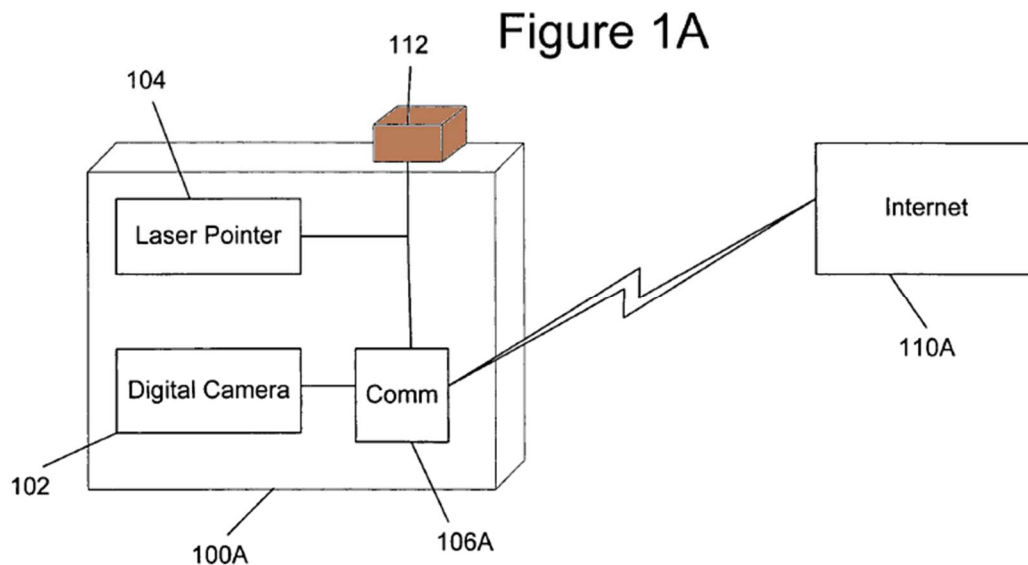
Oami, Fig. 1. A POSITA would have thus been aware that multiple objects could exist in an image containing the target object. Lopresti, ¶113. Further, a POSITA would have recognized this as a potential problem, as *Oami*'s object recognition system can only return four likely matches in a format that remains useful for the user. *See Oami*, § 4.2. A POSITA would have recognized that using a reticle—such as *Darrell*'s—would allow *Oami*'s object recognition system to narrow its image processing focus and return more relevant results. Lopresti, ¶114. For example, by using *Darrell*'s reticle to point at a target object, *Oami*'s object recognition system could ignore some of the other objects contained in the camera's frame to return more accurate results. Lopresti, ¶114.

Applying *Darrell*'s known reticle to *Oami*'s known camera-equipped mobile phones would yield predictable results, including aiding the user in aiming at a desired target object. Lopresti, ¶115. Moreover, *Oami* and *Darrell* would naturally occur to a POSITA as complementary, combinable technologies. Both references relate to object recognition using a camera-equipped device. Lopresti, ¶115; *supra* §§ III.A (*Oami*) and III.C (*Darrell*). Furthermore, both references address the same problem: recognizing and obtaining information about a selected object. Lopresti, ¶115; *supra* §§ III.A (*Oami*) and III.C (*Darrell*). A POSITA would have had a reasonable expectation of success in making this combination given the similarities of the references and the commonplace nature of the technology at the time of the alleged invention. Lopresti, ¶116. Further, a POSITA would have recognized that modifying *Oami*'s camera-equipped mobile phone to include, on its screen, a reticle as disclosed in *Darrell* would have required nothing more than the simple modification of computer code, a task well within the POSITA's ability at the time of the alleged invention. Lopresti, ¶116.

**1[b]: at least one actuation means for actuation by the user
when the user points the pointing and identification device
at the object;**

Oami discloses 1[b]. Lopresti, ¶¶125-31. For example, *Oami* discloses that the "camera equipped mobile phone" has a keypad. *Oami*, Fig. 1, §2.1.

Oami discloses both the structure and function of the “actuation means” described in the ’812 patent or the equivalent thereof. *See supra* § II.C. Regarding structure, the ’812 patent describes a button, such as “mouse button 112,” that “the user clicks to select [an] item.” ’812 patent, 6:30-33; *see also id.*, 3:52-56. The ’812 patent’s **mouse button 112** is illustrated in Figure 1A below:



’812 patent, Fig. 1A (annotated). Similarly, *Oami*’s camera-equipped mobile phone has a **keypad with buttons** that, like the mouse button 112 in the ’812 patent, is actuated by the user to select a target object that will be communicated to the MML to obtain match candidates. *Oami*, § 2.2. Figure 1, as annotated below, illustrates *Oami*’s buttons on a keypad:



Oami, Fig. 1 (annotated).

A POSITA would have known that the keypad could be used to take the picture once the camera is pointed at a target object. Lopresti, ¶130. Thus, a POSITA would have recognized the keypad as an *actuation means for actuation by the user when the user points the pointing an identification device at the object*.

Lopresti, ¶130.

1[c]: a digital camera for forming a digital image of the object or of a portion of the object when the user points the pointing and identification device at the object and actuates the at least one actuation means; and

Oami discloses 1[c]. Lopresti, ¶¶132-34. For example, *Oami* discloses the “camera equipped on recent mobile phones.” *Oami*, § 1; see ’812 patent, 2:49-52.

Oami discloses that “the user takes a picture of the unknown object by a camera-equipped mobile phone.” *Oami*, § 2.2. Further, “tak[ing] a picture of the unknown object” is *forming a digital image of the object or of a portion of the object*. *Oami*, § 2.2; Lopresti, ¶133. Additionally, a POSITA would have known that taking a picture using a “camera-equipped mobile phone” could and would involve pointing a camera and pressing a button (*points the pointing and identification device at the object and actuates the at least one actuation means*). Lopresti, ¶133; *see Oami*, § 2.2.

1[d]: a communication device for communicating the digital image to a different location when the user actuates the at least one actuation means;

Oami discloses 1[d]. Lopresti, ¶¶135-37. For example, *Oami* discloses “internet connection functionalities of mobile phones” to upload a picture “to a remote MML server.” *Oami*, §§ 1, 2.2. This is consistent with the ’812 patent’s description of using “a wireless communication component 106A” to communicate information regarding the captured image to “a computer or Internet 110A.” ’812 patent, 5:8-23. *Oami* explains that “[o]bject identification is computationally intensive” and that “it is quite difficult to perform it only using the resources on mobile phones.” *Oami*, § 1. Thus, *Oami* discloses relying “on a server in place of resource-restricted mobile phones.” *Oami*, § 1. Thus, after the user “takes a picture

of the unknown object,” the “picture is uploaded to a remote MML server[.]”

Oami, § 2.2.

Oami further discloses uploading the picture “to a remote MML server” when the user takes the picture by pressing a button (*when the user actuates the at least one actuation means*). *Oami*, §§ 2.1-2.2; Lopresti, ¶136. *Oami* discloses that its “information retrieval application . . . provides information from a picture of an object taken by a camera on mobile phones” and that “[t]aking a picture of a target object is the only required operation.” *Oami*, § 2.1. Thus, once a user takes the photo, the “MML application” automatically uploads the picture “to a remote MML server[.]” *Oami*, § 2.2; Lopresti, ¶136.

1[e]: (b) communicating the digital image to the different location;

Oami discloses 1[e]. Lopresti, ¶¶138-140. For example, *Oami* discloses that, after the user “takes a picture of the unknown object,” the “picture is **uploaded to a remote MML server[.]**” *Oami*, § 2.2. Figure 1, annotated below, highlights this communication of the image to the remote server (*different location*):

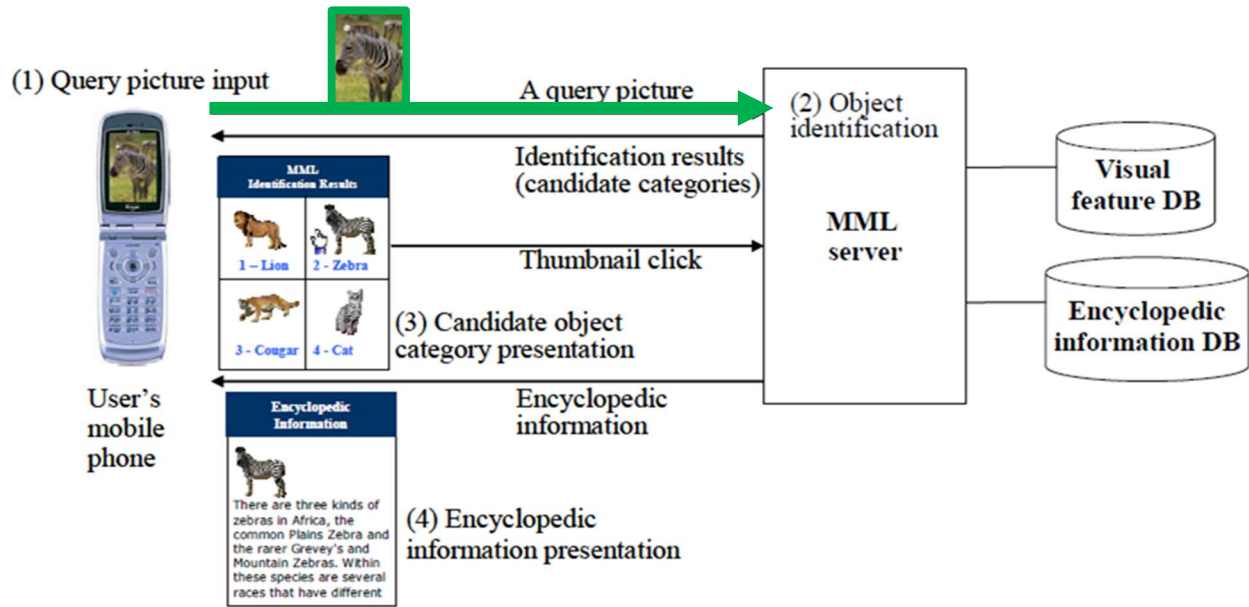


Figure 1: System overview of MML.

Oami, Fig. 1 (annotated).

1[f]: (c) **automatically identifying a list of likely pointed-to objects** from the digital image at the different location to return the list of likely pointed-to objects; and

Oami discloses or in view of *Du* renders obvious 1[f]. Lopresti, ¶¶141-50.

For example, *Oami* discloses that “the query picture is analyzed on the MML server to identify the object in the picture” and, as a result, “**several object categories are derived as candidates based on the identification results.**” *Oami*, § 2.2.

Oami discloses that after the “picture is uploaded to a remote MML server as a query,” “the query picture is analyzed on the MML server” (at the different location) “to identify the object in the picture” (*identifying a list of likely pointed-to objects from the digital image*). *Oami*, § 2.2. *Oami* discloses analyzing color,

texture, and shape of the object and comparing those features to a database of “training images” (i.e., “visual feature DB”) to determine the closest matches.

Oami, §§ 2.2, 3.2-3.4. Figure 1, as annotated below, illustrates this step:

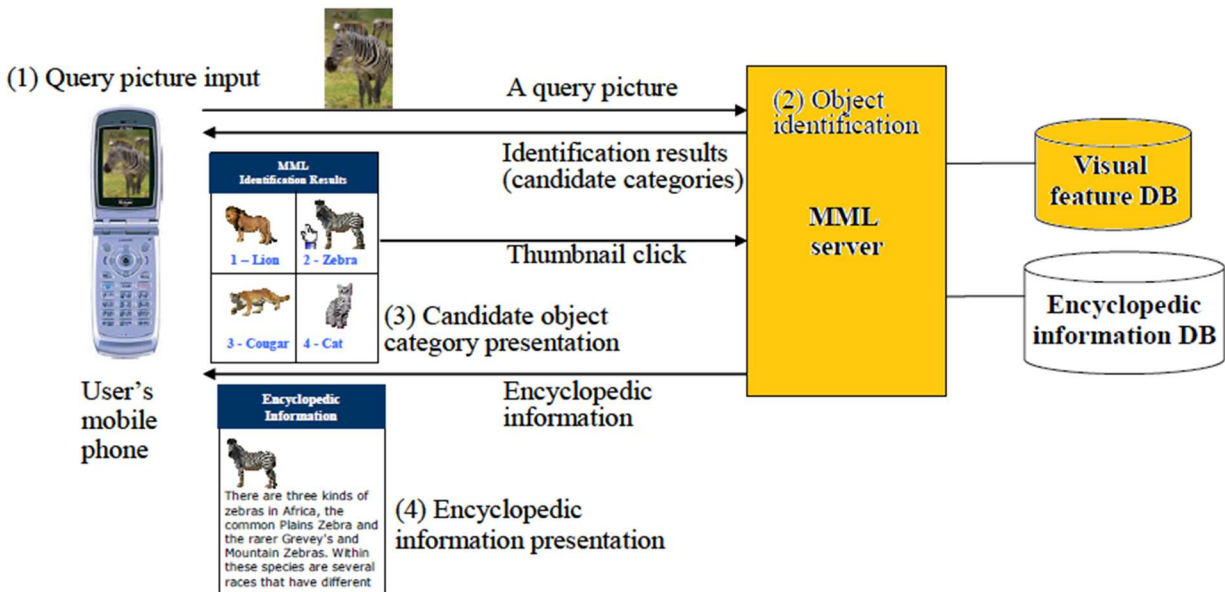


Figure 1: System overview of MML.

Oami, Fig. 1 (annotated). As a result of this identification process, “several object categories are derived as candidates based on the identification results.” *Oami*, § 2.2. These “object categories” are *likely pointed-to objects*. Lopresti, ¶143. The object categories are for the MML server to return to the mobile phone (to return the list of likely pointed-to objects), as explained *infra* § VI.A.1 (element 1[g]).

Oami discloses that this process happens *automatically*, without requiring further user input. Lopresti, ¶144. Again, *Oami* discloses that its “information retrieval application . . . provides information from a picture of an object taken by a camera on mobile phones” and that “[t]aking a picture of a target object is the

only required operation.” *Oami*, § 2.1. Thus, once a user takes the photo, the “MML application” uploads the picture “to a remote MML server” for analysis and identification of “candidates based on the identification results.” *Oami*, § 2.2; Lopresti, ¶144.

Alternatively, it would have been obvious in further view of *Du* to determine which object in an image the user wants to obtain further information on. Lopresti, ¶145. A POSITA would have recognized that an image might contain more than one object. Lopresti, ¶145. In *Oami* itself, the illustrative image—Figure 1—contains two animals and foliage.



Oami, Fig. 1.

A POSITA would have thus been aware that multiple objects could exist in an image containing the target object. Lopresti, ¶146. A POSITA would have recognized this as a potential problem, as *Oami*'s object recognition system can only return four likely matches in a format that remains useful for the user. *See Oami*, § 4.2. A POSITA would have recognized that polling the user when more than one possible target object is detected in the image—as taught by *Du*—would allow *Oami*'s object recognition system to narrow its image processing focus and return more relevant results. Lopresti, ¶146.

In *Du*, “after the image is captured” and “communicated to a server,” the server may determine that the image “contain[s] multiple instances of identifying data,” such as both “a valid UPC” and “text describing a known item.” *Id.*, [0032], [0034]. (*Du* also envisions an image including other identifying data, such as “a graphic design, such as a symbol or trademark.” *Du*, [0034].) When such “conflicting identifying data is obtained, the user of the imaging device may be polled to determine the correct data[.]” A POSITA would have understood such “poll[ing]” of the user to determine which identifying data (*e.g.*, the UPC code or the text) the user intended to capture to teach *identifying a list of likely pointed-to objects from the digital image at the different location to return the list of likely pointed-to objects*.

From *Du*, a POSITA would have been motivated and found it obvious to augment *Oami*'s process to poll a user when the MML server identifies conflicting identifying data in an image. Lopresti, ¶148. First, a POSITA would have recognized that such polling would ensure the MML server retrieves "identification results" for the user's intended item in the image. Lopresti, ¶148.

Applying *Du*'s polling technique to *Oami*'s known camera-equipped mobile phones would yield predictable results, including presenting the user with the potential target objects and receiving a selection from the user. Lopresti, ¶149. By using *Du*'s polling technique to allow the user to identify the intended object within an image, *Oami*'s object recognition system could ignore some of the other objects contained in the camera's frame to return more accurate results. Lopresti, ¶149. Moreover, *Oami* and *Du* would naturally occur to a POSITA as complementary, combinable technologies. Both references relate to object recognition using a camera-equipped device. Lopresti, ¶149; *supra* §§ III.A (*Oami*) and III.B (*Du*). Furthermore, both references address the same problem: recognizing and obtaining information about a selected object. Lopresti, ¶149; *supra* §§ III.A (*Oami*) and III.C (*Darrell*). A POSITA would have had a reasonable expectation of success in making this combination given the similarities of the references and the commonplace nature of the technology at the time of the alleged invention. Lopresti, ¶149. Further, a POSITA would have recognized that

augmenting *Oami*'s object identification process to include polling the user when multiple possible target objects are detected would have required nothing more than the simple modification of computer code, a task well within the POSITA's ability at the time of the alleged invention. Lopresti, ¶149.

1[g]: (d) returning the list of likely pointed-to objects to the user to select one of the likely pointed-to objects;

Oami discloses 1[g]. Lopresti, ¶¶151-155. For example, *Oami* discloses “a web page containing the identification results is presented on the screen of the mobile phone.” *Oami*, § 2.2.

Oami discloses sending the “first four candidate categories” identified by the MML server back to the user's mobile phone as an array of four images. *Oami*, §§ 2.2, 4.2. Sending the “first four candidate categories” to the user's mobile phone is *returning a list of likely pointed-to objects to the user*. Lopresti, ¶152. Figure 1, as annotated below, illustrates sending the user the “first four candidate categories”:

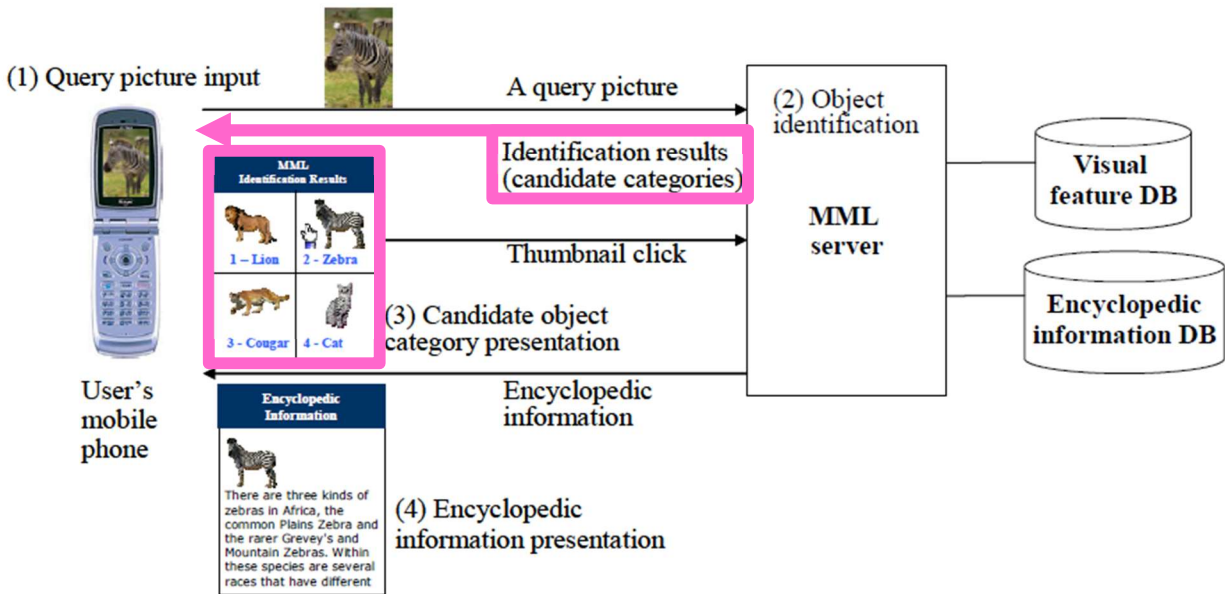


Figure 1: System overview of MML.

Oami, Fig. 1 (annotated), §4.2. *Oami* further discloses that the “identification results” are “presented on the screen of the mobile phone” in a format that “contains representative images corresponding to each candidate object category.” *Oami*, § 2.2.

Oami further discloses that the “one of the candidates is selected by clicking one of the representative images[.]” *Oami*, § 2.2. This disclosure of the user selecting one of the “identification results” by “clicking one of the representative images” is the user *select[ing] one of the likely pointed-to objects*. Lopresti, ¶154. *Oami* discloses that when a user selects “one of the candidates . . . by clicking one of the representative images,” “encyclopedic information about the clicked object category is displayed on the user’s mobile phone.” *Oami*, §§ 2.2, 4.2.

1[h]: wherein the object is at least one of a spot on a displayed image on a display, a subarea of a space on the displayed image on the display, one of a plurality of objects in the displayed image on the display, an object in space, or near an object in space, a subarea of a surface in space, or one of a plurality of objects in space.

Oami discloses or renders obvious 1[h]. Lopresti, ¶¶156-61. For example, *Oami* discloses taking pictures of animals, flowers, buildings, product information, and logos. *Oami*, §§ 2.1, 4.2, 6.

During prosecution the examiner identified this final “wherein” clause as a Markush group. Ex. 1002, 182, 203. When a limitation contains a Markush group, “the entire element is disclosed by the prior art if one alternative in the Markush group is in the prior art.” *Fresenius USA, Inc. v. Baxter Int’l, Inc.*, 582 F.3d 1288, 1298 (Fed. Cir. 2009).

An animal, flower, or building is at least *an object in space, near an object in space, or one of a plurality of objects in space*. Lopresti, ¶157. In *Oami*’s exemplary image, the object is a zebra (*an object in space*), but the image also includes plant life and another zebra (*near an object in space; one of a plurality of objects in space*):



Oami, Fig. 1 (excerpt).

Oami further discloses using “the MML framework” as a “[p]roduct information retrieval system.” *Oami*, § 6. *Oami* discloses, in this system, taking a picture of “a logo of a store.” *Oami*, § 6. A logo is, among other things, *a subarea of a surface in space*. Lopresti, ¶158.

A POSITA would have understood *Oami* to also render obvious *wherein the object is at least one of a spot on a displayed image on a display, a subarea of a space on the displayed image on the display, [or] one of a plurality of objects in*

the displayed image on the display. A POSITA would have recognized that *Oami*'s disclosure in no way limits the types of target objects the user may capture, process, and receive information on. Lopresti, ¶¶86-99, 159. Rather, *Oami* discloses hardware, image-processing, and user feedback techniques that a POSITA would have found broadly applicable and modifiable to fit various applications. Lopresti, ¶¶86. For example, *Oami* discloses using "camera-equipped mobile phones," and a POSITA would have recognized that this broadly refers to well-known camera hardware and software equipped in any "camera-equipped mobile phone." *Oami*, Abstract; Lopresti, ¶¶86. Further, *Oami* discloses an image processing algorithm that analyzes color, texture, and shape vectors extractable from an image without any limitation on the type of image captured. *Oami*, §§ 3.1-3.5, Appendix; Lopresti, ¶¶86. Additionally, *Oami* explicitly recognizes that its disclosure "is widely applicable" to "many application examples," including "[f]lower identifying systems," "[b]uilding/monument guide systems," "[m]useum guide systems," "[p]roduct information retrieval systems with logo identifier," and "[n]avigation systems to a nearest store with logo identifier," *Oami*, § 6, and *Oami*'s disclosure is in no way limited to objects in the real world. Lopresti, ¶¶86. Rather, a POSITA would recognize that *Oami*'s disclosure would naturally be applicable to capturing and processing images of objects on a screen (*on a display*). Lopresti, ¶¶91. These images of objects *on a display* would include images

of, among other things, *a spot on a displayed image on a display, a subarea of a space on the displayed image on the display, and one of a plurality of objects in the displayed image on the display.* Lopresti, ¶158-60.

Even assuming *Oami* did not already disclose the object being “on a display,” this would have been obvious given *Oami*’s disclosure. Because *Oami*’s disclosure in no way limits the types of target objects the user may capture, process, and receive information, *Oami* suggests and teaches to a POSITA that its disclosure is amenable to modification, and a POSITA would have been motivated to modify or otherwise adjust *Oami*’s disclosures. Lopresti, ¶91. For example, a POSITA would have been motivated to use *Oami* to process images captured from a screen because it would broaden the utility of *Oami* to capture and analyze images of objects in a context in which a user is likely to encounter them. Lopresti, ¶91. *Oami*’s disclosure makes it easy for a user to obtain information on a “target object” by simply taking a picture of it. *Oami*, § 2.1. A POSITA would have recognized that people encounter objects on screens frequently. Lopresti, ¶91; *see Bove*, Abstract, 8-10 (disclosing a system for allowing a user to point to and identify objects in a video displayed on a screen). Thus, a POSITA would have been motivated to apply *Oami*’s disclosure to capture and process images of objects displayed on a screen as well as objects in the real world. Applying *Oami*’s known image processing techniques to known camera-equipped mobile phones

would yield predictable results, including allowing the camera-equipped mobile phones to form and process images of any target object. Lopresti, ¶92.

Oami even suggests applying its disclosure to two-dimensional images, similar to images on a screen. For example, *Oami* discloses analyzing an image of “paintings,” which would be comparable to analyzing an image of a screen or display. *Oami*, § 6; Lopresti, ¶93. *Oami* also discloses analyzing an image of a “logo.” *Oami*, § 6; Lopresti, ¶93. Analyzing an image taken of a screen or display would have been obvious to a POSITA given *Oami*’s disclosure.

A POSITA would have had a reasonable expectation of success in extending *Oami*’s disclosure to allow a user to capture and process images displayed on a screen. Lopresti, ¶99. For example, *Oami*’s disclosure of its “algorithm for object identification” is applicable to images displayed on a screen similar to images of objects from the real world. Lopresti, ¶99. A POSITA wanting to analyze an image of an object displayed on a screen would analyze the same color, texture, and shape features that *Oami* describes. Lopresti, ¶99. Moreover, capturing images of objects on a screen was well known at the time of the alleged invention. Lopresti, ¶¶99, 160; *see, e.g., Boncyk*, [0273]. A POSITA would have had a reasonable expectation of success in modifying *Oami*’s disclosure to capture objects that are *a spot on a displayed image on a display, a subarea of a space on the displayed*

image on the display, and one of a plurality of objects in the displayed image on the display. Lopresti, ¶¶99, 160.

2. Claim 2

2[a]: The method of claim 1, wherein step (c) comprises: (i) receiving a designation of the object as designated by the user;

Oami further discloses 2[a]. Lopresti, ¶¶162-64. For example, *Oami* discloses “a target object” that is designated by a user. *Oami*, § 2.1.

In *Oami*, the user targets an object by pointing the camera-equipped mobile phone’s camera at the object. *Oami*, § 2.1; Lopresti, ¶163. The phone’s camera, thus, *receives a designation of the object as designated by the user.* Lopresti, ¶163.

2[b]: (ii) recording a frame image comprising the designated object;

Oami further discloses 2[b]. Lopresti, ¶¶165-68. For example, *Oami* discloses “tak[ing] a picture of the unknown object by a camera-equipped mobile phone.” *Oami*, § 2.1.

When *Oami* discloses that “the user takes a picture of the unknown object,” it discloses that its system is *recording a frame image comprising the designated object.*

Oami further discloses that the taken “picture is uploaded to a remote MML server as a query.” *Oami*, § 2.1. A POSITA would have known that to upload the picture, the picture must have been recorded. Lopresti, ¶167.

**2[c]: (iii) determining a context type of the frame image
from the frame image; and**

Oami further discloses 2[c]. Lopresti, ¶169-73. For example, *Oami* discloses that its system “extracts various visual features from the query picture” and analyzes color, texture, and shape. *Oami*, § 4.1.

The '812 patent explains that the “context type” of an image is the source of the image, for example a “computer screen, game box screen, TV screen, or view of the real world.” '812 patent, 10:64-66. The '812 patent further describes determining a context type by, among other methods, “recognizing the TV’s screen outline or other elements of the TV, computer monitor, etc.” '812 patent, 11:30-34. A POSITA would have known that recognizing such things as a TV screen outline or “other elements of the TV” in an image would involve analyzing the color, shape, and texture of objects in the image. Lopresti, ¶170.

Oami discloses extracting from the image and analyzing eight distinct “visual features”: “the color layout, the color structure, the dominant color,” “the scalable color,” “the edge histogram,” “the homogeneous texture,” “the contour shape, and the region shape.” *Oami*, § 4.1. *Oami* further discloses that its system can be used to “retrieve information of paintings.” *Oami*, § 6. A POSITA would have known that to identify an object in an image as a painting would involve extracting and identifying features indicative of a painting, such as a frame. Lopresti, ¶171. This would be *determining* that the *context-type* is a painting.

Lopresti, ¶171. A POSITA would further have recognized that this disclosure is just as applicable to determine if a taken image is of, for example, a TV screen or a computer monitor. Lopresti, ¶171. Thus, *Oami* discloses *determining a context type of the frame image from the frame image*. Lopresti, ¶171.

2[d]: (iv) identifying the object.

Oami further discloses 2[d]. Lopresti, ¶¶174-175. For example, *Oami* discloses that “the MML server analyzes the query image to identify the object a user wants to know.” *Oami*, § 7; *see supra* § VI.A.1 (element 1[f]).

3. Claim 3: The method of claim 1, further comprising receiving a selection by the user of said one of the likely pointed-to objects as the object and an instruction from the user to obtain information on the object.

Oami further discloses claim 3. Lopresti, ¶176-80. For example, *Oami* discloses that “one of the candidates is selected by clicking one of the representative images, and the encyclopedic information about the clicked object category is displayed on the user’s mobile phone.” *Oami*, § 2.2.

Oami discloses presenting candidate object categories of *likely pointed-to objects* as an array of thumbnail images on the user’s mobile phone. This is illustrated, for example, in Figure 4(a):

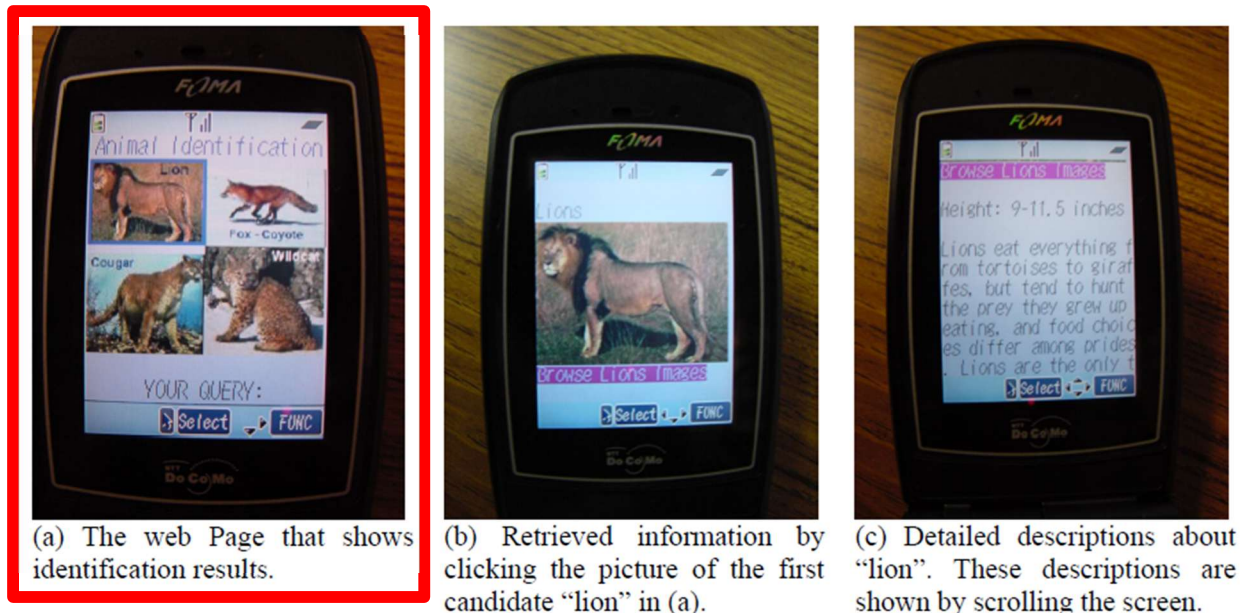


Figure 4: Screen shots presented to a user by the prototype system.

Oami, Fig. 4 (annotated).

Oami further discloses that the user *selects* one of the *likely pointed-to objects* by “clicking one of the representative images.” Oami, § 2.2. Oami further discloses that “just one click action is enough” to select one of the candidates. Oami, § 4.2. This selection by click is illustrated in Figure 1, as annotated below:

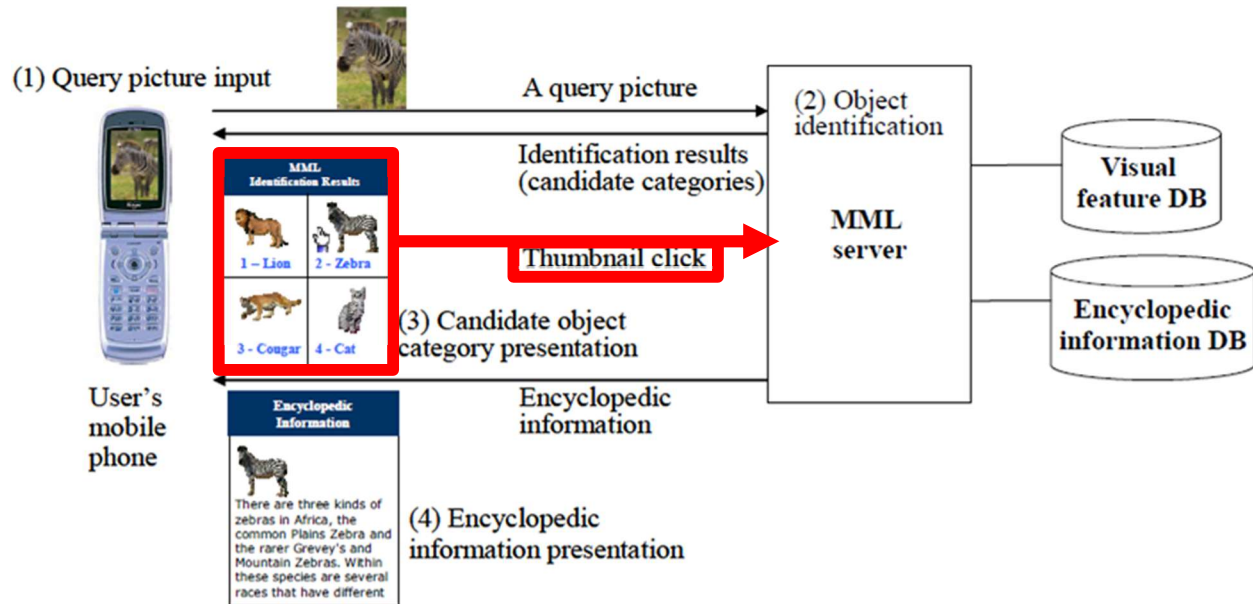


Figure 1: System overview of MML.

Oami, Fig. 1 (annotated).

Oami discloses that by selecting a candidate image the user is sending *an instruction to obtain information on the object*. Lopresti, ¶179. *Oami* discloses that “just one click action is enough to show the information.” *Oami*, § 4.2. According to *Oami*, when “one of the candidates is selected by clicking one of the representative images,” then “the encyclopedic information about the clicked object category is displayed on the user’s mobile phone.” *Oami*, § 2.2. Figures 1 and 4, as illustrated below, illustrate the response to this *instruction to obtain information on the object*:

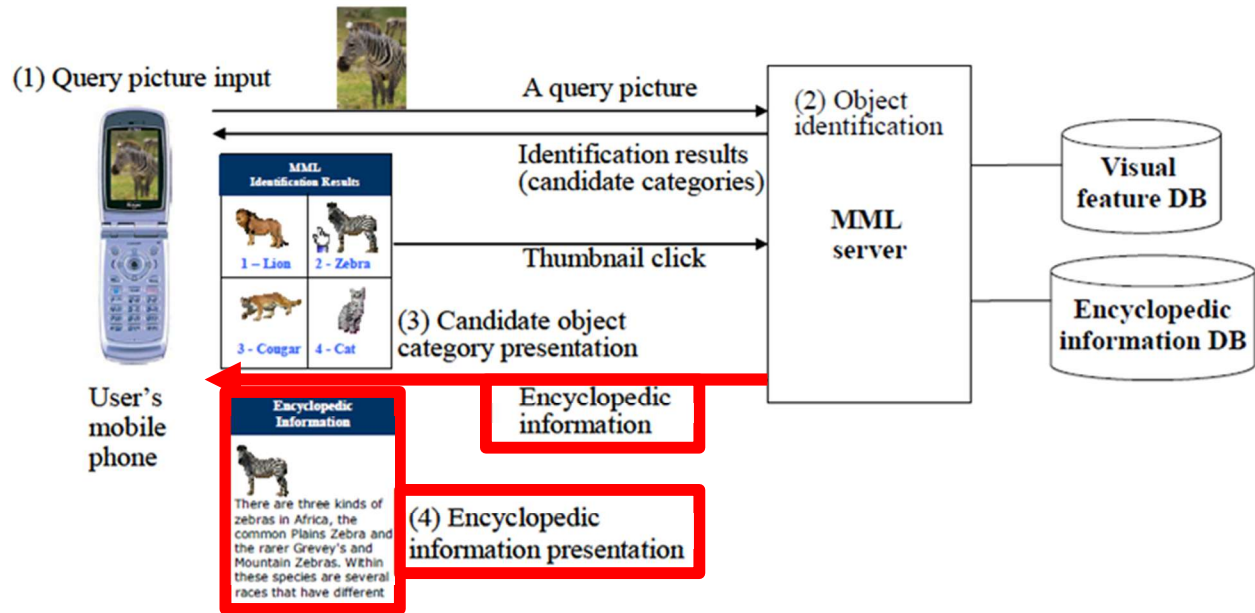


Figure 1: System overview of MML.

Oami, Fig. 1 (annotated).

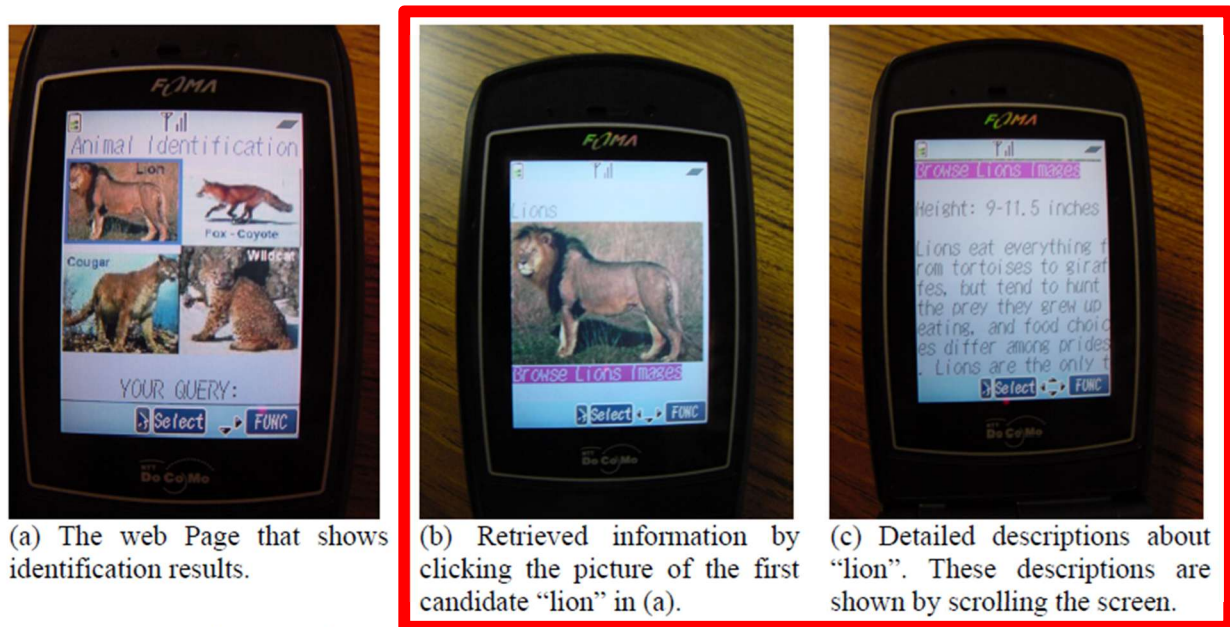


Figure 4: Screen shots presented to a user by the prototype system.

Oami, Fig. 4 (annotated).

4. Claim 4: The method of claim 1, further comprising receiving a selection by the user of said one of the likely pointed-to objects as the object and an instruction from the user to purchase the object.

Oami alone or in view of *Du* renders obvious claim 4. Lopresti, ¶181-86.

Oami discloses *receiving a selection by the user of said one of the likely pointed-to objects as the object. See supra § VI.A.3 (claim 3). Oami renders obvious an instruction from the user to purchase the object.*

Oami is not limited to the exemplary animal identification and further discloses that “the MML framework can effectively be applied” to “product information retrieval systems.” *Oami*, § 6. *Oami* discloses that “one of the candidates is selected by clicking one of the representative images” and using the system to obtain “product information such as price[.]” *Oami*, § 2.2, 6. The “product information retrieval systems” will, for example, take “a logo of a store” in order “to access its product information such as price for each product sold there or electric coupons can be retrieved.” *Oami*, § 6.

It would have been obvious for a POSITA to modify *Oami* to receive *an instruction from the user to purchase the object*. Lopresti, ¶184, 100-102. *Oami* itself notes that the “framework of the MML is widely applicable,” and a POSITA would have been motivated to apply the MML framework commercially. Lopresti, ¶100. *Oami* even implies the commercial applicability of the MML framework by describing the “product information retrieval systems” that retrieve “price” data for

products. Lopresti, ¶100; *Oami*, § 6. Recognizing the commercial applicability suggested in *Oami*, a POSITA would have been motivated to modify the disclosure and provide the user with an option to purchase an object for which *Oami*'s system retrieved the price for the user. Lopresti, ¶100.

Applying *Oami*'s known price-retrieval and user-feedback solicitation techniques to known camera-equipped mobile phones would yield predictable results, including allowing the camera-equipped mobile phones to receive and process an instruction from a user to purchase an object. Lopresti, ¶100-101. A POSITA would also have had a reasonable expectation of success in modifying *Oami* to allow a user to send an instruction to purchase an object, as it discloses the necessary information and functionality to achieve this modification. Lopresti, ¶102. For example, *Oami* discloses that the "product information retrieval system" can give the user the "price for each product sold" at a given store. *Oami*, § 6. *Oami* further discloses the user's ability to select a desired object and to send its selection to the MML server. *Oami*, § 2.2, 4.2.

Giving users an option to purchase an object captured and analyzed using image-processing technology was well known at the time of the alleged invention. *See, e.g., Du*, [0044].

To the extent *Oami* does not render this limitation obvious on its own, it would have been obvious to a POSITA to modify *Oami* to include *Du*'s teaching

of providing the user “with a control that allows the user to immediately purchase the selected item from the online source.” *Du*, [0044]; Lopresti, ¶103-107. *Du* discloses that “when item information is communicated back to the user’s mobile device, the user may be provided with a control that allows the user to immediately purchase the selected item from the online source.” *Du*, [0044]. Modifying *Oami* in this manner would have allowed a user to both identify and purchase the newly identified object using a mobile device. Lopresti, ¶103.

Oami and *Du* would naturally occur to a POSITA as complementary, combinable technologies. Both references relate to object recognition using a camera-equipped device. Lopresti, ¶104; *supra* §§ III.A (*Oami*) and III.B (*Du*). Furthermore, both references address the same problem: recognizing and obtaining information about a selected object. Lopresti, ¶104; *supra* §§ III.A (*Oami*) and III.B (*Du*).

Oami itself suggests reasons for this modification. As explained above, *Oami* notes that the “framework of the MML is widely applicable,” including to retrieving “price” data for products, and a POSITA would have been motivated to apply the MML framework commercially. Lopresti, ¶105; *Oami*, § 6. Recognizing the commercial applicability suggested in *Oami*, a POSITA would have been motivated to modify the disclosure and provide the user with an option to purchase an object for which *Oami*’s system retrieved the price for the user. Lopresti, ¶105.

Du provides explicit disclosure of this next step. Thus, a POSITA would be motivated for at least these reasons to incorporate *Du*'s teaching of allowing a mobile device user to purchase an identified object. Lopresti, ¶105. Applying *Du*'s known purchasing techniques to *Oami*'s known camera-equipped mobile phones would yield predictable results, including allowing the camera-equipped mobile phones to *receiv[e] ... an instruction from the user to purchase an object*. Lopresti, ¶185.

A POSITA would have had a reasonable expectation of success in combining *Du*'s teaching of allowing a user to “purchase the selected item from the online source” with *Oami*'s framework of allowing a user to send a selection to the server in the context of “product information retrieval systems.” Lopresti, ¶107; *Du*, [0044]. As explained above, *Oami* itself discloses the functionality and information necessary to allow a mobile device user to purchase an identified object. Lopresti, ¶107. Further, *Du* itself recognizes that it requires only the application of well-known technology to achieve this combination. *Du* discloses that “[e]xisting software configured to execute electronic commerce purchase transactions can be used to implement such a feature.” *Du*, [0044]. *Du* further recognizes that its invention can be implemented using “one of many standard image recognition and processing programs” and “[e]xisting software configured to execute electronic commerce purchase transactions.” *Du*, [0006], [0044]. Thus,

a POSITA would have been able to incorporate *Du*'s teaching with *Oami*'s system with a reasonable expectation of success. Lopresti, ¶107.

5. Claim 5: The method of claim 1, wherein the object is at least one of the spot on the displayed image on the display, the sub-area of the space on the displayed image on the display, and one of the plurality of objects in the displayed image on the display.

Oami renders obvious these limitations. *See* 1[h].

6. Claim 6: The method of claim 5, wherein the object is the sub-area of the space on the displayed image on the display.

Oami renders obvious these limitations. *See* 1[h].

7. Claim 7: The method of claim 1, wherein the object is a spot on a displayed image on a display.

Oami renders obvious these limitations. *See* 1[h].

8. Claim 8: The method of claim 1, wherein the object is a subarea of a space on the displayed image on the display.

Oami renders obvious these limitations. *See* 1[h].

9. Claim 9: The method of claim 1, wherein the object is one of a plurality of objects in the displayed image on the display.

Oami renders obvious these limitations. *See* 1[h].

10. Claim 10: The method of claim 1, wherein the object is an object in space, or near an object in space.

Oami discloses these limitations. *See* 1[h].

11. Claim 11: The method of claim 1, wherein the object is a subarea of a surface in space.

Oami discloses these limitations. *See* 1[h].

12. Claim 12: The method of claim 1, wherein the object is one of a plurality of objects in space.

Oami discloses these limitations. *See* 1[h].

B. Ground 2: *Du* in view of *Darrell* renders obvious claims 1-12

1. Claim 1

Element 1[pre]:

Du discloses 1[pre]. Lopresti, ¶227-28. For example, *Du* discloses that “a consumer may use the camera of a mobile telephone to capture an image of a desired item,” “communicate” the image “to one or more servers,” and then extract data “identifying the desired item . . . from the image.” *Du*, [0006].

Element 1[a]:

Du discloses or in view of *Darrell* renders obvious 1[a]. Lopresti, ¶¶229-38. In *Du*, “a **mobile device 400** having a camera 413 is used to capture an image 412 of identifying data 402 of an item 401.” *Du*, [0029]; *see infra* § VI.B.1 (element 1[f]). Figure 4A, as annotated below, shows the mobile device:

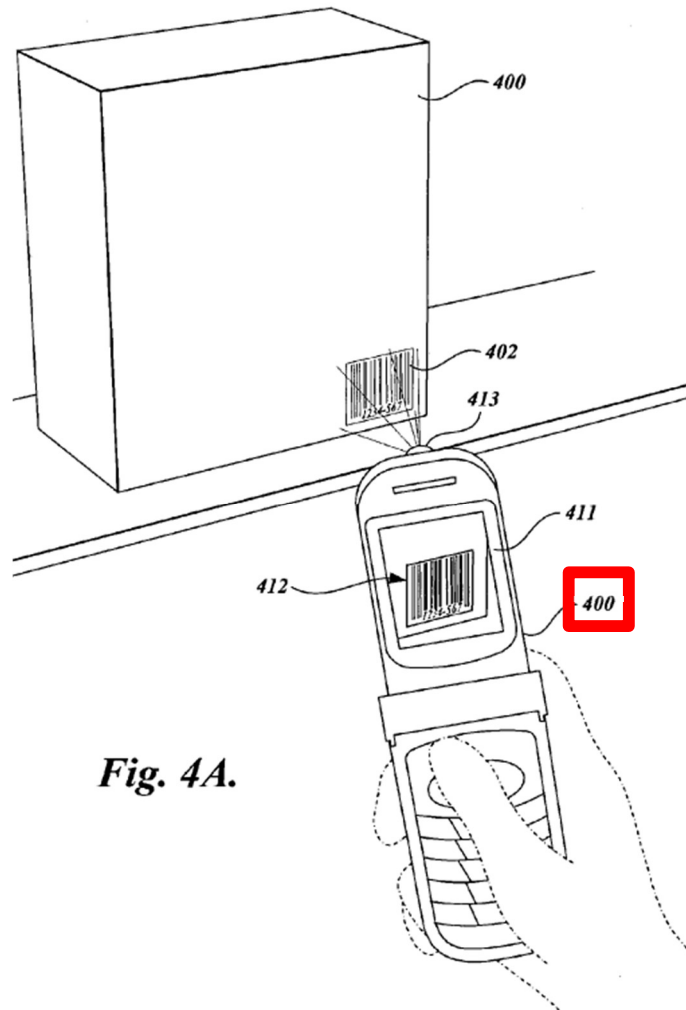


Fig. 4A.

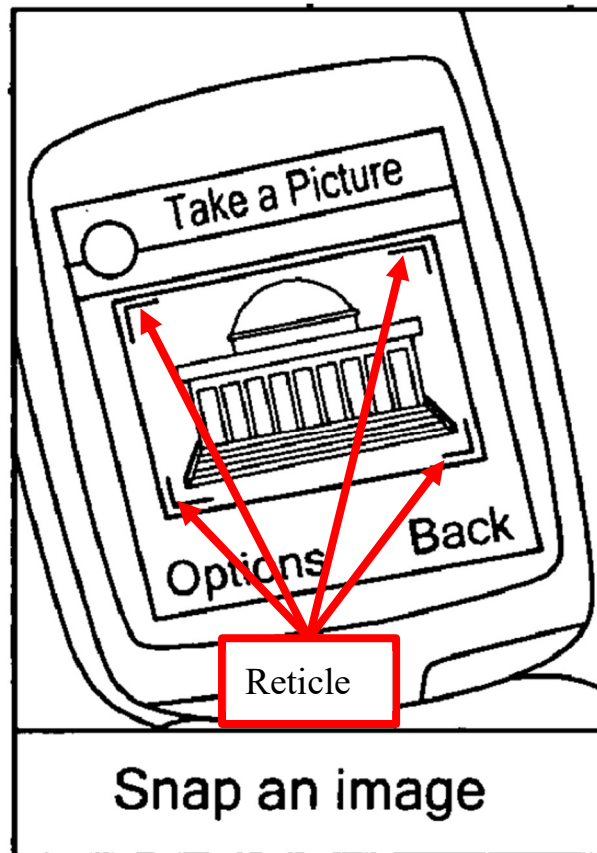
Du, Fig. 4A (annotated).

The '812 patent states that the “pointing and identification device . . . includes a digital camera and one or both of a laser and a reticle for aiming the digital camera.” '812 patent, Abstract; *see* '812 patent, 5:8-23.

To the extent the claimed “pointing and identification device” requires “one or both of a laser and a reticle for aiming the digital camera,” *Du* in view of *Darrell* renders obvious 1[a]. Lopresti, ¶¶220-26. For example, it would have been

obvious to combine *Du*'s camera-equipped mobile device with a reticle, as taught by *Darrell*. Lopresti, ¶234.

Darrell discloses a camera reticle in the form of four reticle corners, as shown in the following image:



Darrell, Fig. 3 (annotated). *Darrell*'s reticle is, like the reticle described in the '812 patent, "any means that the user sees in a viewfinder to designate a spot in the view or a subarea of the view." '812 patent, 9:9-19. A POSITA would have recognized that a reticle was a common aiming aid in digital cameras at the time of the alleged invention. Lopresti, ¶¶237, 220-26; *see, e.g., Darrell*, Fig. 3; *McCune*, §§ 8.2-9.2 (disclosing a mobile phone that uses a crosshair reticle to aid the user in pointing at

an object). Additionally, a POSITA would have known that a motivation existed to use aiming aids—such as reticles or laser pointers—to assist digital cameras in identifying pointed-at objects. Lopresti, ¶223; *see, e.g., Darrell*, Fig. 3; *McCune*, §§ 8.2-9.2; *Bove*, 2, 6 (disclosing a system that uses “a small digital camera” in cooperation with “a laser pointer” to identify an object the laser is pointed at).

A POSITA would have been motivated to incorporate *Darrell*’s reticle with *Du*’s camera-equipped mobile phone. Lopresti, ¶224. First, a POSITA would have recognized that a reticle allows a user to more precisely identify an object in the frame of a camera. Lopresti, ¶224. As *Du* recognizes, an image may include “conflicting identifying data,” such as both a UPC bar code and unrelated text. *Du*, [0034]; Lopresti, ¶224.

A POSITA would have thus been aware that multiple objects could exist in an image containing the intended identifying data. Lopresti, ¶226. A POSITA would have recognized that using a reticle—such as *Darrell*’s—would allow *Du*’s system to narrow its image processing focus to the identifying data (*e.g.*, UPC bar code or text) placed within the reticle and to return more relevant results. Lopresti, ¶226. For example, by using *Darrell*’s reticle to point at identifying data, *Du*’s system could ignore some of the other identifying data contained in the camera’s frame to return more accurate results. Lopresti, ¶224.

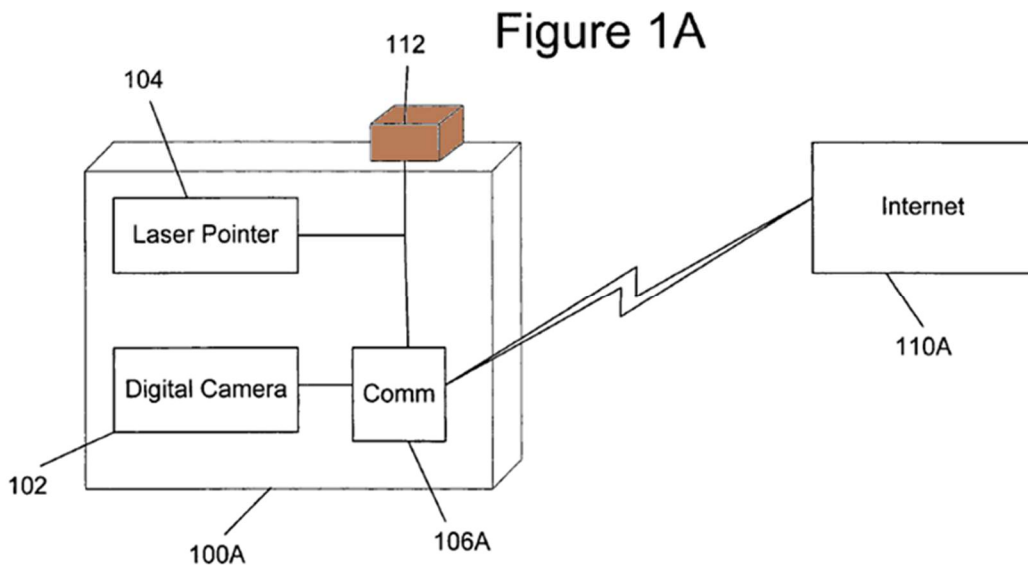
Applying *Darrell*'s known reticle to *Du*'s known camera-equipped mobile phones would yield predictable results, including aiding the user in aiming at a desired target object. Lopresti, ¶225. Moreover, *Du* and *Darrell* would naturally occur to a POSITA as complementary, combinable technologies. Both references relate to object recognition using a camera-equipped device. Lopresti, ¶216; *supra* §§ III.B (*Du*) and III.C (*Darrell*). Furthermore, both references address the same problem: recognizing and obtaining information about a selected object. Lopresti, ¶216; *supra* §§ III.B (*Du*) and III.C (*Darrell*). A POSITA would have had a reasonable expectation of success in making this combination given the similarities of the references and the commonplace nature of the technology at the time of the alleged invention. Lopresti, ¶226. Further, a POSITA would have recognized that modifying *Du*'s camera-equipped mobile phone to include, on its screen, a reticle as disclosed in *Darrell* would have required nothing more than the simple modification of computer code, a task well within the ability of a POSITA at the time of the alleged invention. Lopresti, ¶226.

Element 1[b]:

Du discloses 1[b]. Lopresti, ¶239-44. For example, *Du* discloses a keypad on its “mobile device 400.” *Darrell*, [0021].

Du discloses both the structure and function of the “actuation means” described in the '812 patent or the equivalent thereof. *Supra* § II.C. Regarding

structure, the '812 patent describes a button, such as “mouse button 112,” that “the user clicks to select [an] item.” '812 patent, 6:30-33; *see also id.*, 3:52-56. The '812 patent's mouse button 112 is illustrated in Figure 1A below:



'812 patent, Fig. 1A (annotated). Similarly, *Du* discloses that “mobile device 400” includes **a keypad with buttons** that, like the mouse button 112 in the '812 patent, is actuated by the user “to capture an image of a desired item,” as illustrated in Figure 4A, annotated below. *Du*, [0006]; *see Du*, [0008], [0016].

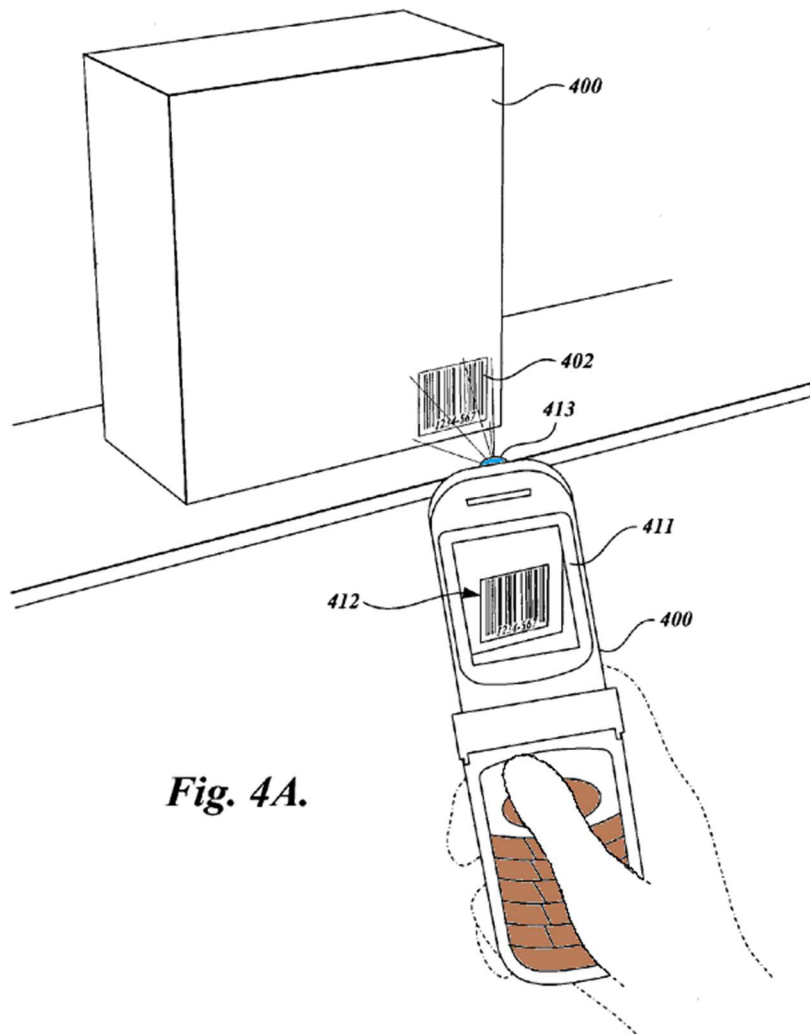


Fig. 4A.

Du, Fig. 4A (annotated).

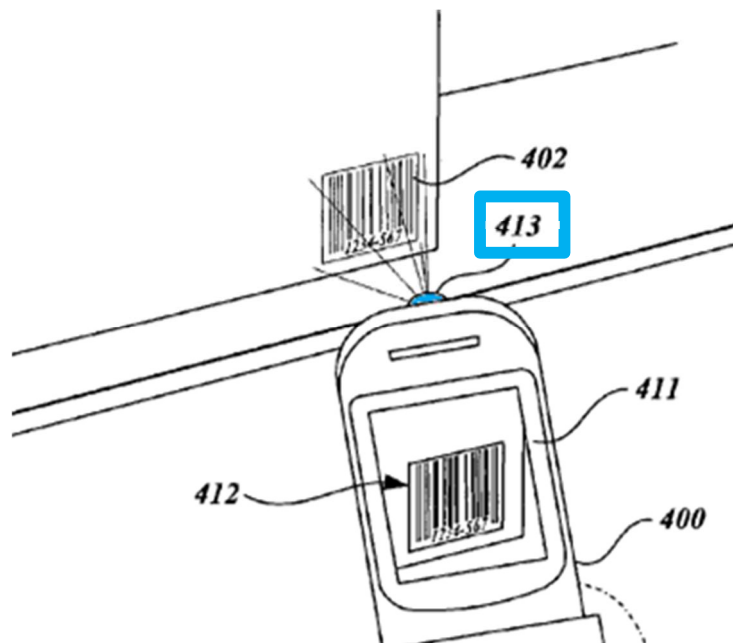
A POSITA would have known that one or more of *Du*'s keypad buttons could be used to capture an image when the camera-equipped mobile phone is pointed at an object. Lopresti, ¶242. Thus, a POSITA would have recognized the keypad with buttons as *at least one actuation means for actuation by the user when the user points the pointing and identification device at the object*. Lopresti, ¶243; *Du*, Fig. 4A (depicting a user's thumb on the keypad while an imaging action is

portrayed); *see also Du*, [0003] (recognizing that “most” “mobile phones” have a “keypad”).

Element 1[c]:

Du discloses 1[c]. Lopresti, ¶¶245-47. For example, *Du* discloses that “a consumer may use the camera of a mobile telephone to capture an image of a desired item.” *Du*, [0006]; *see Du*, [0008], [0016], [0029].

Du discloses that “a mobile phone 400 having a camera 413 is used to capture an image 412 of identifying data 401 of an item 401.” *Du*, [0029]. As shown in the annotated figure below, the user points the camera at the object to capture the image (forming a digital image of the object or of a portion of the object when the user points the pointing and identification device at the object):



Du, Fig. 4A (annotated). A POSITA would have known that capturing an image using a “mobile phone 400 having a camera 413” could and would involve pointing a camera and pressing a button on the keypad (*forming a digital image of the object or of a portion of the object when the user ... actuates the at least one actuation means*). Lopresti, ¶246-47. The “identifying data 401” can be “information such as a Universal Product Code (UPC), text on the box of the product, indicia on the product itself, etc.” *Du*, [0006]. Thus, the captured image is *of the object or of a portion of the object*. Lopresti, ¶246-47.

Element 1[d]:

Du discloses or renders obvious 1[d]. Lopresti, ¶249-254. For example, *Du* discloses that the “imaging device . . . communicates the captured image to the server 103 (operated by an entity that is different than the local retail store).” *Du*, [0019]; *see Du*, [0006], [0008], [0016]. This is consistent with the ’812 patent’s description of using “a wireless communication component 106A” to communicate information regarding the captured image to “a computer or Internet 110A.” ’812 patent, 5:9-19.

Du discloses that its “imaging device . . . communicates the captured image to the server 103.” *Du*, [0019]. *Du* further recognizes that communicating an image to a server is well-known technology and that “communication of the image from the device to the server may be performed using any one of a number of different

communication networks or technologies.” *Du*, [0032]. *Du* notes that “[c]onsumers desiring to search for . . . information remotely must use a wireless-enabled personal digital assistant (PDA) or a mobile telephone.” *Du*, [0003]. A POSITA would have recognized that *Du*’s “imaging device” comprises *a communication device for communicating the digital image to a different location*. Lopresti, ¶250.

Du further discloses that the “captured image” is communicated “to the server 103,” which is *a different location*. Lopresti, ¶251. For example, *Du* discloses embodiments in which “the server operated by the second entity is at a location remote from the location of the first entity.” *Du*, claim 21; *see Du*, claims 11 and 34.

Du also discloses or renders obvious that the “imaging device . . . communicates the captured image to the server 103” *when the user actuates the at least one actuation means*. Lopresti, ¶¶252, 209-14. It would have been obvious to a POSITA to modify *Du* to communicate an image to a server when a user actuates the actuation means and then automatically identify a list of likely pointed-to objects from that image to return to the user. Lopresti, ¶209.

This modification would have been obvious because, a POSITA would have been motivated to minimize the steps necessary to achieve the goal of *Du*’s system: to “capture an image” of an item, extract “identifying data” from the image, and, ultimately, “obtain item information.” Lopresti, ¶210. A POSITA would have

known that it would have been unnecessary to require a user to take multiple actions to capture and send an image. Lopresti, ¶211. Indeed, *Du* suggests that a single actuation operation would both form a digital image and communicate the image because it discloses no scenario in which an image would be captured and *not* then “communicated to one or more servers.” Lopresti, ¶211. Further, *Du* is silent on requiring anything more of a user than simply “us[ing] the camera of a mobile telephone to capture an image of a desired item.” *Du*, [0006]; Lopresti, ¶211.

Applying *Du*’s known image processing techniques to *Du*’s known object recognition system would yield predictable results, including allowing the camera on the mobile telephone to automatically communicate a formed image to a server. Lopresti, ¶213.

A POSITA would have had a reasonable expectation of success modifying *Du* as described. Lopresti, ¶214. POSITA would have recognized that these modifications would have required nothing more than the simple modification of computer code, a task well within the ability of a POSITA at the time of the alleged invention. Lopresti, ¶214.

Element 1[e]:

Du discloses 1[e]. Lopresti, ¶255-57. For example, *Du* discloses that the “imaging device . . . communicates the captured image to the server 103 (operated by an entity that is different than the local retail store).” *Du*, [0019]; see *Du*,

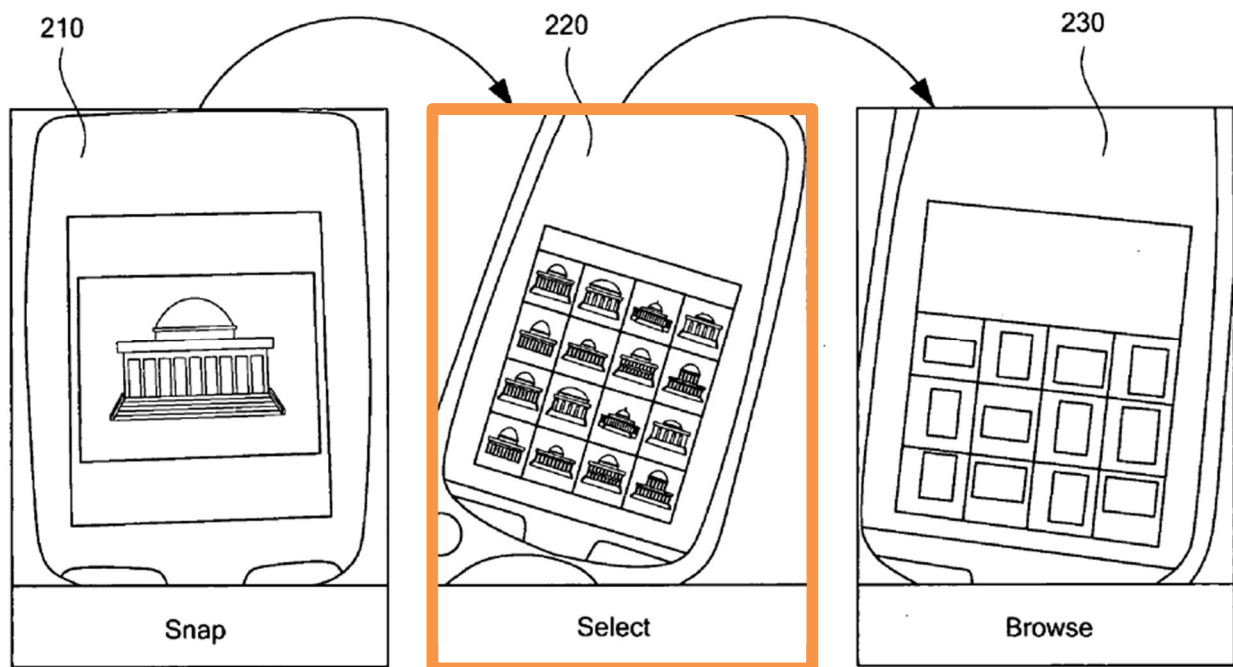
[0006], [0008], [0016]. *Du* further discloses embodiments in which the “the server operated by the second entity is at a location remote from the location of the first entity [(the local retail store)].” *Du*, claim 21; *see Du*, claims 11 and 34.

Element 1[f]:

Du discloses 1[f], or it would have been obvious from the combination of *Du* and *Darrell*. Lopresti, ¶¶258-264. For example, *Du* discloses cases in which “conflicting identifying data is obtained” and the user is “polled to determine the correct data.” *Du*, [0034]. One such case occurs when “a barcode recognition program” identifies a barcode corresponding to one product and an OCR (optical character recognition) program identifies text corresponding to a different product. *Du*, [0034]; Lopresti, ¶259. (*Du* also envisions the image including other identifying data, such as “a graphic design, such as a symbol or trademark.” *Du*, [0034].) In cases where “conflicting identifying data is obtained, the user of the imaging device may be polled to determine the correct data[.]” *Du*, [0034]. A POSITA would have known that for a user to be “polled,” *Du* would have had to identify (or it would have been obvious to identify) *a list of likely pointed-to objects* (such as the product corresponding to the barcode and the product corresponding to conflicting text). Lopresti, ¶259.

Alternatively, a POSITA would have found 1[f] obvious from the combination of *Du* and *Darrell*. Lopresti, ¶263. *Darrell* further discloses that “the

handheld device 10 communicates the captured image to a web server 24 to find images similar to the captured image.” *Darrell*, [0028]; *see, e.g., Darrell*, [0021], [0024], [0026]. *Darrell* further discloses that the “mobile device 10 displays the result by arranging candidate images into a thumbnail mosaic 220 (FIG. 2).” The “user, as the final judge, can easily identify what sub-images are ‘really relevant.’” *Darrell*, [0025]. Figure 2, as annotated below, shows the “thumbnail mosaic 22” from which the user can “easily identify what sub-images are ‘really relevant’”:



Darrell, Fig. 2 (annotated).

A POSITA would have understood such displaying of the candidate images for the user to teach *identifying a list of likely pointed-to objects from the digital image at the different location to return the list of likely pointed-to objects.*

Lopresti, ¶263.

From *Darrell*, a POSITA would have been motivated and found it obvious to augment *Du*'s process to display images corresponding to the identifying data captured in the image. Lopresti, ¶217. Where the identifying data is "a graphic design, such as a symbol or trademark" (*Du*, [0034]), for instance, *Du*'s process could be augmented to display thumbnails of images similar to the graphic design (e.g., candidate symbols or candidate trademarks) in the image, as taught by *Darrell*. Lopresti, ¶217. A POSITA would have been motivated to augment *Du*'s process in this way because, as *Darrell* explains, these sort of thumbnails of candidate images enable the user to "easily identify what sub-images are 'really relevant,'" ensuring the user is able to retrieve accurate information. Lopresti, ¶217; *Darrell*, [0025]; *Du*, [0006]. *Darrell*'s candidate images could be implemented alongside *Du*'s polling of the user, and the combination would further ensure the user is able to retrieve accurate information about something in the image. Lopresti, ¶217. For example, where an image includes both a graphic design and a barcode, the user would first be polled to determine whether the graphic design or the barcode is the identifying data intended by the user. Then, once the user selects the intended identifying data (e.g., the graphic design), the server would identify candidate images similar to the graphic design and present them to the user as thumbnails, as taught by *Darrell*. Lopresti, ¶217.

Applying *Darrell*'s known presentation of candidate images to *Du*'s known camera-equipped mobile phones would yield predictable results, including aiding the user in obtaining information for a desired target object. Lopresti, ¶218. Moreover, *Du* and *Darrell* would naturally occur to a POSITA as complementary, combinable technologies, as explained *supra* § VI.B.1 (element 1[a]). A POSITA would have had a reasonable expectation of success in making this combination given the similarities of the references and the commonplace nature of the technology at the time of the alleged invention. Lopresti, ¶218. Indeed, *Du* itself envisions that, when its server analyzes, the graphic image is “extracted and correlated with known graphic designs stored in a database to identify the selected item.” *Du*, [0034]. And *Du*'s camera-equipped mobile phone already includes a display capable of displaying images. *Du*, [0026]-[0027], [0030], Fig. 4A; *see also* [0016]. Thus, a POSITA would have recognized that modifying *Du*'s process to identify and present candidate images to the user, as disclosed in *Darrell*, would have required nothing more than the simple modification of computer code, a task well within the ability of a POSITA at the time of the alleged invention. Lopresti, ¶219.

Finally, *Du* discloses or renders obvious that its identification process occurs *automatically*, without requiring further user input. Lopresti, ¶260-61, 209-14. For example, *Du* discloses that after the consumer “capture[s] an image of a desired

item,” the “image is communicated to one or more servers, and data identifying the desired item is extracted from the image.” *Du*, [0006]. At no point does *Du* require the user to take any action beyond simply capturing the image. Thus, a POSITA would have recognized that the communication occurs *automatically*. Lopresti, 261.

Even assuming *Du* does not disclose that the identification process occurs *automatically*, it would have been obvious to modify *Du* to make this process occur automatically. A POSITA would have been motivated to do this to minimize the steps necessary to achieve *Du*’s goal: to “capture an image” of an item, extract “identifying data” from the image, and, ultimately, “obtain item information.” Lopresti, ¶209-14. Further, *Du* is also silent on requiring anything more of a user than simply “us[ing] the camera of a mobile telephone to capture an image of a desired item.” *Du*, [0006]; Lopresti, ¶212. Thus, it would have been obvious to a POSITA that *Du* discloses identifying an item *automatically* once the image of the item has been communicated to a server.

A POSITA would have recognized that the server’s analyzing of the image received from the mobile phone in *Du* would occur *automatically*. Lopresti, ¶209-14. *Du* is silent on requiring anything more of a user than simply “us[ing] the camera of a mobile telephone to capture an image of a desired item.” *Du*, [0006]; Lopresti, ¶212. Thus, it would have been obvious to a POSITA that *Du* discloses

the server analyzing the image and identifying related information *automatically* once the image has been communicated to the server. Lopresti, ¶212.

Allowing *Du*'s camera-equipped mobile phone to automatically communicate a formed image to a server where the server could identify a list of likely pointed-to objects would yield predictable results. Lopresti, ¶213. A POSITA would have had a reasonable expectation of success modifying *Du* as described. Lopresti, ¶214. POSITA would have recognized that these modifications would have required nothing more than the simple modification of computer code, a task well within the ability of a POSITA at the time of the alleged invention. Lopresti, ¶214.

Element 1[g]:

Du discloses or in view of *Darrell* renders obvious 1[g]. Lopresti, ¶¶265-69. For example, *Du* discloses cases in which “conflicting identifying data is obtained” and the user is “polled to determine the correct data.” *Du*, [0034].

As explained above, *Du* discloses using barcode recognition and OCR programs to extract “data identifying the desired item . . . from the image.” *Du*, [0006], *see Du*, [0016], [0019]; *see supra* §VI.B.1 (element 1[f]). Further, as explained above, cases exist in which the system extracts “conflicting identifying data,” for example, a barcode corresponding to one product and an OCR (optical

character recognition) program identifies text corresponding to a different product.

Du, [0034]; Lopresti, ¶266; *see supra* §VI.B.1 (element 1[f]).

When “conflicting identifying data is obtained, **the user of the imaging device may be polled to determine the correct data[.]**” *Du*, [0034]. A POSITA would have known that for a user to be “polled,” the user must be able to select between options. Lopresti, ¶267. Thus, a POSITA would have recognized that *Du* disclosed *returning, to the user, a list of likely pointed-to objects* so that the user could *select one of the likely pointed-to objects*. Lopresti, ¶267.

It would have been obvious to a POSITA to modify *Du*’s system to include *Darrell*’s teaching of returning a list of likely pointed-to objects to the user for the user to select one of the objects. Lopresti, ¶¶268, 215-19. *Darrell* discloses that “the handheld device 10 communicates the captured image to a web server 24 to find images similar to the captured image.” *Darrell*, [0028]; *see, e.g., Darrell*, [0021], [0024], [0026]. *Darrell* explains that “there is no need to look for a perfect match” because “[m]oderately good results arranged as a thumbnail mosaic . . . allows any user to swiftly identify just what images are relevant.” *Darrell*, [0024]. *Darrell* further discloses returning results to the user’s mobile device “by arranging candidate images into a thumbnail mosaic 220 (FIG. 2).” In addition, *Darrell* discloses that the “user, as the final judge, can easily identify what sub-images are ‘really relevant’” and that when “a thumbnail is selected, the source URL is

retrieved and the content from that URL is shown on the mobile device 10.”

Darrell, [0025].

Du and *Darrell* would naturally occur to a POSITA as complementary, combinable technologies. Both references relate to object recognition using a camera-equipped device. Lopresti, ¶216; *supra* §§ III.B (*Du*) and III.C (*Darrell*). Furthermore, both references address the same problem: recognizing and obtaining information about a selected object. Lopresti, ¶216; *supra* §§ III.B (*Du*) and III.C (*Darrell*).

Du suggests reasons for this modification. For example, *Du* discloses an embodiment in which “conflicting identifying data” is obtained and “the user of the imaging device may be polled to determine the correct data[.]” *Du*, [0024]. *Du*’s disclosure of a “poll[.]” suggests a number of likely pointed-to objects. Lopresti, ¶217. Applying *Darrell*’s known techniques for identifying a list of likely pointed to objects, returning the list to the user for the user to select one of the objects, and obtaining information on the user-selected object to a *Du*’s known image capture and identification system would yield predictable results. Lopresti, ¶217.

A POSITA would have had a reasonable expectation of success in combining *Darrell*’s teaching of identifying a list of likely pointed to objects and returning the list to the user for the user to select one of the objects with *Du*’s

image capture and identification framework. Lopresti, ¶218-19. As explained above, *Du* discloses the functionality and information necessary to allow a mobile device, working with a remote server, to identify a list of likely pointed to objects and return the list to the user for the user to select one of the objects. Lopresti, ¶219. Further, *Du* recognizes that its invention can be implemented using “one of many standard image recognition and processing programs.” *Du*, [0006]. Thus, a POSITA would have been able to incorporate *Darrell*’s teaching with *Du*’s system with a reasonable expectation of success. Lopresti, ¶219.

Element 1[h]:

Du discloses or renders obvious 1[h]. Lopresti, ¶¶270-75. For example, *Du* discloses taking pictures of any “identifying data,” to include “information such as a Universal Product Code (UPC), text on the box of the product, indicia on the product itself, etc.” *Du*, [0006]; *see Du*, [0029].

Du discloses that *the object* can be *an object in space, near an object in space, or one of a plurality of objects in space*. Lopresti, ¶271. For example, *Du* discloses case in which “the camera 413 may be directed to the product itself.” *Du*, [0029]. A POSITA would have recognized that a product is one or more of *an object in space, near an object in space, or one of a plurality of objects in space*. Lopresti, ¶271.

Next, *Du* discloses that *the object* can be *a subarea of a surface in space*.

Lopresti, ¶272. For example, *Du* discloses that the image can include a Universal Product Code (UPC), text on the box of the product, indicia on the product itself, etc.” *Du*, [0006]. *Du* further discloses that “the camera 413 may be directed to a label, sign, or advertising that includes identifying data of a selected item.” *Du*, [0029]. A POSITA would have recognized that a UPC code, text on a box, a label, a sign, and advertising can all be *a subarea of a surface in space*. Lopresti, ¶272.

In addition, *Du* discloses that *the object* can be *a spot on a displayed image on a display, a subarea of a space on the displayed image on the display, or one of a plurality of objects in the displayed image on the display*. Lopresti, ¶273. *Du* explains its disclosure is applicable to capture and analyze “identifying data” (such as a barcode, alphanumeric code, text, or “graphic design, such as a symbol or trademark”) on any “label, sign, or advertising that includes identifying data of a selected item.” *Du*, [0029], [0033]; Lopresti, ¶273. A POSITA would have recognized that displays in, for example, stores and shopping malls, regularly contain displays or screens with such “label[s], sign[s], or advertising.” Lopresti, ¶273; *Du*, [0029]. Thus, a POSITA would have understood *Du* to envision, for instance, capturing an image of identifying data on a display, such as a trademark displayed on a screen in a store, and analyzing it to provide related information to a user. In such a situation, the object (the trademark) would be *a spot on a displayed*

image on a display, a subarea of a space on the displayed image on the display, or one of a plurality of objects in the displayed image on the display. Lopresti, ¶273.

Even if *Du* did not expressly disclose the object being *a spot on a displayed image on a display, a subarea of a space on the displayed image on the display, and one of a plurality of objects in the displayed image on the display*, it would have been obvious to a POSITA from *Du*'s disclosure. *Du*'s disclosure is not limited to objects in the real world. Lopresti, ¶¶274, 203-208. Rather, a POSITA would recognize that *Du*'s camera-equipped mobile phone could be used to take pictures of objects on a screen (*on a display*) and process them in the same way that it would process a picture of an object in the real world. Lopresti, ¶¶203, 274. *Du* discloses the necessary hardware and software to capture and process images of objects in the real world or displayed on a screen. Lopresti, ¶203, 274. For example, *Du* discloses using “the camera of a mobile telephone” along with “one of many standard image recognition and processing programs.” *Du*, [0006]. A POSITA would have recognized this hardware and software was well known and conventional. Lopresti, ¶203, 274.

Du suggests this application of its disclosure when it says that its disclosure is applicable whether “the selected item is physically present or not.” *Du*, [0029]. *Du* discloses its disclosure is applicable to capture and analyze “identifying data” (such as a barcode or text) on “a label, sign, or advertising that includes identifying

data of a selected item.” *Du*, [0029]; Lopresti, ¶¶204-205. A POSITA would have known that techniques necessary to capture and process images containing “identifying data” on “a label, sign, or advertising” would be applicable to capture and process images displayed on a screen. Lopresti, ¶¶205. Further, a POSITA would have recognized that people encounter objects on screens as much (if not more) than they do in the real world. Lopresti, ¶¶205. Thus, a POSITA would have been motivated to apply *Du*’s disclosure to capture and process images of objects displayed on a screen as well as objects in the real world. Lopresti, ¶¶205.

Applying *Du*’s known image processing techniques to a known camera on a mobile telephone would yield predictable results, including allowing the camera on the mobile telephone to form and process images of any target object. Lopresti, ¶¶205-207.

A POSITA would have had a reasonable expectation of success in extending *Du*’s disclosure to allow a user to capture and process images displayed on a screen. Lopresti, ¶¶208. For example, *Du*’s disclosure of “one of many standard image recognition and processing programs” is just as applicable to images displayed on a screen as it is to images of objects in the real world. *Du*, [0006]; Lopresti, ¶¶208. *Du* even suggests that its system could analyze images displayed on a screen when it discloses that it can analyze images of objects displayed on “a label, sign, or advertising” regardless of whether “the selected item is physically

present or not.” *Du*, [0029]; Lopresti, ¶208. Moreover, capturing images of objects on a screen was well known at the time of the alleged invention. Lopresti, ¶274; *see, e.g., Boncyk*, [0273].

2. Claim 2

Element 2[a]:

Du further discloses 2[a]. Lopresti, ¶276-78. For example, *Du* discloses pointing a mobile phone’s camera at a “a desired item.” *Du*, [0006].

A POSITA would have recognized pointing a mobile phone camera at the “desired item” is *a designation of the object as designated by the user*. Lopresti, ¶277.

Element 2[b]:

Du further discloses 2[b]. Lopresti, ¶279-82. For example, *Du* discloses using “the camera of a mobile telephone to capture an image of a desired item.” *Du*, [0006].

When *Du* discloses that the user “capture[s] an image,” it discloses that its system is *recording a frame image comprising the designated object*.

Du further discloses that the taken “image is communicated to one or more servers.” *Du*, [0006]. A POSITA would have known that to communicate the picture, the picture must have been recorded. Lopresti, ¶281.

Element 2[c]:

Du further discloses 2[c]. Lopresti, ¶283-86. For example, *Du* discloses that “data identifying the selected item is extracted from the image” and that “the extraction process will depend upon the type of image that is received.” *Du*, [0033].

The '812 patent explains that the “context type” of an image is the source of the image, for example a “computer screen, game box screen, TV screen, or view of the real world.” '812 patent, 10:64-66. As one example, the '812 patent discloses “Print Space” as a “Context-type.” '812 patent, 12:56-13:8. In the “Context-type” of a “Print Space,” the '812 patent discloses using OCR to analyze text. '812 patent, 12:56-13:8. The '812 patent further explains that its system can analyze a barcode or logo. '812 patent, 12:56-13:8.

Du discloses that it can determine the *context type of the frame* when it discloses selection and “extraction process depend[ing] upon the type of image that is received.” *Du*, [0033]. For example, *Du* discloses that when “the image contains a barcode, an image processing program using edge detection routines is used to locate the barcode and convert the barcode to the equivalent numeric or other code that is represented by the barcode.” *Du*, [0033]. As another example, if “the image contains text (such as the cover of a book), a commercially available optical character recognition (OCR) program may be used to parse the image and extract

such identifying data as the name of the author, title of the work, etc.” *Du*, [0033].

Further, *Du* discloses that other “commercially available video or image recognition programs may be used in the alternative to obtain the identifying data, which may include, for example, numeric or alphanumeric data containing a UPC or SKU associated with the selected item.” *Du*, [0033]. *Du* discloses that its system, just like the system claimed in the ’812 patent, can determine what “type of image . . . is received” and can identify the sort of “extraction process” appropriate to the “type of image” (*context type of the frame*). Lopresti, ¶285.

Element 2[d]:

Du further discloses 2[d]. Lopresti, ¶287-88. For example, *Du* discloses “process[ing] an image to identify the item identified by the image.” *Du*, [0028]; *see supra* § VI.B.1 (element 1[f]).

3. Claim 3:

Du further discloses or in view of *Darell* renders obvious claim 3. Lopresti, ¶289-94. For example, *Du* discloses that in cases where “conflicting identifying data” exists regarding the identity of the item in an image, “the user of the imaging device may be polled to determine the correct data.” *Du*, [0034].

Du’s goal is to provide a user with information on a product after identifying the product. Lopresti, ¶290. For example, *Du* discloses “the information processing

method 300 processes an image to identify the item identified by the image and returns to the user information pertaining to the identified item.” *Du*, [0028].

Du discloses that, in cases where “conflicting identifying data” exists regarding the identity of the item in an image, “the user of the imaging device may be polled to determine the correct data.” *Du*, [0034]. Again, this polling presents the user with the likely identified items from which to select. *See supra*, 1[f]. The user’s response to the “poll” is a *selection by the user of said one of the likely pointed-to objects*. Lopresti, ¶291.

A POSITA would have recognized that the purpose of the polling, given the rest of *Du*’s disclosure, is so that the system can provide the user with “information pertaining to the identified item.” *Du*, [0028]; Lopresti, ¶292. Thus, when the user’s responds to the poll by selecting one of the options provided in the poll, the user’s response is simultaneously *an instruction from the user to obtain information on the object*. Lopresti, ¶292.

It would have been obvious to a POSITA to modify *Du*’s system to include *Darrell*’s teaching of receiving the user’s selection and then returning information to the user on the selected object. Lopresti, ¶215-219. *Darrell* discloses that “the handheld device 10 communicates the captured image to a web server 24 to find images similar to the captured image.” *Darrell*, [0028]; *see, e.g., Darrell*, [0021], [0024], [0026]. *Darrell* explains that “there is no need to look for a perfect match”

because “[m]oderately good results arranged as a thumbnail mosaic . . . allows any user to swiftly identify just what images are relevant.” *Darrell*, [0024]. *Darrell* further discloses returning results to the user’s mobile device “by arranging candidate images into a thumbnail mosaic 220 (FIG. 2).” In addition, *Darrell* discloses that the “user, as the final judge, can easily identify what sub-images are ‘really relevant’” and that when “a thumbnail is selected, the source URL is retrieved and the content from that URL is shown on the mobile device 10.” *Darrell*, [0025].

Du and *Darrell* would naturally occur to a POSITA as complementary, combinable technologies. Both references relate to object recognition using a camera-equipped device. Lopresti, ¶216; *supra* §§ III.B (*Du*) and III.C (*Darrell*). Furthermore, both references address the same problem: recognizing and obtaining information about a selected object. Lopresti, ¶216; *supra* §§ III.B (*Du*) and III.C (*Darrell*).

Du itself suggests reasons for this modification. A POSITA would have been motivated obtain information on the selected object because *Du*’s goal is to provide a user with information on an object after identifying the object. Lopresti, ¶216. For example, *Du* discloses “the information processing method 300 processes an image to identify the item identified by the image and returns to the user information pertaining to the identified item.” *Du*, [0028]. Applying *Darrell*’s

known techniques for obtaining information on the user-selected object to a *Du*'s known image capture and identification system would yield predictable results. Lopresti, ¶218.

A POSITA would have had a reasonable expectation of success in combining *Darrell*'s teaching obtaining information on the user-selected object with *Du*'s image capture and identification framework. Lopresti, ¶219. As explained above, *Du* discloses the functionality and information necessary to allow a mobile device, working with a remote server, to obtain information on the user-selected object. Lopresti, ¶219. Further, *Du* recognizes that its invention can be implemented using “one of many standard image recognition and processing programs.” *Du*, [0006]. Thus, a POSITA would have been able to incorporate *Darrell*'s teaching with *Du*'s system with a reasonable expectation of success. Lopresti, ¶219.

4. Claim 4:

Du further discloses claim 4. Lopresti, ¶295-296. For example, *Du* discloses that “when item information is communicated back to the user’s mobile device, the user may be provided with a control that allows the user to immediately purchase the selected item from the online source.” *Du*, [0034]; *see e.g.*, *Du*, [0006], [0007], [0016], [0019], [0023].

5. Claim 5:

Du renders obvious these limitations. *See* 1[h].

6. Claim 6:

Du renders obvious these limitations. *See* 1[h].

7. Claim 7:

Du renders obvious these limitations. *See* 1[h].

8. Claim 8:

Du renders obvious these limitations. *See* 1[h].

9. Claim 9:

Du renders obvious these limitations. *See* 1[h].

10. Claim 10:

Du discloses these limitations. *See* 1[h].

11. Claim 11:

Du discloses these limitations. *See* 1[h].

12. Claim 12:

Du discloses these limitations. *See* 1[h].

VII. DISCRETIONARY DENIAL IS IMPROPER

Discretionary denial here would be improper. *See Apple Inc. v. Fintiv, Inc.*, IPR2020-00019, Paper 11 (PTAB Mar. 20, 2020).

VIII. MANDATORY NOTICES

A. Real Party-in-Interest Under 37 C.F.R. § 42.8(b)(1)

The real party-in-interest is Google LLC.¹

B. Related Matters Under 37 C.F.R. § 42.8(b)(2)

To the best of Petitioner's knowledge, the '812 patent is, or has been, involved in the following district court litigations and Board proceedings.

Name	Number	Court	Filed
<i>Pointwise Ventures LLC v. Alibaba Group Holding Limited</i>	2-24-cv-00182	EDTX	Mar. 15, 2024
<i>Pointwise Ventures LLC v. Microsoft Corporation</i>	6-24-cv-00139	WDTX	Mar. 15, 2024
<i>Pointwise Ventures LLC v. Blippar Ltd.</i>	2-24-cv-00183	EDTX	Mar. 15, 2024
<i>Pointwise Ventures LLC v. Farfetch UK Limited</i>	2-24-cv-00184	EDTX	Mar. 15, 2024
<i>Pointwise Ventures LLC v. Glority Global Group Limited</i>	2-24-cv-00187	EDTX	Mar. 16, 2024
<i>Pointwise Ventures LLC v. IKEA North America Services, LLC</i>	2-24-cv-00188	EDTX	Mar. 16, 2024
<i>Pointwise Ventures LLC v. Roadget Business Pte. Ltd.</i>	2-24-cv-00189	EDTX	Mar. 16, 2024
<i>Pointwise Ventures LLC v. SAMSUNG ELECTRONICS AMERICA, INC.</i>	2-24-cv-00190	EDTX	Mar. 17, 2024
<i>Pointwise Ventures LLC v. Syte Visual Conception Ltd.</i>	2-24-cv-00191	EDTX	Mar. 17, 2024

¹ Google LLC is a subsidiary of XXVI Holdings Inc., which is a subsidiary of Alphabet Inc. XXVI Holdings Inc. and Alphabet Inc. are not real parties-in-interest to this proceeding.

<i>Pointwise Ventures LLC v. Google LLC</i>	6-24-cv-00293	WDTX	May 30, 2024
<i>Pointwise Ventures LLC v. Amazon.com, Inc.</i>	2-24-cv-00751	WDW A	May 30, 2024
<i>Pointwise Ventures LLC v. Walmart, Inc.</i>	2-24-cv-00443	EDTX	June 14, 2024
<i>Pointwise Ventures LLC v. Meta Platforms, Inc. f/k/a Facebook, Inc.</i>	6-24-cv-00320	WDTX	June 14, 2024
<i>Pointwise Ventures LLC v. Salesforce, Inc. f/k/a salesforce.com, Inc.</i>	6-24-cv-00321	WDTX	June 14, 2024
<i>Pointwise Ventures LLC v. eBay Inc.</i>	1-24-cv-00706	DDE	June 17, 2024
<i>Pointwise Ventures LLC v. Pinterest, Inc.</i>	1-24-cv-00708	DDE	June 17, 2024
<i>Pointwise Ventures LLC v. Snap One Holdings Corp.</i>	1-24-cv-00709	DDE	June 17, 2024
<i>Pointwise Ventures LLC v. Etsy, Inc.</i>	1-24-cv-04706	EDNY	July 5, 2024
<i>Pointwise Ventures LLC v. Snap Inc.</i>	6-24-cv-00396	WDTX	July 24, 2024
<i>Pointwise Ventures LLC v. Lowe's Companies, Inc.,</i>	2-24-cv-00808	EDTX	Oct. 7, 2024
<i>Pointwise Ventures LLC v. YOOX Net-a-Porter Group SpA</i>	2-24-cv-00809	EDTX	Oct. 7, 2024
<i>Pointwise Ventures LLC v. ASOS.Com Limited</i>	2-24-cv-00810	EDTX	Oct. 7, 2024
<i>Pointwise Ventures LLC v. Wayfair Inc.</i>	2-24-cv-00811	EDTX	Oct. 7, 2024
<i>Pointwise Ventures LLC v. Hennes & Mauritz AB</i>	2-24-cv-00812	EDTX	Oct. 7, 2024
<i>Pointwise Ventures LLC v. Penney OpCo LLC d/b/a JCPenney</i>	2-24-cv-00813	EDTX	Oct. 7, 2024
<i>Pointwise Ventures LLC v. Macy's, Inc.</i>	2-24-cv-00814	EDTX	Oct. 7, 2024
<i>Pointwise Ventures LLC v. Urban Outfitters, Inc.</i>	2-24-cv-00816	EDTX	Oct. 7, 2024
<i>Pointwise Ventures LLC v. Boohoo.com UK Limited</i>	2-25-cv-00219	EDTX	Feb. 19, 2025

<i>Pointwise Ventures LLC v. The Home Depot, Inc.</i>	2-25-cv-00220	EDTX	Feb. 19, 2025
<i>Inter IKEA Systems BV v. Pointwise Ventures LLC</i>	IPR2025-00798	PTAB	April 1, 2025

C. Lead and Back-Up Counsel Under 37 C.F.R. § 42.8(b)(3)

Lead Counsel	Back-Up Counsel
Erika Harmon Arner (Reg. No. 57,540) erika.arners@finnegan.com Finnegan, Henderson, Farabow, Garrett & Dunner, LLP 1875 Explorer Street Suite 800 Reston, VA 20190-6023 Tel: 571-203-2700 Fax: 202-408-4400	Daniel C. Cooley (Reg. No. 59,639) daniel.cooley@finnegan.com Finnegan, Henderson, Farabow, Garrett & Dunner, LLP 1875 Explorer Street Suite 800 Reston, VA 20190-6023 Tel: 571-203-2700 Fax: 202-408-4400

D. Service Information Under 37 C.F.R. § 42.8(b)(4)

Please address all correspondence to counsel at the addresses above. Petitioner consent to service by email at the above addresses.

IX. GROUNDS FOR STANDING

Petitioner certifies that the '812 patent is available for IPR and that Petitioner is not barred or estopped from requesting IPR challenging the patent claims on the grounds herein.

X. CONCLUSION

Petitioner requests that the Board institute *inter partes* review and cancel each challenged claim.

Respectfully submitted,

Dated: June 6, 2025

By: /Erika H. Arner/
Erika H. Arner, Reg No. 57,540

CERTIFICATION UNDER 37 C.F.R. § 42.24(d)

Pursuant to 37 C.F.R. § 42.24(d), the undersigned hereby certifies that the foregoing Petition contains 13,970 words, excluding parts of this Petition exempted under § 42.24(a), as measured by the word-processing system used to prepare this paper.

Respectfully submitted,

Dated: June 6, 2025

By: /Erika H. Arner/
Erika H. Arner, Reg No. 57,540

CERTIFICATE OF SERVICE

The undersigned certifies that the foregoing **Petition for *Inter Partes* Review, the associated Power of Attorney, and Exhibits 1001 through 1015** were served on June 6, 2025, via FedEx Priority Overnight® on the correspondence address of record indicated in the Patent Office's Patent Center system for U.S. Patent No. 8,471,812. A courtesy copy has also been sent via electronic mail to litigation counsel for Patent Owner at the below address.

Correspondence Address of Record	Litigation Counsel
Greenbaum, Michael BLANK ROME LLP 1825 Eye Street NW Washington, DC 20006-5403 United States	Isaac Rabicoff isaac@rabilaw.com Rabicoff Law LLC 4311 N Ravenswood Ave, Suite 315 Chicago, Illinois 60613

Dated: June 6, 2025

/Daniel E. Doku/
Daniel E. Doku
Senior Litigation Legal Assistant
FINNEGAN, HENDERSON, FARABOW,
GARRETT & DUNNER, L.L.P.