



- (51) International Patent Classification:  
G06T 7/00 (2006.01) H04N 5/232 (2006.01)
- (21) International Application Number:  
PCT/US2013/075717
- (22) International Filing Date:  
17 December 2013 (17.12.2013)
- (25) Filing Language: English
- (26) Publication Language: English
- (71) Applicant: INTEL CORPORATION [US/US]; 2200 Mission College Boulevard, Santa Clara, California 95054 (US).
- (72) Inventor; and
- (71) Applicant : ANDERSON, Glen J. [US/US]; 16140 NW Somerset Drive, Beaverton, Oregon 97006 (US).
- (74) Agents: WATSON, Mark L. et al.; Blakely, Sokoloff, Taylor & Zafinan LLP, 1279 Oakmead Parkway, Sunnyvale, California 94085 (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, BN, 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, IR, IS, JP, KE, KG, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: CAMERA ARRAY ANALYSIS MECHANISM

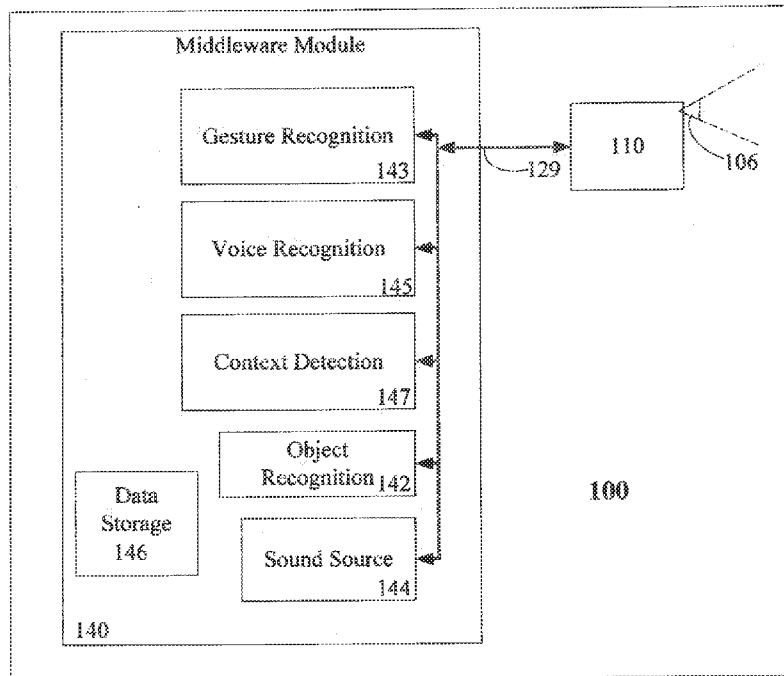


Figure 1

[Continued on next page]

WO 2015/094182 A1

**Declarations under Rule 4.17:**

— *of inventorship (Rule 4.17(iv))*

**Published:**

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

---

**(57) Abstract:** A method is described. The method detecting that a media capture device is prepared to capture media data of a scene, capturing data associated with the scene, analyzing the scene data to identify and classify behavior of one or more objects in the scene and adjusting the media capture device based on the scene data analysis to optimize the capture of the media data.

## CAMERA ARRAY ANALYSIS MECHANISM

### FIELD

5           Embodiments described herein generally relate to perceptual computing, and more particularly, to subject monitoring using camera arrays.

### BACKGROUND

          Camera arrays feature multiple lenses that permit the simultaneous capture of multiple versions of a single scene. For instance, each lens in an array may be set to a different focal  
10   distance, thus enabling the capture of image data with depth information for different objects in the scene. However, when some objects in the scene are behaving differently (e.g., some objects are moving while others are stationary), applying a similar timing and analysis to the entire array may not be optimal.

### BRIEF DESCRIPTION OF THE DRAWINGS

15           Embodiments are illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings in which like reference numerals refer to similar elements.

          Figure 1 is a block diagram illustrating one embodiment of a system for capturing media data.

          Figure 2 is a flow diagram illustrating one embodiment of a camera array analysis  
20   process.

          Figure 3 is an illustrative diagram of an exemplary system.

          Figure 4 is an illustrative diagram of an exemplary system.

### DETAILED DESCRIPTION

          In the following description, numerous specific details are set forth. However,  
25   embodiments, as described herein, may be practiced without these specific details. In other instances, well-known circuits, structures and techniques have not been shown in details in order not to obscure the understanding of this description.

          Throughout this document, terms like "logic", "component", "module", "framework", "engine", "store", or the like, may be referenced interchangeably and include, by way of  
30   example, software, hardware, and/or any combination of software and hardware, such as firmware.

          While the concepts of the present disclosure are susceptible to various modifications and alternative forms, specific embodiments thereof have been shown by way of example in the

drawings and will be described herein in detail. It should be understood, however, that there is no intent to limit the concepts of the present disclosure to the particular forms disclosed, but on the contrary, the intention is to cover all modifications, equivalents, and alternatives consistent with the present disclosure and the appended claims.

5           References in the specification to “one embodiment,” “an embodiment,” “an illustrative embodiment,” etc., indicate that the embodiment described may include a particular feature, structure, or characteristic, but every embodiment may or may not necessarily include that particular feature, structure, or characteristic. Moreover, such phrases are not necessarily referring to the same embodiment. Further, when a particular feature, structure, or characteristic  
10 is described in connection with an embodiment, it is submitted that it is within the knowledge of one skilled in the art to effect such feature, structure, or characteristic in connection with other embodiments whether or not explicitly described.

          The disclosed embodiments may be implemented, in some cases, in hardware, firmware, software, or any combination thereof. The disclosed embodiments may also be implemented as  
15 instructions carried by or stored on a transitory or non-transitory machine-readable (e.g., computer-readable) storage medium, which may be read and executed by one or more processors. A machine-readable storage medium may be embodied as any storage device, mechanism, or other physical structure for storing or transmitting information in a form readable by a machine (e.g., a volatile or non-volatile memory, a media disc, or other media device).

20           In the drawings, some structural or method features may be shown in specific arrangements and/or orderings. However, it should be appreciated that such specific arrangements and/or orderings may not be required. Rather, in some embodiments, such features may be arranged in a different manner and/or order than shown in the illustrative figures. Additionally, the inclusion of a structural or method feature in a particular figure is not meant to  
25 imply that such feature is required in all embodiments and, in some embodiments, may not be included or may be combined with other features.

          Some portions of the detailed descriptions provide herein are presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. Unless specifically stated otherwise, as apparent from the following discussion, it is appreciated  
30 that throughout the description, discussions utilizing terms such as “calculating,” “computing,” “determining” “estimating” “storing” “collecting” “displaying,” “receiving,” “consolidating,” “generating,” “updating,” or the like, refer to the action and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as

physical (electronic) quantities within the computer system's circuitry including registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

While the following description sets forth various implementations that may be  
5 manifested in architectures such system-on-a-chip (SoC) architectures for example,  
implementation of the techniques and/or arrangements described herein are not restricted to  
particular architectures and/or computing systems and may be implemented by any architecture  
and/or computing system for similar purposes. For example, various architectures employing, for  
example, multiple integrated circuit (IC) chips and/or packages, and/or various computing  
10 devices and/or consumer electronic (CE) devices such as set top boxes, smart phones, etc., may  
implement the techniques and/or arrangements described herein. Further, while the following  
description may set forth numerous specific details such as logic implementations, types and  
interrelationships of system components, logic partitioning/integration choices, etc., claimed  
subject matter may be practiced without such specific details. Furthermore, some material such  
15 as, for example, control structures and full software instruction sequences, may not be shown in  
detail in order not to obscure the material disclosed herein.

Systems, apparatus, articles, and methods are described below including operations for  
managing and accessing personal media data based on a perception of an operator of the device  
that captured the media data, as determined during a capture event.

20 Figure 1 is a block diagram illustrating one embodiment of a system 100 for capturing  
media data. System 100 includes media capture device 110 having one or more media capture  
sensors 106. In various embodiments, media capture device 110 includes, or is a component of,  
a mobile computing platform, such as a wireless smartphone, tablet computer, ultrabook  
computer, or the like. In other embodiments, media capture device 110 may be any wearable  
25 device, such as a headset, wrist computer, etc., or may be any immobile sensor installation, such  
as a security camera, etc. In still other embodiments, media capture device 110 may be an  
infrastructure device, such as a television, set-top box, desktop computer, etc.

Generally, media capture sensor 106 is any conventional sensor or sensor array capable  
of collecting media data. In certain embodiments, for example where the capture device 110 is a  
30 mobile computing platform, the media capture sensor 106 has a field of view (FOV) that is  
oriented to capture media data pertaining to a subject. In certain other embodiments, for  
example where the capture device 110 is immobile, the media capture sensor 106 has a field of  
view (FOV) that is oriented to also capture media data including an operator.

In one embodiment, media capture sensor 106 includes a plurality of integrated and/or discrete sensors that form part of a distributed sensor array. In such an embodiment, media capture sensor 106 includes a camera array. In a further embodiment, the sensor data collected includes one or more fields of perception sensor data 129 transmitted via conventional  
5 wireless/wired network/internet connectivity.

System 100 further includes middleware module 140 to receive perception sensor data 129 having potentially many native forms (e.g., analog, digital, continuously streamed, aperiodic, etc.). In one embodiment, middleware module 140 functions as a hub for multiplexing and/or de-multiplexing a plurality of streams of sensor data 129. Generally,  
10 middleware module 140 may be a component of the media capture device 110, or may be part of a separate platform.

Middleware module 140 may employ one or more sensor data processing modules, such as object recognition module 142, sound source module 144, gesture recognition module 143, voice recognition module 145, context determination module 147, etc., each employing an  
15 algorithm to analyze the received data to produce perception data. According to one embodiment, the perception data is a low level parameterization, such as a voice command, gesture, or facial expression (e.g., smile, frown, grimace, etc.), or a higher level abstraction, such as a level of attention, or a cognitive load (e.g., a measure of mental effort) that may be inferred or derived indirectly from one or more fields in sensor data 129. For example, a level of  
20 attention may be estimated based on eye tracking and/or on a rate of blinking, etc., while a cognitive load may be inferred based on pupil dilation and/or a heart rate-blood pressure product. Embodiments of the invention are not limited with respect to specific transformations of sensor data 129 into perception data, which is stored in data storage.

In a further embodiment, each item of perception data may correspond to sensor data 129  
25 collected in response to, or triggered by, a media capture event. In such an embodiment, perception data stored in data storage 146 may be organized as files with contemporaneous perception-media data associations then generated by linking a perception data file with a media data file having approximately the same file creation time.

According to one embodiment, system 100 performs pre-image analysis to enhance  
30 media capture performance. In such an embodiment, scene data from media capture sensor 106 is implemented to identify objects and classify behavior prior to capture of the media. In response, commands are transmitted back to media capture device 110 to optimize settings for sensor 106. In one embodiment, the pre-image analysis process treats imagery from at least one

lens in sensor 106 different from others based on information received regarding behavior and/or identification of objects in a scene. In such an embodiment, the information is obtained by devoting some lenses to obtaining data for scene analysis as opposed to simply varying the lenses across typical dimensions such as focal length. In a further embodiment, microphones, or other sensors, are used to provide information for scene analysis relative to the array.

**Figure 2** is a flow diagram illustrating one embodiment of a pre-image analysis process performed at system 100. At processing block 200, media capture device 110 detects that a user is framing a scene for media capture. At processing block 220, media capture sensor 106 captures scene analysis data. For instance, media capture sensor 106 may capture light, sound, vibration and movement data from a scene to be captured. At processing block 230, the captured data is received and analyzed at middleware module 140. According to one embodiment, object recognition module 142 analyzes the data to identify objects and or/ faces in the scene, while sound source module 144 determines a source of sound input from a scene based on data received from microphones and video analysis of mouth movement.

Similarly, gesture recognition module 143, voice recognition module 145, context determination module 147 is implemented to provide additional contextual analysis (e.g., classify the behavior/activity of objects, determine levels of motion in individual objects and determine levels of overall emotion of subjects). In one embodiment, lenses assigned to subject behavior tracking may perform the task before, during or after image capture by the array.

At processing block 240, media capture sensor 106 is optimized to improve media capture. According to one embodiment, lenses of a camera array component of media capture sensor 106 are optimized to improve image capture. In a further embodiment, associated data may be stored in data storage 146 as image metadata for subsequent post-processing analysis, and enable new features in post-processing (e.g., selecting imagery from corresponding lenses to automatically creating the best image, or for presenting to a user for choices in browsing or user interface editing). At processing block 250, the media data of the scene is captured.

Based on the above-described pre-image analysis process, scenarios in which sensor input (or context determinations made via middleware module algorithms) causes the settings of media capture device 110 to be optimized are determined by the corresponding algorithms. Table 1 illustrates one embodiment of array adjustment optimization rules performed based on sensor detection and corresponding algorithm determination.

Table 1

Sensor/Algorithm	Algorithm Determination	Array Adjustment Rule
------------------	-------------------------	-----------------------

<b>Detection</b>		
Object distance change	Subject is changing distance relative to camera	Adjust focal depths to optimize for subjects changing distance
Speech input	Subject is talking	Use a lens to video record subject at optimal focal depth
Object motion	Subject is moving	Adjust array settings for a range of lower exposures
Object or face recognition	Historical correlations of subject behavior such as “tends to move”	Adjust array settings for a range of lower exposures
Face detection	Tendency of a larger number of faces to require more shots for good group photo	Set array to take more quick-sequence images for later adjustments
Ambient sound level	High ambient sound correlates with more action	Set array to take more quick-sequence images and for low exposure alternatives for later adjustments
Light level	Indoor detection	Set array for higher range of exposure levels
Ambient sound	Detection of sporting event	Set array for “sport setting” range of exposure

As discussed above, pre-image analysis process enables various enhancements to be realized. For instance, if an object and its behavior are classified as being in motion and having high potential for motion (e.g., a dog being identified and classified as moving), an exposure setting for a camera array may be shorter. In another example, focal length depth settings across the camera array may be adjusted to increase odds of a good image in a proper direction (e.g., closer focal depths if a skate-boarder is moving closer to the camera) if a detected object is moving closer or farther away. Moreover, one of the lenses may be focused on a speaker and automatically set to video, if speech is detected, so that video of the speaker can be captured for

later use. In yet another example, a differential focus level of the camera array may be used to automatically capture additional photos of a subject known to be difficult to capture (e.g., a child).

Figure 3 is an illustrative diagram of an exemplary system 300, in accordance with  
5 embodiments. System 300 may implement all or a subset of the various functional blocks depicted in Figure 1. For example, in one embodiment the system 100 is implemented by system 300. System 300 may be a mobile device although system 300 is not limited to this context. For example, system 300 may be incorporated into a laptop computer, ultra-laptop computer, tablet, touch pad, portable computer, handheld computer, palmtop computer, cellular telephone, smart  
10 device (e.g., smart phone, smart tablet or mobile television), mobile internet device (MID), wearable computing device, messaging device, data communication device, and so forth. System 300 may also be an infrastructure device. For example, system 300 may be incorporated into a large format television, set-top box, desktop computer, or other home or commercial network device.

15 In various implementations, system 300 includes a platform 302 coupled to a HID 320. Platform 302 may receive captured personal media data from a personal media data services device(s) 330, a personal media data delivery device(s) 340, or other similar content source. A navigation controller 350 including one or more navigation features may be used to interact with, for example, platform 302 and/or HID 320. Each of these components is described in greater  
20 detail below.

In various implementations, platform 302 may include any combination of a chipset 305, processor 310, memory 312, storage 314, graphics subsystem 315, applications 316 and/or radio 318. Chipset 305 may provide intercommunication among processor 310, memory 312, storage 314, graphics subsystem 315, applications 316 and/or radio 318. For example, chipset 305 may  
25 include a storage adapter (not depicted) capable of providing intercommunication with storage 314.

Processor 310 may be implemented as a Complex Instruction Set Computer (CISC) or Reduced Instruction Set Computer (RISC) processors; x86 instruction set compatible processors, multi-core, or any other microprocessor or central processing unit (CPU). In various  
30 implementations, processor 310 may be a multi-core processor(s), multi-core mobile processor(s), and so forth. In one exemplary embodiment, processor 310 invokes or otherwise implements process 200 and the various modules described as components of middleware 140.

Memory 312 may be implemented as a volatile memory device such as, but not limited to, a Random Access Memory (RAM), Dynamic Random Access Memory (DRAM), or Static RAM (SRAM).

Storage 314 may be implemented as a non-volatile storage device such as, but not limited to, a magnetic disk drive, optical disk drive, tape drive, an internal storage device, an attached storage device, flash memory, battery backed-up SDRAM (synchronous DRAM), and/or a network accessible storage device. In various implementations, storage 314 may include technology to increase the storage performance enhanced protection for valuable digital media when multiple hard drives are included, for example.

Graphics subsystem 315 may perform processing of images such as still or video media data for display. Graphics subsystem 315 may be a graphics processing unit (GPU) or a visual processing unit (VPU), for example. An analog or digital interface may be used to communicatively couple graphics subsystem 315 and display 320. For example, the interface may be any of a High-Definition Multimedia Interface, Display Port, wireless HDMI, and/or wireless HD compliant techniques. Graphics subsystem 315 may be integrated into processor 310 or chipset 705. In some implementations, graphics subsystem 315 may be a stand-alone card communicatively coupled to chipset 305.

The perception-media data associations and related media data management and accessing techniques described herein may be implemented in various hardware architectures.

For example, graphics and/or video functionality may be integrated within a chipset. Alternatively, a discrete graphics and/or video processor may be used. As still another embodiment, the methods and functions described herein may be provided by a general purpose processor, including a multi-core processor. In further embodiments, the methods and functions may be implemented in a purpose-built consumer electronics device.

Radio 318 may include one or more radios capable of transmitting and receiving signals using various suitable wireless communications techniques. Such techniques may involve communications across one or more wireless networks. Example wireless networks include (but are not limited to) wireless local area networks (WLANs), wireless personal area networks (WPANs), wireless metropolitan area network (WMANs), cellular networks, and satellite networks. In communicating across such networks, radio 318 may operate in accordance with one or more applicable standards in any version.

In various implementations, HID 320 may include any television type monitor or display. HID 320 may include, for example, a computer display screen, touch screen display, video

monitor, television-like device, and/or a television. HID 320 may be digital and/or analog. In various implementations, HID 320 may be a holographic display. Also, HID 320 may be a transparent surface that may receive a visual projection. Such projections may convey various forms of information, images, and/or objects. For example, such projections may be a visual  
5 overlay for a mobile augmented reality (MAR) application. Under the control of one or more software applications 316, platform 302 may display user interface 322 on HID 320.

In various implementations, personal media services device(s) 330 may be hosted by any national, international and/or independent service and thus accessible to platform 302 via the Internet, for example. Personal media services device(s) 330 may be coupled to platform 702  
10 and/or to display 320. Platform 302 and/or personal services device(s) 330 may be coupled to a network 760 to communicate (e.g., send and/or receive) media information to and from network 360. Personal media delivery device(s) 340 also may be coupled to platform 302 and/or to HID 320.

In various implementations, personal media data services device(s) 330 may include a  
15 cable television box, personal computer, network, telephone, Internet enabled devices or appliance capable of delivering digital information and/or content, and any other similar device capable of unidirectionally or bidirectionally communicating content between a media data provider and platform 302, via network 360 or directly. It will be appreciated that the content may be communicated unidirectionally and/or bidirectionally to and from any one of the  
20 components in system 700 and a provider via network 360. Examples of personal media include any captured media information including, for example, video, music, medical and gaming information, and so forth.

Personal media data services device(s) 330 may receive content including media information with examples of content providers including any cable or satellite television or  
25 radio or Internet content providers. The provided examples are not meant to limit implementations in accordance with the present disclosure in any way.

In various implementations, platform 302 may receive control signals from navigation controller 350 having one or more navigation features. The navigation features of controller 350 may be used to interact with user interface 322, for example. In embodiments, navigation  
30 controller 350 may be a pointing device that may be a computer hardware component (specifically, a human interface device) that allows a user to input spatial (e.g., continuous and multi-dimensional) data into a computer. Many systems such as graphical user interfaces (GUI),

and televisions and monitors allow the user to control and provide data to the computer or television using physical gestures.

Movements of the navigation features of controller 350 may be replicated on a display (e.g., HID 320) by movements of a pointer, cursor, focus ring, or other visual indicators  
5 displayed on the display. For example, under the control of software applications 316, the navigation features located on navigation controller 350 may be mapped to virtual navigation features displayed on user interface 322, for example. In embodiments, controller 350 may not be a separate component but may be integrated into platform 302 and/or HID 320. The present disclosure, however, is not limited to the elements or in the context shown or described herein.

10 In various implementations, drivers (not shown) may include technology to enable users to instantly turn on and off platform 302 like a television with the touch of a button after initial boot-up, when enabled, for example. Program logic may allow platform 302 to stream content to media adaptors or other personal media services device(s) 330 or personal media delivery device(s) 340 even when the platform is turned “off.” In addition, chipset 305 may include  
15 hardware and/or software support for 8.1 surround sound audio and/or high definition (7.1) surround sound audio, for example. Drivers may include a graphics driver for integrated graphics platforms. In embodiments, the graphics driver may comprise a peripheral component interconnect (PCI) Express graphics card.

In various implementations, any one or more of the components shown in system 300  
20 may be integrated. For example, platform 302 and personal media data services device(s) 330 may be integrated, or platform 302 and captured media data delivery device(s) 340 may be integrated, or platform 302, personal media services device(s) 330, and personal media delivery device(s) 340 may be integrated, for example. In various embodiments, platform 302 and HID 320 may be an integrated unit. HID 320 and content service device(s) 330 may be integrated, or  
25 HID 320 and personal media delivery device(s) 340 may be integrated, for example. These examples are not meant to limit the present disclosure.

In various embodiments, system 300 may be implemented as a wireless system, a wired system, or a combination of both. When implemented as a wireless system, system 300 may include components and interfaces suitable for communicating over a wireless shared media,  
30 such as one or more antennas, transmitters, receivers, transceivers, amplifiers, filters, control logic, and so forth. An example of wireless shared media may include portions of a wireless spectrum, such as the RF spectrum and so forth. When implemented as a wired system, system 300 may include components and interfaces suitable for communicating over wired

communications media, such as input/output (I/O) adapters, physical connectors to connect the I/O adapter with a corresponding wired communications medium, a network interface card (NIC), disc controller, video controller, audio controller, and the like. Examples of wired communications media may include a wire, cable, metal leads, printed circuit board (PCB),  
5 backplane, switch fabric, semiconductor material, twisted-pair wire, co-axial cable, fiber optics, and so forth.

Platform 302 may establish one or more logical or physical channels to communicate information. The information may include media information and control information. Media information may refer to any data representing content meant for a user. Examples of content  
10 may include, for example, data from a voice conversation, videoconference, streaming video, electronic mail (“email”) message, voice mail message, alphanumeric symbols, graphics, image, video, text and so forth. Data from a voice conversation may be, for example, speech information, silence periods, background noise, comfort noise, tones and so forth. Control information may refer to any data representing commands, instructions or control words meant  
15 for an automated system. For example, control information may be used to route media information through a system, or instruct a node to process the media information in a predetermined manner. The embodiments, however, are not limited to the elements or in the context shown or described in Figure 3.

As described above, system 300 may be embodied in varying physical styles or form  
20 factors. Figure 4 illustrates embodiments of a small form factor device 400 in which system 300 may be embodied. In embodiments, for example, device 400 may be implemented as a mobile computing device having wireless capabilities. A mobile computing device may refer to any device having a processing system and a mobile power source or supply, such as one or more batteries, for example.

As described above, examples of a mobile computing device may include a personal  
25 computer (PC), laptop computer, ultra-laptop computer, tablet, touch pad, portable computer, handheld computer, palmtop computer, personal digital assistant (PDA), cellular telephone, combination cellular telephone/PDA, television, smart device (e.g., smart phone, smart tablet or smart television), mobile internet device (MID), messaging device, data communication device,  
30 and so forth.

Examples of a mobile computing device also may include computers configured to be worn by a person, such as a wrist computer, finger computer, ring computer, eyeglass computer, belt-clip computer, arm-band computer, shoe computers, clothing computers, and other wearable

computers. In various embodiments, for example, a mobile computing device may be implemented as a smart phone capable of executing computer applications, as well as voice communications and/or data communications. Although some embodiments may be described with a mobile computing device implemented as a smart phone by way of example, it may be appreciated that other embodiments may be implemented using other wireless mobile computing devices as well. The embodiments are not limited in this context.

As shown in Figure 4, device 400 may include a housing 402, a display 404, an input/output (I/O) device 406, and an antenna 408. Device 400 also may include navigation features 412. Display 404 may include any suitable display unit for displaying information appropriate for a mobile computing device. I/O device 406 may include any suitable I/O device for entering information into a mobile computing device. Examples for I/O device 406 may include an alphanumeric keyboard, a numeric keypad, a touch pad, input keys, buttons, switches, rocker switches, microphones, speakers, voice recognition device and software, and so forth. Information also may be entered into device 400 by way of microphone (not shown). Such information may be digitized by a voice recognition device (not shown). The embodiments are not limited in this context.

Embodiments may be implemented as any or a combination of: one or more microchips or integrated circuits interconnected using a parentboard, hardwired logic, software stored by a memory device and executed by a microprocessor, firmware, an application specific integrated circuit (ASIC), and/or a field programmable gate array (FPGA). The term "logic" may include, by way of example, software or hardware and/or combinations of software and hardware.

Embodiments may be provided, for example, as a computer program product which may include one or more machine-readable media having stored thereon machine-executable instructions that, when executed by one or more machines such as a computer, network of computers, or other electronic devices, may result in the one or more machines carrying out operations in accordance with embodiments described herein. A machine-readable medium may include, but is not limited to, floppy diskettes, optical disks, CD-ROMs (Compact Disc-Read Only Memories), and magneto-optical disks, ROMs, RAMs, EPROMs (Erasable Programmable Read Only Memories), EEPROMs (Electrically Erasable Programmable Read Only Memories), magnetic or optical cards, flash memory, or other type of media/machine-readable medium suitable for storing machine-executable instructions.

Moreover, embodiments may be downloaded as a computer program product, wherein the program may be transferred from a remote computer (e.g., a server) to a requesting computer

(e.g., a client) by way of one or more data signals embodied in and/or modulated by a carrier wave or other propagation medium via a communication link (e.g., a modem and/or network connection).

As used in the claims, unless otherwise specified the use of the ordinal adjectives “first”, “second”, “third”, etc., to describe a common element, merely indicate that different instances of like elements are being referred to, and are not intended to imply that the elements so described must be in a given sequence, either temporally, spatially, in ranking, or in any other manner.

The following clauses and/or examples pertain to further embodiments or examples. Specifics in the examples may be used anywhere in one or more embodiments. The various features of the different embodiments or examples may be variously combined with some features included and others excluded to suit a variety of different applications. Examples may include subject matter such as a method, means for performing acts of the method, at least one machine-readable medium including instructions that, when performed by a machine cause the machine to performs acts of the method, or of an apparatus or system for facilitating content-morphism and distribution of advertisement content and user content according to embodiments and examples described herein.

Some embodiments pertain to Example 1 that includes a method comprising detecting that a media capture device is prepared to capture media data of a scene, capturing data associated with the scene, analyzing the scene data to identify and classify behavior of one or more objects in the scene and adjusting the media capture device based on the scene data analysis to optimize the capture of the media data.

Example 2 includes the subject matter of Example 1 and further comprising capturing media data of the scene after adjusting the media capture device.

Example 3 includes the subject matter of Example 1 and wherein capturing data associated with the scene comprises capturing light, sound, vibration and movement data from the scene.

Example 4 includes the subject matter of Example 1 and wherein capturing data comprises a camera array having two or more lenses to capture the data.

Example 5 includes the subject matter of Example 4 and wherein a first of the two or more lenses to capture the scene data.

Example 6 includes the subject matter of Example 1 and wherein analyzing the scene data comprises analyzing the scene data to identify one or more objects in the scene.

Example 7 includes the subject matter of Example 1 and wherein analyzing the scene data comprises analyzing the scene data to recognize one or more faces in the scene.

Example 8 includes the subject matter of Example 5 and wherein analyzing the scene data comprises determining a source of sound input from a scene.

5 Example 9 includes the subject matter of Example 8 and wherein the source of sound input from the scene is determined based on sound data received from microphones and video analysis of mouth movement from data captured by the one or more lenses.

Example 10 includes the subject matter of Example 4 and wherein adjusting the media capture device comprises optimizing the two or more lenses to improve image capture.

10 Some embodiments pertain to Example 1 that includes a media data management system comprising a media capture device to capture data associated with a scene prior to capturing media data for the scene and a middleware module to analyze the scene data to identify and classify behavior of one or more objects in the scene and adjust the media capture device based on the scene data analysis to optimize the capture of the media data.

15 Example 12 includes the subject matter of Example 11 and wherein the media capture device captures media data of the scene after the adjustment.

Example 13 includes the subject matter of Example 11 and wherein the media capture device captures light, sound, vibration and movement data from the scene.

20 Example 14 includes the subject matter of Example 11 and wherein the media capture device comprises a camera array having two or more lenses to capture the data.

Example 15 includes the subject matter of Example 14 and wherein a first of the two or more lenses to capture the scene data.

25 Example 16 includes the subject matter of Example 15 and wherein the middleware module comprises an object recognition module to identify one or more objects and in one or more faces in the scene.

Example 17 includes the subject matter of Example 11 and wherein the media capture device comprises one or more microphones.

30 Example 18 includes the subject matter of Example 11 and wherein the middleware module further comprises a sound source module to determine a source of sound input from the scene.

Example 19 includes the subject matter of Example 18 and wherein the source of sound input from the scene is determined based on sound data received from microphones and video analysis of mouth movement from data captured by the one or more lenses.

Example 20 includes the subject matter of Example 14 and wherein adjusting the media capture device comprises optimizing the two or more lenses to improve image capture.

Some embodiments pertain to Example 21 that includes a media capture device comprising media capture sensors to data capture associated with a scene prior to capturing  
5 media data for the scene and a middleware module to analyze the scene data to identify and classify behavior of one or more objects in the scene and adjust one or more of the media capture sensors based on the scene data analysis to optimize the capture of the media data.

Example 22 includes the subject matter of Example 21 and wherein the media capture sensors comprise a camera array having two or more lenses to capture the data and one or more  
10 microphones.

Example 23 includes the subject matter of Example 22 and wherein the middleware module comprises an object recognition module to identify one or more objects and in one or more faces in the scene and a sound source module to determine a source of sound input from the scene.

Example 24 includes the subject matter of Example 23 and wherein the source of sound input from the scene is determined based on sound data received from microphones and video analysis of mouth movement from data captured by the one or more lenses.

Some embodiments pertain to Example 25 that includes a machine-readable medium comprising a plurality of instructions that in response to being executed on a computing device,  
20 causes the computing device to carry out operations according to any one of claims 1-10.

Some embodiments pertain to Example 26 that includes a system comprising a mechanism to carry out operations according to any one of claims 1 to 10.

Some embodiments pertain to Example 27 that includes means to carry out operations according to any one of claims 1 to 10.

Some embodiments pertain to Example 28 that includes a computing device arranged to carry out operations according to any one of claims 1 to 10.

Some embodiments pertain to Example 29 that includes a communications device arranged to carry out operations according to any one of claims 1 to 10.

Some embodiments pertain to Example 30 that includes a machine-readable medium comprising a plurality of instructions that in response to being executed on a computing device,  
30 causes the computing device to carry out operations comprising detecting that a media capture device is prepared to capture media data of a scene, capturing data associated with the scene, analyzing the scene data to identify and classify behavior of one or more objects in the scene and

adjusting the media capture device based on the scene data analysis to optimize the capture of the media data.

Example 31 includes the subject matter of Example 30 and further comprising capturing media data of the scene after adjusting the media capture device.

5 Example 32 includes the subject matter of Example 30 and wherein capturing data associated with the scene comprises capturing light, sound, vibration and movement data from the scene.

Example 33 includes the subject matter of Example 30 and wherein capturing data comprises a camera array having two or more lenses to capture the data.

10 Example 34 includes the subject matter of Example 33 and wherein a first of the two or more lenses to capture the scene data.

Example 35 includes the subject matter of Example 33 and wherein analyzing the scene data comprises analyzing the scene data to identify one or more objects in the scene.

15 Example 36 includes the subject matter of Example 33 and wherein analyzing the scene data comprises analyzing the scene data to recognize one or more faces in the scene.

Example 37 includes the subject matter of Example 34 and wherein analyzing the scene data comprises determining a source of sound input from a scene.

20 Example 38 includes the subject matter of Example 37 and wherein the source of sound input from the scene is determined based on sound data received from microphones and video analysis of mouth movement from data captured by the one or more lenses.

Example 39 includes the subject matter of Example 33 and wherein adjusting the media capture device comprises optimizing the two or more lenses to improve image capture.

25 Some embodiments pertain to Example 40 that includes an apparatus comprising means for detecting that a media capture device is prepared to capture media data of a scene, means for capturing data associated with the scene, means for analyzing the scene data to identify and classify behavior of one or more objects in the scene and means for adjusting the media capture device based on the scene data analysis to optimize the capture of the media data.

Example 41 includes the subject matter of Example 40 and further comprising means for capturing media data of the scene after adjusting the media capture device.

30 Example 42 includes the subject matter of Example 40 and wherein capturing data associated with the scene comprises means for capturing light, sound, vibration and movement data from the scene.

Example 43 includes the subject matter of Example 40 and wherein capturing data comprises a camera array having two or more lenses to capture the data.

Example 44 includes the subject matter of Example 43 and wherein a first of the two or more lenses to capture the scene data.

5 Example 45 includes the subject matter of Example 43 and wherein the means for analyzing the scene data comprises means for analyzing the scene data to identify one or more objects in the scene.

10 Example 46 includes the subject matter of Example 43 and wherein the means for analyzing the scene data comprises means for analyzing the scene data to recognize one or more faces in the scene.

Example 47 includes the subject matter of Example 44 and wherein the means for analyzing the scene data comprises means for determining a source of sound input from a scene.

15 Example 48 includes the subject matter of Example 47 and wherein the source of sound input from the scene is determined based on sound data received from microphones and video analysis of mouth movement from data captured by the one or more lenses.

Example 49 includes the subject matter of Example 43 and wherein the means for adjusting the media capture device comprises means for optimizing the two or more lenses to improve image capture.

20 The drawings and the forgoing description give examples of embodiments. Those skilled in the art will appreciate that one or more of the described elements may well be combined into a single functional element. Alternatively, certain elements may be split into multiple functional elements. Elements from one embodiment may be added to another embodiment. For example, orders of processes described herein may be changed and are not limited to the manner described herein. Moreover, the actions any flow diagram need not be implemented in the order shown;  
25 nor do all of the acts necessarily need to be performed. Also, those acts that are not dependent on other acts may be performed in parallel with the other acts. The scope of embodiments is by no means limited by these specific examples. Numerous variations, whether explicitly given in the specification or not, such as differences in structure, dimension, and use of material, are possible. The scope of embodiments is at least as broad as given by the following claims.

30

CLAIMS

What is claimed is:

1. A computer-implemented method for performing pre-image analysis comprising:  
detecting that a media capture device is prepared to capture media data of a scene;  
5 capturing data associated with the scene;  
analyzing the scene data to identify and classify behavior of one or more objects in the scene; and  
adjusting the media capture device based on the scene data analysis to optimize the capture of the media data.
- 10 2. The method of claim 1 further comprising capturing media data of the scene after adjusting the media capture device.
3. The method of claim 1 wherein capturing data associated with the scene comprises capturing light, sound, vibration and movement data from the scene.
4. The method of claim 1 wherein capturing data comprises a camera array having two or  
15 more lenses to capture the data.
5. The method of claim 4 wherein a first of the two or more lenses to capture the scene data.
6. The method of claim 1 wherein analyzing the scene data comprises analyzing the scene data to identify one or more objects in the scene.
7. The method of claim 1 wherein analyzing the scene data comprises analyzing the scene  
20 data to recognize one or more faces in the scene.
8. The method of claim 5 wherein analyzing the scene data comprises determining a source of sound input from a scene.
9. The method of claim 8 wherein the source of sound input from the scene is determined based on sound data received from microphones and video analysis of mouth movement from  
25 data captured by the one or more lenses.
10. The method of claim 4 wherein adjusting the media capture device comprises optimizing the two or more lenses to improve image capture.
11. A media data management system comprising:

a media capture device to capture data associated with a scene prior to capturing media data for the scene; and

a middleware module to analyze the scene data to identify and classify behavior of one or more objects in the scene and adjust the media capture device based on the scene data analysis to  
5 optimize the capture of the media data.

12. The media data management system of claim 11 wherein the media capture device captures media data of the scene after the adjustment.

13. The media data management of claim 11 wherein the media capture device captures light, sound, vibration and movement data from the scene.

10 14. The media data management of claim 11 wherein the media capture device comprises a camera array having two or more lenses to capture the data.

15. The media data management of claim 14 wherein a first of the two or more lenses to capture the scene data.

15 16. The media data management of claim 15 wherein the middleware module comprises an object recognition module to identify one or more objects and in one or more faces in the scene.

17. The media data management of claim 16 wherein the media capture device comprises one or more microphones.

18. The media data management of claim 17 wherein the middleware module further comprises a sound source module to determine a source of sound input from the scene.

20 19. The media data management of claim 18 wherein the source of sound input from the scene is determined based on sound data received from microphones and video analysis of mouth movement from data captured by the one or more lenses.

20. The media data management of claim 14 wherein adjusting the media capture device comprises optimizing the two or more lenses to improve image capture.

25 21. A media capture device comprising:  
media capture sensors to data capture associated with a scene prior to capturing media data for the scene; and

a middleware module to analyze the scene data to identify and classify behavior of one or more objects in the scene and adjust one or more of the media capture sensors based on the scene data analysis to optimize the capture of the media data.

22. The media capture device of claim 11 wherein the media capture sensors comprise:  
5 a camera array having two or more lenses to capture the data; and  
one or more microphones.
23. The media capture device of claim 15 wherein the middleware module comprises:  
an object recognition module to identify one or more objects and in one or more faces in  
the scene; and  
10 a sound source module to determine a source of sound input from the scene.
24. The media data management of claim 23 wherein the source of sound input from the scene is determined based on sound data received from microphones and video analysis of mouth movement from data captured by the one or more lenses.
25. A machine-readable medium comprising a plurality of instructions that in response to  
15 being executed on a computing device, causes the computing device to carry out operations according to any one of claims 1-10.

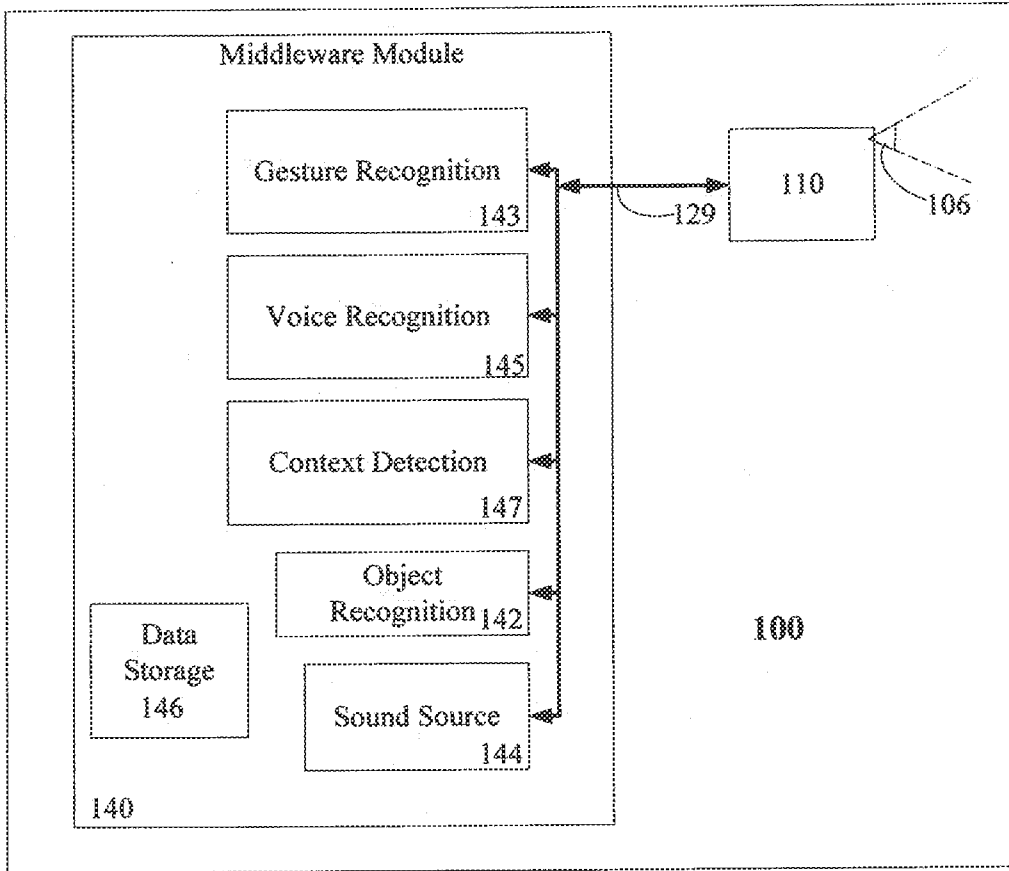


Figure 1

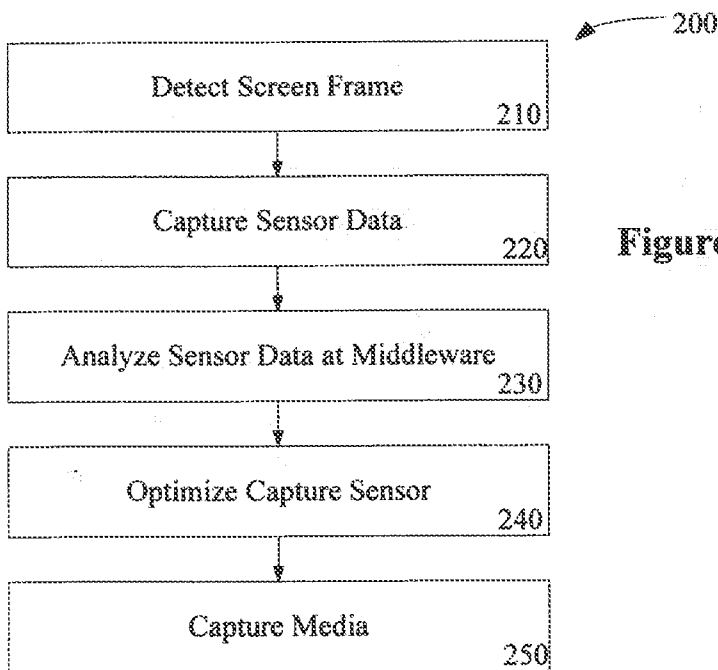


Figure 2

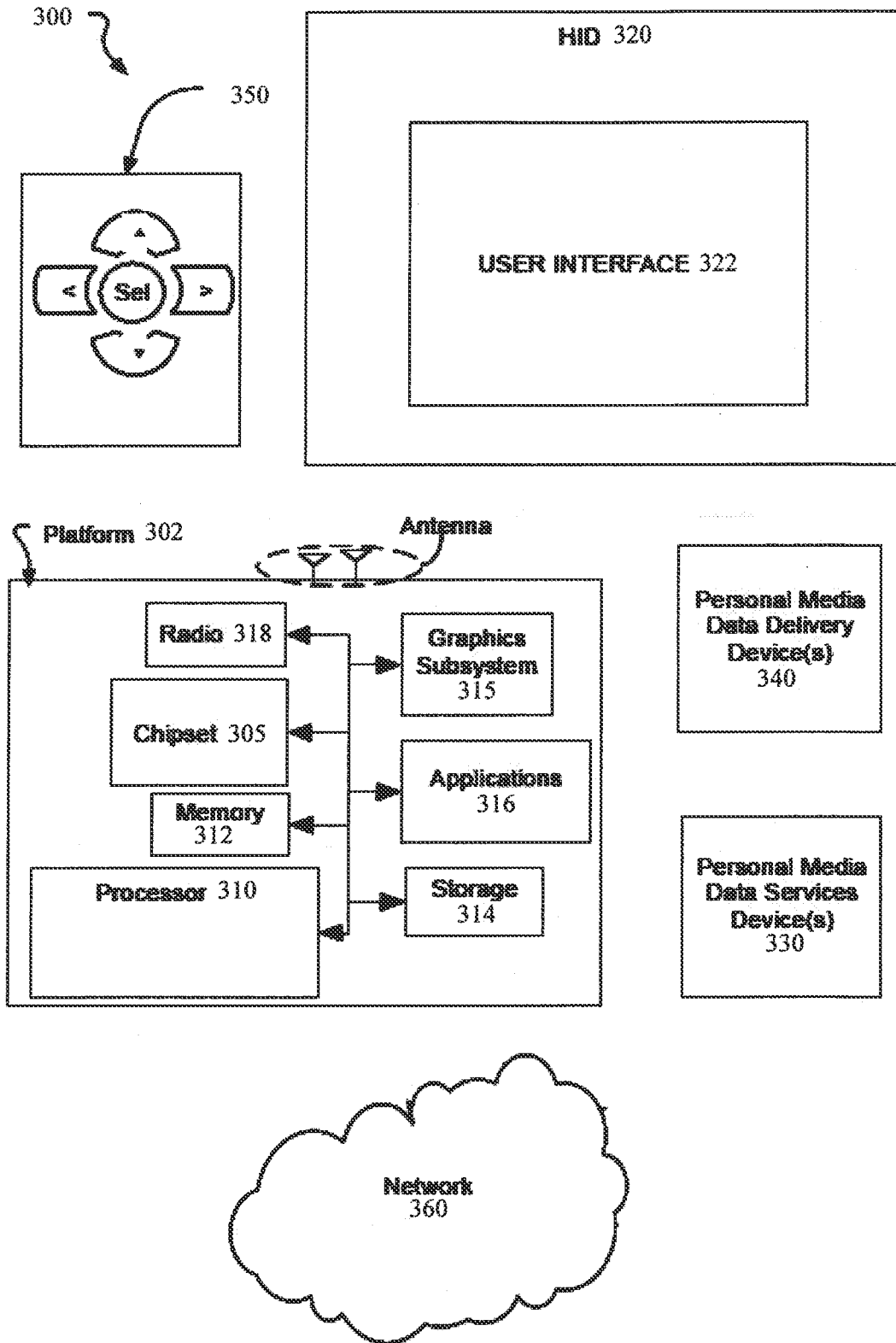


Figure 3

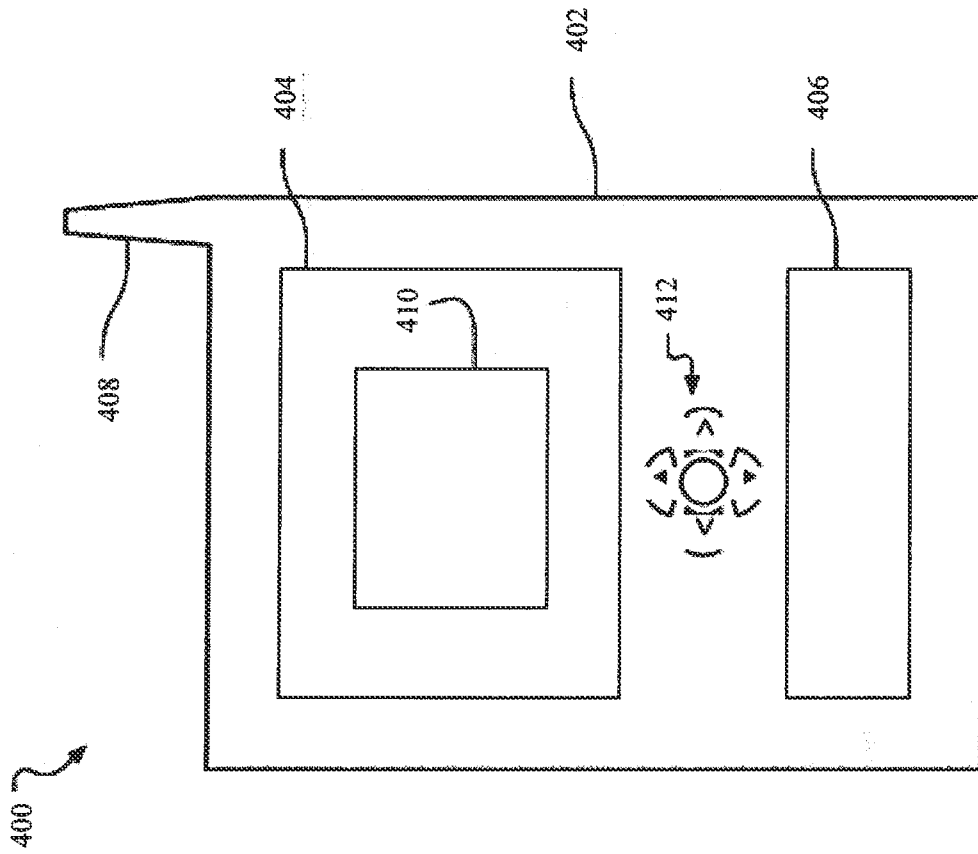


Figure 4

## INTERNATIONAL SEARCH REPORT

International application No.  
**PCT/US2013/075717****A. CLASSIFICATION OF SUBJECT MATTER****G06T 7/00(2006.01)i, H04N 5/232(2006.01)i**

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)

G06T 7/00; H04N 1/46; H04N 5/225; H04N 1/00; G06T 17/00; G06K 9/00; H04N 5/232

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

Korean utility models and applications for utility models

Japanese utility models and applications for utility models

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

eKOMPASS(KIPO internal) &amp; Keywords: camera array, identify, object, moving, adjust, capture device

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	US 2010-0097476 A1 (RICHARD L. MARKS) 22 April 2010 See paragraphs [0026]-[0028], [0039]; claim 1; and figures 1, 6.	1-2, 6-7, 11-12, 21 , 25 3-5, 8-10, 13-20 , 22-24
A	US 2003-0179418 A1 (TIMOTHY WENGENDER et al.) 25 September 2003 See paragraphs [0006]-[0010]; claim 1; and figure 5.	1-25
A	WO 2006-026688 A2 (SARNOFF CORP.) 09 March 2006 See paragraph [0005]; and claim 1.	1-25
A	US 2009-0059009 A1 (FERNANDO MARTINS et al.) 05 March 2009 See paragraphs [0062]-[0065]; and figure 15.	1-25
A	US 7995055 B1 (XIAOXU MA et al.) 09 August 2011 See column 5, lines 15-29; claim 1; and figure 4.	1-25

 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 application or patent 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

"&amp;" document member of the same patent family

Date of the actual completion of the international search

17 September 2014 (17.09.2014)

Date of mailing of the international search report

**17 September 2014 (17.09.2014)**

Name and mailing address of the ISA/KR

International Application Division  
Korean Intellectual Property Office  
189 Cheongsu-ro, Seo-gu, Daejeon Metropolitan City, 302-701,  
Republic of Korea

Facsimile No. +82-42-472-7140

Authorized officer

AHN, Jeong Hwan

Telephone No. +82-42-481-8440



**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No.

**PCT/US2013/075717**

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2010-0097476 A1	22/04/2010	CN 1926851 A	07/03/2007
		CN 1926851 B	05/10/2011
		EP 1704711 A1	27/09/2006
		EP 2290949 A2	02/03/2011
		EP 2290949 A3	25/09/2013
		JP 2007-518357 A	05/07/2007
		KR 10-0810844 B1	06/03/2008
		KR 10-2007-0006692 A	11/01/2007
		TW 266248 A	11/11/2006
		US 2005-0157204 A1	21/07/2005
		US 7663689 B2	16/02/2010
		US 8085339 B2	27/12/2011
		WO 2005-074257 A1	11/08/2005
		US 2003-0179418 A1	25/09/2003
WO 2006-026688 A2	09/03/2006	US 2005-0232491 A1	20/10/2005
		US 2007-237398 A1	11/10/2007
		US 7466860 B2	16/12/2008
		US 7672514 B2	02/03/2010
		WO 2005-022077 A2	10/03/2005
		WO 2005-022077 A3	09/06/2005
		WO 2005-086078 A1	15/09/2005
		WO 2005-086079 A1	15/09/2005
		WO 2006-026688 A3	20/07/2006
US 2009-0059009 A1	05/03/2009	US 2004-0201719 A1	14/10/2004
		US 2005-0093984 A1	05/05/2005
		US 6952224 B2	04/10/2005
		US 7453497 B2	18/11/2008
		US 7884858 B2	08/02/2011
US 7995055 B1	09/08/2011	None	