



(51) International Patent Classification:

H04N 1/00 (2006.01) *H04N 1/32* (2006.01)
H04N 1/21 (2006.01) *H04N 5/232* (2006.01)

(21) International Application Number:

PCT/US2010/028206

(22) International Filing Date:

23 March 2010 (23.03.2010)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

12/412,663 27 March 2009 (27.03.2009) US

(71) Applicant (for all designated States except US): **MOTOROLA, INC.** [US/US]; 1303 East Algonquin Road, Schaumburg, Illinois 60196 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **TRIPAK, Thomas Michael**, [US/US]; 1310 Pine Street, Glenview, Illinois

60025 (US). **ATHALE, Anant**, [US/US]; 131 Stirling Lane, Schaumburg, Illinois 60194 (US).

(74) Agents: **VAAS, Randall S.** et al.; 600 North U.S. Highway 45, Libertyville, Illinois 60048 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE,

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR IMAGE SELECTION AND CAPTURE PARAMETER DETERMINATION

(57) Abstract: An apparatus and method for identifying image capture opportunities. Expert photo agents (526) that process metadata extracted from a captured image are obtained (706) at a user -device (102). Metadata (214) characterizing visual content of a captured image is produced (310) by processing the visual content of the captured image. Image capture opportunities are selected (410) based upon processing of the metadata with the at least one expert photo agent. The image capture opportunity selection specifies content to capture in a subsequent image.

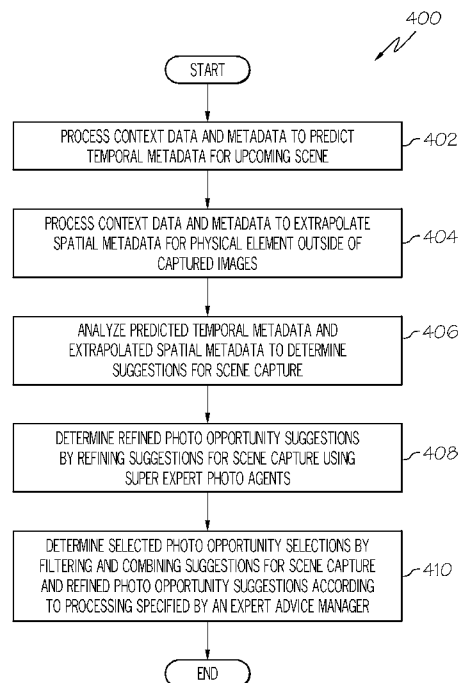


FIG. 4



ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
ML, MR, NE, SN, TD, TG).

— *before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments (Rule 48.2(h))*

Published:

— *with international search report (Art. 21(3))*

SYSTEM AND METHOD FOR IMAGE SELECTION AND CAPTURE

PARAMETER DETERMINATION

Field of the Invention

[0001] The present invention relates generally to the field of determining multimedia
5 content capture opportunities, and more particularly relates to selecting photo capture
parameters and identifying interesting scenes.

Background of the Invention

[0002] With the availability of multimedia capture capabilities on a wide array of
devices, communication is becoming more visual, and content sharing more social.
10 Problems traditionally faced by photojournalists and movie directors have become the
problems of billions of communication device users world-wide, namely: Where can I
get the “best shot”? Once I’m “on location”, how do I compose the best shot? When
should I shoot, to get the best shot? How can I collaborate with others on a shoot?

[0003] Although electronic cameras are able to incorporate automatic exposure
15 control, a user is still left to manually determine which scenes are of interest, and how
to best capture an image of them. Therefore, a need exists to improve upon the prior
art.

Summary of the Invention

[0004] Briefly, in accordance with the present invention, disclosed is a method for
20 identifying image capture opportunities including obtaining, at a user-device, at least
one expert photo agent adapted to process metadata extracted from a captured image.
The method also includes producing metadata characterizing visual content of a

captured image. The metadata are produced by processing the visual content of the captured image. The method also includes determining, in response to producing the metadata, an image capture opportunity selection based upon processing of the metadata with the at least one expert photo agent. The image capture opportunity
5 selection specifies content to capture in a subsequent image.

[0005] Also disclosed is an image capture opportunity detector that includes a memory and a processor that is communicatively coupled to the memory. The processor is adapted to obtain, at a user-device, at least one expert photo agent adapted to process metadata extracted from a captured image. The processor is also
10 adapted to produce metadata characterizing visual content of a captured image. The metadata is produced by processing the visual content of the captured image. The processor is further adapted to determine, in response to producing the metadata, an image capture opportunity selection based upon processing of the metadata with the at least one expert photo agent. The image capture opportunity selection may specify
15 content to capture in a subsequent image.

[0006] Further disclosed is an image capture opportunity image capturing device including a camera, at least one environmental sensor, a memory and a processor. The processor is communicatively coupled to the memory, the camera, and the environmental sensor. The processor is adapted to obtain, at a user-device, at least
20 one expert photo agent adapted to process metadata extracted from a captured image. The processor is also adapted to produce metadata characterizing visual content of a captured image. The metadata are produced by processing the visual content of the

captured image. The processor is further adapted to determine, in response to producing the metadata, an image capture opportunity selection based upon processing of the metadata with the at least one expert photo agent. The image capture opportunity selection specifies content to capture in a subsequent image. The
5 processor is also adapted to configure the camera to capture, with the at least one capture parameters defined by the image capture opportunity selection, the subsequent image

Brief Description of the Drawings

[0007] The accompanying figures, where like reference numerals refer to identical or
10 functionally similar elements throughout the separate views and which together with the detailed description below are incorporated in and form part of the specification, serve to further illustrate various embodiments and to explain various principles and advantages all in accordance with the present invention.

[0008] FIG. 1 illustrates a multiple user-device environment, in accordance with one
15 embodiment of the present invention.

[0009] FIG. 2 illustrates a block diagram of a user-device data flow, in accordance with one embodiment of the present invention.

[0010] FIG. 3 illustrates a diagram of an image processing flow, in accordance with one embodiment of the present invention.

20 [0011] FIG. 4 illustrates an opportunity determination processing flow, in accordance with one embodiment of the present invention.

[0012] FIG. 5 illustrates a user-device component block diagram, in accordance with one embodiment of the present invention.

[0013] FIG 6 illustrates an identified image capture processing flow, in accordance with one embodiment of the present invention.

5 [0014] FIG. 7 illustrates an expert agent selection and definition processing flow, in accordance with one embodiment of the present invention.

Detailed Description

[0015] While the specification concludes with claims defining the features of the invention that are regarded as novel, it is believed that the invention will be better
10 understood from a consideration of the following description in conjunction with the drawing figures, in which like reference numerals are carried forward.

[0016] One embodiment of the present invention includes a system that monitors an individual's activities and identifies contexts for capturing interesting media such as photos, movies, and the like. One embodiment of the present invention provides
15 processing to use shared or user-defined modules, as selected by a user, that work with sensors and actuators/displays to help guide the user to take, for example, memorable photographs. Example scenarios include, for example, automatically capturing media such as photos, movies and the like, to document highlights of an individual's active lifestyle, e.g., photographs of a mother bear and cub that were
20 unexpectedly encountered at close range while hiking in Rocky Mountain National Park.

[0017] One embodiment of the present invention includes a user-customizable system for context-driven media capture and event handling on mobile devices. Examples of processing performed by one embodiment of the present invention include: automated activity monitoring through, for example, image analysis of a scene that is being
5 currently captured by a camera or other image capture device. One embodiment further performs fusion processing of the results of the automated activity monitoring with other sensor data and guides the camera to the most interesting scene to monitor. The processing of one embodiment differentiates between normal and abnormal events that are detected by image analysis, e.g., break-in vs. normal opening of a car
10 door, via Outlier Detection techniques that have been developed for Data Mining applications.

[0018] In one embodiment, an image capture capability, such as a camera, is adapted to enable it to capture content that is identified as desirable. Based on the type of content captured, additional actions are taken. In one embodiment, additional image
15 processing is performed to determine if an emergency situation occurs, and a "Help" MMS message is sent to a friend, or a phone call is placed to an In Case of Emergency (ICE) contact.

[0019] FIG. 1 illustrates a multiple user-device environment 100, in accordance with one embodiment of the present invention. The multiple user-device environment 100
20 illustrates a number of user-devices, e.g., user 1's device 102, user 2's device 104 and user Z's device 106, that are in wireless communications with a base station 110 and/or with one another. In one embodiment, the user-devices all have similar

processing functionality, although further embodiments are able to operate with different types of user-devices. In the following description, user 1's device 102 is used as an example of the multiple user-devices that are able to be used by various embodiments of the present invention. It is to be understood that further
5 embodiments of the present invention are able to use user-devices that have different capabilities and functionalities from one another.

[0020] User-devices such as user 1's device 102 are able to communicate with other user-devices over any suitable communications medium. A wireless network incorporating a wireless communications base station 110 is illustrated in FIG. 1.
10 One example of a wireless communications base station 110 is a cellular base station that is part of a cellular network. User-devices are able to communicate with the base station 110 through, for example, cellular communications signals 120.

[0021] Further embodiments of the present invention are able to incorporate communications systems that allow wireless or wired communications directly
15 between user-devices or a combination of communications directly between user-devices and a central base station. In one embodiment, user-devices are also able to directly communicate with other user-devices via, for example, a Bluetooth connection 122 and/or networked services.

[0022] As described in further detail below, images, environmental data, expert photo
20 agents and other data or processing definitions used by a user-device is able to be communicated between a user-device and one or more user-devices to facilitate collaborative image collection and/or image collection techniques.

[0023] One embodiment of the present invention includes an agent database 112 that is a sharable library storing data that is shared, for example, among user-devices. As is described in further detail below, the agent database 112 stores a central repository of expert photo agents adapted to process metadata extracted from a captured image.

5 The expert photo agents stored in the agent database 112 define processing used to identify image capture opportunities or to control image capture parameters. Image capture parameters, such as exposure and the like, are able to be determined by particular expert photo agents based upon characteristics of the image presented to an image capture device, such as a camera, within a user-device. Some expert agents
10 process image characteristics to adjust image capture parameters so as to implement, for example, a desired photographic style.

[0024] FIG. 2 presents a block diagram of a user-device data flow 200, in accordance with one embodiment of the present invention. FIG. 3 presents a diagram of an image processing flow 300, in accordance with one embodiment of the present invention.

15 FIG. 4 illustrates an opportunity determination processing flow 400, in accordance with one embodiment of the present invention. The processing performed by a user-device 250, which in one example is a version of user 1's device 102, is described below with respect to both the user-device data flow 200, the image processing flow 300, and the opportunity determination processing flow 400.

20 [0025] The user-device data flow 200 shows a user-device 250 that includes several elements depicted therein. In one embodiment, user-device 250 includes a data transceiver 206 that communicates, for example, with other user-devices, such as user

2's device 104, through base station 110 or directly through a Bluetooth link 122. In one embodiment, data transceiver 206 receives data from a central database 202, which is part of the agent database 112. Data transceiver 206 is also able to receive data, such as images, photographic style models, environmental data, and the like, from associates 204 who have another user-device. Additionally, the data transceiver 206 of a user-device 250 is able to include circuitry to support communications with peer user-devices directly through, for example, Bluetooth connection 122. In one embodiment, data transceiver 206 is able to transmit observed environmental data to other user-devices. In one embodiment, data transceiver 206 is further able to transmit expert agents developed by a user of a particular user-device 250 to other user-devices or to a user defined agents database 240, such as would be included in agent database 112.

[0026] With reference to image processing flow 300 and user-device data flow 200, one embodiment of the present invention activates, at step 302, an image capture capability 210 of the user-device 250 and captures an image. User-device 250 of one embodiment, as is described in further detail below, is able to periodically or at various pre-determined times automatically activate the image capture capability 210 to capture a series of images.

[0027] In one embodiment, environmental sensors 208 within the user-device, such as sensors to detect sound, acceleration, temperature, location, and other environmental context data, are also polled and processed, at step 304, to gather data that is able to be integrated with data or metadata determined by processing of images, as is

described below. The sensors within the user-device of one embodiment are able to be periodically polled and processed without capturing an image in order to enhance image capture opportunity detection. In one embodiment, additional environmental context data are collected, at step 306, from peer-devices such as user 2's device 104
5 or any other environmental sensor through the data transceiver 206. In one embodiment, additional environmental data is exchanged with other devices via a Bluetooth connection and/or networked services, if and when such connections and data are available.

[0028] The collected environmental context data and processed image data in one
10 embodiment is stored in a decision input data store 214 where, in one embodiment, the data are organized into categories, e.g., location information (GPS, location relative to known objects, etc.) and capture device settings (focal length, aperture, zoom, flash, lighting, etc.).

[0029] Captured images are processed, at step 310, by one or more image processing
15 algorithms 212 to produce metadata characterizing visual content of a captured image by processing the visual content of the captured image. Metadata characterizing the captured images is stored in one embodiment within the decision input data 214. Processing algorithms 212 that are able to be applied determine image capture parameters are algorithms that determine, for example, parameters used to select
20 images as desirable, such as image illumination, color balance, focus, etc. These image capture parameters may have been selected according to a predefined image

collection parameter profile or according to image capture parameters that were set by the user.

[0030] One embodiment of the present invention is able to also apply further image recognition algorithms 212 to the captured images in order to generate, at step 312, additional metadata characterizing, for example, the presence of certain shapes/patterns in the content, and/or identifying certain objects in the content, e.g., Millennium Park in Chicago. In one embodiment, the metadata is produced by processing the visual content of the captured image with more than one expert photo agent, where each expert photo agent produces a respective set of initial metadata.

[0031] One embodiment of the present invention assembles the metadata determined by the image analysis processing 212 that is stored in the decision input data 214 in order to support detecting opportunities to capture images that are of interest to the user. The decision input data 214 is processed in one embodiment by extrapolation algorithms 215, and by one or more expert photo agents according to an expert advice manager 222. The extrapolation algorithms 215 of one embodiment produce a set of extrapolated metadata characterizing extrapolated visual content based upon information contained within captured images. The metadata of one embodiment is produced by processing the visual content of captured images. In one embodiment, expert advice manager 222 manages and applies various expert photo agents, including, for example, expert photo agents 218 and super expert photo agents 220, as are described in more detail below. The processing of assembled metadata in one

embodiment is performed according to the opportunity determination processing flow 400.

[0032] The processing of assembled metadata to determine image capture opportunities includes the periodic evaluation, at step 402, by a Temporal
5 Extrapolation Engine (TEE) of the assembled metadata that was determined by analysis of one or more captured images and/or detected environmental data. The temporal extrapolation engine of one embodiment is part of the extrapolation algorithms 215 and determines further image capture opportunities for a user of a particular device. The temporal extrapolation engine of one embodiment outputs
10 vectors of temporally extrapolated metadata characterizing image capture opportunities that it predicts will occur in the future.

[0033] Examples of processing and predictions performed by the temporal extrapolation engine of one embodiment of the present invention include: 1) identifying motion trajectories of objects currently in the scene, 2) using the Doppler
15 Effect to “chase” moving objects that have transitioned from approaching to leaving the scene, 3) matching images to a database of time-tagged recurring events, e.g., Buckingham Fountain spouts water the highest during the first 5 minutes of each hour, and 4) applying Laws of Physics to estimate what will happen next to the objects in a scene, e.g., gravity, collisions, explosions, etc.

20 [0034] The processing of assembled metadata to determine image capture opportunities includes periodic evaluation, at step 404, by a Spatial Extrapolation Engine (SEE) of the assembled metadata that was determined by analysis of one or

more captured images and/or detected environmental data. The spatial extrapolation engine of one embodiment is part of the extrapolation algorithms 215 and produces spatially extrapolated metadata characterizing estimates of other interesting image capture opportunities that may exist in the user's current vicinity. The spatial
5 extrapolation engine of one embodiment outputs vectors of spatially extrapolated metadata for scenes that the spatial extrapolation engine predicts to exist, based upon extrapolation of known metadata, if the image capture device were to be re-oriented in a specified direction.

[0035] One embodiment of the present invention implements a spatial extrapolation
10 engine that is able to: 1) compare the current scene to a database of known images, and extrapolate the "missing portion." 2) Integrate scene prediction with data from other environmental sensors, e.g., directional audio. For example, if a bird call is heard from a given direction, then there is a good probability that a bird can be photographed by aiming the image capture device in that direction. 3) Calculate the
15 location of a light source in the current scene, and subsequently estimating the location of a region of maximal (best) illumination or backlighting (if this is the desired effect). 4) Use thermal imaging to identify temperature gradients in the image, and follow the gradient to points outside of the current scene to image other heat-producing entities. In one embodiment, the processing includes collecting
20 environmental data from environmental sensors that are associated with the captured image, and producing a set of environmentally extrapolated metadata characterizing extrapolated visual content based upon the environmental data. Such embodiments

determine an image capture opportunity selection based further upon the environmentally extrapolated metadata.

[0036] The extrapolated metadata produced by the spatial extrapolation engine and temporal extrapolation engine within the extrapolation algorithms 215 are analyzed, at step 406, to determine suggestions to provide to a user for image capture, such as photographs or movies to capture. One embodiment of the present invention further uses one or more Expert Photo Agents (EPAs) 218 to process, at step 408, the aggregated image, image metadata and context data that was gathered and/or determined for a current media capture scene. In one embodiment, the one or more Expert Photo Agents (EPAs) 318 are contained within a shared agent library 216. One embodiment of the present invention is able to store expert photo agents in the agent library 216 that consist of one or more of pre-configured expert photo agents that are configured into the user-device 250, expert photo agents that are defined by a user of the user-device 250, or expert photo agents that are downloaded from the agent database 112 through, for example, data transceiver 206.

[0037] The expert photo agents of one embodiment are able to further process metadata and other information determined for extrapolated scenes, e.g., scenes that are either spatially or temporally extrapolated based upon image and environmental data that was received/determined from captured data. Individual expert photo agents of one embodiment represent and/or determine one or more perspectives on the quality of the scene.

[0038] One embodiment of the present invention includes expert photo agents to determine, for example, image content balance and/or symmetry, identification of the presence of a dog or of the device owner's dog, identification of straight edges within the image, identification of shadows within the image, and the like.

- 5 [0039] One embodiment of the present invention includes one or more expert photo agents that identify various objects or characteristics. Examples of some expert photo agents include, but are not limited to, agents that determine:

[0040] 1) When to take a photo of a human subject, such as: focus on subject's head, head and shoulders, head to waist, full body. Capture an image whenever someone is
10 in a scene. Capture image of profile or facing-forward subject. Wait until subject smiles. Selecting one or more of these characteristics is able to be defined by the user or according to a sharable profile.

[0041] 2) Desired lighting to capture an image, e.g., where is/are the light sources, relative to the image capture device?

- 15 [0042] 3) Determination of object edges in a scene. Characterizations include determining a number of edges in the scene. Other characterization to determine when to capture desirable images include "Avoid the arrow," i.e., do any of the edges in the image form arrows? Do they form arrows when viewed with the corner of an image?

[0043] 4) Object Centering: How much distance is there from the centroid of the largest/most important objects/people in the image? Who is in the center of the picture? (Most important person?)

[0044] 5) Contrast: a) image level, and/or 2) Object level, e.g., “6 black ducklings vs.
5 1 yellow duckling”

[0045] 6) Foreground/background: which one is in focus? Are both in focus? Number of objects in foreground/background. Image lighting.

[0046] 7) Similarity of objects in image: repeated shapes, repeated gestures (as a set of shapes), what is the perceived axis of symmetry? Profile or user preference
10 determines how this is computed: e.g., by color, shape or other logical content. Reflected shapes and repeated entities in different forms, e.g., a dog and a sign with a dog, etc.

[0047] 8) Themes: e.g., some well-understood event, e.g., gestures indicating or detected sound saying “Hey that’s mine. No. It’s mine,” emotion, repeated entities in
15 different forms, e.g., two dogs, a dog and a sign with a dog, etc.

[0048] 9) Humor: audible detection of defined comical phrases, image detection of inanimate objects/animals doing human things, e.g., “Enjoying the scent of a flower”. (sensorial).

[0049] 10) Thought provoking, e.g., “a good picture poses a good question.”

[0050] 11) Detected motion in a sequence of images: stop-action, still life, number of independent object motions in a scene, number of parallel/dependent object motions in a scene.

[0051] 12) Uniqueness of images: check captured image with images stored in a local
5 image database, check with a global image database. Checking with, for example, pre-scanned images that are readily characterized and analyzed through a neural network.

[0052] 13) Logical Completeness: e.g., “One fish for every claw,” each human has 2 legs, each animal has 4 legs, expected combinations.

10 [0053] 14) Evocation of other senses: e.g., image about enjoying the scent of a flower, image with motion, implies touching something, image may imply loud music, advertising images for food products imply taste and smell - desire.

[0054] 15) Imbalance: e.g. detected irony in an image such as “moose vs. fighter jet,”
“horses have only 1 head,” unexpected combinations, e.g., a camel and snowman in
15 the same scene.

[0055] 16) Geometric Fit, e.g., “the cat fits in the shoe.”

[0056] 17) Orientation of items/people in the image: direction of gaze, orientation of primary axis.

[0057] 18) Determination of embedded shapes/signs: “Dog’s legs form the letter T.”

Determination is able to be by, for example, Hough Transforms to efficiently identify shapes in images.

[0058] In one embodiment, images are periodically captured and processed, along

5 with other available data, to determine if conditions exist to satisfy a rule or model that indicates the captured image is an interesting image that should be retained.

If/when the context data associated with a given scene are sufficient to trigger the rules/model within a given expert photo agents, the expert photo agents of one embodiment is able to provide an output indicating this status. For example, an

10 output may have the following form: “photo opportunity” information:

[0059] {EPA ID, scene ID, interestingness score, <context summary> vector, <photo expert advice parameter set> vector}.

[0060] In order to allow a user to better capture the detected image of interest, one embodiment of the present invention is able to produce photo expert advice to assist

15 the user to better capture the image. Photo expert advice is able to include: zoom, focus, flash, contrast, red eye reduction, filtering, and timing for a specific event.

[0061] One embodiment of the present invention includes a library of Super-Expert Photo Agents (SEPAs) 220 that process, at step 408, instances of “photo opportunity” information output by the expert photo agents selected to process metadata and image

20 data. In one embodiment, the super expert photo agents implement one or more methods of combining the recommendations from multiple expert photo agents 218 regarding particular scene identification. For example, a super expert photo agent 220

is able to be defined to create hybrid or compound scores that are combinations of the outputs of the expert photo agents. One example of a super expert photo agent creates an arithmetic average of the outputs of multiple expert photo agents or a pre-defined weighted average of outputs of multiple expert photo agents. For example, a super expert photo agent may combine the output of three expert photo agents, identified as output quantity A1, A2, and A3, respectively as $(A1+A2+A3)/3$. Another super expert photo agent is able to combine these three outputs according to the equation $((0.5*A1)+(0.2*A2)+(0.3*A3))$ to produce an output that is evaluated to determine selected photo capture opportunities. In one embodiment, determining the image capture opportunity selection by the super photo agents 220 is based upon such a combination of each of the respective set of initial metadata that is according to pre-defined criteria into a composite recommendation.

[0062] The data produced by the Expert Photo Agents 218 and the Super-Expert Photo Agents 220 are received and analyzed, at step 410, by an expert advice manager 222. The above analysis steps, in one embodiment, determine an image capture opportunity selection based upon processing of the metadata characterizing the captured image with at least one expert photo agent, where the image capture opportunity selection specifies content to capture in a subsequent image.

[0063] The expert advice manager 222 of one embodiment monitors, filters, and combines the instances of “photo opportunity” information according to the recommendations provided by the expert photo agents 218 and super expert photo agents 220. In one embodiment, the output of the super expert photo agent 220 is

compared to a threshold to determine selected photo capture opportunities. The determining of the image capture opportunity selection of one embodiment includes filtering a plurality of composite recommendations, such as a time-sequence of composite recommendations, according to pre-defined criteria to produce the image
5 capture opportunity selection.

[0064] FIG. 5 illustrates a user-device component block diagram 500, in accordance with one embodiment of the present invention. The user-device block diagram 500 shows several components included within a user-device, such as user 1's device 102. Various processing components within the user-device block diagram 500 are able to
10 communicate via a data communications bus 550.

[0065] A camera 510 of the user-device block diagram captures images to be processed by other components. Camera 510 of one embodiment is able to capture a sequence of images to operate as, in one example, a movie camera. Images captured by camera 510 are processed by an image processor 504. Image processor 504 is able
15 to include any suitable image processing architecture, such as a programmable microprocessor, configurable digital signal processing hardware, or any suitable combination of those two or other processing architectures. Image processor 504 includes, for example, feature/shape detection processing, temporal extrapolation engines, spatial extrapolation engines, and any other image processing used to support
20 expert photo agent processing.

[0066] Data produced by either one or both of camera 510 and image processor 504 is able to be stored in the data storage 520, as described below. Sensors 514 are able to

determine environmental information, such as sounds, temperature, location, and the like.

[0067] A user interface 508 allows a user of the user-device to configure and/or control the operation of the user-device. User interface 508 is able to accept input
5 from a user and is also able to display prompts and other information to the user. Camera controller 502 of one embodiment controls operation of the camera 510. Camera controller 502 is controlled either in response to inputs from the user via the user interface 508, or according to processing defined for the user-device according to, for example, expert systems or information determined by the image processor
10 504.

[0068] A data transceiver 512 of one embodiment is able to communicate with peer devices of associates of a user of a user-device. Data exchanged over the data transceiver 512 is discussed above and includes, for example, environmental information detected by sensors 514 and/or image capture related information
15 extracted by the image processor 504 and/or expert system processor 506.

[0069] Expert system processor 506 processes, for example, metadata extracted or determined by the image processor 504 for one or more images captured by the camera 510 or environmental information detected by environmental sensors 514. Examples of expert systems processing, including expert photo agents, super expert
20 photo agents, and expert agent managers, as is performed by one embodiment of the present invention is described above. Suggestions determined by the Expert system

processor 506 of one embodiment of the present invention are able to be provided to a user via the user interface 508 or to camera controller 502 to capture images.

[0070] Data storage 520 includes a memory to store various items used by a user-device of one embodiment of the present invention. Data storage 520 includes an image database 522 to store captured images captured by camera 510. The image database 522 of one embodiment is able to store a sequence of capture images to support, for example, time extrapolation.

[0071] Data storage 520 further includes an opportunity history database 524 that is used to store image capture opportunities identified by the expert system processor 506. Data storage 520 further includes an expert agent definition database 526 that is used to store definitions of expert agents that are to be implemented by the expert system processor 506. Data storage 520 further includes a super expert agent definition database 528 that is used to store definitions of expert agents that are also to be implemented by the expert system processor 506. Expert agent definitions and super expert agent definitions that are stored in the expert agent definition database 526 and super expert agent definition database 528, respectively, are able to be one or more of pre-defined expert agents that are programmed into the user-device, user defined expert agents, and/or expert agents received from other user-devices or a remote agent database 112 through data transceiver 512.

[0072] Data storage 520 further includes expert advice management definitions 530 that define expert agent management algorithms that implement the expert advice

manager 222 that manages and combines the output of various expert agents and super expert agents.

[0073] FIG 6 illustrates an identified image capture processing flow 600, in accordance with one embodiment of the present invention. The identified image capture processing flow 600 directs the capture of images that were identified by the expert advice manager 222 as possibly interesting according to the above described processing. The expert advice manager 222 of one embodiment identifies, at step 602, selected photo opportunity selections by comparing scores produced by the expert agent manager 222 to a threshold for minimum interestingness score, as specified, for example, by the user or a default value. If the score exceeds the threshold, the photo opportunity is identified as a selected photo opportunity. The “photo opportunity” instances selected by the expert advice manager 222, in one embodiment, are communicated to the User Interface (UI) Manager 508 and other application software controlling the media capture devices.

[0074] In one mode of operation, the “photo opportunity” instances determined by the expert advice manager 222 are configured to be “automatically captured,” at step 604. If “auto capture” is configured, the processing advances to capture, at step 606, the image specified by the selected photo opportunity. These captured images are then stored in the automatic capture database 224 without further user intervention. The processing then automatically analyzes, at step 608, the context summary for a “photo opportunity” instance with a Personal Safety Monitoring Engine to determine if pre-specified personal safety threat conditions are detected. The processing determines,

at step 610 if there is a person safety threat and if so, the corresponding action is taken, e.g., a call is placed to 911, at step 612.

- [0075] In the case that auto capture is determined, at step 604, to not be enabled, the processing continues by providing, at step 614, metadata describing one or more current “photo opportunity” instances are displayed on the user interface 508. The expert advice parameter set for one or more “photo opportunity” instances determined by the expert advice manager 222 are able to be further analyzed by a user cue generator 230 to generate user cues based on the time, location, and/or orientation of potentially interesting scenes.
- [0076] The user cues derived based upon analysis of the “photo opportunity” instances determined by the expert advice manager 222 are able to be presented via one or more parts of the user interface 508, which is able to include a display, viewfinder, speaker, haptic interface, and the like. Further ways of providing user cues include, for example, a user interface 508 that includes a speaking voice, e.g., “Move left 30 degrees for a better view,” guidance in the border of the viewfinder, the device shakes when you’re not pointing it in a good direction, and it shakes less as you get closer to preferred direction, motorized mirror assembly to automatically track the best scene, e.g., a mechanical system that mimics how a frog’s eyeballs bulge out and move a few degrees, even though the frog’s head is motionless. In the case of providing user cues to change an image capture parameter, such as zoom, focus, and the like, the viewfinder is able to, for example, indicate zooming out by

shrinking the displayed image and indicate zooming in by shading the border of the displayed image.

[0077] The processing determines, at step 616, if the user selects to capture one of these identified photo opportunities. If the user selects to capture one of the identified photo opportunities, the associated metadata is interpreted and used to capture, at step 5 622, media for the selected “photo opportunity.” In one embodiment, the identified photo opportunities include an image capture opportunity selection that contains image capture parameters to be used to capture a subsequent image. The interpretation of metadata and media capture of one embodiment captures, with the at least one capture parameters defined by the image capture opportunity selection, the 10 subsequent image.

[0078] In a case of a user not selecting to capture, at step 616, the processing determines, at step 618, if a timeout occurred. If a time out did not occur, the processing returns to determining, at step 616, if the user selects to capture. If a 15 timeout does occur, the “photo opportunities” that are not selected by the user within the pre-specified time of the time out are removed from the current list and stored, at step 620, in an opportunity history data base 226. The opportunity history data base 226 of one embodiment may be reviewed by the user at a future time for, e.g., training purposes and/or for reconstructing potentially interesting scenes.

[0079] After a selected photo opportunity is stored, at step 620, or after an automatically captured photo is stored and analyzed, at step 610, the processing is able to send, at step 624, the selected photo opportunity selection capture parameters to 20

user-devices used by associates of the user of this user-device. In one embodiment, selected photo opportunity selection capture parameters are sent or not sent to associates according to a pre-configured or user defined parameters. In one embodiment, the selected photo opportunity selection capture parameters are
5 determined, for example, by the expert agent manager 222 alone or in combination with user input.

[0080] FIG. 7 illustrates an expert agent selection and definition processing flow 700, in accordance with one embodiment of the present invention. One embodiment of the present invention allows a user to decide, at step 702, to initiate a search of external
10 sources for one or more desired expert photo agents or super expert photo agents. In one embodiment, expert photo agents are available from external data bases, such as agent database 112, that operate to automatically implement photo capture parameters in a manner that emulates certain professional photographers or photojournalists. An operator of agent database 112 is able to charge users to download particular expert
15 photo agents. A user is also able to query associate's user-devices to obtain an expert photo agent or super expert photo agent possessed and/or defined by that associate.

[0081] If the user decides to search for an expert photo agent, the processing searches, at step 704, external sources for expert photo agents. In one embodiment, the user forms a request to receive photo agents that match a user's specifications.
20 That request is then transmitted to the sharable library, such as the agent database 112. The application software implementing intelligent media capture in one embodiment of the present invention allows sending and receiving expert photo

agents and super expert photo agents in a standard representation or format. Expert photo agents or super expert photo agents are able to be communicated through any suitable medium, such as via Multimedia Message (MMS). Once the search is completed, the processing receives, at step 706, at least one expert photo agent or super expert photo agent that match the user's specification from the sharable library. In one embodiment, receiving at least one expert photo agent or super expert photo agent is obtaining, at a user-device, at least one expert photo agent adapted to process metadata extracted from a captured image. The user is given the option to apply a received expert photo agents or super expert photo agents to currently captured images and environmental data, or the user is able to select storing it (as inactive) for reference. A user is also able to simply discard received agents.

[0082] The processing then gives the user an opportunity to select, at step 708, one or more expert photo agents or super expert photo agents to use to capture images. If the user selects an agent, the selection is processed, at step 710. If the user did not decide to select an agent, the processing selects a pre-defined or default agent to use to capture images. The processing then proceeds to use the selected agent to select and/or process, at step 714, images that were captured with the camera 510.

[0083] Once an image is captured, the processing gives the user an option, at step 716, to modify image capture parameters. If the user opts to modify capture parameters, the processing accepts, at step 718, user defined capture parameters. In one embodiment, user defined capture parameters are accepted through user interface 508. Once user defined parameters are accepted, one embodiment of the present

invention develops, at step 720, an expert model for the user's style based on the user defined capture parameters. The developed expert model is able to contain one or more expert photo agents and/or super expert photo agents that will yield consistent image capture results based upon the user's photographic style as determined by his
5 or her manually defined capture parameters.

[0084] In one embodiment, Image and/or Video Mining techniques and Reinforcement Learning methods are used, during periods of low processor load, to create and/or improve the image capture model for a given user. In this way, it is possible to automatically discover style characteristics for a given user, based on his
10 or her media captured. In one embodiment, the user's model is sent to a searchable, networked repository, such as agent database 112. In this manner, the sharable library comprises expert photo agents developed by users of peer systems.

[0085] A user of one embodiment is able to share expert photo agents or super expert photo agents with associates or even a more general audience. One embodiment of
15 the present invention includes user-devices that give the user an option, at step 722, to send the user-developed model, which consists of photo expert agents and/or super expert photo agents developed to mirror the user's style, to associates. In one embodiment, the user is able to send a whole model or to compress the model to a certain size by manual or automatic techniques, such as by using the Text Mining
20 approach of Adaptive Text Summarization, and/or by omitting certain features of the model.

[0086] The processing then continues to present an option to the user, at step 726, to share the captured image. If the user selects to share the image, the image is sent, at step 728, to associates. If the user opts to share the captured image or not, the image is stored, at step 730, and the processing ends.

5 [0087] The terms program, software application, and the like as used herein, are defined as a sequence of instructions designed for execution on a computer system. A program, computer program, or software application may include a subroutine, a function, a procedure, an object method, an object implementation, an executable application, an applet, a servlet, a source code, an object code, a shared
10 library/dynamic load library and/or other sequence of instructions designed for execution on a computer system.

[0088] Reference throughout the specification to "one embodiment" means that a particular feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment of the present invention. Thus, the
15 appearances of the phrases "in one embodiment" in various places throughout the specification are not necessarily all referring to the same embodiment. Furthermore, the particular features, structures, or characteristics may be combined in any suitable manner in one or more embodiments. Moreover these embodiments are only examples of the many advantageous uses of the innovative teachings herein. In
20 general, statements made in the specification of the present application do not necessarily limit any of the various claimed inventions. Moreover, some statements may apply to some inventive features but not to others. In general, unless otherwise

indicated, singular elements may be in the plural and visa versa with no loss of generality.

[0089] While the various embodiments of the invention have been illustrated and described, it will be clear that the invention is not so limited. Numerous
5 modifications, changes, variations, substitutions and equivalents will occur to those skilled in the art without departing from the spirit and scope of the present invention as defined by the appended claims.

What is claimed is:

CLAIMS

1. A method for identifying image capture opportunities, the method comprising:
obtaining, at a user-device, at least one expert photo agent adapted to process
5 metadata extracted from a captured image;
producing metadata characterizing visual content of a captured image, the
metadata being produced by processing the visual content of the captured image; and
determining, in response to producing the metadata, an image capture
opportunity selection at the user-device based upon processing of the metadata with the
10 at least one expert photo agent, the image capture opportunity selection specifying
content to capture in a subsequent image.
2. The method of claim 1, wherein the image capture opportunity selection
contains image capture parameters to be used to capture the subsequent image, the
15 method further comprising:
capturing, with the at least one capture parameters defined by the image capture
opportunity selection, the subsequent image.

3. The method of claim 1, wherein

producing the metadata comprises processing the visual content of the captured image with a plurality of expert photo agents, wherein each expert photo agent within the plurality of expert photo agents produces a respective set of initial metadata,

5 wherein the determining comprises determining the image capture opportunity selection based upon a combination of each of the respective set of initial metadata according to a pre-defined criteria.

4. The method of claim 1, wherein

10 producing the metadata comprises:

processing the visual content of the captured image with a plurality of expert photo agents, wherein each expert photo agent within the plurality of expert photo agents produces a respective set of initial metadata;

15 combining each of the respective set of metadata into a plurality of composite recommendations according to a respective pre-defined criteria for each composite recommendation; and

generating a plurality of image capture recommendations the image capture opportunity selection based upon a combination of each of the respective set of metadata according to a pre-defined criteria; and

20 determining comprises filtering the plurality of composite recommendations according to pre-defined criteria to produce the image capture opportunity selection.

5. The method of claim 1, further comprising:

producing a set of extrapolated metadata characterizing extrapolated visual content based upon information contained within a captured image, the metadata being produced by processing the visual content of the captured image,

5 wherein the determining the image capture opportunity selection is further based upon the extrapolated metadata.

6. The method of claim 5, wherein the set of extrapolated metadata comprises spatially extrapolated metadata determined by processing of the captured image and

10 temporally extrapolated metadata determined by processing of the captured image.

7. The method of claim 1, further comprising:

collecting environmental data associated with the captured image; and

producing a set of environmentally extrapolated metadata characterizing

15 extrapolated visual content based upon the environmental data,

wherein the determining the image capture opportunity selection is further based upon the environmentally extrapolated metadata.

8. The method of claim 7, wherein the set of environmentally extrapolated

20 metadata comprises spatially extrapolated metadata determined by processing of the captured image and temporally extrapolated metadata determined by processing of the captured image.

9. The method of claim 1, wherein the obtaining comprises:
forming a request to receive photo agents that match a user's specifications;
transmitting the request to a sharable library; and
receiving, from the sharable library, the at least one expert photo agent that
5 match the user's specification.

10. The method of claim 1, wherein the obtaining obtains the at least one expert
photo agent from a sharable library that comprises a central repository of expert photo
agents.

10

11. The method of claim 1, wherein the obtaining obtains the at least one expert
photo agent from a sharable library that sharable library comprises expert photo agents
developed by users of peer systems.

15 12. The method of claim 1, wherein the obtaining obtains the at least one expert
photo agent from a sharable library, the method further comprising:

maintaining the sharable library, wherein the sharable library is
communicatively accessible to a plurality of user-devices; and
charging, in response to the obtaining, a user for use of the at least one expert
20 photo agent.

13. An image capture opportunity detector, comprising:

a memory;

a processor, communicatively coupled to the memory, the processor adapted to:

obtain, at a user-device, at least one expert photo agent adapted to

5 process metadata extracted from a captured image;

produce metadata characterizing visual content of a captured image, the
metadata being produced by processing the visual content of the captured
image; and

determine, in response to producing the metadata, an image capture
10 opportunity selection at the user-device based upon processing of the metadata
with the at least one expert photo agent, the image capture opportunity selection
specifying content to capture in a subsequent image.

14. The image capture opportunity detector of claim 13, wherein the processor is

15 further adapted to:

produce the metadata by processing comprising processing the visual content of
the captured image with a plurality of expert photo agents, wherein each expert photo
agent within the plurality of expert photo agents produces a respective set of initial
metadata; and

20 determine the image capture opportunity selection by processing comprising
determining the image capture opportunity selection based upon a combination of each
of the respective set of initial metadata according to a pre-defined criteria.

15. The image capture opportunity detector of claim 13, wherein the processor is further adapted to:

produce the metadata by processing comprising:

5 processing the visual content of the captured image with a plurality of expert photo agents, wherein each expert photo agent within the plurality of expert photo agents produces a respective set of initial metadata;

combining each of the respective set of metadata into a plurality of composite recommendations according to a respective pre-defined criteria for each composite recommendation; and

10 generating a plurality of image capture recommendations the image capture opportunity selection based upon a combination of each of the respective set of metadata according to a pre-defined criteria; and
determine the image capture opportunity selection by processing comprising filtering the plurality of composite recommendations according to pre-defined criteria
15 to produce the image capture opportunity selection.

16. The image capture opportunity detector of claim 13, wherein the processor is further adapted to:

20 produce a set of extrapolated metadata characterizing extrapolated visual content based upon information contained within a captured image, the metadata being produced by processing the visual content of the captured image; and

determine the image capture opportunity selection based further upon the extrapolated metadata.

17. The image capture opportunity detector of claim 13, further comprising:
at least one environmental sensor adapted to collect environmental data
associated with the captured image, and

5 wherein the processor is further adapted to:

producing a set of environmentally extrapolated metadata characterizing
extrapolated visual content based upon the environmental data; and

determine the image capture opportunity selection is based further upon
the environmentally extrapolated metadata.

10

18. An image capture opportunity image capturing device, comprising:

a camera;

at least one environmental sensor;

a memory;

5 a processor, communicatively coupled to the memory, the camera, and the environmental sensor, the processor adapted to:

obtain, at a user-device, at least one expert photo agent adapted to process metadata extracted from a captured image;

10 produce metadata characterizing visual content of a captured image, the metadata being produced by processing the visual content of the captured image;

determine, in response to producing the metadata, an image capture opportunity selection at the user device based upon processing of the metadata with the at least one expert photo agent, the image capture opportunity selection specifying content to capture in a subsequent image; and

15 configure the camera to capture, with the at least one capture parameters defined by the image capture opportunity selection, the subsequent image.

19. The image capture opportunity image capturing device of claim 18, further comprising:

20 a data transceiver, communicatively coupled to the memory, the data transceiver adapted to receive the at least one expert photo agent.

20. The image capture opportunity image capturing device of claim 18, wherein the processor is adapted to produce metadata by a process comprising:

5 processing the visual content of the captured image with a plurality of expert photo agents, wherein each expert photo agent within the plurality of expert photo agents produces a respective set of initial metadata;

 combining each of the respective set of metadata into a plurality of composite recommendations according to a respective pre-defined criteria for each composite recommendation; and

10 generating a plurality of image capture recommendations the image capture opportunity selection based upon a combination of each of the respective set of metadata according to a pre-defined criteria, and

 wherein the processor is adapted to determine an image capture opportunity selection by a process comprising filtering the plurality of composite recommendations according to pre-defined criteria to produce the image capture opportunity selection.

1 / 6

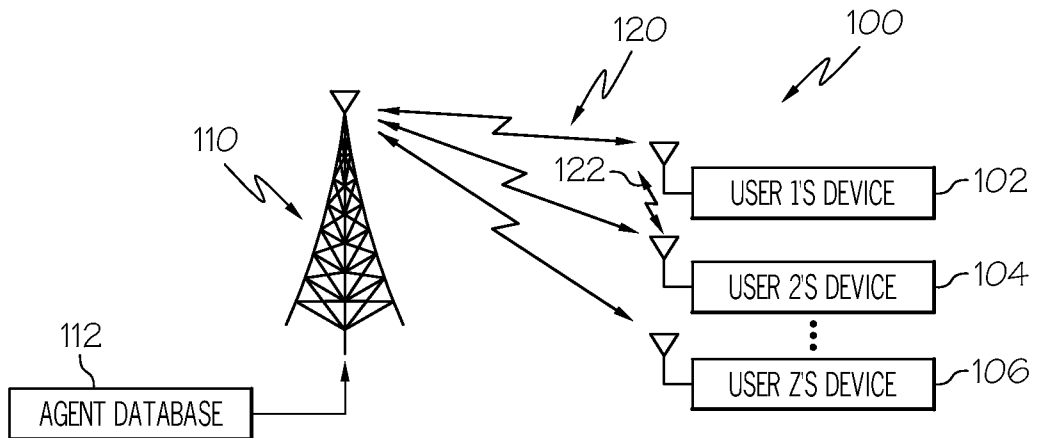


FIG. 1

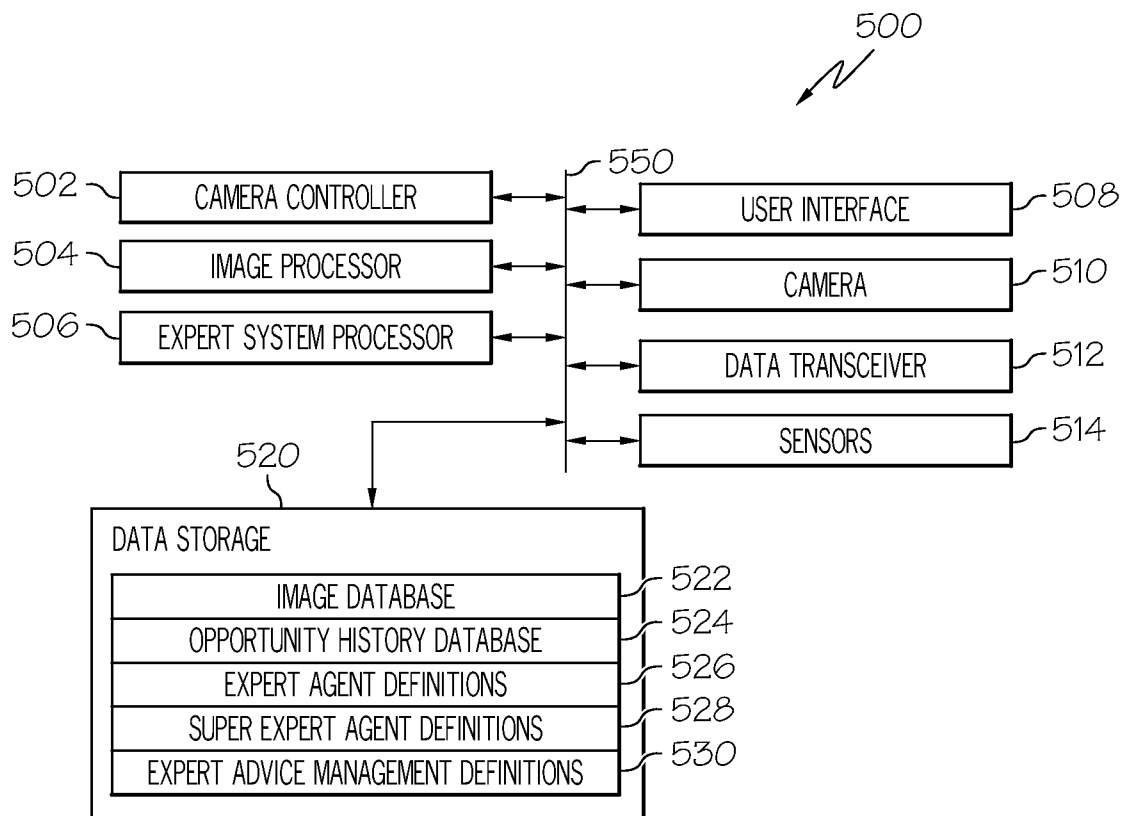


FIG. 5

2 / 6

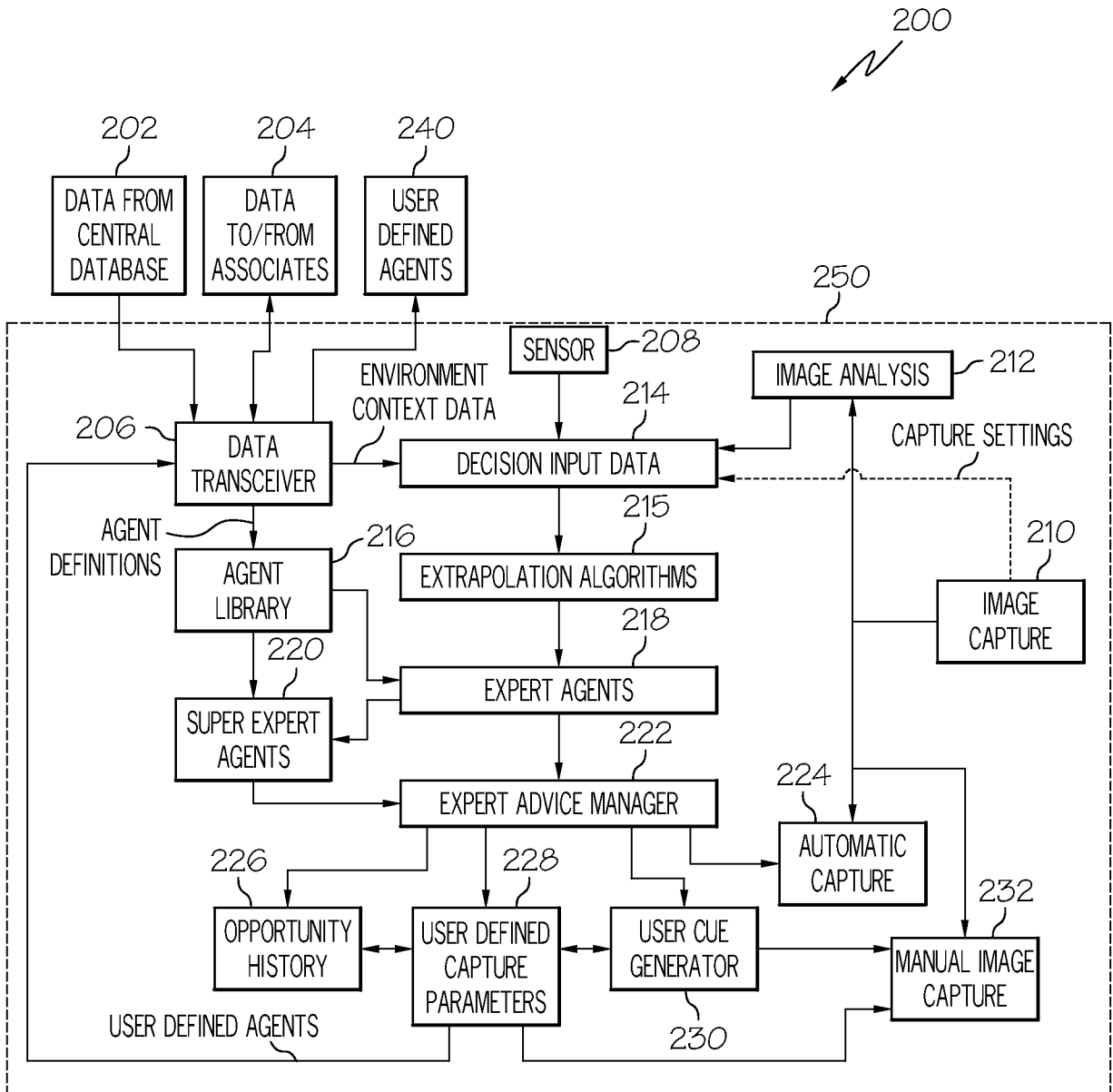


FIG. 2

3 / 6

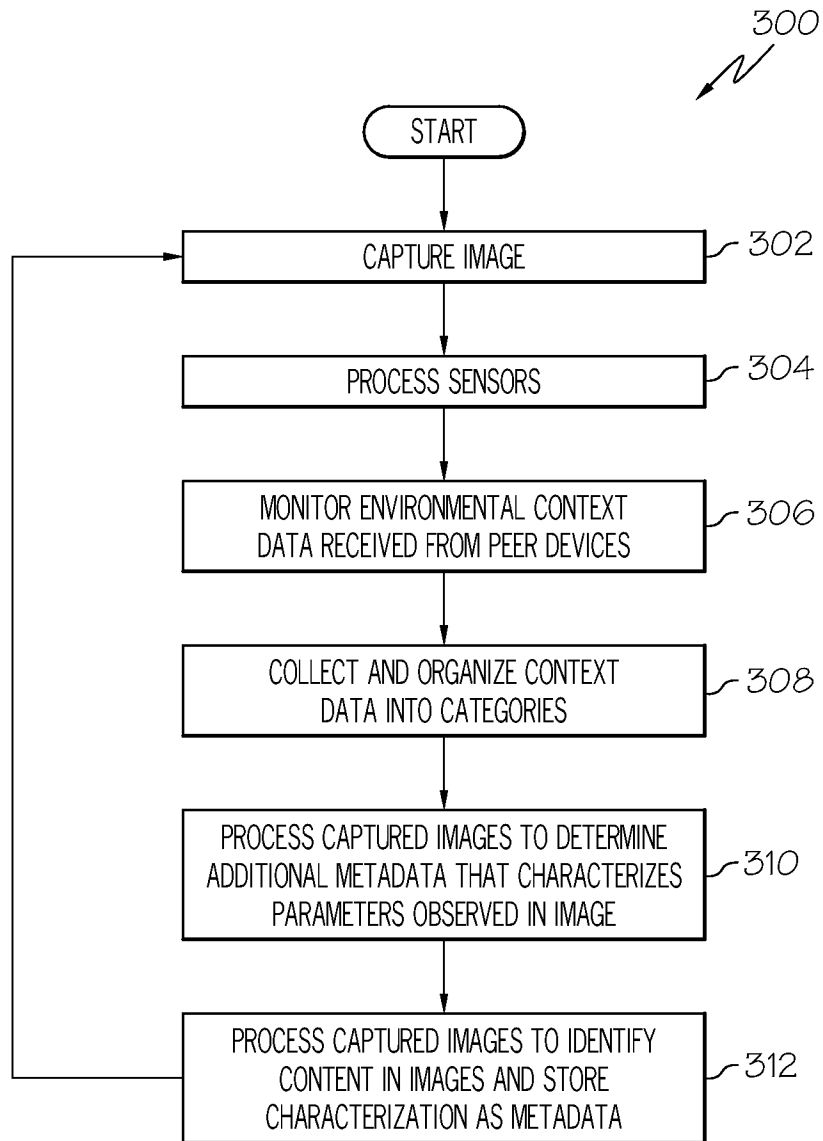


FIG. 3

4 / 6

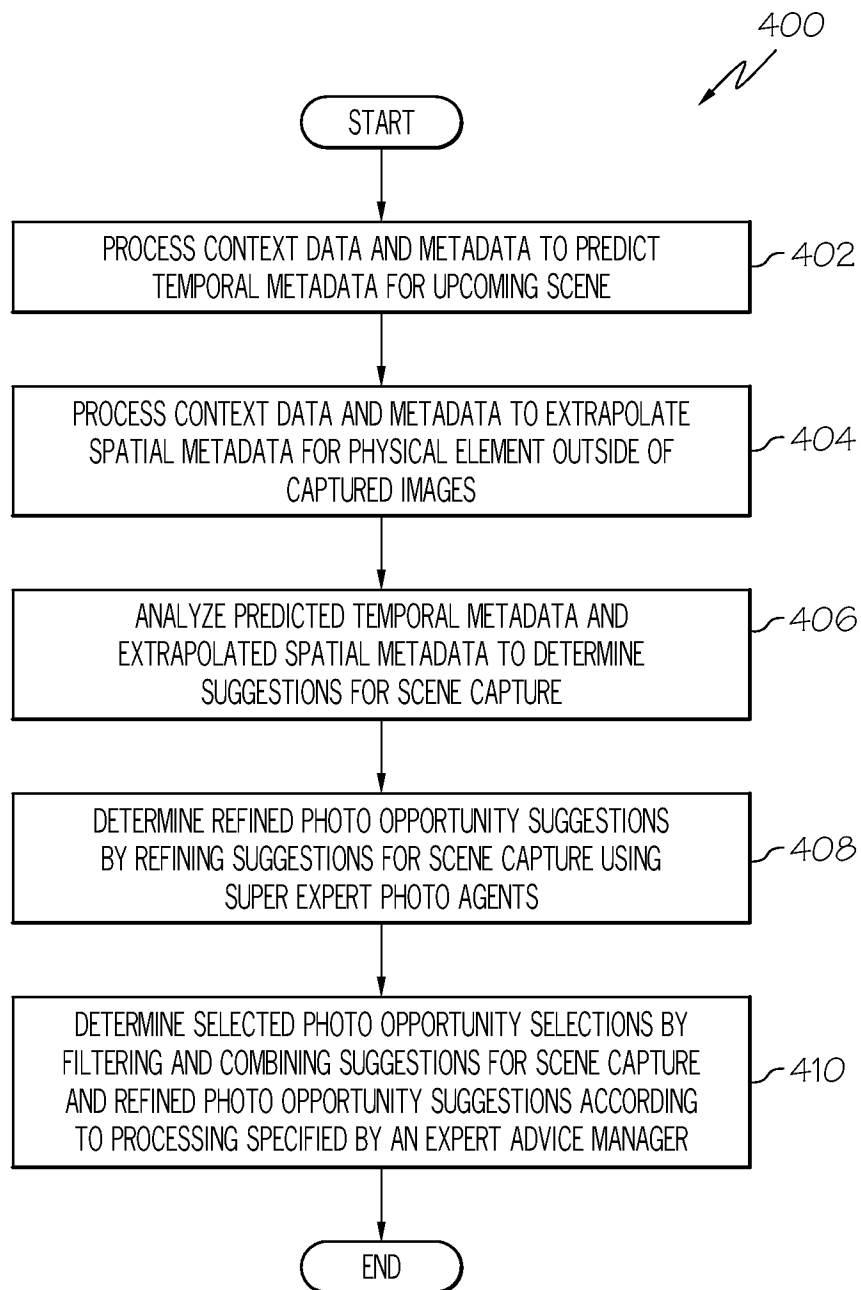


FIG. 4

5 / 6

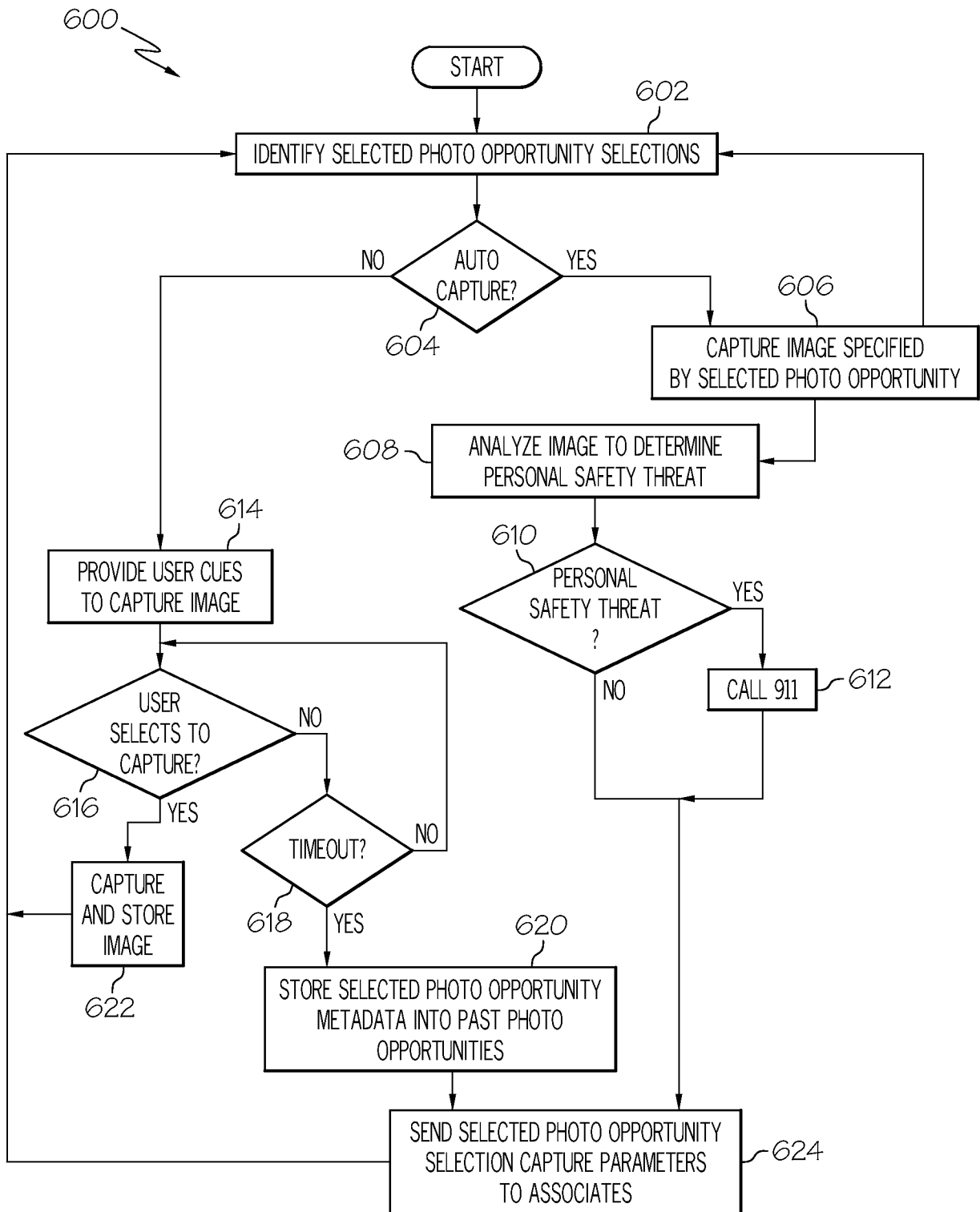
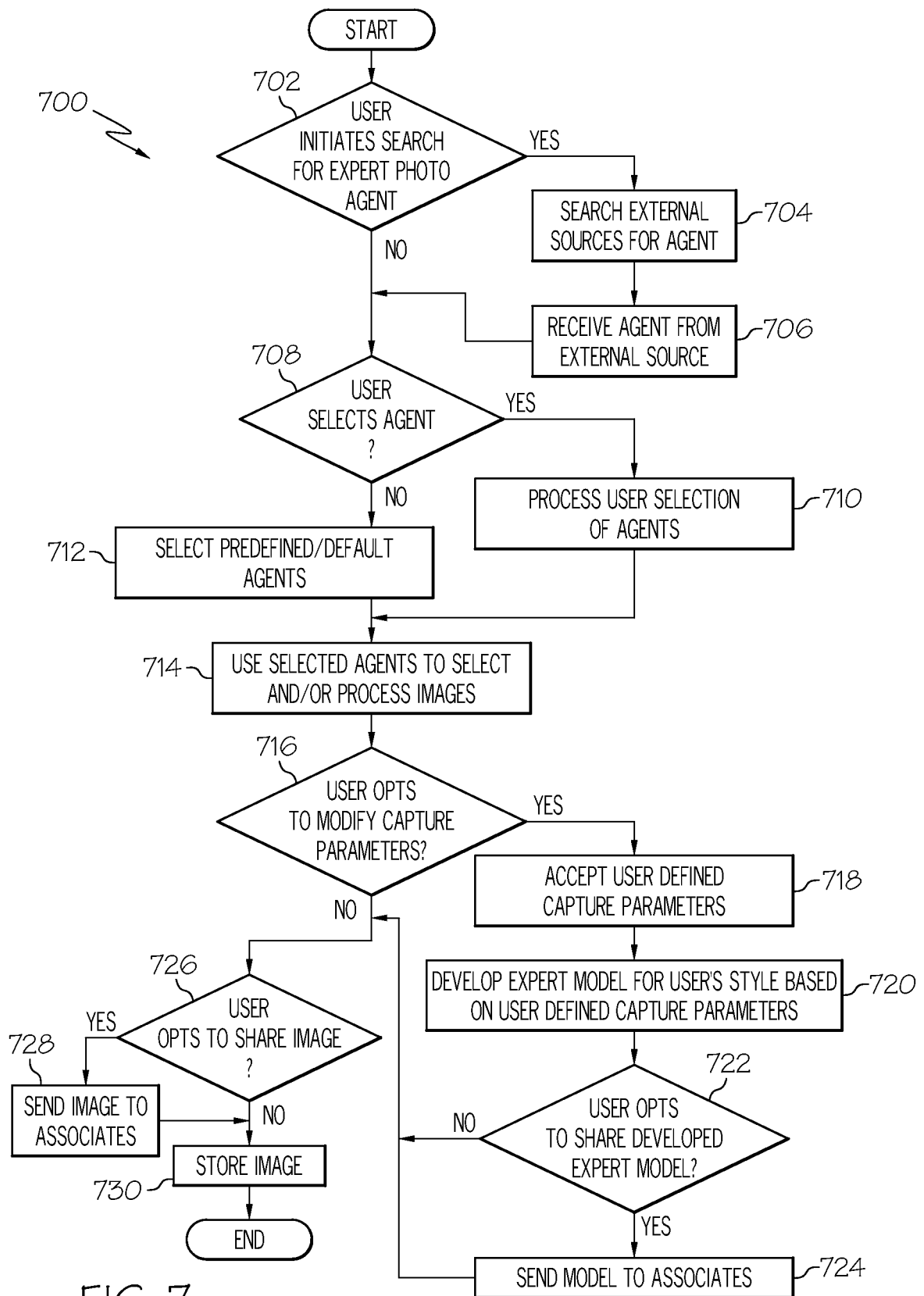


FIG. 6

6 / 6



INTERNATIONAL SEARCH REPORT

International application No

PCT/US2010/028206

A. CLASSIFICATION OF SUBJECT MATTER

INV. H04N1/00

H04N1/21

H04N1/32

H04N5/232

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GB 2 428 927 A (HEWLETT PACKARD DEVELOPMENT CO [US]) 7 February 2007 (2007-02-07) the whole document	1,2,4-8, 13, 15-18,20
X	US 2005/172147 A1 (EDWARDS ERIC [US] ET AL) 4 August 2005 (2005-08-04) the whole document	1,2,5,7, 13,16-18
X	US 2005/168588 A1 (FISHER CLAY [US] ET AL) 4 August 2005 (2005-08-04)	1,2,13, 18
Y	the whole document	2,9-12, 19
X	US 2005/007468 A1 (STAVELY DONALD J [US] ET AL) 13 January 2005 (2005-01-13) the whole document	1,2,4, 13,15,20
Y		3,14
	----- -/--	

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

10 August 2010

Date of mailing of the international search report

18/08/2010

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Moorhouse, David

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2010/028206

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2009/006949 A1 (SONY ERICSSON MOBILE COMM AB [SE]; GARCIA ENRIQUE ALONSO [SE]) 15 January 2009 (2009-01-15) the whole document	1,2,4, 13,15, 18,20
Y		2,3, 9-12,14, 19
X	US 6 608 650 B1 (TORRES DAN [US] ET AL) 19 August 2003 (2003-08-19) the whole document	1,2,4, 13,15,20
Y		2,9-12, 19
X	US 2008/056706 A1 (BATTLES AMY E [US] ET AL) 6 March 2008 (2008-03-06) the whole document	1,2,4, 13,15, 18,20
Y		2,3,14
X	US 2004/201699 A1 (PARULSKI KENNETH A [US] ET AL) 14 October 2004 (2004-10-14) the whole document	1-4, 13-15, 18,20
Y		2,3,14
Y	US 2004/252217 A1 (BATTLES AMY E [US] ET AL) 16 December 2004 (2004-12-16) the whole document	2

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2010/028206

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
GB 2428927	A	07-02-2007	US 2007030363 A1	08-02-2007
US 2005172147	A1	04-08-2005	CN 1914853 A	14-02-2007
			EP 1712033 A1	18-10-2006
			JP 2007520830 T	26-07-2007
			KR 20060130642 A	19-12-2006
			WO 2005076524 A1	18-08-2005
US 2005168588	A1	04-08-2005	EP 1728383 A2	06-12-2006
			JP 2007527663 T	27-09-2007
			KR 20060132679 A	21-12-2006
			WO 2005076896 A2	25-08-2005
US 2005007468	A1	13-01-2005	US 2008030599 A1	07-02-2008
WO 2009006949	A1	15-01-2009	CN 101690164 A	31-03-2010
			EP 2177021 A1	21-04-2010
			US 2009015702 A1	15-01-2009
US 6608650	B1	19-08-2003	NONE	
US 2008056706	A1	06-03-2008	NONE	
US 2004201699	A1	14-10-2004	EP 1280340 A2	29-01-2003
			JP 2003078813 A	14-03-2003
			US 2009284637 A1	19-11-2009
US 2004252217	A1	16-12-2004	DE 102004007649 A1	13-01-2005
			JP 2005006330 A	06-01-2005
			US 2005212955 A1	29-09-2005
			US 2010123805 A1	20-05-2010