(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 February 2009 (12.02.2009)

PCT

(10) International Publication Number
WO 2009/018665 A1

(51) International Patent Classification:
H04N 7/15 (2006.01) *H04M 3/22* (2006.01)
G06F 3/14 (2006.01) *H04M 3/56* (2006.01)
G06Q 30/00 (2006.01) *H04N 7/14* (2006.01)
H04L 29/02 (2006.01)

(21) International Application Number:
PCT/CA2008/001440

(22) International Filing Date: 7 August 2008 (07.08.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/964,088 8 August 2007 (08.08.2007) US
12/186,264 5 August 2008 (05.08.2008) US

(71) Applicant *(for all designated States except US)*: **QNX SOFTWARE SYSTEMS (WAVEMAKERS), INC.** [CA/CA]; Suite 302, 134 Abbott Street, Vancouver, British Columbia V6B 2K4 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HETHERINGTON, Phillip A.** [CA/CA]; 23 Fernway Drive, Port Moody, British Columbia V3H 5H7 (CA). **VAIRA, Mario** [CA/CA]; Suite 302, 134 Abbott Street, Vancouver, British Columbia V6B 2K4 (CA). **ESCOTT, Alex** [CA/CA]; 506-55 Water Street, Vancouver, British Columbia V6B 1A1 (CA).

(74) Agents: **SUE, Hilton W.C.** et al.; Suite 480 - The Station, 601 West Cordova Street, Vancouver, British Columbia V6B 1G1 (CA).

(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, 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, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, 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, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: VIDEO PHONE SYSTEM

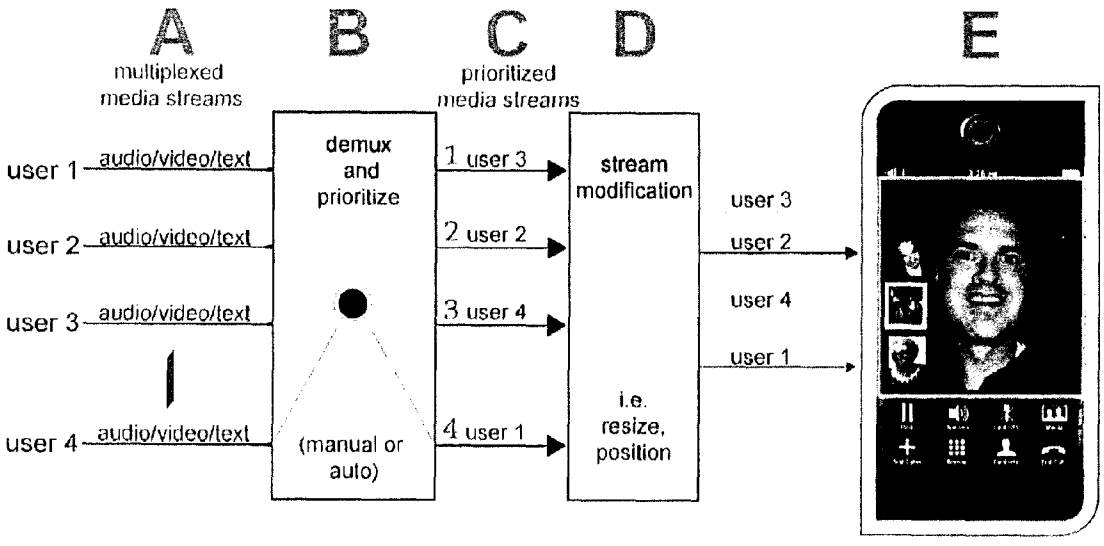

FIGURE 1

(57) Abstract: A system allocates channel bandwidth based on the data received from a plurality of remote sources. A de-multiplexer / priority circuit separates two or more different data streams into their components parts. A stream modification driver modifies one or more characteristics of the data received from the de-multiplexer / priority circuit based on a priority assigned to the data by the de-multiplexer / priority circuit. The de-multiplexer / priority circuit determines the data transfer rates for each of the different data streams based on the assigned priority.

WO 2009/018665 A1

Published:
— *with international search report*

Video Phone System

INVENTORS:

Phillip A. Hetherington
Alex Escott
Mario Vaira

BACKGROUND OF THE INVENTION

1. Priority Claim.

[0001] This application claims the benefit of priority from U.S. Provisional Application No. 60/964,088, filed August 8, 2007, which is incorporated by reference.

2. Technical Field.

[0002] This disclosure relates to data management, and more particularly to an automated interface that manages media.

3. Related Art.

[0003] In an audio or video conference users connect to others through a data network. Each user broadcasts media streams to other users whose systems receive and render sound or images.

[0004] To maintain interactions, systems receive and render data from each user. As the number of users increase, the audio speaker or video display that renders the media may distort the output or lose the detail delivered by a data stream. In a video application, the limited display area of a monitor may distribute images in equal areas about a screen hiding the fine details of some transmitted images. In an audio system, sound may converge from all of the loud speakers of the system. The convergence may distort the audio or render the media unintelligible.

SUMMARY

[0005] A system allocates channel bandwidth based on the data received from a plurality of remote sources. A de-multiplexer / priority circuit separates two or more different data streams into their components parts. A stream modification driver modifies one or more characteristics of the data received from the de-multiplexer / priority circuit based on a priority assigned to the data by the de-multiplexer / priority circuit. The de-multiplexer / priority circuit determines the data transfer rates for each of the different data streams based on the assigned priority or priorities.

[0006] Other systems, methods, features, and advantages will be, or will become, apparent to one with skill in the art upon examination of the following figures and detailed description. It is intended that all such additional systems, methods, features and advantages be included within this description, be within the scope of the invention, and be protected by the following claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] The system may be better understood with reference to the following drawings and description. The components in the figures are not necessarily to scale, emphasis instead being placed upon illustrating the principles of the invention. Moreover, in the figures, like referenced numerals designate corresponding parts throughout the different views.

[0008] Figure 1 is a logic diagram of a multi-user priority system.

[0009] Figure 2 is a visual display of circular frames.

[0010] Figure 3 is a visual display of a rotating cube.

[0011] Figure 4 is a visual display highlighting a geometric shape.

[0012] Figure 5 is a spatial diagram of multiple audio streams.

[0013] Figure 6 is a bandwidth allocation process.

[0014] Figure 7 a local multi-input user interface process.

[0015] Figure 8 a distributed multi-input user interface process.

[0016] Figure 9 is an automated user interface process.

[0017] Figure 10 is an automated panning process.

[0018] Figure 11 is an automated cropping process.

[0019] Figure 12 is a process that integrates a virtual reality element with an optional image element.

[0020] Figure 13 is an information placement process.

[0021] Figure 14 is a second information placement process.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0022] An automated system improves the fidelity of audio, video, text, or data received from remote sources. Through a priority (or priorities) assigned by one or more metrics, the system optimally distributes a limited bandwidth among connecting nodes of a network. Signal characteristics that may include image size, position, crop margins, skew, brightness, and/or layer order (z-depth) of a video and/or the volume level, pan, equalization, reverberation of a signal may be modified according to the priority (or priorities). Priority may be assigned by a detection event (e.g., occurrence or magnitude of voice, motion, etc.), a manual event (e.g., a gesture detected by an input that captures and/or tracks images or on a single or multi-input touch sensitive screen), and/or an automated event (e.g., a facial recognition, a sound recognition, etc.).

[0023] Figure 1 is a system diagram of a multi-user device. The interface logic allows users to send and receive information from remote sources across a network. The network may transfer data related to similar or different media or data elements. The media may comprise audio, video, text, and/or other data that are pushed from a source or pulled by a user device. The media may be received continuously or in discontinuous time periods. The nature of the media and rate of delivery may affect the frequency and priority (or priorities) in which the media or data is rendered.

[0024] In Figure 1, a communication device receives one or more signals through multiple channels (shown in A). The signals may be a combination of media originating from a common source or from multiple sources. One or more remote concentrators may combine several different streams of data such as audio, video, text, and/or other data and transmit the different or segregated concentrations of data across a common or distributed communication lines. In Figure 1, four separate channels deliver the multiplexed media to an inverse multiplexer or de-multiplexer and priority circuit or logic (de-multiplexer / priority circuit). The entire bandwidth at

the client connection may be equally or unequally distributed before the media is processed by the de-multiplexer / priority circuit. Additionally, more or less media communication lines (telephone lines or communication channels) may be processed by the multi media device.

[0025] The de-multiplexer / priority circuit analyzes or processes each of the media streams before allocating bandwidth among the output user channels. After separating each media stream into its component parts, the de-multiplexer / priority circuit analyzes one or more metrics of the received media. Some systems detect voice (e.g., through a voice detector or voice detection process), video (e.g., through a video detector or video detection process), or noise (e.g., through a noise detector or a noise detection process); some systems use a manual determination; some systems apply automated systems or methods, and other systems use a combination embodied in logic, devices, or circuitry to assign priority (or priorities) to the user inputs or prioritized media streams. Bandwidth allocations are assigned before or after the channels are transmitted to a stream modification driver. In each channel, the audio/video/text/etc. may be re-combined into one (as shown) or more signals before they are transmitted to the stream modification driver through a unidirectional or bidirectional bus. The de-multiplexer / priority circuit and stream modification driver may be separate from or a unitary part of the multi media user device (e.g., end user media device).

[0026] Using the derived priority from the de-multiplexer / priority circuit, bandwidth is allocated. Bandwidth allocation may occur continuously during a user session or may lack a temporal relationship; it may occur asynchronously. Based on the priority, the de-multiplexer / priority circuit may communicate with the originating source or a surrogate source (that may be linked to or in communication with an originating source) through a transceiver to assign data transfer rates to each input channel. In some systems the communication may comprise a request to reduce transfer rates.

[0027] In some systems, the cumulative data transfer rates of the separate channels may be limited to the bandwidth of the de-multiplexer / priority circuit interface. In other systems, a de-multiplexer / priority circuit may adjust the transmission speed or rendering rates through data selections, truncations, or buffering that may record audio, video, and/or messaging in a local memory (e.g., a flash memory, hard drive,

digital media/video recorded memory, etc.) that may or may not be rendered to a user in session (at a user's direction or this may occur automatically). The local memory may comprise a memory module or one or more elements that may be contributed to a shared addressable memory space that interfaces one or more nodes in communication with the end user media device (e.g., client device or client media device). In some systems the total data transfer rates allocated to the channels may be fixed while communication connections are sustained or may vary during a user's session. A channel may be allocated a relatively low bandwidth when little or no information is received, such as when no activity or data is received or process. During that same session or period another channel may have increased activity or data transfer rates and may be allocated a relatively higher bandwidth through a flexible and/or dynamic adjustment.

[0028] In one multimedia session, each user may have a different bandwidth allocation (e.g., a static allocation or a dynamic that may change during a session such as a dynamic temporal allocation) based on the metrics analyzed by the client media device. Once allocated and processed by the de-multiplexer / priority circuit, the signals are transmitted to the stream modification driver that may resize, position, amplify, pan, and/or etc., the media before sending it to the end user media device that renders it. In Figure 1, the stream modification driver may position the images of the users (shown by end user media device E).

[0029] The end user media device (E) may render the media in many forms. Calls may be received and made in many manners. The end user media device may convert voice, other sound signals and data into one or more forms that may be transmitted to one or more remote locations (e.g., teleconferencing) that receive and reconvert the content into an original or other perceptible forms through devices that translate the media from one perceptible form to another (e.g., a voice-to-text converter, text-to-voice converter, video-to-voice converter, etc.). A call may include a combination of video, audio, and/or text. In some systems, the diverse media may not be transmitted or stored in the end-user media device as separate files until the complete file is received. In these systems, the multimedia may be encoded in a single format that may be rendered in real time 'on-the-fly' even when the communication content is not completely received (e.g., downloaded).

[0030] The hardware or software of the end user media device may present the media through virtual frames shown in Figure 2. Higher priority images or data may be rendered in larger diameter bubbles, circles, or spheres (e.g., geometric shapes). In some systems, the display may simulate physical motion by rendering images that "bounce" into each other. As the activity associated with an image increases audio volume may be adjusted and panned to identify the more active channels. A user may move image around through a wireless absolute or relative pointing method or device (sensed on a single or multi-touch sensitive screen or gesture sensitive input or screen). When images appear in moving frames, a gesture may be used to "flick" an image. It may send an image spinning about the display region. In some systems the acceleration, speed, and/or direction of the gesture may determine the acceleration, speed, and direction of the movement of the image. The movement may be rendered locally or may be disturbed to the local and the geographically separated remote users so that the users may view a customized display or share a common view.

[0031] Figure 3 shows a three dimensional rotating virtual cube that may display media across the visual face(s) of an object. In this interface, higher priority or high bandwidth media (e.g., such as advertising) may rotate towards a user. When two users speak simultaneously or almost simultaneously, they may appear on an "edge" of the virtual cube facing the user. In some applications the two other faces of the virtual cube may separately render each video stream. When three users speak simultaneously or almost simultaneously, they may appear on separate "edges" of the virtual cube to from a corner. Image size, audio volume, textual size, color, font, and/or other characteristics may be adjusted to highlight one or more active or relatively more active channels. A user may navigate about the virtual cube through an absolute or relative input (e.g., through an absolute or relative pointing method or device), a gesture captured on a multi-touch or single touch screen or gesture capturing user interface, or by some other circuits, devices, or methods. An absolute pointing device or method may associate a user's selection or position to a position on a display. A relative pointing device or method may associate a user's or device's movement to an on-screen movement but not to a position.

[0032] Figure 4 shows a virtual geometric object. The media may be sectioned in geometric shapes. Higher priority media may be circumscribed with a highlight. In

some systems higher bandwidth channels may be positioned near a prominent position of the geometric shape such as near the top of a pyramid, the top of a triangle, the center of a sphere or the center of a circle, for example, and may have a higher adjusted volume (or more media) delivered by a source closest to the end user. In Figure 2, 3, and 4, the images may be rendered in full screen by a video driver and the video may be selectable and modified in real time. The volume and data may also be adjusted to call out or highlight the active channel or speaker.

[0033] Bandwidth allocation and priority may also be used in audio systems or social networks that receive several audio streams that are rendered into aural sound for each user. Multiple audio streams may be routed to different speakers ("panned") in a surround sound field as shown in Figure 5 according to the priority (or priorities) assigned or relative bandwidth allocated by the de-multiplexer / priority circuit. If a user is speaking, an audio stream may be routed to a center channel of a system exclusively, for example, and may be amplified to reinforce their presence. Ancillary sources (e.g., other users' audio streams) may not be delivered through a direct communication path to the end user. Ancillary sound may include background noise, or may capture sporadic conversations. For example, acknowledgements from other sources expressed as a "ya".. "of course"... "okay" may be routed to rear side, or lower priority loud speakers that may not deliver audio directly to the end user. In some systems, such ancillary communication or noise may be attenuated. In some alternative systems noise removal circuitry or logic may process the information rendered by the end user media device or delivered by channels when a low signal-to-noise ratio is detected ("SNR"). In each of the systems or in alternative systems an automatic-gain-control ("AGC") may be applied to channels to bring each of the channels to a relatively consistent or different (but perceptible) call or volume level. In these systems, the noise removal circuitry or logic and/or automatic gain control may be a unitary part or separate from the end user media device.

[0034] Customized graphical images or skins may be disposed within, near, or behind the visual displays. A standard background, one or more advertising messages, or corporate skin may be displayed (e.g., scrolled). In some systems, the media may be positioned behind a more traditional display. In a "teen" mode, images such as head shots may be mapped onto circle, bubble, sphere, cube, or other objects that move

around in front of, behind, or near a graphical image. The elements may interact (e.g., move, strike, collide with, dissolve, attach to, surround, etc.) with each other or may become animated in a game-like or interactive game-like manner. Members may be invited to join a "video/textual chat room." In a "family" mode, members may be invited into a common element such as an image of a "house." The skins may maintain one or more visual impression characteristics of the content provider, an advertiser, a source, or may identify the selected mode or function. The visual impression characteristics may comprise one or more color(s), font(s), styles (s), or symbol(s) that may be associated with or may differentiate a content provider, source, product, or user. A visual impression characteristic may be a distinctive characteristic by which a product, service, or response comes to be known, becomes associated with a source, or is well known (e.g., a trademark or trademark like).

[0035] Since the media streams sent between users may contain any type of media or data, the system may interface an audio-visual system or file that renders an accompaniment to a song that may be performed by singing or following the words on the video screen. One or multiple users may broadcast a data stream that includes audio, text (e.g., timed text, possibly embedded in a multi-track media container such as *.mp4), data and/or possibly video to all the other or selected end user transceivers in a session or a call. Each user or multiple users may assume the spotlight by performing the song which may be heard and seen by all or in some systems, by selected users in the session or call.

[0036] Some systems support separate sub-sessions within a user session or a multi-user videophone call. If a user wishes to speak privately to a number of other users, they may transmit a request to those users to "join a private room" through a transceiver coupled to the de-multiplexer / priority circuit. A user may request a private conversation through a graphical user interface. In one system, a user may select an object associated with a user through a gesture input, pointing device, and/or touch sensitive screen and may drag or move the objects to another object or location on the screen. A user may release the object (e.g., image) onto another object (e.g., an image or icon) that may appear as a door icon or another image. Such a movement may create a sub-group that would allow users to automatically invite others to join a call. If an invitee accepts the invitation, users may speak privately or share media

(e.g., text, video, images, media conversions such as speech-to-text, etc.) with the invitee(s) through separate channel(s) or a call line "on the side" without disturbing (or releasing access to) the others that may continue the earlier session or call. The system may maintain multiple communication links, and some systems may automatically mute one or more audio or video output channels in response to a user's direction or through an activity detector or method when a side-bar connection occurs. Some systems may allocate bandwidth among sub-session participants as described herein. In alternative systems, the bandwidth allocation may occur through an aggregate of all if the active multi media calls or sessions that the end user media device supports.

[0037] An adaptive bandwidth allocation may automatically adjust the transmission speed (e.g., bandwidth) or rending rates of the end user media device based on one or more metrics. Some processes have or provide the capability of increasing throughput in increments such as the process shown in Figure 6. The process reads data at a node that interfaces one or more communication channels linked to the end user media device at 602. The communication rates of some or each of the communication channels may be estimated or measured. The speed at which information is received (or the amount), the content (voice, video, textual, or data content), signal quality, the magnitude, or one or more other criteria or metrics are processed to allocate transmission speeds or rendering rates at 606. As one or more of the criteria or metrics change an adaptive bandwidth or rendering adjustment may occur at 608. The channels and adjustments may be associated with user interface elements that may comprise images, real time processed and/or rendered video streams, pictures, photographs, animations, caricatures, avatars (of users, animals, objects, etc,) and/ or static or animated objects at 610. Each user interface element may represent a static or dynamic identity element or object that may be rendered by an end user media device and shown on a display or transmitted through an audio output such as an audio speaker.

[0038] A pressure and/or optical sensitive input (may include touch sensitive membranes that include pressure sensors and/or gesture sensitive inputs that capture images) in Figure 7 may allow a user to enter multiple inputs through one or more nodes or touches. When a user makes a selection an input is detected at 702.

Through a motion or one or more contacts (e.g., touches with an input) the process or a driver receives data from the user. The process or circuitry (e.g., an input/output controller) monitors the input that may comprise a single selection through a single input or one or more selections that may be made through a gesture at 704. As the input is tracked and classified at 704 and 706, the user may activate programs and select options resident to a memory module. If the user interactions are classified as a gesture (e.g., more than a single interaction with an input or an image) it may be compared to a set of objects that have meaning through a gesture recognizing device at 708. Based on the comparison, a computing device may execute the pre-coded objects (e.g., computer executable modules) generated by the recognizing device. When executed, the pre-coded objects may activate programs stored in a client-side or end user memory. If the user gestures are classified as an activation (e.g., a single user interaction), the user interaction may cause a processor (resident to the end user media device) to execute pre-coded objects that may activate a program stored in the client side memory.

[0039] The display elements may be modified or changed in real time based on substantially simultaneous input from a local user and a remote source. Visual display elements may be rendered on an input or display (e.g., single point or multipoint pressure or optical sensitive membrane) in Figure 8 at 802. The display or a portion of it may be available or accessible to a single user or linked through a multi-user network to multiple geographically separated users. One or more gestures or activations may be received over or across a visual element such as an image of one or more users that may be linked to the end user media device at 804. The inputs may comprise a single interaction or multiple interactions that may occur simultaneously, substantially simultaneously, or may occur in series that may be sensed one after another in time. Remote input may also be received through a communication channel in unison with or at anytime during the process at 808. Based on the local and/or remote input, display control codes may be transmitted to the display driver to modify or manipulate the appearance or movement of a video, an image, picture, animation, avatar, and/or element that may depict one or more user's identity (referred to as visual element(s)). In some systems the movement, changes, or manipulations may occur and/or correspond to the inertia, speed, and/or direction

of the input (e.g., the gesture or activation). For example, a counterclockwise rotation may cause an image or object to rotate in a counterclockwise direction. In some processes, the rate and/or direction (e.g., drift) of the rotating image or object on the display may correspond to the rate and/or direction of the user's gesture. In some systems, the acoustical pitch of an input (e.g., a fundamental frequency of a person's voice) detected through an acoustical pitch detector may vary the horizontal and/or vertical movement (e.g., the angular movement) of one or more visual elements rendered through a pitch controller, driver, or processor. A pitch detection and tracking method, detector, driver or processor may track a source's acoustical pitch and tilt a visual element such as a three dimensional avatar's head forward and backward in relation to the detected acoustical pitch. The controller, driver, or processor may correlate an aural spectrum to a visual movement (referred to as a visual pitch control).

[0040] In some processes the changes, manipulations, and/or modifications may be shown on a single end user media device or shared with multiple users through a multiple user network at 810. When shared, the computer executable control codes may be transmitted to the other display drivers so each or some of the user's end user media devices may render the modified display elements at 812 (e.g., a common interactive display) in real time or in substantially real time.

[0041] The rate of change or an image or position of an image, picture, animation, avatar, and/or identity (that may participate in a video and/or audio conference) may depend on the dynamic or static allocated bandwidth. An input/output controller that may comprise a pressure and/or optical sensitive device or process (that may include touch sensitive membranes that may have pressure sensors and/or gesture sensitive inputs that capture images) may render display elements in Figure 9 at 902. Based on the received communication channel transmission rates, a derived priority (or priorities) may allocate bandwidth between the participants of a call or session at 904. Based on the allocation, the rate at which the display elements change may correspond to the rate of change of the user's input and the assigned bandwidth allocation at 906. In some processes the rate of change of a display element (e.g., an identity element) may correspond exclusively to the assigned bandwidth or transmission rates. In some processes, the rates may be fixed or may be dynamic. In

alternative processes, a processor executed tween may control changes of visual elements such as processes by which one visual element (e.g., an image) may change or gradually transform into another visual element. The process may create an illusion of a metamorphosis that may occur in a short period of time. In some uses, the processor executed tween may transform a static element into an animated element (e.g., gradually transform a static image into an image that appears to be moving or talking). When combined with a visual pitch control, the animated element may move in two (e.g., a two coordinate system), three (e.g., a three coordinate system), or four (e.g., a fourth coordinate system where one coordinate may comprise time) dimensions.

[0042] A display process that may automatically bring off screen extensions of an image into view may also be retained in a memory of the end user media device (e.g., media player). At 1002 the process of Figure 10 may capture an image. A digital or video camera may store electronic images in the memory. When an image is not centered or miss positioned, the end user media device may process or scan the image or a portion of the image (a head in a photograph for example) at 1004. The measured distances between a reference point (e.g., a midpoint) and an outer boundary may be compared. Through a visible or transparent movable display window, the image or object may be repositioned automatically at 1008 and 1010. In some processes, the centered or repositioned object may maintain the same aspect ratio as the unmodified object. As the object is rescaled, an auto-sizing process may accept the image at one resolution and display the image at a different resolution. The object may be enlarged or reduced to fit an allocated space or position.

[0043] Some processes may automatically remove the outer portions (or unwanted areas or irrelevant details) of an image to improve framing, accentuate subject matter, improve a composition and/or change the aspect ratios. In the automated process of Figure 11, the end user media device may remove unwanted areas of an object in a horizontal and vertical plane. If unwanted areas are positioned near a left or right portion of an image or object, a horizontal crop signal is generated at 1102. If unwanted areas are positioned near an upper and/or lower portion of an image or object, a vertical crop signal is generated at 1104. The two dimensional cropping may occur substantially simultaneously to improve the placement, centering, or focus of an

object. In some alternative processes, an object may be brought into focus by generating vignette signals in one or more dimensions (e.g., height, width, depth, etc.). Vignetting may reduce an object's (e.g., image) brightness or saturation at the object's periphery compared to the object's (or image) center.

[0044] In some processes objects or images may be generated or modified by placing or layering a graphical representation, animation, image, and/or object near or on another object. Following the image capture and processing of 1002 – 1004 (previously described), the process of Figure 12 examines facial features by detecting patterns of dark and light at 1202. Once the end user media device has detected one or more patterns, template recognition methods that pattern match the detected patterns to stored sets of templates in the end user media device occur at 1204. Once matched, a position point may be identified. In some processes, a sharp gradient may identify a position point at 1206. An avatar, generic picture, animation, caricature, or other object may be positioned at 1206. The avatar may be automatically selected by the process. In alternative processes, the avatar may be chosen by a user.

[0045] In some processes the display elements may be shown with information that attracts one or more user's attention to a product, service, or business. The audio and/or visual messages may overlay or supplement virtual or interactive elements such as those shown in Figures 1 – 4 (Figure 13 at 1302). The product, service, or business information may be identified by a user interface process (or user interface). The user interface process may identify a product that may be of interest through a contextual recognition (e.g., recognition of audio, visual, and/or textual content through an audio and/or visual recognition device) or by an association with a location. In some processes the local or remote contextual recognition may automatically recognize demographic elements or physical features (e.g., the gender, age, race, and physical features (glasses, beard, etc.) of a source and/or associate (e.g., one or more contacts that may be linked in a social network) to deliver a targeted message or marketing. The process may communicate with a remote advertising module that may retain files or stream media for distribution to mobile media clients. In response to a request (or on its own imitative, e.g. target marketing) the advertising module may transmit multimedia product, service, or business information to the end user media device. The end user media device may receive local or universal

advertising information that may be retained in separate files or encoded with, or linked to objects such as one or more frames that may bound or outline the geographic elements at 1304. A computing device (within the end user media device) may execute an application that renders the contextual and/or targeted advertising at 1306. The advertising may be transmitted among the various geographically separated participants of a call or session. In alternative processes, separate advertising may be distributed to some or each of the participants. In some processes the customized distribution may be based on the content a user originates or transmits, a user's geographic location, a user's interactions with the conferencing media, the multi media the user selects, etc.

[0046] In Figure 14 a second receiver (or transceiver in the end user media device) may track location through navigation signals that may comprise a GPS (global positioning system) protocol, a differential GPS protocol, a trilateraleralism of external encoded signals (e.g., may be in the radio frequency range), protocols that monitor continuously transmitted coded signals, or other locating protocols or systems at 1402. Based on the geographic location of one or more participant, the process may communicate with a remote advertising module that may create and maintain files or stream media for distribution to end user media devices. In response to a request (or on its own initiative) the advertising module may transmit multimedia product, service, or business information that may be processed as described at 1304 and 1306.

[0047] The systems, methods, and descriptions of Figures 1 – 14 may be encoded in a portable or static signal bearing medium, a computer readable medium, a computer executable module, or a computer readable storage medium such as a memory that may comprise unitary or separate logic, programmed within a device such as one or more integrated circuits, or processed by a controller or a computer. If the methods are performed by software, the software or logic may reside in a memory resident to or interfaced to one or more portable or static processors or controllers, a wireless communication interface, a wireless system, a powertrain controller, an entertainment and/or comfort controller of a vehicle or non-volatile or volatile memory remote from or resident to the end user or mobile media device. The memory may retain an ordered listing of executable instructions for implementing logical functions. A

logical function may be implemented through or comprise digital circuitry, source code, analog circuitry, and/or through an analog source such as through an analog electrical, or audio signals. The software may be embodied in any computer-readable medium or signal-bearing medium, for use by, or in connection with an instruction executable system or apparatus resident to a vehicle or a hands-free or wireless communication system. Alternatively, the software may be embodied in media players (including portable media players) and/or recorders. Such a system may include a computer-based system, one two or more processor-containing system that includes an input and output interface that may communicate with an automotive or wireless communication bus through any hardwired or wireless communication protocol, combinations, or other hardwired or wireless communication protocols to a local or remote destination, server, or cluster.

[0048] A computer-readable medium, machine-readable medium, propagated-signal medium, and/or signal-bearing medium may comprise any medium that contains, stores, communicates, propagates, or transports software for use by or in connection with an instruction executable system, apparatus, or device. The machine-readable medium may selectively be, but not limited to, an electronic, magnetic, optical, electromagnetic, infrared, or semiconductor system, apparatus, device, or propagation medium. A non-exhaustive list of examples of a machine-readable medium would include: an electrical or tangible (e.g., physical) connection having one or more links, a portable magnetic or optical disk, a volatile memory such as a Random Access Memory "RAM" (electronic), a Read-Only Memory "ROM," an Erasable Programmable Read-Only Memory (EPROM or Flash memory), or an optical fiber. A machine-readable medium may also include a tangible medium upon which software may be electronically stored as an image or in another format (e.g., through an optical scan), then compiled by a controller, and/or interpreted or otherwise processed. The processed medium may then be stored in a local or remote computer and/or a machine memory.

[0049] Each of the systems described above may improve the quality of audio, video, text, and/or data received from remote sources. By assigning a priority or priorities based on one or more metrics, the system (e.g., a mobile or portable system) optimally distributes a limited bandwidth among connecting nodes of a network or within the

processing environment of the end user media device. Signal characteristics that may include image size, position, crop margins, skew, brightness, and/or layer order (z-depth) of a video or the volume level, pan, equalization, reverberation of a signal, etc. may be modified by an assigned priority within or separate from the end-user or rendering device. One or more priorities may be assigned by an event such as a detection event (e.g., voice, motion, etc.), a manual event (e.g., a gesture on a touch screen), and/or an automated event (e.g., facial recognition, sound recognition, etc.).

[0050] While various embodiments of the invention have been described, it will be apparent to those of ordinary skill in the art that many more embodiments and implementations are possible within the scope of the invention. Accordingly, the invention is not to be restricted except in light of the attached claims and their equivalents.

CLAIMS

We claim:

1. A system that allocates channel bandwidth to data received from a plurality of remote sources comprising:

a de-multiplexer / priority circuit configured to separate a plurality of different data streams into their components parts; and

a stream modification driver that modifies one or more characteristics of data received from the de-multiplexer / priority circuit based on a priority assigned to the data by the de-multiplexer / priority circuit;

where the de-multiplexer / priority circuit determines a plurality of data transfer rates for each of the different data streams based on the assigned priority.

2. The system of claim 1 where the priority is based on a voice detection.

3. The system of claim 1 where the priority is based on a manual gesture sensed by a touch screen.

4. The system of claim 1 where the priority is based on a facial recognition circuit in communication with the de-multiplexer / priority circuit.

5. The system of claim 1 where the de-multiplexer / priority circuit adjusts the data transfer rates for each of the data streams by communicating a request to reduce transfer rates with an originating source.

6. The system of claim 1 where the de-multiplexer / priority circuit adjusts the data transfer rates for each of the data streams by communicating a request to a surrogate to reduce transfer rates.

7. The system of claim 1 where the de-multiplexer / priority circuit adjust the data transfer rates for each of the data streams by selectively storing data through a media recorder that is later rendered in a user session that transmitted the data.

8. The system of claim 1 further comprising a display driver that renders the data stream into interactive geometric shapes that interact with each other.

9. The system of claim 8 further comprising a user interface that associates a user's movement to an on screen movement but not to a position.

10. The system of claim 8 further comprising a user interface that associates a user's position to an on screen position on a display.

11. The system of claim 1 further comprising a user interface that identifies product or service information that may be of interest to a user through a contextual recognition device, where the user interface communicates with a remote advertising module to receive streaming multi-media product, business, or service information that is rendered by a display.

12. The system of claim 1 where the system comprises a multi-media portable device that supports a plurality of separate teleconferences that occur at a common time.

13. The system of claim 12 further comprising an activity detector that mutes an audio or video output when an activity is detected on one of the separate teleconferences.

14. A method of allocating channel bandwidth to data received from a plurality of remote sources comprising:

detecting multiplexed media streams from a plurality of input channels;

separating each of the multiplexed media streams into their component parts;

assigning a priority to each of the multiplexed media streams based on one or more metrics used to measure data within the multiplexed media streams;

determining a plurality of data transfer rates for each of the multiplexed media streams that will highlight one or more active input channels; and

transmitting bandwidth requests from an end user device to sources of the multiplexed media streams.

15. The method of claim 14 further comprising detecting a voice input and assigning the priority based on the detected voice input.

16. The method of claim 14 further comprising detecting a gesture input and assigning the priority based on the detected gesture input.

17. The method of claim 14 further comprising storing selective data through a media recorder that is later rendered in a user session that the data was received.

18. The method of claim 14 further comprising rendering the data streams into interactive geometric shapes.

19. The method of claim 18 further comprising receiving a user input that causes at least one of the interactive geometric shapes to move in a direction of the input.

20. The method of claim 14 further comprising rendering a plurality of multi-media teleconferences.

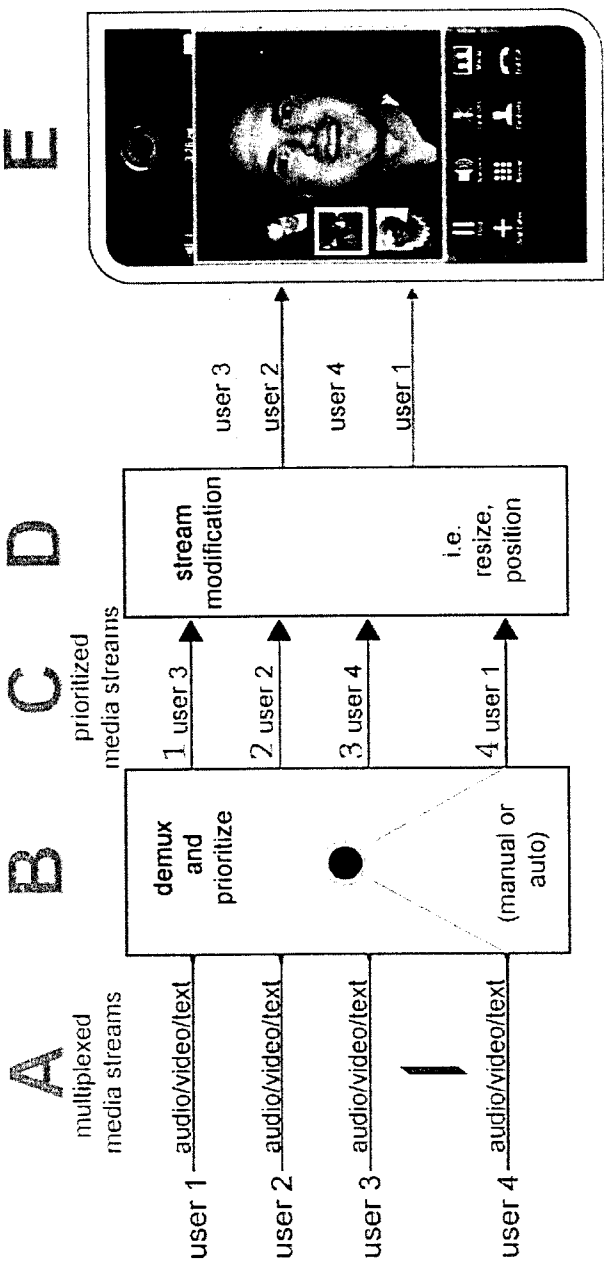
A
B
C
D
E
multiplexed media streams
prioritized media streams
user 1
audio/video/text
user 2
audio/video/text
user 3
audio/video/text
user 4
audio/video/text
demux and prioritize
(manual or auto)
1 user 3
2 user 2
3 user 4
4 user 1
stream modification
i.e. resize, position
user 3
user 2
user 4
user 1

FIGURE 1

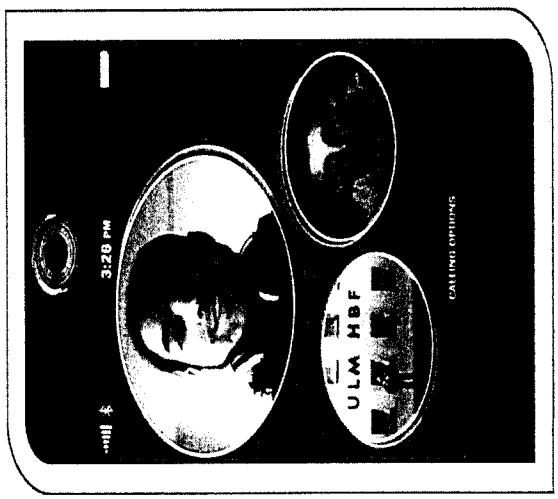

FIGURE 2


3:26 PM
John Cornell
Clarence Chu
Mario Valra
CALLING OPTIONS

FIGURE 3

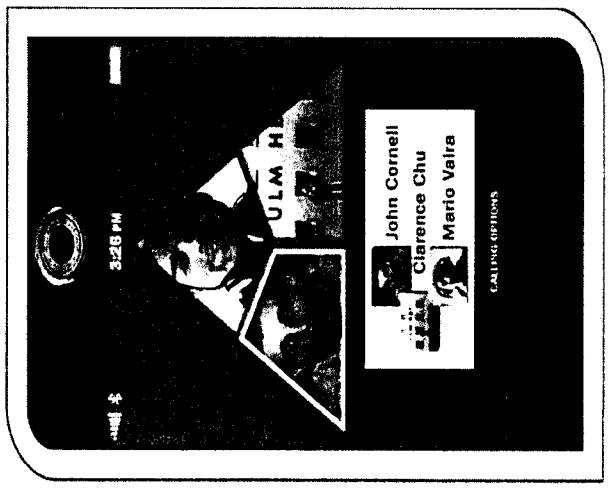

FIGURE 4

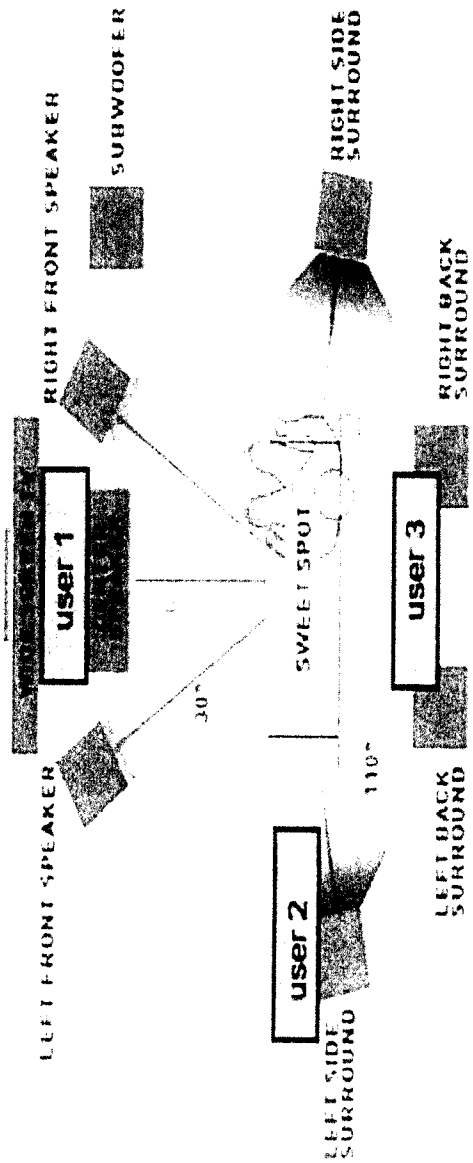
RIGHT FRONT SPEAKER
SUBWOOFER
RIGHT SIDE SURROUND
RIGHT BACK SURROUND
user 1
user 3
user 2
LEFT FRONT SPEAKER
LEFT SIDE SURROUND
LEFT BACK SURROUND

FIGURE 5

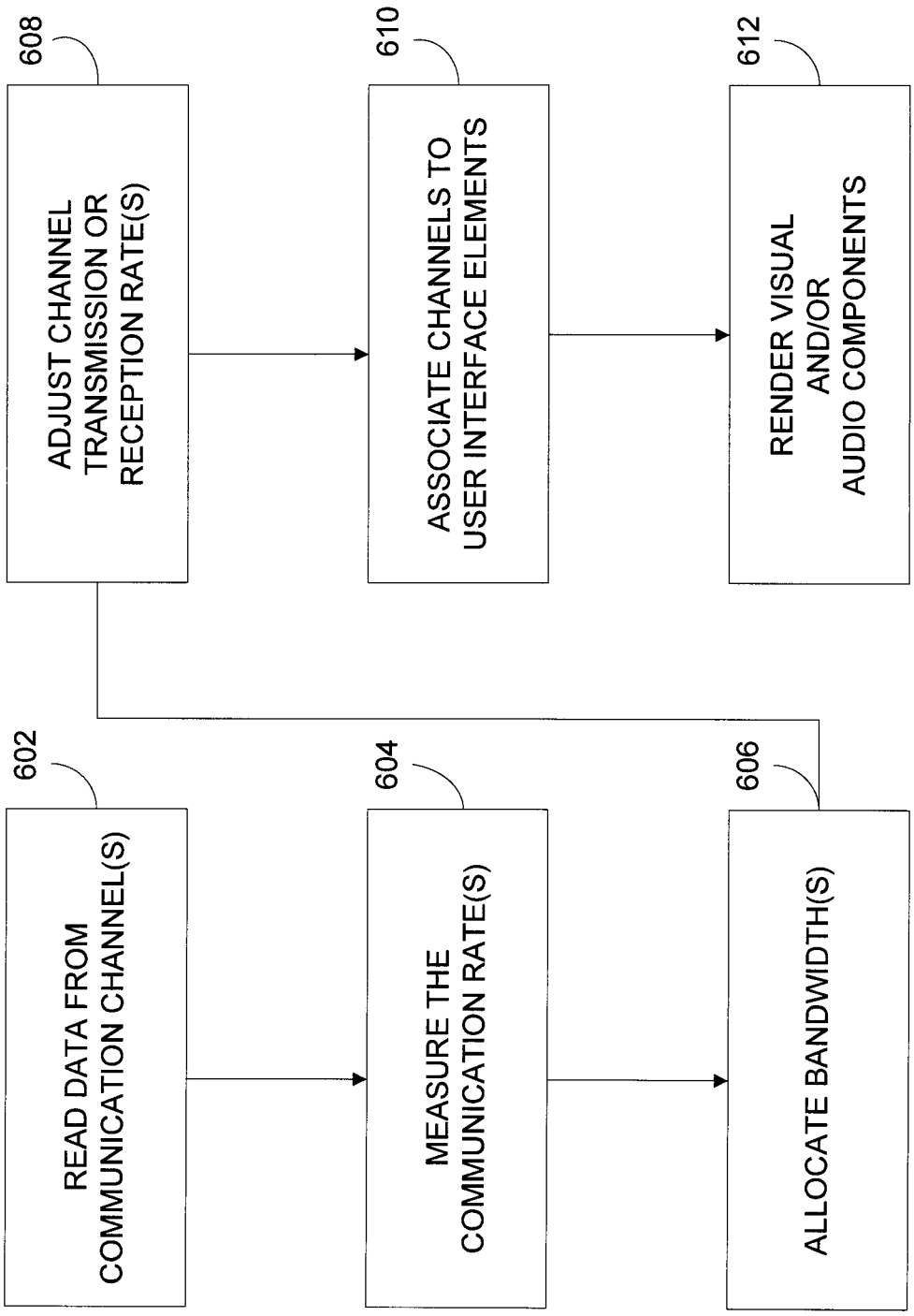
602
READ DATA FROM COMMUNICATION CHANNEL(S)
604
MEASURE THE COMMUNICATION RATE(S)
606
ALLOCATE BANDWIDTH(S)
608
ADJUST CHANNEL TRANSMISSION OR RECEPTION RATE(S)
610
ASSOCIATE CHANNELS TO USER INTERFACE ELEMENTS
612
RENDER VISUAL AND/OR AUDIO COMPONENTS

FIGURE 6

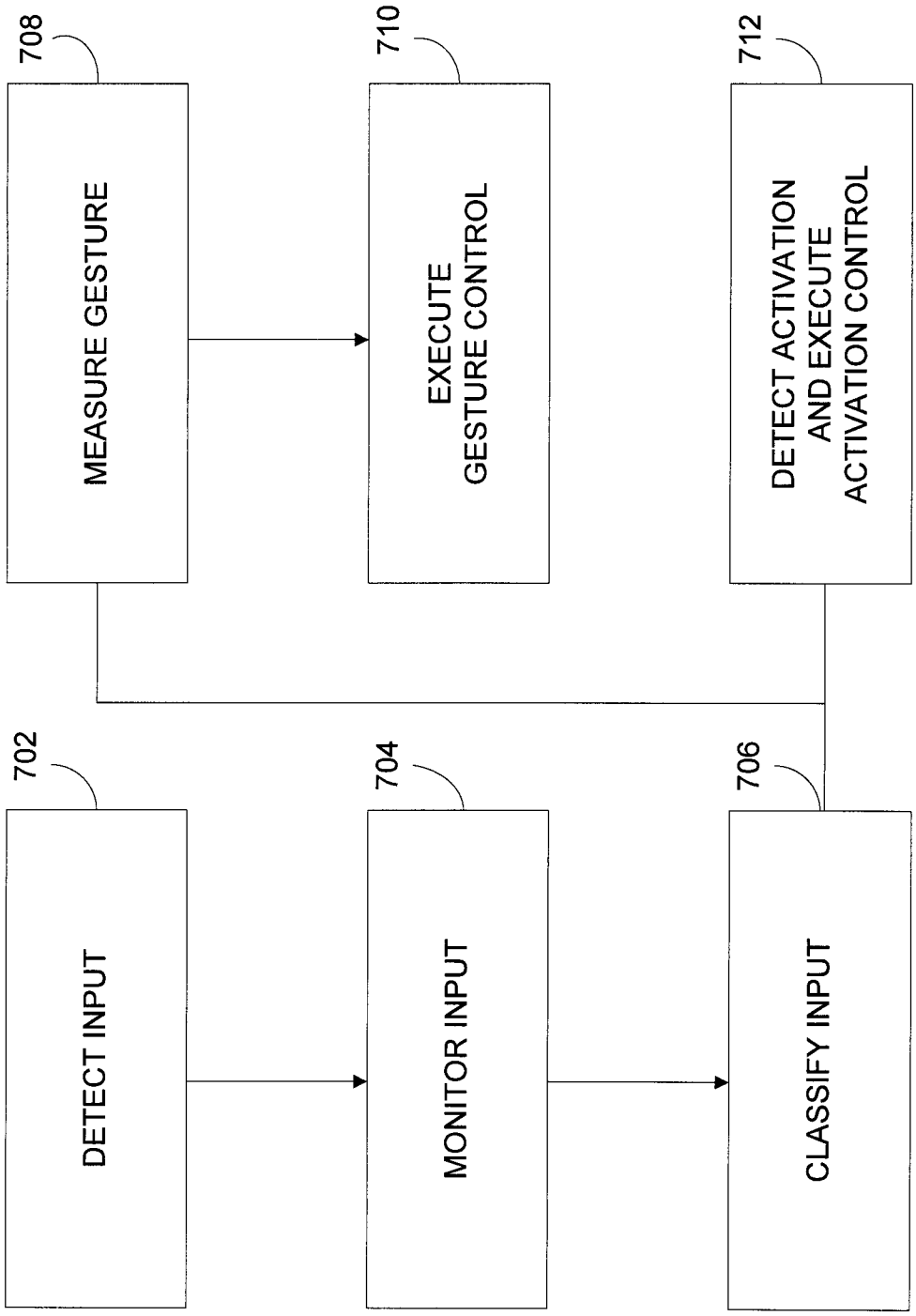
708
MEASURE GESTURE
710
EXECUTE GESTURE CONTROL
712
DETECT ACTIVATION AND EXECUTE ACTIVATION CONTROL
702
DETECT INPUT
704
MONITOR INPUT
706
CLASSIFY INPUT

FIGURE 7

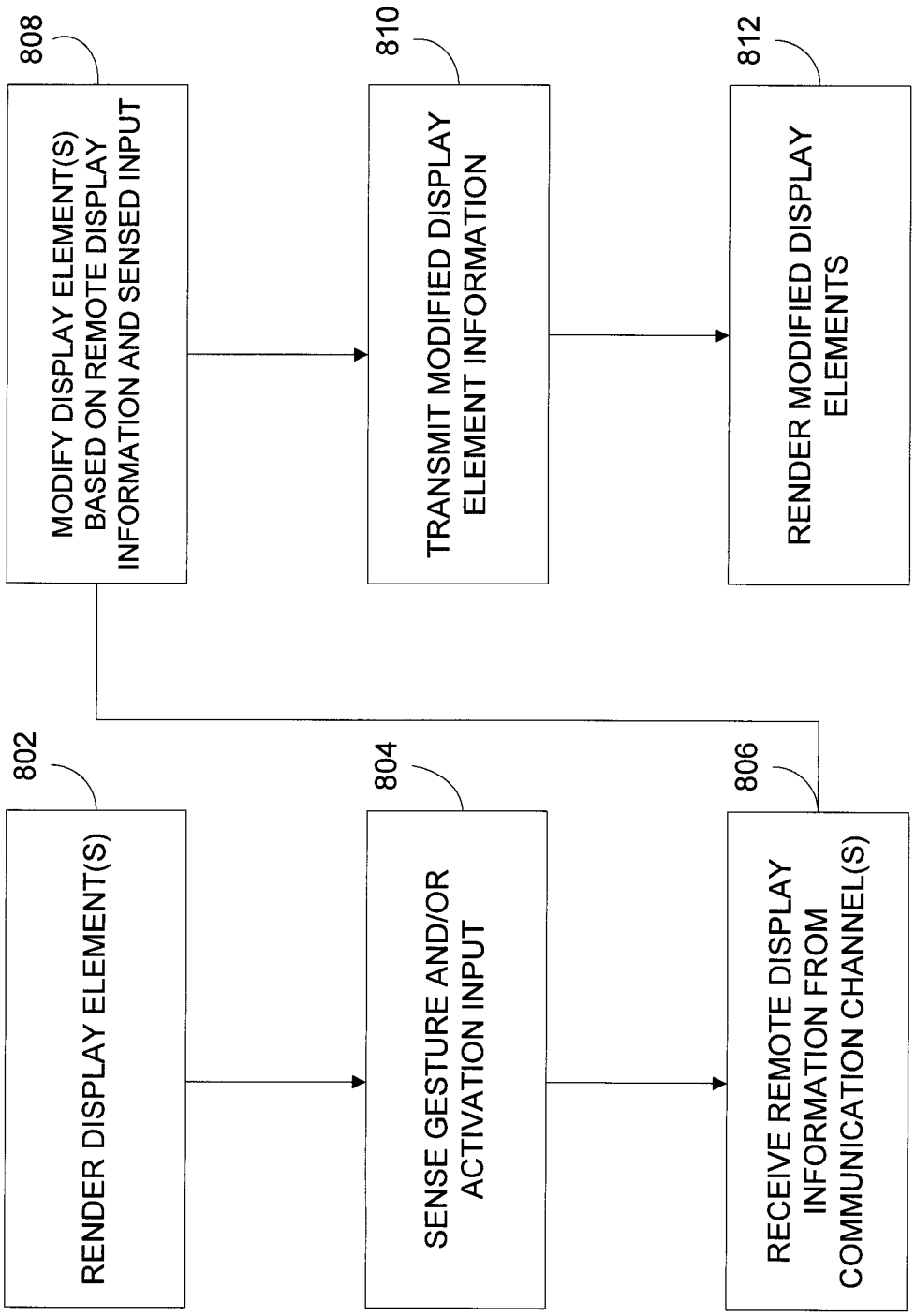
808
MODIFY DISPLAY ELEMENT(S) BASED ON REMOTE DISPLAY INFORMATION AND SENSED INPUT
810
TRANSMIT MODIFIED DISPLAY ELEMENT INFORMATION
812
RENDER MODIFIED DISPLAY ELEMENTS
802
RENDER DISPLAY ELEMENT(S)
804
SENSE GESTURE AND/OR ACTIVATION INPUT
806
RECEIVE REMOTE DISPLAY INFORMATION FROM COMMUNICATION CHANNEL(S)

FIGURE 8

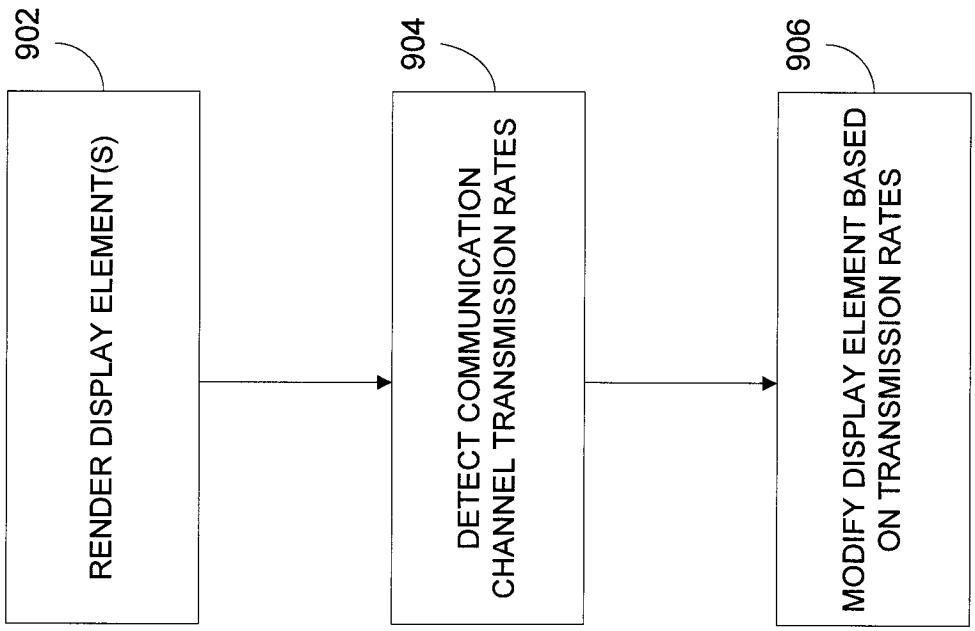
902
RENDER DISPLAY ELEMENT(S)
904
DETECT COMMUNICATION CHANNEL TRANSMISSION RATES
906
MODIFY DISPLAY ELEMENT BASED ON TRANSMISSION RATES

FIGURE 9

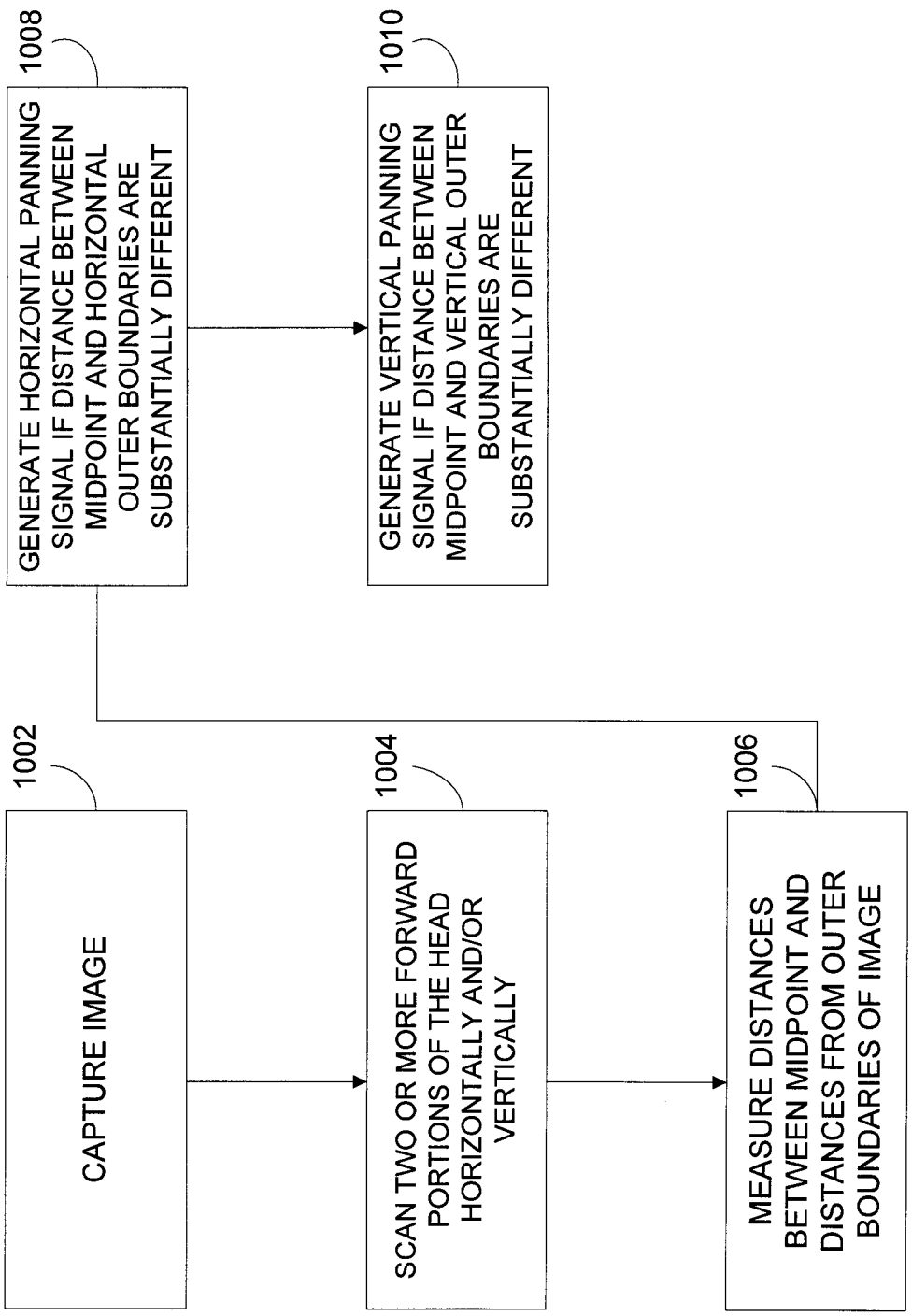
1002
CAPTURE IMAGE
1004
SCAN TWO OR MORE FORWARD PORTIONS OF THE HEAD HORIZONTALLY AND/OR VERTICALLY
1006
MEASURE DISTANCES BETWEEN MIDPOINT AND DISTANCES FROM OUTER BOUNDARIES OF IMAGE
1008
GENERATE HORIZONTAL PANNING SIGNAL IF DISTANCE BETWEEN MIDPOINT AND HORIZONTAL OUTER BOUNDARIES ARE SUBSTANTIALLY DIFFERENT
1010
GENERATE VERTICAL PANNING SIGNAL IF DISTANCE BETWEEN MIDPOINT AND VERTICAL OUTER BOUNDARIES ARE SUBSTANTIALLY DIFFERENT

FIGURE 10

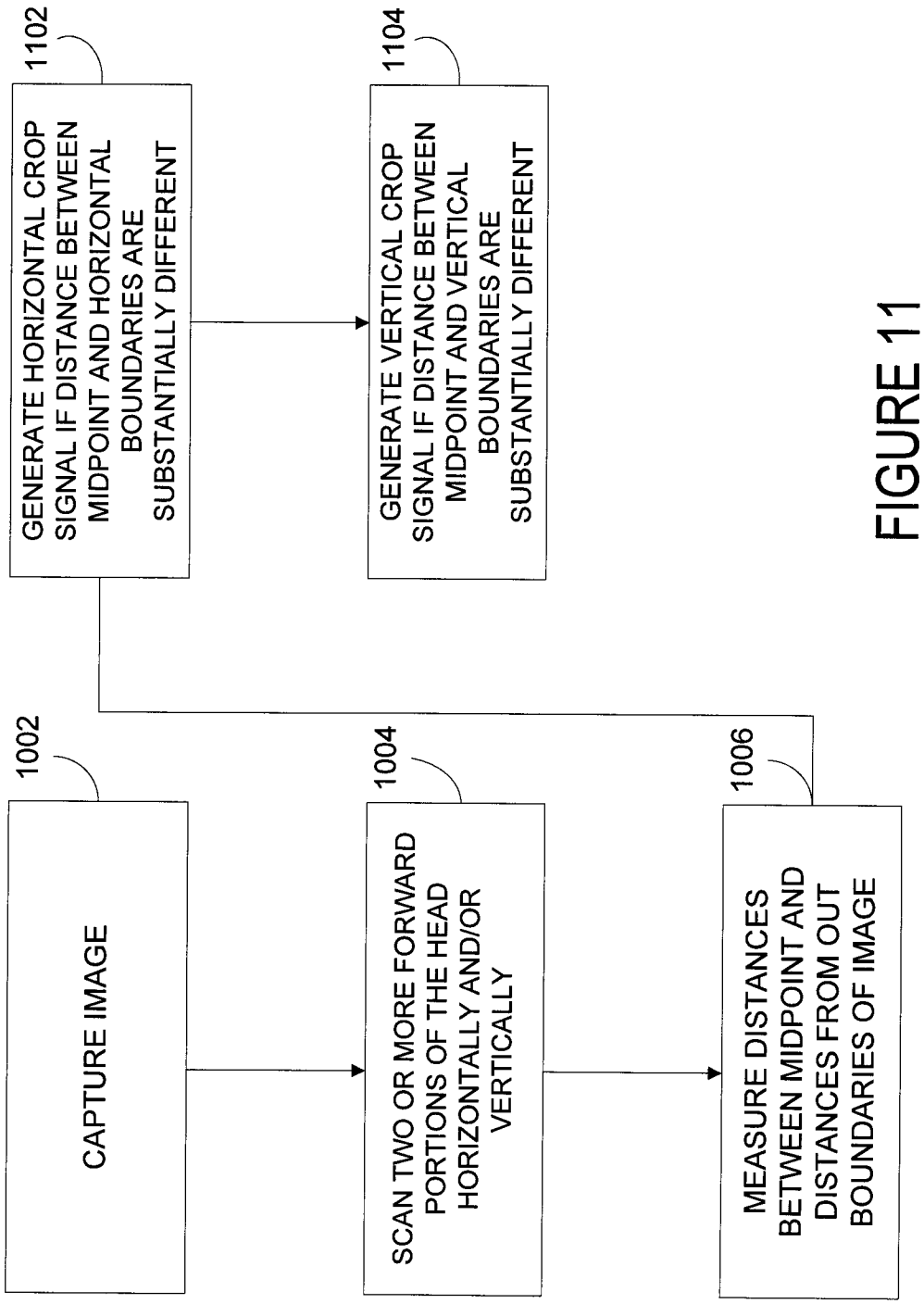
1102
GENERATE HORIZONTAL CROP SIGNAL IF DISTANCE BETWEEN MIDPOINT AND HORIZONTAL BOUNDARIES ARE SUBSTANTIALLY DIFFERENT
1104
GENERATE VERTICAL CROP SIGNAL IF DISTANCE BETWEEN MIDPOINT AND VERTICAL BOUNDARIES ARE SUBSTANTIALLY DIFFERENT
1002
CAPTURE IMAGE
1004
SCAN TWO OR MORE FORWARD PORTIONS OF THE HEAD HORIZONTALLY AND/OR VERTICALLY
1006
MEASURE DISTANCES BETWEEN MIDPOINT AND DISTANCES FROM OUT BOUNDARIES OF IMAGE

FIGURE 11

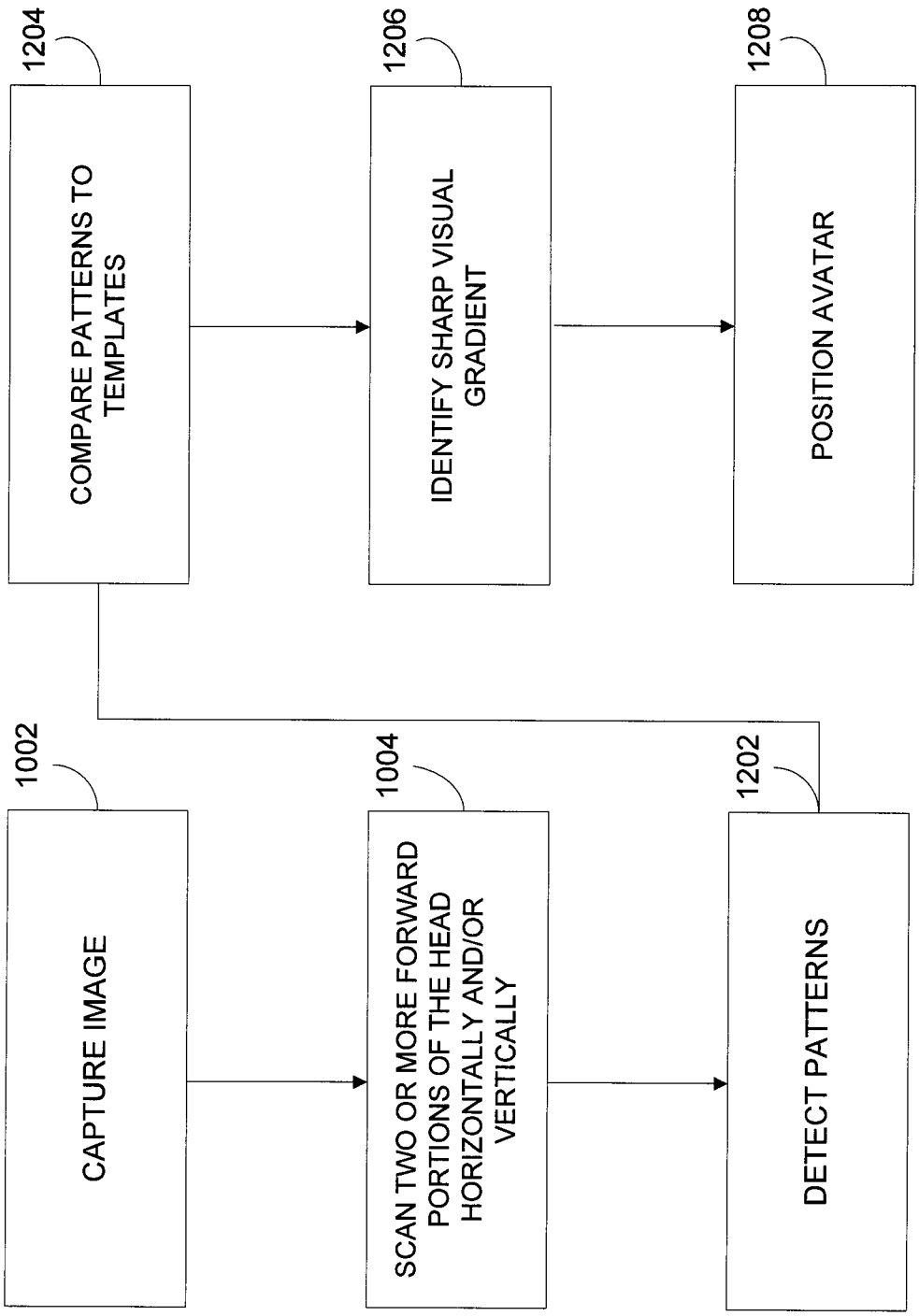
1002
CAPTURE IMAGE
1004
SCAN TWO OR MORE FORWARD PORTIONS OF THE HEAD HORIZONTALLY AND/OR VERTICALLY
1202
DETECT PATTERNS
1204
COMPARE PATTERNS TO TEMPLATES
1206
IDENTIFY SHARP VISUAL GRADIENT
1208
POSITION AVATAR

FIGURE 12

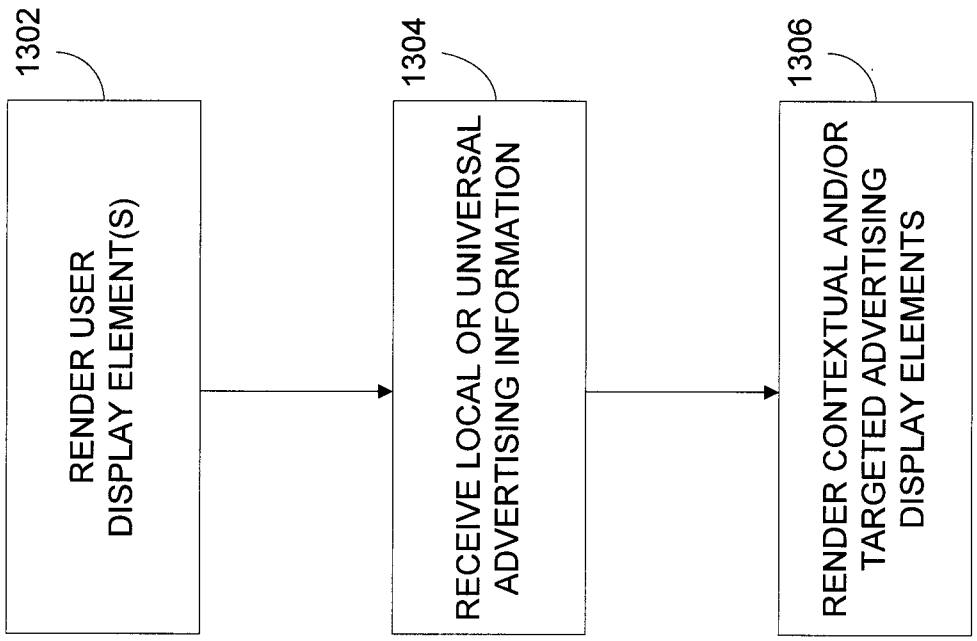
1302
RENDER USER DISPLAY ELEMENT(S)
1304
RECEIVE LOCAL OR UNIVERSAL ADVERTISING INFORMATION
1306
RENDER CONTEXTUAL AND/OR TARGETED ADVERTISING DISPLAY ELEMENTS

FIGURE 13

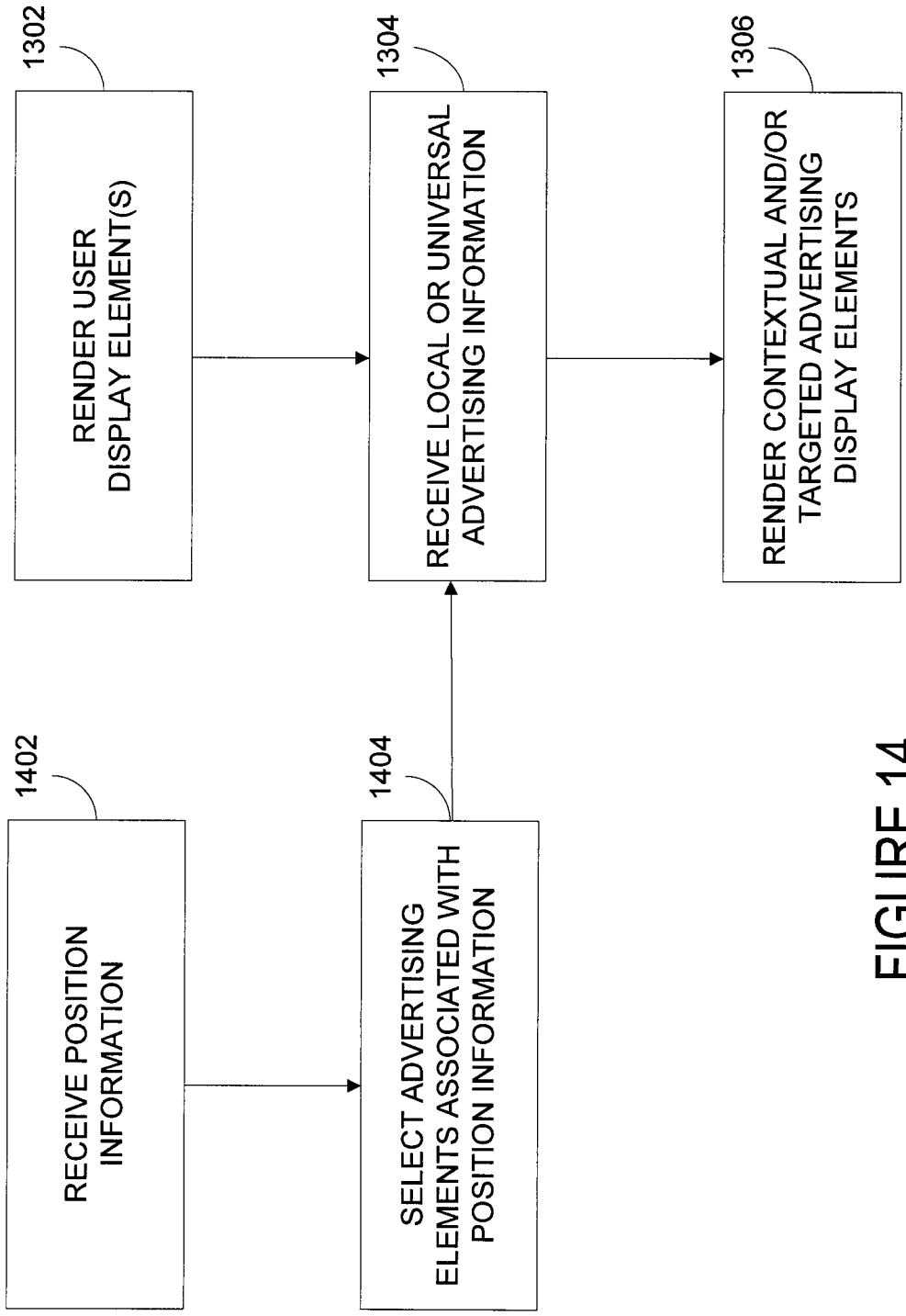
1302
RENDER USER DISPLAY ELEMENT(S)
1304
RECEIVE LOCAL OR UNIVERSAL ADVERTISING INFORMATION
1306
RENDER CONTEXTUAL AND/OR TARGETED ADVERTISING DISPLAY ELEMENTS
1402
RECEIVE POSITION INFORMATION
1404
SELECT ADVERTISING ELEMENTS ASSOCIATED WITH POSITION INFORMATION

FIGURE 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2008/001440

A. CLASSIFICATION OF SUBJECT MATTER
IPC: ***H04N 7/15*** (2006.01) , ***G06F 3/14*** (2006.01) , ***G06Q 30/00*** (2006.01) , ***H04L 29/02*** (2006.01) , ***H04M 3/22*** (2006.01) , ***H04M 3/56*** (2006.01) , ***H04N 7/14*** (2006.01)
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)
IPC: H04N, G06F, G06Q, H04L, H04M

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

Electronic database(s) consulted during the international search (name of database(s) and, where practicable, search terms used)
Databases: Delphion, ECLA, Canadian Patent Database
Keywords: channel, bandwidth, priority, data, stream, demultiplexer, modify, characteristic, rate

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 02/052869 A2 (Satt et al.) 4 July 2002 (04-07-2002) entire document	1-20
A	US 2005/0021621 A1 (Welch et al.) 27 January 2005 (27-01-2005) entire document	1-20
A	US 2003/0189943 A1 (Gorti et al.) 9 October 2003 (09-10-2003) entire document	1-20
A	WO 2004/084006 A2 (Iwamura) 30 September 2004 (30-09-2004) entire document	1-20

[ ] Further documents are listed in the continuation of Box C.

[X] 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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 October 2008 (23-10-2008)	17 November 2008 (17-11-2008)

Name and mailing address of the ISA/CA	Authorized officer
Canadian Intellectual Property Office Place du Portage I, C114 - 1st Floor, Box PCT 50 Victoria Street Gatineau, Quebec K1A 0C9 Facsimile No.: 001-819-953-2476	S. Choudhury 819- 934-6751

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CA2008/001440

Patent Document Cited in Search Report	Publication Date	Patent Family Member(s)	Publication Date
WO02052869	04-07-2002	AU2002216279 A1	08-07-2002
		EP1413151 A2	28-04-2004
		US2004248583 A1	09-12-2004
US2005021621	27-01-2005	AU5321300 A	18-12-2000
		AU5596000 A	18-12-2000
		DE60008102D D1	11-03-2004
		DE60008102T T2	23-12-2004
		DE60028967D D1	03-08-2006
		DE60028967T T2	15-02-2007
		EP1183820 A1	06-03-2002
		EP1183834 A1	06-03-2002
		JP2003501881T T	14-01-2003
		JP2003501886T T	14-01-2003
		US6735633 B1	11-05-2004
		US6850987 B1	01-02-2005
		US2004139150 A1	15-07-2004
		WO0074312 A1	07-12-2000
		WO0074322 A1	07-12-2000
US2003189943	09-10-2003	US7385997 B2	10-06-2008
		US2007081456 A1	12-04-2007
		US2008232386 A1	25-09-2008
WO2004084006	30-09-2004	US6947409 B2	20-09-2005
		US2006034330 A1	16-02-2006