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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 January 2014 (16.01.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/011709 A2

(51) International Patent Classification:
H04N 7/15 (2006.01)

(21) International Application Number:
PCT/US2013/049821

(22) International Filing Date:
10 July 2013 (10.07.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
13/548,438 13 July 2012 (13.07.2012) US

(71) Applicant: **MICROSOFT CORPORATION** [US/US]; One Microsoft Way, Redmond, Washington 98052-6399 (US).

(72) Inventors: **TANDON, Ankit**; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). **AICHNER, Robert**; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). **CUTLER, Ross**; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). **VELAYUTHAM, Senthil**; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (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, 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, 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: VIDEO DISPLAY MODIFICATION FOR VIDEO CONFERENCING ENVIRONMENTS

(57) Abstract: Provided herein are systems, methods, and software for facilitating a video conference environment. In at least one implementation, video captured of a scene is received. A modification to display of the video is identified based at least in part on a share of a display space associated with an object in the scene relative to a share of another display space associated with another object in another scene. The modification to the display of the video is then initiated. In another implementation, the modification may also be based on a target share indicated by way of an interactive graphic.

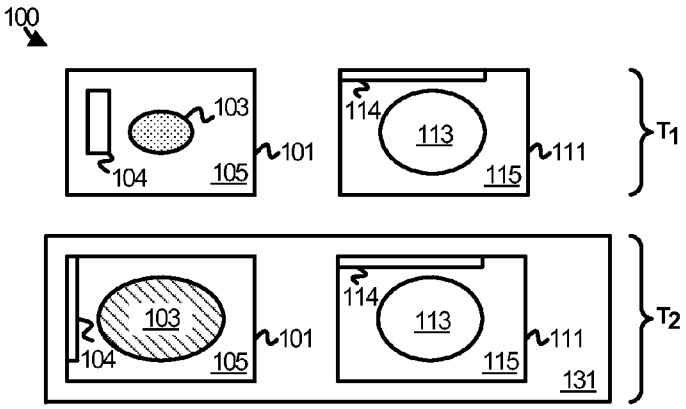


FIGURE 1

WO 2014/011709 A2

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

VIDEO DISPLAY MODIFICATION FOR VIDEO CONFERENCING ENVIRONMENTS

TECHNICAL BACKGROUND

[0001] Video conferencing involves the interconnection of two or more parties to communicate via an exchange of video. A wide variety of conferencing applications have developed to allow for video conferences that are easy to establish and manage. Participants can join prescheduled conferences but can also initiate impromptu conferences. Most conferencing applications allow participants to connect with each other using phone numbers, email addresses, and service handles, as well as many other suitable mechanisms.

[0002] As the feasibility of video conferencing has increased, so too have the ways and environments in which it can be delivered. For example, conference participants may engage in a video session using traditional desktop or laptop computers, as well as tablets, mobile phones, gaming systems, dedicated conferencing systems, or any other suitable communication device. Different architectures can be employed to deliver video conferencing solutions, including centrally managed and peer to peer architectures.

[0003] Many video conferencing solutions display video of each participant within display windows arrayed within the application. The presentation of video within each display window may vary considerably depending upon a number of factors, including the capture resolution employed at the source, the data rate available for transmission of the captured video, and the display resolution available on the presenting device. Other factors include the relative proportion of the display space formed by the display windows that is occupied by the objects corresponding to the participants.

OVERVIEW

[0004] Provided herein are systems, methods, and software for facilitating a video conference environment. In at least one implementation, video captured of a scene is received. A modification to display of the video is identified based at least in part on a share of a display space associated with an object in the scene relative to a share of another display space associated with another object in another scene. The modification to the display of the video is then initiated. In another implementation, the modification may also be based on a target share indicated by way of an interactive graphic.

[0005] This Overview is provided to introduce a selection of concepts in a simplified form that are further described below in the Technical Disclosure. It should be understood

that this Overview is not intended to identify key features or essential features of the claimed subject matter, nor is it intended to be used to limit the scope of the claimed subject matter.

BRIEF DESCRIPTION OF THE DRAWINGS

[0006] Many aspects of the disclosure can be better understood with reference to the following drawings. While several implementations are described in connection with these drawings, the disclosure is not limited to the implementations disclosed herein. On the contrary, the intent is to cover all alternatives, modifications, and equivalents.

[0007] Figure 1 illustrates a video conference scenario in an implementation.

[0008] Figure 2 illustrates a video modification process in an implementation.

[0009] Figure 3 illustrates a computing system in an implementation.

[0010] Figure 4 illustrates a video conference architecture in an implementation.

[0011] Figure 5A illustrates video captured for display within a video conference environment in an implementation.

[0012] Figure 5B illustrates video captured for display within a video conference environment and modified based on a target share in an implementation.

[0013] Figure 6 illustrates two instances of a video conference environment with video modified based on a target share in an implementation.

[0014] Figure 7 illustrates two instances of a video conference environment with video modified based on a target share in an implementation.

[0015] Figure 8 illustrates several views of video as initially captured and then modified for display within a video conference environment in an implementation.

[0016] Figure 9 illustrates video conference operations in an implementation.

[0017] Figure 10 illustrates video conference operations in an implementation.

[0018] Figure 11 illustrates video conference operations in an implementation.

TECHNICAL DISCLOSURE

[0019] Implementations described herein provide for improved video conferencing. In at least one implementation, video is captured of participants in a video conference. The captured video includes objects corresponding to the participants. The share of a display space occupied by each object may be analyzed to determine a modification to make to the video such that each object is generally the same size as the other objects. In other words, the share of each display space occupied by each object may generally be the same as the share occupied by each other object. With respect to the visual aesthetics of the video

conference, at least some of the conference participants may be represented in visually proportional ways.

[0020] Figures 1-11, discussed in more detail below, generally depict various scenarios, systems, processes, architectures, and operational sequences for carrying out various implementations. With respect to Figures 1-3, Figure 1 illustrates a video conference scenario in which a video modification process, illustrated in Figure 2, is employed. Figure 3 illustrates a computing system for carrying out the video modification process.

[0021] With respect to Figures 4-11, Figure 4 illustrates a video conference architecture, while Figure 5A and Figure 5B illustrate video captured by elements of the video conference architecture. Figure 6 and Figure 7 each illustrate two instances of video conference environments rendered by video conference applications running on elements of the video conference architecture. Figure 8 illustrates several views of the video provided within the video conference environments. Lastly, Figure 9, Figure 10, and Figure 11 illustrate various operational sequences possibly implemented with respect to the video conference architecture.

[0022] Turning now to Figure 1, video conference scenario 100 illustrates an operational scenario whereby conference video is handled according to modification process 200. The following discussion of Figure 1 will therefore proceed with reference to the steps of modification process 200 illustrated in Figure 2.

[0023] Initially (at time T1) video is received of objects corresponding to conference participants (step 201). In particular, video 101 includes object 103 corresponding to one participant, while video 111 includes object 113 corresponding to another conference participant. Video 101 also includes object 104 and video 111 includes object 114. Object 104 and object 114 are optional and are illustrated merely for exemplary purposes. The objects may be, for example, head shapes, body shapes, faces, or some other object corresponding to a participant.

[0024] Object 103 occupies a share of display space 105. For example, the share of display space 105 occupied by object 103 may be considered in terms of a ratio of the size of object 103 to the size of display space 105. The size of display space 105 may correspond to the size of the underlying frames of video 101, the size of a viewing window within which video 101 may be played, or some other measure or dimension against which the size of object 103 may be compared. Object 104 also occupies a share of display space 105.

[0025] Object 113 occupies a share of display space 115. For example, the share of display space 115 occupied by object 113 may be considered in terms of a ratio of the size of object 113 to the size of display space 115. The size of display space 115 may correspond to the size of the underlying frames of video 111, the size of a viewing window within which video 111 may be played, or some other measure or dimension against which the size of object 113 may be compared. Object 114 also occupies a share of display space 115.

[0026] Next (at time T2) a modification to at least video 101 is identified based at least in part on the share of display space 105 occupied by object 103 relative to the share of display space 115 occupied by object 113 (step 203). Upon identifying the modification, the modification to video 101 is initiated (step 205). Video 101 may be displayed within video conference environment 131 in its modified form. Video 111 may also displayed within video conference environment 131.

[0027] The modification performed on objects may be carried out in a number of ways. In one implementation, a digital zoom in or out may be performed on the video such that the dimension of an object within the video, relative to a display space, changes. In some implementations, an instruction may be provided to a capture device instructing the capture device to zoom in or out on a scene, thereby changing the relative size of an object to the remainder of a captured scene or display space. A variety of other mechanisms may be possible and should be considered within the scope of this disclosure.

[0028] In this scenario, the modification made to video 101 results in an increase in the share of display space 105 occupied by object 103. However, it should be understood that the modification may result in a decrease in the share of display space 105 occupied by object 103. The increase in the share of display space occupied by object 103 is evident in the larger size of object 103 at time T2 relative to the initial size of object 103 at time T1. In addition, assuming that the size of display space 105 is unchanged, the entirety of object 104 is no longer visible or otherwise completely contained within display space 105. Thus, the share of display space 105 occupied by both object 103 and object 104 has increased.

[0029] Also evident from Figure 1 is that, from time T1 to time T2, the size of object 103 relative to the size of object 113 has changed. Namely, whereas object 103 and object 105 appeared with considerably different sizes initially, upon modifying video 101, the size object 103 appears generally equal to or substantially similar to the size of object 113.

At the very least, upon modifying video 101, the difference in the size of object 103 with respect to object 113 has decreased.

[0030] Optionally, video 111 may also be analyzed and modified. However, in this scenario it is assumed for exemplary purposes that video 111 is not modified, or at least not modified substantially. Thus, the size of object 113 and object 114 at time T2 relative to their respective sizes at time T1 remain the same.

[0031] From the perspective of a conference participant viewing video conference environment 131, object 103 and object 113 will appear as similarly sized objects. In contrast, had no modification been performed, object 103 and object 113 would appear as differently sized objects, even though both correspond to conference participants. Thus, modification process 200, as described with respect to video conference scenario 100, achieves an aesthetic benefit whereby object corresponding to conference participants appear to a viewer as similarly sized objects.

[0032] Referring now to Figure 3, a computing system suitable for implementing a video modification process is illustrated. Computing system 300 is generally representative of any computing system or systems on which modification process 200 may be suitably implemented. Optionally, computing system 300 may also be suitable for implementing video conference environment 131. Examples of computing system 300 include server computers, client computers, virtual machines, distributed computing systems, personal computers, mobile computers, media devices, Internet appliances, desktop computers, laptop computers, tablet computers, notebook computers, mobile phones, smart phones, gaming devices, and personal digital assistants, as well as any combination or variation thereof.

[0033] Computing system 300 includes processing system 301, storage system 303, software 305, and communication interface 307. Computing system 300 also includes user interface 309, although this is optional. Processing system 301 is operatively coupled with storage system 303, communication interface 307, and user interface 309. Processing system 301 loads and executes software 305 from storage system 303. When executed by computing system 300 in general, and processing system 301 in particular, software 305 directs computing system 300 to operate as described herein for video modification process 200. Computing system 300 may optionally include additional devices, features, or functionality not discussed here for purposes of brevity and clarity.

[0034] Referring still to Figure 3, processing system 301 may comprise a microprocessor and other circuitry that retrieves and executes software 305 from storage

system 303. Processing system 301 may be implemented within a single processing device but may also be distributed across multiple processing devices or sub-systems that cooperate in executing program instructions. Examples of processing system 301 include general purpose central processing units, application specific processors, and logic devices, as well as any other type of processing device, combinations of processing devices, or variations thereof.

[0035] Storage system 303 may comprise any storage media readable by processing system 301 and capable of storing software 305. Storage system 303 may include volatile and nonvolatile, removable and non-removable media implemented in any method or technology for storage of information, such as computer readable instructions, data structures, program modules, or other data. Storage system 303 may be implemented as a single storage device but may also be implemented across multiple storage devices or sub-systems. Storage system 303 may comprise additional elements, such as a controller, capable of communicating with processing system 301.

[0036] Examples of storage media include random access memory, read only memory, magnetic disks, optical disks, flash memory, virtual memory, and non-virtual memory, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to store the desired information and that may be accessed by an instruction execution system, as well as any combination or variation thereof, or any other type of storage media. In some implementations, the storage media may be a non-transitory storage media. In some implementations, at least a portion of the storage media may be transitory. It should be understood that in no case is the storage media a propagated signal.

[0037] Software 305 may be implemented in program instructions and among other functions may, when executed by computing system 300, direct computing system 300 to receive video captured of an object in a scene corresponding to a participant in a video conference, identify a modification to a display of the video within a video conference environment based on a share of a display space associated with the object relative to a share of another display space associated with another object corresponding to another participant in the video conference, and initiate the modification to the display of the video within the video conference environment.

[0038] Software 305 may include additional processes, programs, or components, such as operating system software or other application software. Software 305 may also

comprise firmware or some other form of machine-readable processing instructions capable of being executed by processing system 301.

[0039] In general, software 305 may, when loaded into processing system 301 and executed, transform processing system 301, and computing system 300 overall, from a general-purpose computing system into a special-purpose computing system customized to facilitate a video conference environment as described herein for each implementation. Indeed, encoding software 305 on storage system 303 may transform the physical structure of storage system 303. The specific transformation of the physical structure may depend on various factors in different implementations of this description. Examples of such factors may include, but are not limited to the technology used to implement the storage media of storage system 303 and whether the computer-storage media are characterized as primary or secondary storage.

[0040] For example, if the computer-storage media are implemented as semiconductor-based memory, software 305 may transform the physical state of the semiconductor memory when the program is encoded therein. For example, software 305 may transform the state of transistors, capacitors, or other discrete circuit elements constituting the semiconductor memory. A similar transformation may occur with respect to magnetic or optical media. Other transformations of physical media are possible without departing from the scope of the present description, with the foregoing examples provided only to facilitate this discussion.

[0041] It should be understood that computing system 300 is generally intended to represent a computing system with which software 305 is deployed and executed in order to implement modification process 200 and optionally render video conference environment 131. However, computing system 300 may also represent any computing system on which software 305 may be staged and from where software 305 may be distributed, transported, downloaded, or otherwise provided to yet another computing system for deployment and execution, or yet additional distribution.

[0042] Referring again to Figure 1, through the operation of computing system 300 employing software 305, transformations may be performed with respect to video 101 and video 111. As an example, video 101 could be considered transformed from one state to another when subject to video modification process 200. In a first state, object 103 occupies a certain share of display space 105. Upon modifying video 101, object 103 occupies a different share of display space 105, thereby changing video 101 to a second, different state.

[0043] Referring again to Figure 3, communication interface 307 may include communication connections and devices that allow for communication between computing system 300 other computing systems not shown over a communication network or collection of networks. Examples of connections and devices that together allow for inter-system communication include network interface cards, antennas, power amplifiers, RF circuitry, transceivers, and other communication circuitry. The aforementioned network, connections, and devices are well known and need not be discussed at length here.

[0044] User interface 309 may include a mouse, a voice input device, a touch input device for receiving a gesture from a user, a motion input device for detecting non-touch gestures and other motions by a user, and other comparable input devices and associated processing elements capable of receiving user input from a user, such as a camera or other video capture device. Output devices such as a display, speakers, printer, haptic devices, and other types of output devices may also be included in user interface 309. The aforementioned user input devices are well known in the art and need not be discussed at length here. User interface 309 may also include associated user interface software executable by processing system in support of the various user input and output devices discussed above. Separately or in conjunction with each other and other hardware and software elements, the user interface software and devices may be considered to provide a graphical user interface, a natural user interface, or any other kind of user interface.

[0045] Figure 4 illustrates video capture architecture 400 in an implementation. Video capture architecture 400 includes client device 411, client device 413, client device 415, client device 417, and conference system 431. Client devices 411, 413, 415, and 417 may exchange at least some conference communications with each other via conference system 431 over communication network 430. In this implementation, client devices 411, 413, 415, and 417 are operated by or otherwise associated with users 401, 403, 405, and 407 respectively to engage in a video conference with each other. Towards that end, each client device 411, 413, 415, and 417 includes running thereon video conference applications 421, 423, 425, and 427 respectively.

[0046] Client devices 411, 413, 415, and 417 may each be any type of computing system capable of hosting a video conference application. The video conference application may be locally installed and executed, hosted remotely and accessibly through a web browser, or even streamed for local execution, as well as provisioned or delivered according to any number of other techniques, combinations of techniques, or variations thereof. Examples of client devices include desktop or laptop computers, tablets, mobile

phones, gaming systems, dedicated conferencing systems, or any other suitable communication device.

[0047] Conference system 431 may be any system of collection of systems capable of hosting a video conference service that allows for participation in video conferences by client devices. While conference system 431 is illustrated as a stand-alone of centralized system, it should be understood that conference system 431 may be distributed among multiple computing systems. In some implementations, the functionality provided by conference system 431 may be distributed amongst peers, allowing the video conference service to be provisioned according to a peer-to-peer model. Other combinations or variations thereof are possible.

[0048] Communication network 430 may be any network or collection of networks capable of carrying video conference communications. Examples of communication network 430 include internets, intranets, local area networks, wide area networks, wireless networks, and wired networks, as well as any combination or variation thereof.

[0049] In operation, users may arrange to participate in a video conference with each other. In some cases, the users may initiate invitations, exchange links, dial each other, or in some other manner notify each other of a desire to participate in a conference. Many well- known systems and software provide such conference capabilities and need not be discussed at length here. Once the conference is established, video is captured of the participants by their respective video capture devices and provide to conference system 431 for distribution to the other conference participants.

[0050] One particular example video conference involving users 401, 403, 405, and 407 is illustrated in Figures 5-11. In this example, the video conference is generally described from the perspective of user 407 and client device 417. For example Figure 5A illustrates video that may be received by client device 417 and presented to user 407 by way of video conference application 427. Figure 6 and Figure 7 illustrate video conference environments that may be generated by video conference application 427, within which the conference video is displayed. Figure 8 illustrates in more detail the modification of video within a video conference environment. Figures 9-11 illustrates several ways in which video may be exchanged and handled in order to provide the video conference.

[0051] Referring now to Figure 5A, video 511, 513, and 515 is illustrated. Video 511, 513, and 515 is representative of video captured by client devices 411, 413, and 415 illustrated in Figure 4, and communicated to client device 417 by way of conference system 431. Video 511 includes object 501, corresponding to user 401. User 401 is a

participant in the video conference with users 403, 405, and 407. Video 511 also includes object 531. Object 501 and object 531 each occupy a share of display space 521. Video 513 includes object 503, corresponding to user 403. User 403 is also a participant in the video conference with users 401, 405, and 407. Video 513 also includes object 533. Object 503 and object 533 each occupy a share of display space 523. Video 515 includes object 505, corresponding to user 405. User 405 is another participant in the video conference with users 401, 403, and 407. Object 505 occupies a share of display space 525.

[0052] As will be discussed in more detail with respect to Figures 6-8, video 511, 513, and 515 may be processed and rendered within a video conference environment. Specifically, client device 417 executing video conference application 427 processes and potentially modifies portions of video 511, 513, and 515 for display within a video conference environment, such as video conference environment 641 and video conference environment 741 discussed in more detail below. The video may be modified based on a target share of display share accorded to each object in the video. Figure 5B illustrates with particularity how an object may change based on variations in a target share.

[0053] Referring to Figure 5B, video 515 captured by client device 415 is illustrated for exemplary purposes. As target share 502 is increased or decreased, the size of object 505 changes with respect to display space 525. For example, at the upper most end of target share 502, object 505 has a vertical dimension nearly equal to the vertical dimension of displace space 525. In contrast, towards the lower most end of target share 502, object 505 has dimensions generally much smaller than the dimensions of display space 525. Note that object 535 is entirely visible within display space 525 towards the lower end of target share 502. When set towards the middle range of target share 502, object 505 is smaller than when the target share is set towards the upper end, but larger than when the target share is set towards the lower end. Note that object 535 is partially visible within display space 525 when the target share is set towards the middle range.

[0054] In other words, as target share 502 is adjusted, the share of display space 525 occupied by object 505 varies generally proportionally. Reducing target share 502 results in a reduced share of display space 525 occupied by object relative to when target share 502 is increased. It should be understood that target share 502 may be applied to video 511 and video 522 as well as video 515.

[0055] Figure 6 provides one such illustration of a video conference environment 641 whereby video may be modified based on a target share, as demonstrated in Figure 5B.

Namely, video may be modified such that the relative share of display space occupied by objects captured in the video varies depending upon how a target share is defined or set.

[0056] Video conference environment 641 includes video module 643, chat module 645, and white board module 647. Video conference environment 641 also includes interactive graphic 648 and interactive graphic 649, each representative of presentation context A and presentation context B respectively, that may be selected by a user input. Depending upon the target share selected or otherwise indicated by the user input, some modifications may be made to video 511, 513, and 515 such that objects 501, 503, and 505 each occupy a share of display spaces 521, 523, and 525 in accordance with the target share.

[0057] In this example, presentation context A and presentation context B correspond to different target shares selectable by the user input. In this example, presentation context A corresponds to a target share value less than the target share corresponding to presentation context B. Thus, when presentation context A is selected, the share of display spaces within video module 643 occupied by objects is less than the share of the display spaces occupied by objects when presentation context B is selected. Presentation context A may correspond to, for example, a wide presentation context where the display intent is to display a wide view of participants in the video conference. In contrast, presentation context B may correspond to, for example, a close-in presentation context where the display intent is to display a close-in view of the participants in the video conference.

[0058] To demonstrate how varying a target share may result in different modifications to video 511, 513, and 515, presentation context A is initially selected at time T1, while presentation context B is selected at time T2. When presentation context A is selected, the share of display space 521 occupied by object 501 is relatively unchanged compared to the share of display space 521 occupied by object 501 when initially captured in video 511, as illustrated in Figure 5A. However, the share of display space 523 occupied by object 503, upon selection of presentation context A, is reduced compared to when initially captured in video 513, as illustrated Figure 5A. Likewise, the share of display space 525 occupied by object 505, upon selection of presentation context A, is reduced compared to when initially captured in video 515, as also illustrated Figure 5A.

[0059] In contrast, upon selection of presentation context B at time T2, the share of display spaces 521, 523, and 525 occupied by objects 501, 503, and 505 respectively changes compared to when presentation context A is selected at time T1. The size of object 501 is larger than when presentation context A is selected, as is the size of object 503 and 505. In other words, the share of display spaces 521, 523, and 525 occupied by

objects 501, 503, and 505 at time T2 is greater than the share at time T1. Note also that, when presentation context b is selected, the share of display space 521 occupied by object 501 is increased compared to the share of display space 521 occupied by object 501 when initially captured in video 511, as illustrated in Figure 5A. The share of display space 523 occupied by object 503, upon selection of presentation context A, is generally unchanged compared to when initially captured in video 513, as also illustrated Figure 5A. The share of display space 525 occupied by object 505, upon selection of presentation context A, is reduced compared to when initially captured in video 515, as also illustrated Figure 5A.

[0060] Figure 7 provides another illustration of a video conference environment 741 whereby video may be modified based on a target share, as demonstrated in Figure 5B. Namely, video may be modified such that the relative share of display space occupied by objects captured in the video varies depending upon how a target share is defined or set.

[0061] Video conference environment 741 includes video module 743, chat module 745, and white board module 747. Video conference environment 741 also includes an interactive graphic comprised of a range graphic 748 and a selector graphic 749. Range graphic 748 is representative of a range of target share values that may be selected by a user input by moving or otherwise positioning selector graphic 749. Depending upon the target share selected or otherwise indicated by the user input, some modifications may be made to video 511, 513, and 515 such that objects 501, 503, and 505 each occupy a share of display spaces 521, 523, and 525 in accordance with the input target share.

[0062] In this example, selector graphic 749 may be slid or otherwise moved through the range of target share values represented by range graphic 748. In this example, the left most portion of range graphic 748 corresponds to a target share value less than the target share corresponding to the right most portion of range graphic 748. Thus, when selector graphic 749 is positioned toward the left of range graphic 748, the share of display spaces within video module 743 occupied by objects is less than the share of the display spaces occupied by objects when selector graphic 749 is positioned toward the right of range graphic 784. The target share values selectable on range graphic 748 may correspond to, for example, a wide presentation context where the display intent is to display a wide view of participants in the video conference or a close-in presentation context where the display intent is to display a close-in view of the participants in the video conference.

[0063] To demonstrate how varying a target share may result in different modifications to video 511, 513, and 515, selector graphic 749 is initially positioned toward the left end of range graphic 748 at time T1, selector graphic 749 is positioned toward the ride end of

range graphic 748 at time T2. At time T1, the share of display space 521 occupied by object 501 is relatively unchanged compared to the share of display space 521 occupied by object 501 when initially captured in video 511, as illustrated in Figure 5A. However, the share of display space 523 occupied by object 503 is reduced compared to when initially captured in video 513, as illustrated Figure 5A. Likewise, the share of display space 525 occupied by object 505 is reduced compared to when initially captured in video 515, as also illustrated Figure 5A.

[0064] In contrast, upon positioning selector graphic 749 toward the right end of range graphic 748 at time T2, the share of display spaces 521, 523, and 525 occupied by objects 501, 503, and 505 respectively changes compared to the scenario illustrated with respect to time T1. The size of object 501 is larger at time T1 than at time T2, as is the size of object 503 and object 505. In other words, the share of display spaces 521, 523, and 525 occupied by objects 501, 503, and 505 at time T2 is greater than the share at time T1. Note also that at time T2 the share of display space 521 occupied by object 501 is increased compared to the share of display space 521 occupied by object 501 when initially captured in video 511, as illustrated in Figure 5A. The share of display space 523 occupied by object 503 is generally unchanged compared to when initially captured in video 513, as also illustrated Figure 5A. The share of display space 525 occupied by object 505 is reduced compared to when initially captured in video 515, as also illustrated Figure 5A.

[0065] Figure 8 illustrates view 801 and view 802 of video 511 and 513 respectively as initially captured by client device 411 and client device 413 respectively. Figure 8 also illustrates view 803 of video 511 and video 513 as initially displayed within a video conference environment. Finally, Figure 8 illustrates view 804 of video 511 and video 533 as displayed upon identifying an initiating a modification to at least a portion of the video. It should be understood that the pixel dimensions provided in Figure 8 are merely illustrative and could vary in any implementation.

[0066] Referring to view 801, video 511 is defined or characterized in terms of pixel dimensions. Video 511 includes 798 horizontal pixels and 600 vertical pixels. Object 501 is approximately 266 pixels wide. Thus, the total horizontal dimension of video 511 includes the 266 pixels of object 501 plus two 266 pixel sections to the left and right. In terms of a simple ratio, object 501 occupies approximately 1/3 of display space 521.

[0067] Referring to view 802, video 513 is also defined or characterized in terms of pixel dimensions. Video 513 includes 798 horizontal pixels and 600 vertical pixels,

similar to video 511. Object 503 is approximately 534 pixels wide when initially captured, or twice as large as object 50`. Thus, the total horizontal dimension of video 513 includes the 534 pixels of object 503 plus two 132 pixel sections to either side. In terms of a simple ratio, object 503 occupies approximately 2/3 of display space 521.

[0068] Referring now to view 803, illustrated are video 511 and video 513 upon being transmitted by client device 411 and client device 413 respectively, to possible conference system 431 or another client device, such as client device 417. Due to data transmission constraints or some other factor, the total pixel size count of video 511 and video 513 has been reduced relative to their respective pixel counts at the time of capture. Video 511 now has just 120 horizontal pixels. Likewise, video 513 now has just 120 pixels. To maintain scale, object 501 is rendered with 40 pixels and is therefore approximately 1/3 the horizontal width of video 511. Object 503 is rendered with 80 pixels or 2/3 the width of video 513. Note that object 501 is approximately half the size of object 504 in view 803.

[0069] Referring now to view 804, video 511 and video 513 are illustrated upon modification to video 511. In particular, video 511 has been modified based on the share of display space 521 occupied by object 501 relative to the share of display space 523 occupied by object 503 in view 803. In view 803, object 501 occupies approximately 1/3 of display space 521, whereas object 503 occupies 2/3 of display space 523. It is therefore determined that the share of display space 521 occupied by object 501 be increased in order to substantially equalize the relative shares.

[0070] As can be seen from view 804, the size of object 501 is increased such that its horizontal pixel count is now 80 pixels, or approximately equal to the horizontal pixel count of object 503. Note that the fill characteristic of object 501 changes from view 803 to view 804. This is intended to demonstrate that the resolution of object 501 may decrease from view 803 to view 804 as its size is expanded.

[0071] Figure 9 illustrates an operational sequence with respect to video conference architecture 400 illustrated in Figure 4. Video conference application 421, executing on client device 411, captures video and transmits the video to conference system 431. Similarly, video conference application 423 and video conference application 425, executing on client device 413 and client device 415 respectively, capture video and transmit the video to conference system 431. The video captured by video conference application 421 includes at least object 501; the video captured by video conference

application 423 includes at least object 503; and the video captured by conference application 425 includes at least object 505.

[0072] In this implementation, conference system 431 provides the video, possibly as separate video streams or files or as an integrated video stream or file, to video capture application 427 executing on client device 417. It should be understood that conference system may provide the video or portions thereof to any or all of the other video capture applications 421, 423, and 425. It should be understood that video conference application 427 may also capture video and provide its captured video to conference system 431 for distribution to video conference applications 421, 423, and 425.

[0073] Upon receiving the video from conference system 431, video conference application 427 applies a modification process, such as modification process 200, so that objects 501, 503, and 505 may have similar sizes and proportions relative to display spaces associated with the objects when rendered and displayed within a video conference environment. As discussed previously, the modification process may result in increases or decreases to some or all of objects 501, 503, and 505.

[0074] Figure 10 illustrates another operational sequence with respect to video conference architecture 400 illustrated in Figure 4. Video conference application 421, executing on client device 411, captures video and transmits the video to conference system 431. Similarly, video conference application 423 and video conference application 425, executing on client device 413 and client device 415 respectively, capture video and transmit the video to conference system 431. The video captured by video conference application 421 includes at least object 501; the video captured by video conference application 423 includes at least object 503; and the video captured by conference application 425 includes at least object 505.

[0075] In this implementation, conference system 431 applies a modification process, such as modification process 200, so that objects 501, 503, and 505 may have similar sizes and proportions relative to display spaces associated with the objects when rendered and displayed within a video conference environment. As discussed previously, the modification process may result in increases or decreases to some or all of objects 501, 503, and 505.

[0076] Upon modifying the video, conference system 431 provides the video, possibly as separate video streams or files or as an integrated video stream or file, to video conference application 427 executing on client device 417. Video conference application 427 renders the video within a video conference environment. It should be understood

that conference system may provide the video or portions thereof to any or all of the other video capture applications 421, 423, and 425. It should also be understood that video conference application 427 may also capture video and provide its captured video to conference system 431 for distribution to video conference applications 421, 423, and 425.

[0077] Figure 12 illustrates yet another operational sequence with respect to video conference architecture 400 illustrated in Figure 4. Video conference application 421, executing on client device 411, captures video and transmits the video to conference system 431. Similarly, video conference application 423 and video conference application 425, executing on client device 413 and client device 415 respectively, capture video and transmit the video to conference system 431. The video captured by video conference application 421 includes at least object 501; the video captured by video conference application 423 includes at least object 503; and the video captured by video conference application 425 includes at least object 505.

[0078] In this implementation, conference system 431 applies a portion of modification process, such as modification process 200, to analyze whether or not, and to what extent, objects 501, 503, and 505 may be modified. As discussed previously, the modification process may result in increases or decreases to some or all of objects 501, 503, and 505. Conference system 431 determines to what extent each object 501, 503, and 505 should be modified, if at all, and communicates the video and a modification instruction to video conference application 427. Conference system 431 may provide the video possibly as separate video streams or files or as an integrated video stream or file to video conference application 427.

[0079] Upon receiving the video from conference system 431, video conference application 427 applies another portion of the modification process so that objects 501, 503, and 505 may have similar sizes and proportions relative to display spaces associated with the objects when rendered and displayed within a video conference environment. It should be understood that conference system 431 may provide the modification instruction and the video or portions thereof to any or all of the other video capture applications 421, 423, and 425. It should also be understood that video conference application 427 may also capture video and provide its captured video to conference system 431 for distribution to video conference applications 421, 423, and 425.

[0080] In a brief operational scenario, a process for facilitating a video conference environment includes receiving video captured of a first scene including a first object corresponding to a first participant in a video conference, identifying a modification to a

display of the video within the video conference environment based on display information including a first share of a first display space associated with the first object relative to a second share of a second display space associated with a second object corresponding to a second participant in the video conference, and initiating the modification to the display of the first video within the video conference environment.

[0081] In some implementations, the modification to the display of the first video includes a reduction in a difference between the first share of the first display space associated with the first object and the second share of the second display space associated with the second object. In some implementations, the modification to the display of the first video includes an increase in a size of the first object such that the first share of the first display space associated with the first object increases upon implementing the modification. In some implementations, the modification to the display of the first video includes a decrease in a size of the first object such that the first share of the first display space associated with the first object decreases upon implementing the modification.

[0082] In an implementation, the process may also include, prior to identifying the modification to the display of the video, receiving an indication of a preferred share of each of a plurality of display spaces to be associated with a plurality of objects corresponding to participants in the video conference and identifying an initial modification to the display of the video within the video conference environment based at least in part on an initial second share of the second display space associated with the second object and the preferred share. The process may also include initiating the initial modification to the display of the video within the video conference environment.

[0083] In yet another implementation, the process may include, prior to identifying the modification to the display of the video, receiving a selection of the second display space from multiple display spaces on which to base, at least in part, the identifying of the modification to the display of the video. The display information may be representative of minimum acceptable size criteria for the first object.

[0084] In another brief example, a video conference environment may be rendered by a computing system. The computing system receives video captured of scenes including objects corresponding to participants in the video conference. The computing system, executing a video conference application, generates and displays an interactive share graphic configured to receive an input representative of at least a target share of each of multiple display spaces to be occupied by each of the objects. The computing system modifies at least a portion of the video based on display information including the target

share and a share of each of the multiple display spaces occupied by each of the objects. Upon modifying at least the portion of the video, the computing system displays the video within the video conference environment.

[0085] In some implementations, the interactive share graphic may be a graphical representation of a first presentation context and a second presentation context, wherein the input may be a selection of one of the first presentation context and second presentation context. In some implementations, the first presentation context is a seated presentation context and the second presentation context is a standing presentation context.

[0086] The functional block diagrams, operational sequences, and flow diagrams provided in the Figures are representative of exemplary architectures, environments, and methodologies for performing novel aspects of the disclosure. While, for purposes of simplicity of explanation, the methodologies included herein may be in the form of a functional diagram, operational sequence, or flow diagram, and may be described as a series of acts, it is to be understood and appreciated that the methodologies are not limited by the order of acts, as some acts may, in accordance therewith, occur in a different order and/or concurrently with other acts from that shown and described herein. For example, those skilled in the art will understand and appreciate that a methodology could alternatively be represented as a series of interrelated states or events, such as in a state diagram. Moreover, not all acts illustrated in a methodology may be required for a novel implementation.

[0087] The included descriptions and figures depict specific implementations to teach those skilled in the art how to make and use the best mode. For the purpose of teaching inventive principles, some conventional aspects have been simplified or omitted. Those skilled in the art will appreciate variations from these implementations that fall within the scope of the invention. Those skilled in the art will also appreciate that the features described above can be combined in various ways to form multiple implementations. As a result, the invention is not limited to the specific implementations described above, but only by the claims and their equivalents.

CLAIMS

1. A method for facilitating a video conference environment comprising:

receiving video captured of a first scene comprising a first object corresponding to a first participant in a video conference;

identifying a modification to a display of the video within the video conference environment based on display information comprising a first share of a first display space associated with the first object relative to a second share of a second display space associated with a second object corresponding to a second participant in the video conference; and

initiating the modification to the display of the first video within the video conference environment.

2. The method of claim 1 wherein the modification to the display of the first video comprises a reduction in a difference between the first share of the first display space associated with the first object and the second share of the second display space associated with the second object.

3. The method of claim 2 wherein the modification to the display of the first video comprises an increase in a size of the first object such that the first share of the first display space associated with the first object increases upon implementing the modification.

4. The method of claim 2 wherein the modification to the display of the first video comprises a decrease in a size of the first object such that the first share of the first display space associated with the first object decreases upon implementing the modification.

5. The method of claim 1 wherein the method further comprises:

prior to identifying the modification to the display of the video, receiving an indication of a preferred share of each of a plurality of display spaces to be associated with a plurality of objects corresponding to a plurality of participants in the video conference; and

identifying an initial modification to the display of the video within the video conference environment based at least in part on an initial second share of the second display space associated with the second object and the preferred share;

initiating the initial modification to the display of the video within the video conference environment.

6. One or more computer readable media for facilitating a video conference environment, wherein the program instructions, when executed by a computing system, direct the computing system to at least:

receive video captured of a plurality of scenes comprising a plurality of objects corresponding to a plurality of participants in the video conference;

generate and display an interactive share graphic configured to receive an input representative of at least a target share of each of a plurality of display spaces to be occupied by each of the plurality of objects;

modify at least a portion of the video based on display information comprising the target share and a share of each of the plurality of display spaces occupied by each of the plurality of objects; and

upon modifying at least the portion of the video, displaying the video within the video conference environment.

7. The one or more computer readable media of claim 6 wherein the interactive share graphic comprises a graphical representation of a first presentation context and a second presentation context, wherein the input comprises a selection of one of the first presentation context and second presentation context.

8. The one or more computer readable media of claim 7 wherein the first presentation context comprises a seated presentation context and the second presentation context comprises a standing presentation context.

9. The one or more computer readable media of claim 6 wherein to modify at least the portion of the video, the program instructions direct the computing system to reduce a difference between the share of each of the plurality of display spaces occupied by each of the plurality of objects.

10. The one or more computer readable media of claim 6 wherein the display information further comprises minimum acceptable size criteria for the plurality of objects and wherein each of the plurality of objects comprises a head shape.

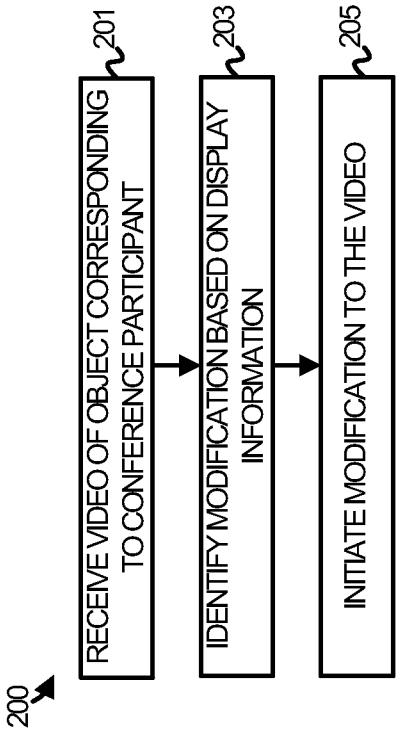


FIGURE 2

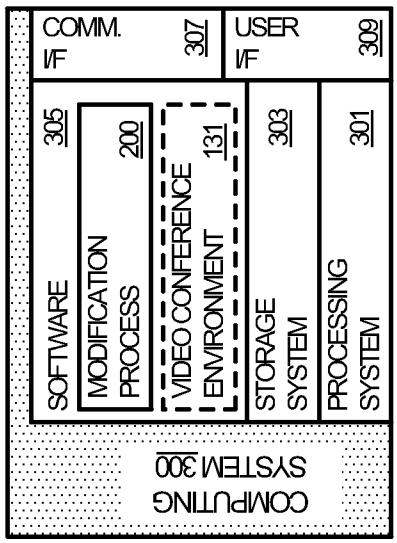


FIGURE 3

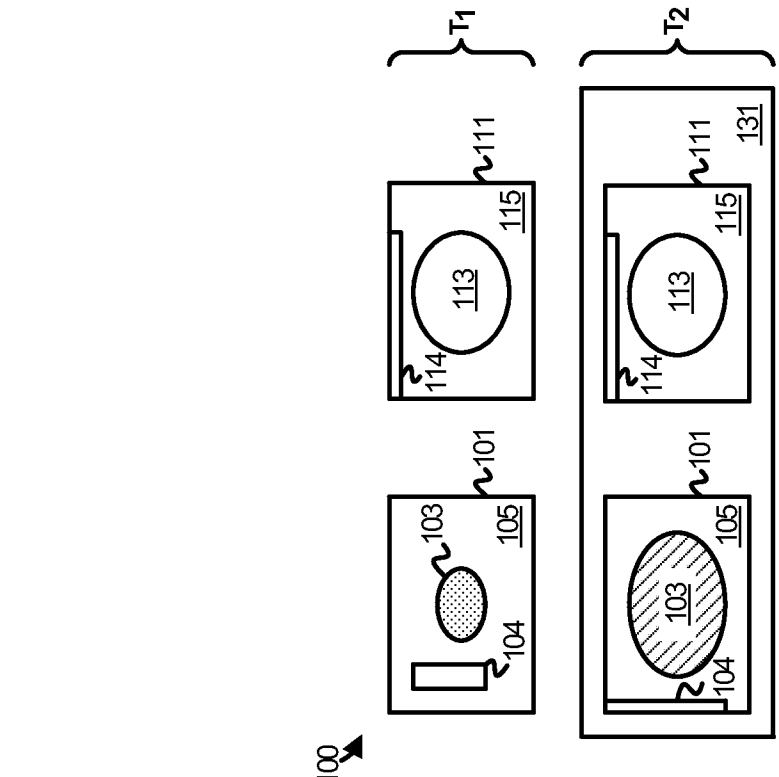


FIGURE 1

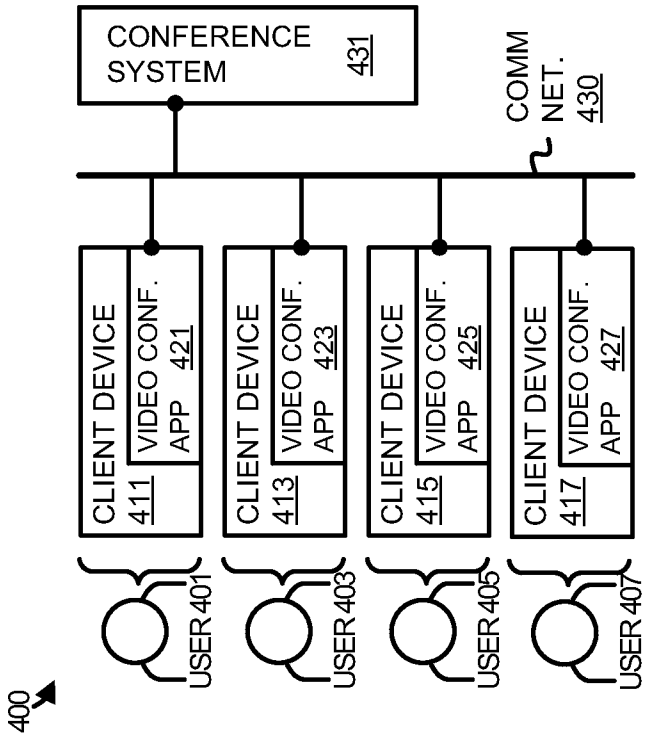

CONFERENCE SYSTEM 431
COMM NET. 430
CLIENT DEVICE 411
VIDEO CONF. APP 421
CLIENT DEVICE 413
VIDEO CONF. APP 423
CLIENT DEVICE 415
VIDEO CONF. APP 425
CLIENT DEVICE 417
VIDEO CONF. APP 427
USER 401
USER 403
USER 405
USER 407
400

FIGURE 4

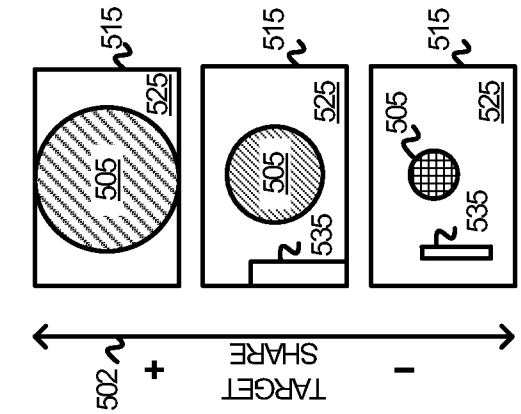


FIGURE 5B

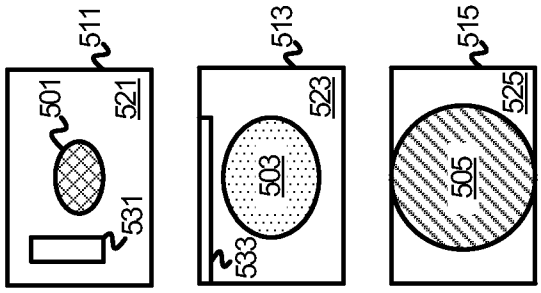


FIGURE 5A

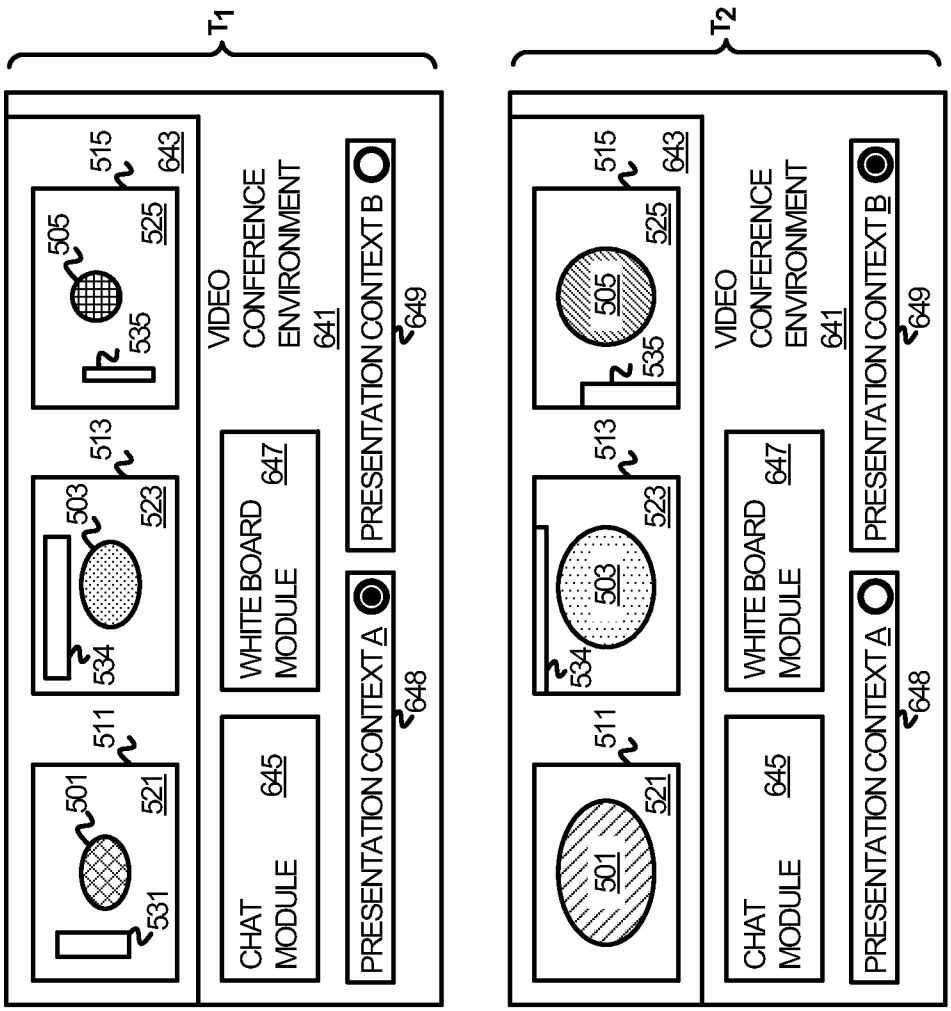

T1
501
511
521
531
503
513
523
534
505
515
525
535
643
VIDEO CONFERENCE ENVIRONMENT 641
WHITE BOARD MODULE 647
CHAT MODULE 645
PRESENTATION CONTEXT A 648
PRESENTATION CONTEXT B 649
T2
501
511
521
503
513
523
534
505
515
525
535
643
VIDEO CONFERENCE ENVIRONMENT 641
WHITE BOARD MODULE 647
CHAT MODULE 645
PRESENTATION CONTEXT A 648
PRESENTATION CONTEXT B 649

FIGURE 6

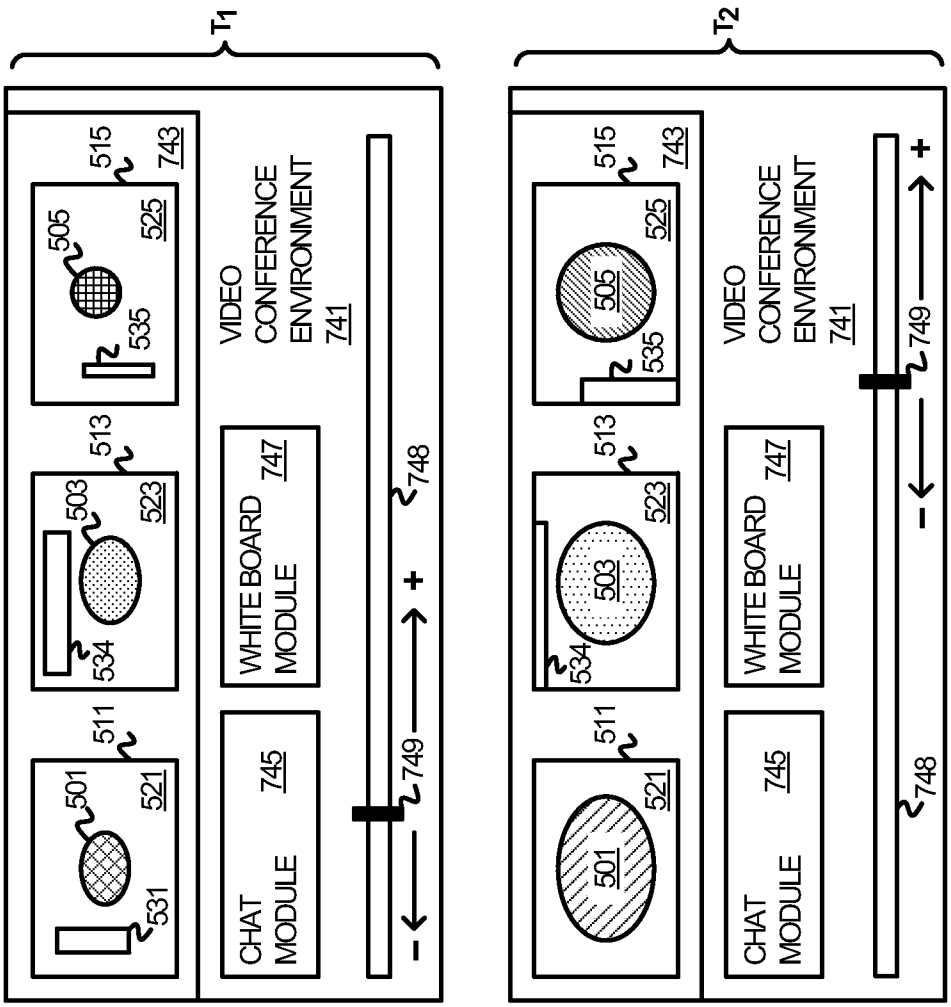

T1
T2
501
511
521
531
503
513
523
534
505
515
525
535
743
VIDEO CONFERENCE ENVIRONMENT 741
WHITE BOARD MODULE 747
CHAT MODULE 745
748
749
-
+

FIGURE 7

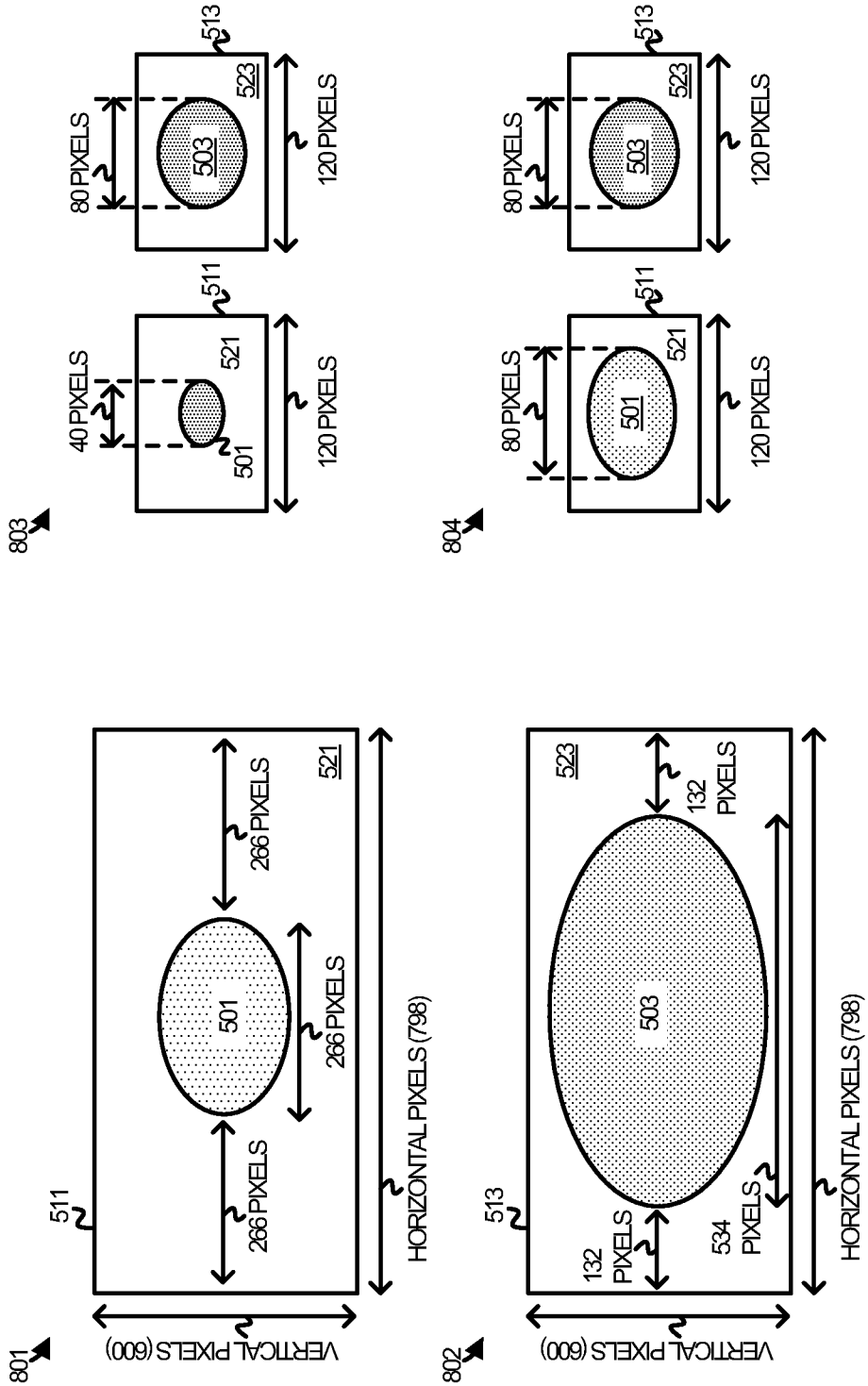

801
802
803
804
511
513
521
523
501
503
VERTICAL PIXELS (600)
HORIZONTAL PIXELS (798)
266 PIXELS
132 PIXELS
534 PIXELS
40 PIXELS
80 PIXELS
120 PIXELS

FIGURE 8

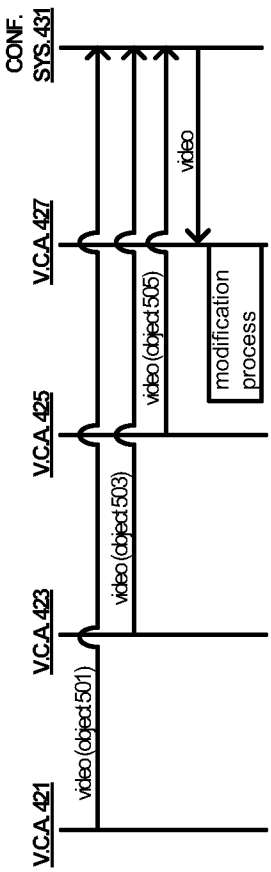


FIGURE 9

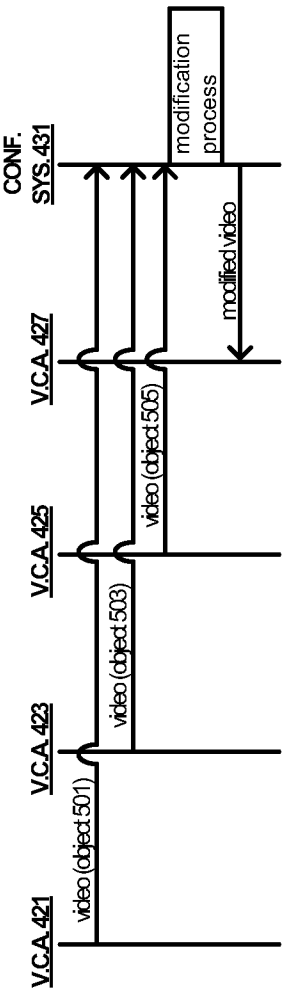


FIGURE 10

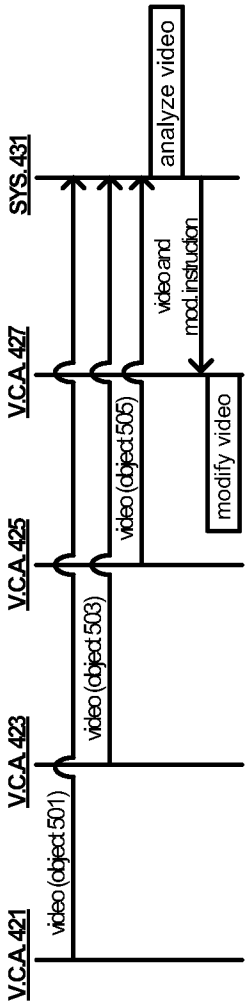


FIGURE 11