

Petitioner's Demonstratives

Google LLC

v.

138 East LCD Advancements Limited

IPR2024-00960

U.S. Patent No. 7,486,807

September 30, 2025

U.S. Patent No. 7,486,807

U.S. Patent No. 7,486,807
“Image Retrieving Device, Method
for Adding Keywords in Image
Retrieving Device, and Computer
Program Therefor”

US007486807B2
Patent No.: US 7,486,807 B2
Date of Patent: Feb. 3, 2009

(54) IMAGE RETRIEVING DEVICE, METHOD FOR ADDING KEYWORDS IN IMAGE RETRIEVING DEVICE, AND COMPUTER PROGRAM THEREFOR

(75) Inventor: Toshinori Nagahashi, Tatsuno-machi (JP)

(73) Assignee: Seiko Epson Corporation, Tokyo (JP)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 676 days.

(21) Appl. No.: 10/762,450

(22) Filed: Jan. 23, 2004

(65) Prior Publication Data
US 2004/0213553 A1 Oct. 28, 2004

(30) Foreign Application Priority Data
Jan. 29, 2003 (JP) 2003-020805

(51) Int. Cl. G06K 9/00 (2006.01)

(52) U.S. Cl. 382/118; 382/305; 382/232; 382/243; 382/276; 382/234; 382/120; 382/126; 382/115; 382/307; 707/104.1; 707/1; 707/3; 707/100; 704/2; 704/3; 704/4; 704/277; 704/251; 704/255; 704/257; 235/487; 358/448; 358/450; 358/403; 358/453; 358/527

(58) Field of Classification Search 382/305; 382/118, 209; 348/231.2; 386/69; 707/1; 707/102, 3, 104, 5, 512; 355/27; 434/217; 704/2, 3, 4, 5, 6, 7, 8, 9, 10, 277, 251, 252, 704/255, 257; 358/403, 404

See application file for complete search history.

References Cited

U.S. PATENT DOCUMENTS

5,608,387 A * 3/1997 Davies 349:5,27

6,408,301 B1 * 6/2002 Patton et al. 707:102

6,785,429 B1 * 8/2004 Senoh 382:305

6,826,316 B2 * 11/2004 Luo et al. 382:305

(Continued)

FOREIGN PATENT DOCUMENTS

CN 1245317 A 2/2000 (Continued)

OTHER PUBLICATIONS

"Keyword and Face Image Retrieval Based on Latent Semantic Indexing", Koshimizu et al., IEEE, 2004 *

(Continued)

Primary Examiner—Samir A. Ahmed
Assistant Examiner—Nancy Hilar
(74) Attorney, Agent, or Firm—Oliff & Berridge, PLC

(57) ABSTRACT

An image retrieving device for classifying and retrieving an image by detecting an object in the image and adding a keyword comprises an image storing section for storing the image which is supposed to be classified and retrieved together with the keyword in a database, an object acknowledging section for acknowledging a predetermined object in the image which is inputted, a keyword proposing section for proposing the keyword which relates to the object which is acknowledged by the object acknowledging section, and an object information inputting section for confirming, adding, and correcting the keyword which is proposed by the keyword proposing section. By doing this, the object in the image is detected for classifying the image; thus, it is possible to provide an image retrieving device in which the keywords are added to the image efficiently.

14 Claims, 8 Drawing Sheets

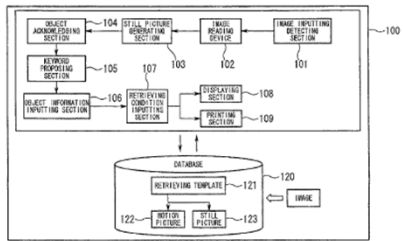


FIG. 8

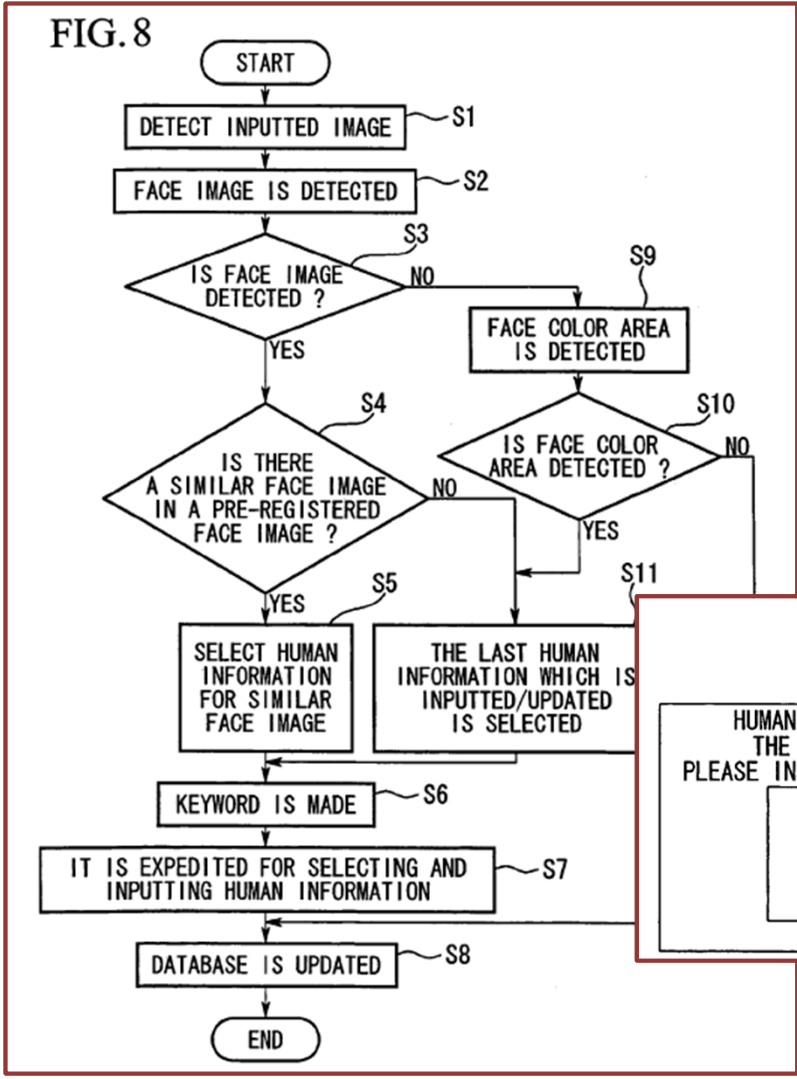


FIG. 3



Ex. 1001, FIG. 3

Ex. 1001, FIG. 8

See, e.g., Petition 1-4; Pet'r Reply 6

Independent claim 1

1. An image retrieving device for classifying and retrieving an image by detecting an object in the image and adding a keyword, the image retrieving device comprising:
 - an image storing section for storing the image which is supposed to be classified and retrieved together with the keyword in a database and an object of the image being previously contained in the database;
 - an image inputting detecting section that detects an inputted image that is newly inputted to the image retrieving device;
 - an object acknowledging section for acknowledging an inputted object in the inputted image that has been detected by the image inputting detecting section;
 - a keyword proposing section for proposing the keyword on a display, the keyword which relates to the inputted object which is acknowledged by the object acknowledging section; and
 - an object information inputting section for confirming by the user, adding, and correcting the keyword which is proposed by the keyword proposing section when the inputted object acknowledged by the object acknowledging section is similar to the object of the image previously contained in the database.

See, e.g., Ex. 1001, 13:32-54 (as-corrected); generally Patent Owner's Response; Pet'r Reply; PO Sur-Reply

Grounds of unpatentability

Ground	Claims	Basis	Prior Art
1	1-14	§§ 102/103	Stubler
2	2-5, 9-12	§ 103	Stubler and Wang (not disputed)

- Patent Owner argues single issue for Ground 1, independent claim 1
 - Construction of “similar to the object of the image previously contained in the database”
- Patent Owner does not argue against Ground 2

See, e.g., Petition 1, Pet'r Reply 1-2, 26-27

Prior art: Stubler – Determining objects and metadata, searching database for similar objects, proposing object keywords

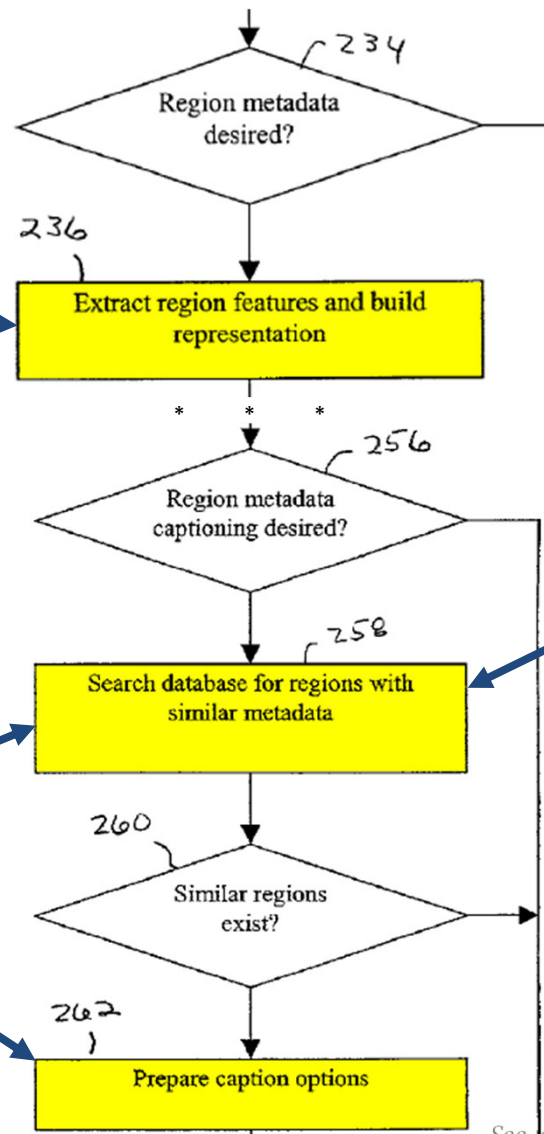
[0022] This invention uses image similarity based upon objectively measurable or determinable metadata to provide a mechanism for assigning captions or semantic labels to multiple images simultaneously, and to automatically assign these captions or labels to new images based upon shared metadata, thereby reducing the effort required of the user. Typical metadata includes low-level measurable features such as color and texture or image-accompanying data such as date and time of capture. In a preferred embodiment, such

or part of the image in which they are found. While the measurements and region identification is derived from low-level features, the processing(including knowledge based operations) may yield a region with a mixture of high-level feature description (clear sky, cloudy sky, mixed sky) and low-level features (color representation, area coverage, location). Similarities between such regions would be calculated on the basis of a combination of high-level and low-level features.

Stubler [0022]

for extension to the acquired images. If captioning or semantic labeling of the acquired images based on region metadata is desired (256), the database is searched (258) for any images with region metadata (e.g., face, sky, grass, skin, open space, etc.) similar to that of the acquired images, and repeated as needed (260) for similar regions until all images are analyzed. Then the captions or semantic labels for the regional aspects of the retrieved and matched images are prepared as options (262) for extension to the acquired images. The result thus obtained (264) comprises one or more images with a collection of potential captions or semantic labels, and/or region annotations. Finally, the cap-

Stubler [0040]



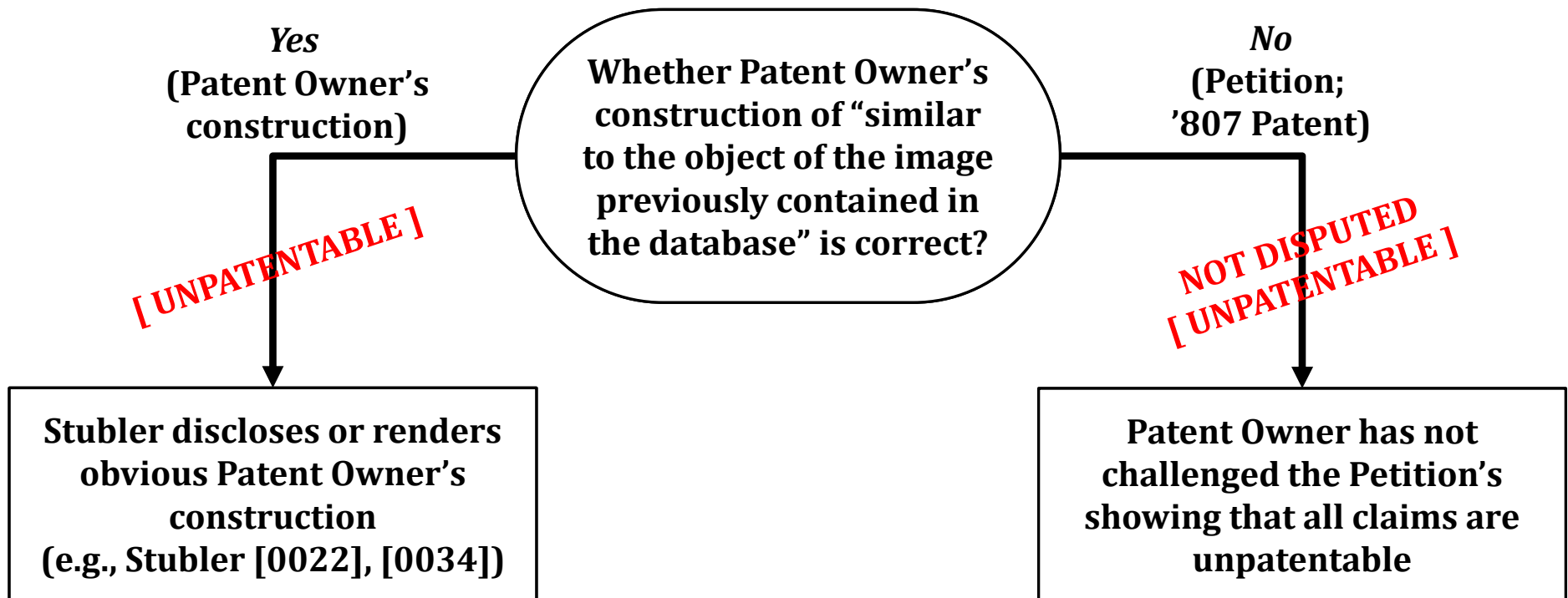
[0034] For other types of metadata, the data may already be represented in a form suitable for clustering. For instance, in the preferred embodiment time/date information is used and may be directly subjected to cluster analysis. In other cases, for example that of region metadata, the data may need to be processed for similarity by known algorithms. As an example, determination of facial similarity is described in M. Turk and A. Pentland, "Eigenfaces for Recognition", *J. Cognitive Neurosci.* 3(1)1991.

Stubler [0034]

Stubler FIGS. 4-5 (excerpted)

See, e.g., Petition 7, 18, 20, 32-33, 42-43; Pet'r Reply 6-7, 21, 24-25

All claims unpatentable regardless of Board's resolution of claim construction



See, e.g., Pet'r Reply 1-2, 16, 18, 23

**Construction of
“similar to the object of the image
previously contained in the database”**

Patent Owner's construction has multiple flaws

Claim Term	Petitioner's Position	Patent Owner's Construction
"similar to the object of the image previously contained in the database"	No construction necessary	" <u>sufficiently</u> similar to <u>identify</u> the <u>specific</u> object of the image previously contained in the database"

- Claims and specification to not support Patent Owner's narrowing
 - Neither "sufficient" nor "identify" appear in the specification
- Would inject ambiguity into claims

See, e.g., Petition 5; Patent Owner's Response 13-21; Pet'r Reply 2-15; PO Sur-Reply 2-17

Patent Owner would require identifying a unique object, but the '807 Patent does not require uniqueness

According to the two-step process discussed above, when a new image of “Jiro” is detected and ready to be classified, the image retrieving device must first conclude that the image (1) contains an image of a human and (2) contains the *specific person*, “Jiro,” before recommending that that the new image be categorized as containing “Jiro.” *Id.*, ¶ 64. As discussed above, the '807 patent presents the

Patent Owner's Response 19

human.” “The detected human” is the specific, unique human who is contained within the photograph. It is not some generic concept of a human, as Petitioner seems to argue. Indeed, immediately before the cited excerpt, the '807 patent

PO Sur-Reply 9

See, e.g., Patent Owner's Response 19; PO Sur-Reply 9; Pet'r Reply 4-9

The claims do not support Patent Owner's construction requiring unique identification

1. An image retrieving device for classifying and retrieving an image by detecting an object in the image and adding a keyword, the image retrieving device comprising:

an image storing section for storing the image which is supposed to be classified and retrieved together with the keyword in a database and an object of the image being previously contained in the database;

an image inputting detecting section that detects an inputted image that is newly inputted to the image retrieving device;

an object acknowledging section for acknowledging an inputted object in the inputted image that has been detected by the image inputting detecting section;

a keyword proposing section for proposing the keyword on a display, the keyword which relates to the inputted object which is acknowledged by the object acknowledging section; and

an object information inputting section for confirming by the user, adding, and correcting the keyword which is proposed by the keyword proposing section when the inputted object acknowledged by the object acknowledging section is similar to the object of the image previously contained in the database.

2. An image retrieving device according to claim 1, wherein the object acknowledging section includes:

a human detection condition inputting section for setting up conditions for determining whether or not the image contains a human;

a face image detecting section for detecting a face image in the image; and

a face image similarity determining section for detecting a face image which is detected by the face image detecting section so as to detect a similar face image stored in the database according to the detected face image.

5. An image retrieving device according to claim 2, wherein the keyword proposing section proposes the keyword to the image which is inputted last in a case in which the similar face image is not detected by the face image similarity determining section.

See, e.g., Ex. 1001, claim 1 (as-corrected), claim 2, claim 5 (as corrected)

The specification identifies objects as “human” or “car” and does not require Patent Owner’s unique identification

In such a structure, an object which is supposed to be acknowledged (detected) is a human such that the keyword (human information) is confirmed for the detected human and the keyword is inputted for the human if a human is detected in the image.

By doing this, the human in the inputted image is detected and the keywords (human information) which relate to the human in the image is added; thus, it is possible to add the keywords accurately.

Ex. 1001, 4:59-67

keywords in the image storing step. Consequently, a specific object (such as a human or a car) in the newly inputted image is acknowledged (detected) in the object acknowledging step. In addition, if a specific object is detected in the object acknowledging step, the keywords which relate to the object are proposed to the user in the keyword proposing step. Also,

Ex. 1001, 5:56-61

FIG. 3

HUMAN IS DETECTED FROM
THE INPUTTED IMAGE.
PLEASE INPUT HUMAN INFORMATION



Ex. 1001, FIG. 3

FIG. 3 is a view for showing an example for a display for inputting the human information. Here, when a new image is contained in the image retrieving device 100, the image retrieving device 100 tries to detect a human in the contained image. If a human is detected, the image retrieving device 100 expedites a user to input the human information by a display form which is shown in FIG. 3.

Ex. 1001, 9:7-13

See, e.g., Pet'r Reply 4-9

The specification identifies keywords like “family” and “friend” and does not require Patent Owner’s unique identification

by the image storing section in such a structure. Consequently, a new predetermined object which is inputted in the image (such as a human or a car) is acknowledged (detected) by the object acknowledging section. If a predetermined object is detected by the object acknowledging section, the keywords which relate to the object are proposed to the user by the keyword proposing section. Also, in this aspect of the

Ex. 1001, 4:8-14

posed to be retrieved. For example, if an object which is supposed to be retrieved is a human, retrieving conditions (keywords) such as selecting among “family”, and “friend” are inputted.

Ex. 1001, 7:48-51

FIG.4



Ex. 1001, FIG. 4

(2) The keyword is used for retrieving an image which contains a human. Such keywords can be set up in a hierarchical structure. For example, if an object is a human, a “family” and “relative” may be keywords for superior concept. Furthermore, it is possible to name “eldest brother”, “eldest sister”, and “maternal relative” for a subordinate concept.

Ex. 1001, 6:64-7:3

See, e.g., Pet'r Reply 4-9

Patent Owner's additional citations fail to require its construction of unique identification

image by the face image detecting section. Also, when the face image is detected by the face image detecting section, a similar face image is detected with reference to the database by the image similarity determining section according to the detected face image. Consequently, if the similar face image is detected, the keyword which relates to the face image is proposed to the user so that the user may confirm, add, or correct the keyword.

Ex. 1001, 4:40-47

image in the detecting/classifying device for the image. Also, the keywords which relate to the object in the image is added; thus, it is possible to retrieve the image which contains a specific object among numerous images accurately and efficiently.

Ex. 1001, 4:20-24

FIG.6

SUPERIOR KEYWORD 1	SUBORDINATE KEYWORD 2	NAME	BIRTHDAY	IMAGE
FAMILY	SECND ELDEST BROTHER	JIRO	JANUARY 1, 2003	IMAGE 1:TAKEN ON JANUARY 1, 2003 IMAGE 2:TAKEN ON FEBRUARY 1, 2003 IMAGE 3:TAKEN ON MARCH 1, 2003

Ex. 1001, FIG. 6

If it is determined that the face image coincides to the face image which is contained previously or that the face image is very similar to the face image which is contained previously, the information for such a human is proposed so as to expedite the user to confirm.

Ex. 1001, 9:18-22

See, e.g., Pet'r Reply 4-9

Patent Owner's construction would make the claim ambiguous

None of “sufficiently similar,” “sufficient,” or “sufficiently” appears in the ’807 patent or claims. Neither does “identify” or any of its derivatives. As such, there is no basis for a POSITA to understand that “similar” should be modified by “sufficiently” or require “sufficient[.]” “identification” under Patent Owner’s construction, nor would a POSITA be able to determine a level of “sufficiently” or “identification” objectively from the specification, as the law requires.

Patent Owner’s descriptor “to identify the specific object of the image previously contained in the database” also does not clarify the scope of “sufficiently similar.” POR 14-21. As explained above, the ’807 patent describes a “specific object” as a “human” or “car.” A simple example shows the ambiguity Patent Owner seeks to inject. In the context of cars, for example, if the database

Pet’r Reply 13

See, e.g., Pet’r Reply 12-14

Patent Owner cannot say what is “sufficiently similar to identify the specific object” under its construction and the '807 Patent has no guidance

Database Object (Patent Owner's construction)	Sufficiently Similar to Identify the Specific Object ??
Blue sedan	Blue SUV ?? Blue sports car ?? Purple sedan ??
Grey SUV	Blue SUV ?? Black SUV ?? White SUV ??
Blue Honda Accord sedan	Blue Subaru sedan ?? Blue Honda Civic Sedan ??
Individual woman	Woman's sister ?? Woman with same skin tone ?? Woman with same skin color ?? Woman with same hair color ?? Woman with same skin tone and hair color ??

See, e.g., Pet'r Reply 13-14

Patent Owner's expert cannot rewrite the claims or specification



“[A] court should discount any expert testimony ‘that is clearly at odds with the claim construction mandated by ... the written record of the patent.’”

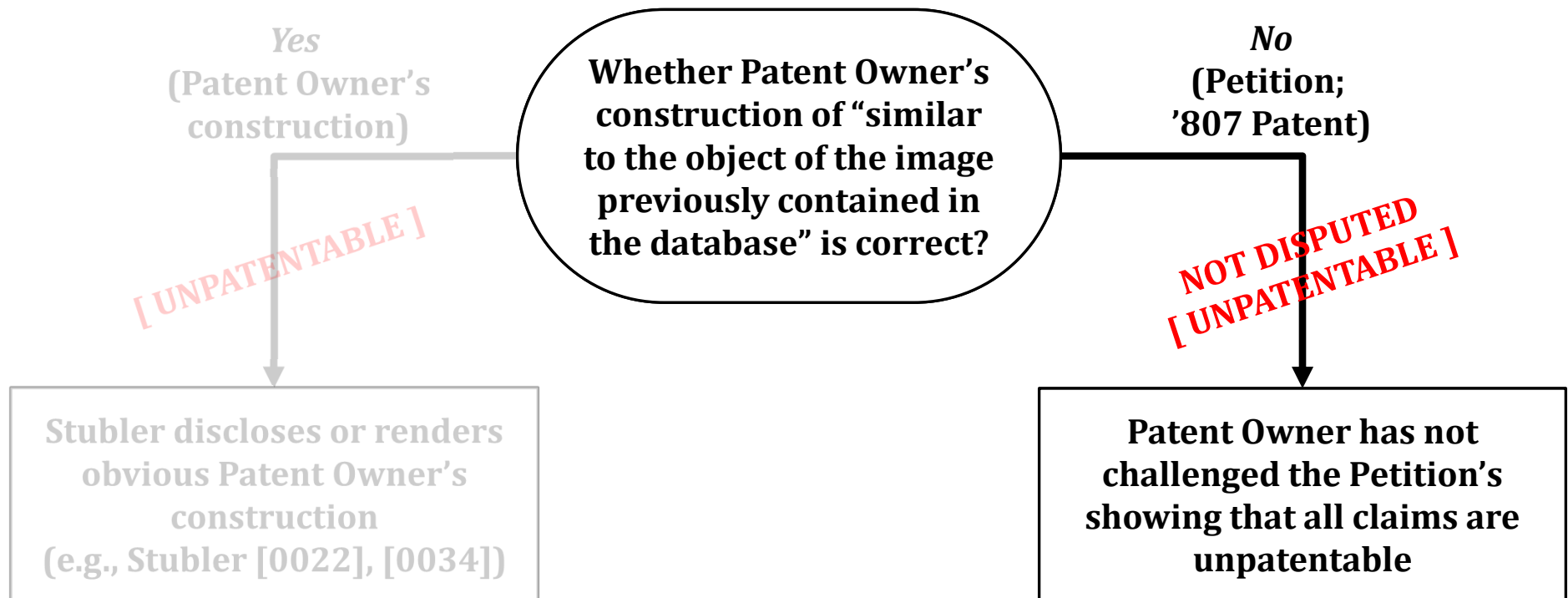
Sequoia Tech., LLC v. Dell, Inc., 66 F.4th 1317, 1325 (Fed. Cir. 2023)

See also Finisar Corp. v. DirecTV Grp., Inc.,
523 F.3d 1323, 1328 (Fed. Cir. 2008)

See, e.g., Pet'r Reply 10 n.3

Stubler Ground

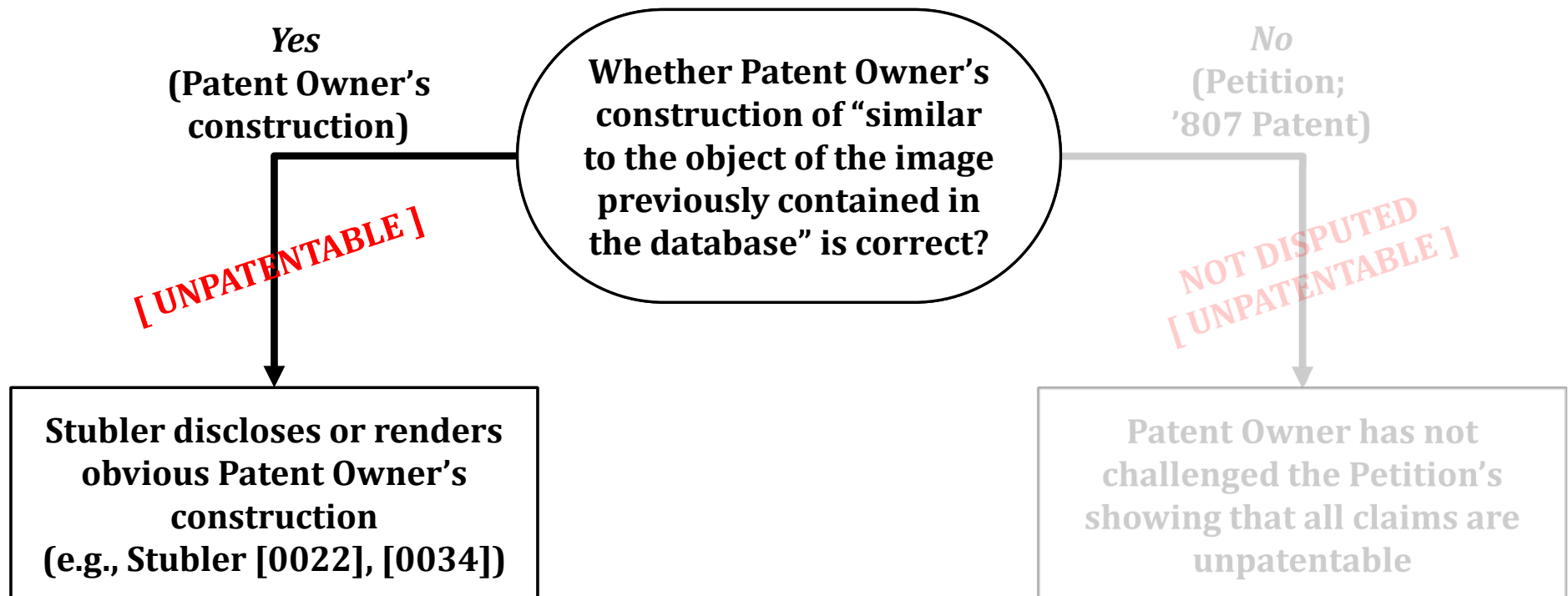
No dispute that Stubler discloses or renders claims obvious once Patent Owner's construction is rejected



See generally Petition 5-59; Patent Owner's Response; PO Sur-Reply; see also Pet'r Reply 15-23

DEMONSTRATIVE EXHIBIT – NOT EVIDENCE

Stubler discloses or renders obvious the claims under Patent Owner's construction



See generally Patent Owner's Response; PO Sur-Reply; *see also* Pet'r Reply 23-27

DEMONSTRATIVE EXHIBIT – NOT EVIDENCE

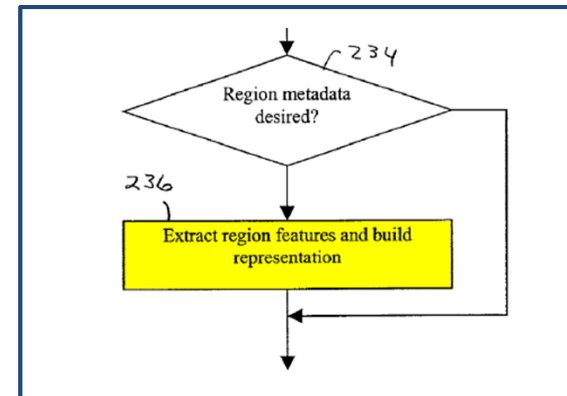
Stubler discloses Patent Owner's construction: Detecting objects and proposing keywords (Stubler)

[0038] The image representation stage 202 is further described in FIG. 4. If global metadata is determined (230) to be desirable, then the global features are extracted (232) in order to build a global representation. As described earlier, the global features include without limitation color histograms, texture features, overall composition and the like. Then, if region metadata is determined (234) to be desirable, the region features are extracted (236) in order to build a region representation. As described earlier, the region features include without limitation face, skin, sky, grass, open space, and the like.

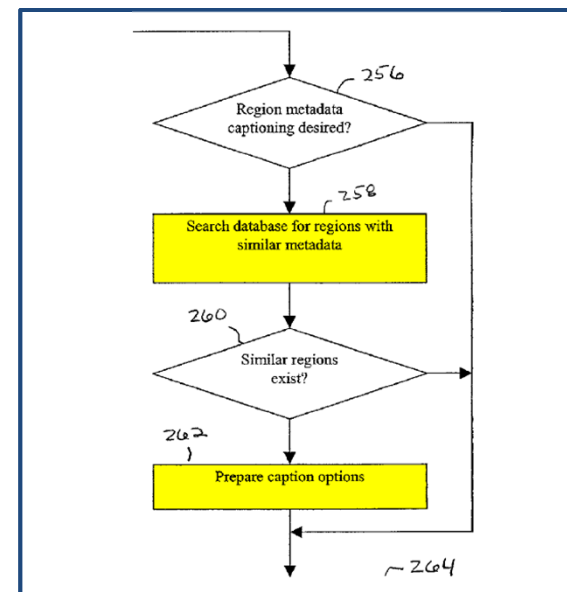
Stubler [0038]

for extension to the acquired images. If captioning or semantic labeling of the acquired images based on region metadata is desired (256), the database is searched (258) for any images with region metadata (e.g., face, sky, grass, skin, open space, etc.) similar to that of the acquired images, and repeated as needed (260) for similar regions until all images are analyzed. Then the captions or semantic labels for the regional aspects of the retrieved and matched images are prepared as options (262) for extension to the acquired images. The result thus obtained (264) comprises one or more images with a collection of potential captions or semantic labels, and/or region annotations. Finally, the cap-

Stubler [0040]



Stubler FIG. 4 (excerpted)



Stubler FIG. 5 (excerpted)

See, e.g., Pet'r Reply 23-26; see also Petition 5-38

Stubler discloses Patent Owner's construction: Determining object similarity and proposing keywords (Stubler)

[0034] For other types of metadata, the data may already be represented in a form suitable for clustering. For instance, in the preferred embodiment time/date information is used and may be directly subjected to cluster analysis. In other cases, for example that of region metadata, the data may need to be processed for similarity by known algorithms. As an example, determination of facial similarity is described in M. Turk and A. Pentland, "Eigenfaces for Recognition", *J. Cognitive Neurosci.* 3(1)1991.

Stubler [0034]

[0030] Region metadata includes such detected regional features as face, skin, sky, grass, open space, background vs. foreground, and the like. Various well-known techniques may be used to detect region metadata. For instance, a typical face detection method is described in K. K. Sung and T. Poggio, "Example-based Learning for View-based Human Face Detection", *IEEE Pattern Analysis and Machine Intelligence*, 20(1), 39 - 51, January, 1998. A typical model for representing skin color in a chromatic space utilizing histogram analysis is described in an article by J. Yang and A. Waibel, "A Real-time Face Tracker, "Proceedings of Third IEEE Workshop on Applications of Computer Vision, WACV 1996, pp. 142 - 147. For the

Stubler [0030]

[0037] FIG. 3 shows a preferred implementation of the method and system shown in FIG. 2. After acquiring the image data and extracting (200) any capture metadata that may have been obtained by the capture device, the image representation is built (202) from the global and region metadata. This will be further explained in relation to FIG. 4. After the image representation is built, a search of the database is performed, and similar images are retrieved (204) from the database, and a set of caption or semantic label options are prepared. The retrieval and preparation

Stubler [0037]

See, e.g., Pet'r Reply 23-26; see also Petition 5-38

Stubler discloses Patent Owner's construction: Object detection (e.g., face, skin, human) (Petitioner's expert)

92. For example, Stubler's program can use "[v]arious well-known techniques," such as algorithms, to detect objects, such as a human face, in the newly acquired image, including those described in Sung (Ex. 1007); U.S. Patent Application No. 09/223,860, titled "Method for Automatic Determination of Main Subjects in Photographic Images," (Luo, Ex. 1008⁷), and Jie Yang and Alex Waibel, *A Real-time Face Tracker*, Proceedings of Third IEEE Workshop on Applications of Computer Vision, WACV, pp. 142-47 (1996) (Yang, Ex. 1009). Stubler, [0030]. Stubler's program can also use, for example, "histogram analysis" to "detect skin color in a chromatic space" corresponding to a face contained the inputted image. Stubler, [0030]. These well-known techniques are incorporated by reference into Stubler, and further discussed in claim 2, *infra*.

Stevenson ¶¶92

101. As explained in [1d], Stubler discloses that,
[i]f captioning or semantic labeling of the acquired images based on region metadata is desired (256), the database is searched (258) for any images with region metadata (e.g., face, sky, grass, skin, open space, etc.) similar to that of the acquired images, and repeated as needed (260) *for similar regions* until all images are analyzed. Then the *captions or semantic labels for the regional aspects* of the retrieved and *matched images are prepared as options*.
Stubler, [0040] (emphases added).

Stevenson ¶¶101

See, e.g., Petition 27, 33; Pet'r Reply 24

Stubler discloses Patent Owner's construction: Detecting an object and object similarity (Petitioner's expert)

113. Detecting a human's face correlates to detecting human skin and faces. For example, facial detection algorithms at the time typically relied on various features, such as facial structure and skin color. Techniques, such as eigenfaces, excelled at identifying facial structure, but did not utilize skin color, and could misidentify some objects as faces that were not. Color-based clustering techniques, such as identifying skin color, could identify skin, but also tended to include other objects that contained similar tones. It was, therefore, common to combine different techniques, such as facial recognition and skin color detection, to provide more robust facial detection, e.g., to detect only areas with correct facial structure, as identified with eigenface approaches, and areas with correct skin colors, as identified by color-based clustering, so that only areas meeting both criterial would have been identified as containing faces. In my opinion, a POSITA would have appreciated the benefits of this implementation to avoid false positives for facial detection, would have understood the impact to improve facial detection, and would have easily been able to implement such approaches, as indicated by Stubler.

Stevenson ¶113

117. As explained for [1preamble] and [1a], Stubler discloses storing previously acquired image objects (e.g., faces) in a database. Stubler, [0035], [0040], [0041], claim 25. For example, Stubler explains that when "captioning or semantic labeling of ... acquired images based on region metadata is desired (256), the database is searched (258) for any images with region metadata (e.g., face, sky, grass, skin, open space, etc.) similar to that of the acquired images, and repeated as needed (260) for similar regions until all images are analyzed." Stubler, [0040] (emphases added), FIG. 5.

Stevenson ¶117

118. Stubler further explains that a "determination of facial similarity" may be ascertained using the algorithms described in Matthew Turk and Alex Pentland, *Eigenfaces for Recognition*, J. Cognitive Neurosci., 3(1), pp. 71-86 (1991) (Turk, Ex. 1010), which Stubler incorporates by reference. Stubler, [0034].

Stevenson ¶118

See, e.g., Petition 27, 40-43; Pet'r Reply 24

Stubler discloses Patent Owner's construction: Searching database for similar objects (Petitioner's expert)

81. Stubler further discloses storing the newly acquired image with a previously contained/previously acquired object of the image because as explained in [1preamble], the newly acquired images are stored and classified with previously contained images containing the same "region[al] metadata" (object). Stubler, [0035], [0041], claim 25. For example, Stubler explains that when "captioning or semantic labeling of ... acquired images based on region metadata is desired (256), the database is searched (258) for any images with region metadata (e.g., face, sky, grass, skin, open space, etc.) similar to that of the acquired images, and repeated as needed (260) for similar regions until all images are analyzed." Stubler, [0040] (emphases added), FIG. 5. That is, the database is searched for an object of the image that is previously contained in the database. When a match is found, the newly acquired image is classified with images containing the same regional metadata. Stubler, [0040], claim 25 (method including "automatically extending the global captions or labels to other images in the group having the same global metadata and the regional captions or labels to other images in the group having the same regional metadata" (emphasis added)). Thus Stubler also discloses storing the acquired image with an object of the image previously contained in the database.

Stevenson ¶81

131. Stubler explains that "[v]arious well-known techniques may be used to detect region metadata" in the acquired image, including "skin color." Stubler, [0030]. A POSITA would have understood that Stubler may employ several of the disclosed techniques to allow multiple types of region metadata to be detected in the image, without limitation. Stubler, [0030], [0038] ("[I]f region metadata is determined (234) to be desirable, the region features are extracted (236) in order to build a region representation. As described earlier, the region features include without limitation face, skin, sky, grass, open space, and the like.") (emphases added), FIG. 4.

Stevenson ¶131

See, e.g., Petition 20-21, 27, 33, 48; Pet'r Reply 24

Stubler discloses the examples in the '807 Patent: Detecting objects in images (face, skin color)

[0030] Region metadata includes such detected regional features as face, skin, sky, grass, open space, background vs. foreground, and the like. Various well-known techniques may be used to detect region metadata. For instance, a typical face detection method is described in K. K. Sung and T. Poggio, "Example-based Learning for View-based Human Face Detection", *IEEE Pattern Analysis and Machine Intelligence*, 20(1), 39 - 51, January, 1998. A typical model for representing skin color in a chromatic space utilizing histogram analysis is described in an article by J. Yang and A. Waibel, "A Real-time Face Tracker, "Proceedings of Third IEEE Workshop on Applications of Computer Vision, WACV 1996, pp. 142 - 147. For the

Stubler [0030]

The face image detecting section 112 detects the fact for whether or not a face image is contained in the image by analyzing the image. There are various methods for detecting the face image such that patterns for the face images are stored in advance, or frequency analysis is performed for the image pattern. It may be acceptable if any one of these methods is performed. Also, it is possible to employ a method which is disclosed in Japanese Unexamined Patent Application, First Publication No. Hei 8-221547. Here, it may be acceptable if the face image detecting section 112 uses a face color area detecting section 113 which is explained in a next paragraph compatibly so as to perform a human detecting operation.

* * *

template. Here, the face color is used. The face color area detecting section 113 determines whether or not the pixel which is read out is within a range of a skin color. For detect-

Ex. 1001, 8:22-48

In such a structure, an object which is supposed to be acknowledged (detected) should be a human. Consequently, a condition (for example, for detecting a skin color area) for determining whether or not a human is contained in the image is set up by a human detection condition inputting section. Consequently, a face image is detected in the newly inputted image by the face image detecting section. Also, when the

Ex. 1001, 4:34-40

See, e.g., Petition 6-7, 9, 26-27, 39-43; Pet'r Reply 18-23

Stubler discloses the examples in the '807 Patent: Determining similarity of detected object

[0034] For other types of metadata, the data may already be represented in a form suitable for clustering. For instance, in the preferred embodiment time/date information is used and may be directly subjected to cluster analysis. In other cases, for example that of region metadata, the data may need to be processed for similarity by known algorithms. As an example, determination of facial similarity is described in M. Turk and A. Pentland, "Eigenfaces for Recognition", *J. Cognitive Neurosci.* 3(1)1991.

Stubler [0034]

(4) Face Image Similarity Determining Section 114

The face image similarity determining section 114 refers to the database 120 so as to check whether or not a similar face image to the detected face image is contained. Here, it is possible to employ a method which is disclosed in Japanese Unexamined Patent Application, First Publication No. Hei. 8-221547 for determining such a similarity by forming a density image of a face so as to classify the category.

Ex. 1001, 8:60-67

See, e.g., Petition 42-43, 64; Pet'r Reply 18-23

Stubler discloses the examples in the '807 Patent: Proposing keywords/captions for detected object

for extension to the acquired images. If captioning or semantic labeling of the acquired images based on region metadata is desired (256), the database is searched (258) for any images with region metadata (e.g., face, sky, grass, skin, open space, etc.) similar to that of the acquired images, and repeated as needed (260) for similar regions until all images are analyzed. Then the captions or semantic labels for the regional aspects of the retrieved and matched images are prepared as options (262) for extension to the acquired images. The result thus obtained (264) comprises one or more images with a collection of potential captions or semantic labels, and/or region annotations. Finally, the cap-

Stubler [0040]; *see also* Stubler [0022], [0035]

Consequently, a face image is detected in the newly inputted image by the face image detecting section. Also, when the face image is detected by the face image detecting section, a similar face image is detected with reference to the database by the image similarity determining section according to the detected face image. Consequently, if the similar face image is detected, the keyword which relates to the face image is proposed to the user so that the user may confirm, add, or correct the keyword.

Ex. 1001, 4:39-47

See, e.g., Petition 7, 9-15, 20-21, 29, 32-33; Pet'r Reply 7, 18-23

Patent Owner is wrong that Turk must be incorporated by reference for Stubler to disclose or render obvious the claims

- Patent Owner does not challenge that Stubler identifies Turk as an example “known algorithm[]” to determine “facial similarity” under Patent Owner’s construction

[0034] For other types of metadata, the data may already be represented in a form suitable for clustering. For instance, in the preferred embodiment time/date information is used and may be directly subjected to cluster analysis. In other cases, for example that of region metadata, the data may need to be processed for similarity by known algorithms. As an example, determination of facial similarity is described in M. Turk and A. Pentland, “Eigenfaces for Recognition”, *J. Cognitive Neurosci.* 3(1)1991.

Stubler [0034]

- Example is permissible to render claims unpatentable



Sherwin-Williams Co. v. PPG Indus., Inc., 2024 WL 3534113, at *3 (Fed. Cir. July 25, 2024) (affirming anticipation finding where prior art referred to, but did not incorporate by reference, other publication, because referral was “unequivocal disclosure”)

See, e.g., Pet’r Reply 20 n.7

Patent Owner's assertion that Stubler teaches "textual" descriptions still discloses an example in the '807 Patent

An object of the invention which is disclosed in Japanese Unexamined Patent Application, First Publication No. Hei 8-221547 is to provide a simple face image model retrieval processing method in which it is possible to extract characteristics of the face area very accurately and effectively in a high speed from the object as a written datum. Consequently,

Ex. 1001, 2:22-27

See, e.g., Pet'r Reply 18-21

Stubler discloses non-textual metadata for comparison

[0022] This invention uses image similarity based upon objectively measurable or determinable metadata to provide a mechanism for assigning captions or semantic labels to multiple images simultaneously, and to automatically assign these captions or labels to new images based upon shared metadata, thereby reducing the effort required of the user. Typical metadata includes low-level measurable features such as color and texture or image-accompanying data such as date and time of capture. In a preferred embodiment, such automatic assessment will be provided at an image region level. For the purpose of this description, the term “image region” will be used when referring to areas of an image that correspond to a significant characteristic low-level feature, such as “sky” or “grass”. These regions may encompass all or part of the image in which they are found. While the measurements and region identification is derived from low-level features, the processing(including knowledge based operations) may yield a region with a mixture of high-level feature description (clear sky, cloudy sky, mixed sky) and low-level features (color representation, area coverage, location). Similarities between such regions would be calculated on the basis of a combination of high-level and low-level features.

Stubler [0022]

[0034] For other types of metadata, the data may already be represented in a form suitable for clustering. For instance, in the preferred embodiment time/date information is used and may be directly subjected to cluster analysis. In other cases, for example that of region metadata, the data may need to be processed for similarity by known algorithms. As an example, determination of facial similarity is described in M. Turk and A. Pentland, “Eigenfaces for Recognition”, *J. Cognitive Neurosci.* 3(1)1991.

Stubler [0034]

ants [4]. Several systems of locating human face have been reported. Eigenfaces, obtained by performing a principal component analysis on a set of faces, are commonly used to identify faces [5]. By moving a window covering a subimage

Yang, 1 (incorporated into Stubler [0030])

See, e.g., Petition 40, 42-43, 64; Pet'r Reply 21-23

Stubler discloses non-textual metadata for comparison

Machine Intelligence, 20(1), 39 - 51, January, 1998. A typical model for representing skin color in a chromatic space utilizing histogram analysis is described in an article by J. Yang and A. Waibel, "A Real-time Face Tracker," *Proceedings of Third IEEE Workshop on Applications of Computer Vision*, WACV 1996, pp. 142 - 147. For the

Stubler [0030]

and may be directly subjected to cluster analysis. In other cases, for example that of region metadata, the data may need to be processed for similarity by known algorithms. As an example, determination of facial similarity is described in M. Turk and A. Pentland, "Eigenfaces for Recognition", *J. Cognitive Neurosci.* 3(1)1991.

Stubler [0034]

ants [4]. Several systems of locating human face have been reported. Eigenfaces, obtained by performing a principal component analysis on a set of faces, are commonly used to identify faces [5]. By moving a window covering a subimage

Yang, 1 (incorporated into Stubler)

[0033] According to Ser. No. 09/291,857, color feature-based representation of an image is based on the assumption that significantly sized coherently colored regions of an image are perceptually significant. Therefore, colors of

Stubler [0033]

level. While the measurements and region identification is derived from low-level features, the processing(including knowledge based operations) may yield a region with a mixture of high-level feature description (clear sky, cloudy sky, mixed sky) and low-level features (color representation, area coverage, location). Similarities between such regions would be calculated on the basis of a combination of high-level and low-level features.

Stubler [0011]

or part of the image in which they are found. While the measurements and region identification is derived from low-level features, the processing(including knowledge based operations) may yield a region with a mixture of high-level feature description (clear sky, cloudy sky, mixed sky) and low-level features (color representation, area coverage, location). Similarities between such regions would be calculated on the basis of a combination of high-level and low-level features.

Stubler [0022]

boring pixels. Furthermore, according to Ser. No. 09/291,857, texture feature-based representation of an image is based on the assumption that each perceptually significantly texture is composed of large numbers of repetitions of the same color transition(s). Therefore, by identifying the fre-

Stubler [0033]

See, e.g., Petition 5-7, 40, 42-43, 64; Pet'r Reply 21-22

Patent Owner's arguments about "the object" rely solely on its improper construction

B. Stubler fails to disclose "an object information inputting section for confirming by the user, adding, and correcting the keyword which is proposed by the keyword proposing section when the inputted object acknowledged by the object acknowledging section is similar to the object of the image previously contained in the database."

As with the "image storing section" discussed above, Petitioner fails to meet its burden of proving that Stubler anticipates or renders obvious the claimed "object information inputting section" because Stubler does not disclose a way to determine that "the inputted object acknowledged by the object acknowledging section is similar to the object of the image previously contained in the database," as a POSITA would understand that phrase. Specifically, Stubler describes only

POR 41-42

See, e.g., Patent Owner's Response 41-42; Pet'r Reply 14-15, 23; PO Sur-Reply 20-23

DEMONSTRATIVE EXHIBIT – NOT EVIDENCE

Patent Owner's arguments for independent claims 7 and 8 rely on the same arguments as claim 1

C. Petitioner Fails to Demonstrate that Independent Claims 7 and 8 Are Unpatentable.

Petitioner also fails to demonstrate that independent claims 7 and 8 are unpatentable for the same reasons as claim 1. Claims 7 and 8 contain similar limitations to those discussed above with respect to claim 1. Specifically, claim 7

* * *

database.” Id., Cl. 8 (emphasis added). Because claims 7 and 8 contain similar limitations as claim 1, Petitioner fails to demonstrate that claims 7 and 8 are unpatentable for the same reasons that Petitioner fails with respect to claim 1.

POR 42-43

See, e.g., Patent Owner's Response 42-43; Pet'r Reply 26

Patent Owner does not dispute that Stubler discloses or renders obvious dependent claims 2-6 and 9-14

D. Petitioner fails to demonstrate that dependent claims 2-6 and 9-14 are unpatentable.

Dependent claims 2-6 each depend from claim 1 and dependent claims 9-14 each depend from claim 8. Therefore, claims 2-6 include all of the limitations of claim 1 and claims 9-14 include all of the limitations of claim 8. Accordingly, Petitioner fails to demonstrate a reasonable likelihood that claims 2-6 and 9-14 are unpatentable for the same reasons as claims 1 and 8.

POR 43

See, e.g., Patent Owner's Response 43; Pet'r Reply 26

Stubler+Wang Ground

Patent Owner does not separately dispute that Stubler and Wang render obvious claims 2-4 and 9-12

VII. Petitioner Fails To Demonstrate that Any Claim is Unpatentable Under Ground 2.

In Ground 2, Petitioner asserts that Stubler in view of Wang renders dependent claims 2-6 and 9-14 obvious. As discussed above, dependent claims 2-6 each depend from claim 1 and dependent claims 9-14 each depend from claim 8. Therefore, claims 2-6 include all of the limitations of claim 1 and claims 9-14 include all of the limitations of claim 8. Because Petitioner does not challenge independent claims 1 and 8 under Ground 2, Petitioner fails to demonstrate that claims 2-6 and 9-14 are unpatentable for the same reasons as claims 1 and 8.

POR 44

See, e.g., Patent Owner's Response 44; Pet'r Reply 26-27

Appendix

Claim 1 [as corrected, element identification added]

1. [1preamble] An image retrieving device for classifying and retrieving an image by detecting an object in the image and adding a keyword, the image retrieving device comprising:

[1a] an image storing section for storing the image which is supposed to be classified and retrieved together with the keyword in a database and an object of the image being previously contained in the database;

[1b] an image inputting detecting section that detects an inputted image that is newly inputted to the image retrieving device;

[1c] an object acknowledging section for acknowledging an inputted object in the inputted image that has been detected by the image inputting detecting section;

[1d] a keyword proposing section for proposing the keyword on a display, the keyword which relates to the inputted object which is acknowledged by the object acknowledging section; and

[1e] an object information inputting section for confirming by the user, adding, and correcting the keyword which is proposed by the keyword proposing section when the inputted object acknowledged by the object acknowledging section is similar to the object of the image previously contained in the database.

Ex. 1001, claim 1 (13:32-54, Cert. of Correction)

Claims 2, 3 [element identification added]

2. [2preamble] An image retrieving device according to claim 1, wherein the object acknowledging section includes:

[2a] a human detection condition inputting section for setting up conditions for determining whether or not the image contains a human;

[2b] a face image detecting section for detecting a face image in the image; and

[2c] a face image similarity determining section for detecting a face image which is detected by the face image detecting section so as to detect a similar face image stored in the database according to the detected face image.

3. An image retrieving device according to claim 2 wherein the object information inputting section serves as a personal information inputting section for confirming, adding, and correcting a personal information.

Claims 4-6 [as corrected]

4. An image retrieving device according to claim 2 further comprising a skin color area detecting section for detecting a skin color area in the image, wherein the skin color area detecting section is used when the human is detected.
5. An image retrieving device according to claim 2, wherein the keyword proposing section proposes the keyword to the image which is inputted last in a case in which the similar face image is not detected by the face image similarity determining section.
6. An image retrieving device according to claim 1 wherein the keyword is added according to a retrieving template which is formed by keywords which have hierarchical structure.

Claim 7 [as corrected, element identification added]

7. [7preamble] A method for adding keywords in an image retrieving device for classifying and retrieving an image by detecting an object in the image and adding a keyword, the method comprising:

[7a] storing the image which is supposed to be classified and retrieved together with the keyword in a database and an object of the image being previously contained in the database;

[7b] detecting an inputted image that is newly inputted to the image retrieving device by an image inputting detection section;

[7c] acknowledging an inputted object in the inputted image that has been detected by the image inputting detecting section;

[7d] proposing the keyword on a display, the keyword which relates to the inputted object which is acknowledged by an object acknowledging section; and

[7e] confirming by the user, adding, and correcting the keyword which is proposed by a keyword proposing section when the inputted object acknowledged by the object acknowledging section is similar to the object of the image previously contained in the database.

Ex. 1001, claim 7 (13:32-54, Cert. of Correction)

Claim 8 [as corrected, element identification added]

8. [8preamble] A computer readable medium storing a computer program for operating an image retrieving device which classifies and retrieves an image by detecting an object in the image and adding a keyword, the storage medium comprising:

[8a] storing the image which is supposed to be classified and retrieved together with the keyword in a database and an object of the image being previously contained in the database;

[8b] detecting an inputted image that is newly inputted to the image retrieving device by an image inputting detecting section;

[8c] acknowledging an inputted object in the image that has been detected by the image inputting detecting section;

[8d]proposing the keyword on a display, the keyword which relates to the inputted object which is acknowledged by an object acknowledging section; and

[8e] confirming by the user, adding, and correcting the keyword which is proposed by a keyword proposing section when the inputted object acknowledged by the object acknowledging section is similar to the object in the image previously contained in the database.

Ex. 1001, claim 8 (13:32-54, Cert. of Correction)

Claims 9-11 [as corrected]

9. An image retrieving device according to claim 2 wherein the keyword is added according to a retrieving template which is formed by keywords which have hierarchical structure.

10. An image retrieving device according to claim 3 wherein the keyword is added according to a retrieving template which is formed by keywords which have hierarchical structure.

11. An image retrieving device according to claim 4 wherein the keyword is added according to a retrieving template which is formed by keywords which have hierarchical structure.

Ex. 1001, claims 9-11 (13:32-54, Cert. of Correction)

Claims 12-14 [as corrected]

12. An image retrieving device according to claim 5 wherein the keyword is added according to a retrieving template which is formed by keywords which have hierarchical structure.

13. A method according to claim 7, wherein the keyword is added according to a retrieving template which is formed by keywords which have a hierarchical structure.

14. A computer readable medium storing a computer program according to claim 8, wherein the keyword is added according to a retrieving template which is formed by keywords which have a hierarchical structure.

Ex. 1001, claims 12-14 (13:32-54, Cert. of Correction)