200003038-7
31 May 2000

SYSTEM FOR PERSONALIZED FIELD OF VIEW IN A BROADCAST ENVIRONMENT

BACKGROUND OF THE INVENTION

1. Field of the Invention

The invention disclosed broadly relates to the field of broadcast systems and more specifically to a broadcast system providing a personalized field of view.

2. Description of the Related Art

Broadcast media today offer a far greater degree of individual choice than ever before. For example, cable television provides a great choice of programming, pay per view allows customers to select among many recent movies or events for reception at a time selected by the user, and the World Wide Web (WWW) provides varied content on a great number of subjects. However, in most of these cases the format and view of the content is controlled by the producer of the content.

It is advantageous for broadcasters to present program material from multiple camera angles. In a simple case, a football game is recorded by many different cameras. An editor working in real-time chooses which of the camera angles will be broadcast along with commentary. The broadcasters know that different camera angles may be of interest to segments of the viewing population. Depending on which team a viewer is most interested in, the angle at which the game is most interesting will be different. Accordingly, there is a need for a broadcast system in which a user has a greater choice or control in the presentation of the broadcasted material or content.

SUMMARY OF THE INVENTION

Briefly, according to the invention, a business model, a methodology, and communication system comprise a video processing personalization system (VPPS) coupled to at least one video source and to at least one end user unit, the VPPS. The VPPS comprises: a receiver for receiving composite video signals representing more than one than one view of an event; an end user input for receiving signals representing a selection of at least one of the more than one view of the video; and
a rendering device for creating a personalized view for presentation to an end user. According to the invention the VPPS can be implemented at a content producer station, along the network distribution chain, or at the end user's unit (e.g., a set-top box or a television set). The video signals may be television signals or streaming video (a sequence of "moving images" that are sent in compressed form over the Internet, or other packet-switched network, and displayed by the viewer as they arrive). The video may include sound. Thus, the invention makes new methods of doing business possible by enabling different choices to be offered to viewers or subscribers to a network.

FIG. 1 is a simplified block diagram showing a known present day broadcast system in which the invention can be implemented.

FIG. 2 is an illustration of a broadcast system including a video processing personalization system or server (VPPS) according to an aspect of the invention.

FIG. 3 shows one possible implementation of a VPPS in accordance with an aspect of the invention.

FIG. 4 is a flow chart illustrating viewer selection of a desired field of view and zoom level with a VPPS according to the present invention.

FIG. 5 shows that when the viewer changes the field of view, a dedicated proxy responds by changing the field of view being transmitted according to an aspect of the invention.

FIG. 6 shows multi-user selections from a composited view.

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

A system according to the invention allows the user to choose his own virtual camera angle and field of view, in effect giving him a virtual camera allowing him to freely view the event. Further, the user may be attached to the network, whether it be wireless, wireline, narrowband or broadband, Internet or other protocol, with some restrictions on the maximum bandwidth, instantaneous or average, available to him.

Personalized view provides a new business model for broadcasters and distributions and set top or TV manufacturers. This invention enables the sale of enhanced entertainment, enhanced education, etc. as a premium service in conjunction with traditional broadcast. We also describe a new concept, the Video Processing Personalization Server, its architecture and its internal composition, to allow the end user to accomplish this using traditional TV control techniques.

A system in accordance with the invention allows the user to "create" his own view of broadcast or other video. Many shows are created by artists who convey a specific view or message with the selection of each camera angle, and shot. These

shows are not necessarily amenable to a user created experience. However, other shows have other goals. Sports events, educational videos, etc. would be of far more value if users could select the area of interest to them. For shows such as sports events, today, editors select the view that will be broadcast; all viewers see this view. Our invention takes advantage of the advent of new technologies, such as digital set-top boxes, and of high speed two way communications technology, such as cable modems, and the Internet, and allows users to request a personalized view of the event or program broadcast. In effect, the user becomes an audience of one.

For some applications, a larger audience may be desired. For example, a teacher in a distance learning scenario may wish to lead a class in viewing a nature film. While the nature film contains footage on all the flora and fauna in an area, the teacher may wish to focus the class on the foliage being shown so as to illustrate a botany lesson. The class may later view the same video, with a focus an the animal life. The class is larger than an audience of one, but smaller than the broadcast audience to whom the broadcast is available.

For purposes of describing an embodiment of the invention, the term "producer" shall be used hereafter to refer to the source of the video material. The producer may be a broadcaster, cable programmer, film programmer, or a non traditional creative entity of the new digital age. The term "event" shall be used to refer to the program being made available to the viewers. This might be a live event, such a sporting event, political rally, or current news, or it may evolve from what is today a film, video, or conventional or Internet television program, or any other program source where multiple inputs can be aggregate into a broader computer audio or video image.

One of the principles of the invention is to capture as much information about the event as possible, and then create personalized views for the subscriber through

video processing. Capturing the information in this manner should lower the burden on the producer for having esthetically gifted operators manning the cameras.

Using technology well known in the art (e.g.,U.S. Patent 5,187,571, relating to a television system for displaying multiple views of a remote location), the output of multiple cameras can be combined into one single view. For example, consider a situation where one camera is capturing the left side of a room and another camera is capturing the right side of the room. In U.S. Patent 5,187,571 the camera views are combined so one can see the middle of the room. The view of the middle of the room will contain images not found in the left camera, and not found in the right camera, but only in the composite picture.

Additional video processing technology, known in the art, allows zooming in and out from a captured digital image or digital video stream. Additional technology, known in the art, allows video compositing, taking multiple overlapping camera images, and creating a composite image. U.S. Patent 5.657,073, Seamless Multi-Camera Panoramic imaging with Distortion Correction and Selectable Field of View (issued to Stuart Henley) teaches this art. U.S. Patent 5,444,478, Image Processing Method and Device for Constructing an Image from Adjacent Images, teaches a method of processing images for constructing a target image from adjacent images. The above patents are hereby incorporated by reference.

Additional technology, (Stern, Willner and Dunn), teaches how a video stream may be optimally processed for transmission through a restricted bandwidth connection. This patent was filed 2/99, Selective Reduction of Video Data (BC9-98-030), Serial No. 09/256,567, hereby incorporated by reference.

For the purposes of example and simplicity, a video private channel (or view) will be described in detail. Video signals and processing are described but those skilled in the art will recognize that similar techniques could be used with audio signals using compositing techniques current used in stereophonic recordings. Many

current stereophonic devices today use signal processing techniques to combine, split and redirect audio signals to emphasize different rounds or spatial locations. The video example was chosen since it represents the most complex case.

The Private Channel Concept

A system according to the invention includes the concept of a private channel for an end user or viewer. Depending on the nature of the access method used by the viewer, a private channel may be materialized at a centralized point for broadband distribution such as a web server on the Internet, at a satellite or CATV headend using technology known in the art for broadband applications for Video on Demand (VOD), or at telephone central office DSLAM (Digital Subscriber Line Access Multiplier) for switched broadband xDSL. A private channel may also be materialized at the viewer's set-top box. One vendor of Video on Demand solutions is Scientific Atlanta/Seachange, with the ability to materialize the private channel being provided by Scientific Atlanta's OpenCable-compliant digital broadband system. Bandwidth and processing tradeoffs must be considered in the implementation of the private channels. A more detailed description of these tradeoffs is found herein.

In one implementation of our invention, the producer records the event from multiple camera angles, each with very high resolution (e.g., HDTV, or higher). For example, the cameras may be at fixed locations, completely recording every aspect of the event at high resolution. If the event is a football game, cameras may, for instance, be placed every 10 yards, and on the sidelines as well as views of the stadium. These cameras produce video streams which faithfully record all events in the arena (raw video). Since many video streams are employed, each being at high resolution, we assume that the totality of this bandwidth is too great to be transmitted to the viewing community at large under existing technology. The

cameras transmissions are directed to a video processing personalization server (VPPS).

The VPPS may be fixed or mobile, located at a broadcast studio, cable headend, or major Internet routing node or server location, sporting arena, or a new specialized facility. The VPPS may be one system, or maybe a distributed system. The VPPS records the output of all the camera angles as well as having the ability to produce large numbers of personalized video outputs, each appropriate to transmit to a viewer. In one implementation, the VPPS "sits" at the location in the network furthest from the user, at which it is possible to materialize a private channel for that user. The VPPS function may occur at the cable headend, or further back in the network. In this case, the VPPS must produce large numbers of personalized video outputs In another implementation, the processing of the multiple video streams to produce the personalized video output may occur at the set-top box. In the case where the VPPS is implemented at the end user's unit (e.g., set-top box) the VPPS only needs to produce one video output. That is, multiple video streams may be directed to the set-top box for personalized processing into a single video stream, and the set-top may provide the VPPS function. With the advent of digital television, the spectrum allocated for a single channel today may provide multiple digital video stream capacity; use of this capacity to provide a personalized view is a novel approach to the "excess" channels available. In another implementation, the VPPS function is divided into a distributed architecture where part of the VPPS is closer to the event, and part of it is closer to the end user.

Default View

The producer may employ an expert to create a pleasing and informative view of the event. This view will resemble today's programming in that it will employ many camera angles, perspectives, overlay drawings (such as the iNFiNiTI family of products by Chyron, http://www.chyron.com/products/index.html), and provide an

exciting and satisfying viewing experience. This may be used as a default viewer experience, so that if a viewer does nothing he will receive these images.

End User Perspective

The viewer, V, watches on a TV, PC, or other device. His initial view of the game is determined by the producer and transmitted by the VPPS. When no longer satisfied with the view, V personalizes the view by selecting different camera angles, Zoom degrees, focus, and other selectable broadcast attributes. Through a user interface such as a keyboard, remote control, IR to the set-top, or other such devices, V chooses what should be shown on the display. Those commands are received by the VPPS, or ancillary processor, and appropriate camera angles composed from the raw material to create the personalized view desired. The VPPS transmits this view for presentation (e.g., display) to V, sending a view compatible with the bandwidth available to V on the network on which he is connected. V may continue to modify the view received with the same user interface, and may elect to return to the producer's default view. Thus, referring to FIG. 6 there is illustrated a set of camera views as presented to various viewers (A-F).

Zoom

An important advantage of our invention is the ease with which views of zoom in and zoom out can be supplied. In the preferred embodiment, the composite picture is at higher resolution than the user receiving device (TV). For example, the cameras may be HDTV format and the user's TV may be NTSC. The cameras may also be NTSC, but many of them are used, each with a close-up, to create the resolution composite. Zoom is achieved by interpreting the users' commands to define a rectangle within the composite image which will be image processed to fill the field of view of the receiving device. This rectangle may be from explicit specification from the user device, or may be derived dynamically so as to include

objects of interest to the user. For example, in a football game, the zoom automatically changes to include the viewer's favorite player and the football. If the selected resolution is lower than that of the composite image, then bits are eliminated by well known averaging techniques to produce a lower resolution image from a higher one. If the selected resolution is higher than that of the composite image, then bits are introduced by well known up-conversion techniques to produce a higher resolution image from a lower one. Examples of this art are embodied in the Snell-Wilcox conversion chips and products (Alchemist fine, Kudos line, etc. which are capable of video conversion. See: http://www-snellwilcox.com).

Tagging Objects of interest

In many cases, a moving object or person is the subject of intense scrutiny. In order to provide a better viewing experience, our invention includes the system of tagging the object or person of interest (OOI) with an identifiable visual tag (such as a unique color or emblem). When a viewer has chosen to follow the OOI, this allows the VPPS to examine the composite and select a subsection which contains the OOI for presentation to the viewer on the personalized channel. In an alternate embodiment, a non-visual tag is used; for example, an RF or infrared tag affixed to the OOI, and the cameras recording the event constructed to respond to the tag by either following the OOI mechanically, or by tagging the video as containing the desired image. Different tags may be used for different OOIs, or a common tag may be used when multiple objects may be considered OOIs.

Object recognition may also be used to locate OOIs in the composite. A purely software approach may therefore be employed to select the subset of the composite containing OOIs.

For example, a user may select a view that includes 10 feet around player 56, an OOI, and the football. As the field of play moves, the VPPS will select a continuously changing view as specified.

Referring to FIG. 1, there are shown multiple video cameras 1-3 on a sporting event, broadcasting a sporting event, as it is done today. Players 10 and 11 and football 12 are on the field of play. Player 10's activity is being captured by camera 1, player 11's activity by camera 2. The football is being tracked by camera 3. The video streams thus captured are visible to the producer on monitors 4, 5, and 6. A video console 15 receives each view and processes it for presentation to the producer 7. The producer 7 selects his or her desired video stream from the streams being captured, and that stream is broadcast through a satellite uplink 9 to the broadcast audience for this event.

Referring to FIG. 2, there is shown a distribution network with VPPS processing, illustrating a broadcast of a sporting event, with an implementation of the invention. Players 10, and 11 and football 12 are on the field of play. The players and the football are tracked as in FIG. 1. The video streams from all three cameras are directed into the network infrastructure of today 70. Attached to the infrastructure 70 are VPPS systems 20, and 21, associated with network distribution points (such as cable headends) 30 and 31 respectively. In our example we use a cable TV network for simplicity, but any 2-way network with sufficient capacity could adequately serve as well. Cable headend 30 has additional inputs from other sources, such as broadcast channels 40 and 50. Headend 31 has additional inputs from other sources such as broadcast channels 41 and 51. Between headend 30 and viewer 80 exists the capacity for a personal channel for viewer 80. Between headend 31 and viewer 81 exists the capacity for a personal channel for viewer 81. Broadcast tree 60 shows the branch and tree structure of a cable network implemented for two way communication, as does tree 61. Attached to a branch of tree 60, viewer 80 selects and views an image composed of portions of the video stream from cameras 1 and 3, and sees images of player 10 and football 12. Attached to a branch of tree 61, viewer 81 selects and views an image composed of portions of the video streams from cameras 1 and 2, and sees images of players

10 and 11. Note that in this example, the VPPS facility is upstream of the branch and tree distribution parts of the network, and closer to the event. In another embodiment, discussed above, the VPSS is in the set-top box or in the TV itself, or distributed.

FIG. 3 shows a VPPS design according to the invention. Blocks 17, 18, and 19 represent video input processors. Through the broadcast network, not shown, the outputs of cameras 1, 2, and 3 are transmitted to the input of video input processors 17, 18, and 19 respectively. The processors convert the compressed video input to uncompressed digital video, and communicating over bus 22, store it in video memory 100 which consists of multiple image pipelines. We show three, for example (one per video input processor shown as blocks 110 - 142). Blocks 110, 120, 130, and 140 represent a pipeline corresponding to camera 1; that is, each block contains one frame of the video so that the pipeline contains the current frame and three previous frames from each video input processor. As each new frame comes in, the oldest frame is overwritten. Memory 100 therefore contains the three most recent frames, and one in progress, from each of the video input processors. Thus, memory 100 acts as a buffer for processing the image data by compositing processors 11-14. This pipelining approach makes each frame available for three "frame times" to facilitate further processing. Video memory 100 is accessible through two high speed parallel buses, 22, and 23. As mentioned, bus 22 supplies connectivity to the video input processors. Bus 23 supplies connectivity to the compositing processors, 11-14. The compositing processors working through the pipeline, take the overlapping frame images and produce one large composite view. Four frames of composite view are stored in composite memory 200. Blocks 210-240 each contain one frame of composite view. Memory 200 contains the three most recent composite frames, and one in progress. As new frames are received, by input processors 17-19, they are processed, and stored in one of the four frame buffers in video memory 100, for example 110, 111, 112. One of the compositing

processors immediately begins to process this frame buffer, for example compositing processor 11. By the time the frame buffer will be overwritten by new input, compositing processor 11 will have formed the composite image and written it into a composite frame buffer, for example frame buffer 210, in composite memory 200. Composite memory 200 is accessible by two high speed parallel buses, 23 and 24. As mentioned above, bus 23 supplies connectivity to the composite processors 11-14, and bus 24 supplies connectivity to the User Proxy Processors (UPPs) 320-325. The UPPs 320-325 are dynamically associated with end users that are currently receiving personalized views. We would expect that in a production headend, supporting 50,000 subscribers, that several hundred UPPs may be attached to a VPPS. Viewers communicate to the VPPS via a communications processor 350. When a viewer first requests a personalized view associated with a broadcast being processed by the VPPS, the two-way cable system transmits this request to the communications processor 350 within the VPPS. The resource management function within this processor assigns an available UPP to service the user. Routing commands are sent to the two-way cable system so that input from the user is routed to the assigned UPP. The two way cable system instructs the user device to tune to the "channel" where the personalized view is available. In the preferred embodiment, the transmission is digital it but could be analog.

As noted earlier, the VPSS could be contained in a facility in the network at various points, contained in the terminal unit, or split into a hierarchical function where the steps through compositing are performed in the centralized location and the steps comprising the User Proxy Processors are in the terminal devices. The VPPS function is performed with bus 24, communication processor 350, and UPPs 320 - 325, each connected to a bus 24.

The bus 24 provides a communication link among UPPs 320-325, the communication processor 350, and an Internet server 352, which is also linked to the Internet 354. This enables the user to select either standard web pages or events

and movies broadcast over the Internet. Thus, the Internet becomes another source of content that does not get composited but can be processed by the UPPs to provide a personalized view.

For example, the viewer might select an additional channel of data, supplied over a data network such as the Internet, to be displayed concurrently with the broadcast program. In this case, the UPPs would produce an overlay or insert area on the selected composite view before sending it to the terminal.

FIGs. 4 and 5 show viewer selection of a desired field of view (perspective) and the zoom level shows the flow for selection of a view. FIG. 4 shows a process 400 of viewer selection of a desired field of view and zoom level. The viewer chooses to watch the sporting event (step 402), and then chooses the desired camera angle or location and focus (selects zoom level) (step 404). The viewer's device transmits this information via the two-way cable plant to the VPPS (step 406). The communication interface 350 associates UPP 320 with viewer 80 and communicates initial viewer choices (step 408). Thus, the VPPS initiates a user proxy process (UPP) 320 to provide service to viewer 80, and initializes it with the appropriate information as to viewer camera angle and zoom selection. The UPP 320 subsets aggregate video into a desired field of view and zoom level (step 410). The UPP 320 determines the tuning location for the personalized channel for viewer 80, creates the desired field of view, informs user 80's device where to receive the personalized channel, and transmits the location to viewer 80's device (step 412). The device receives the information and tunes to the correct channel (step 414).

FIG. 5 shows that when the viewer changes the field of view, the dedicated proxy responds by changing the field of view being transmitted in a process 500. No new viewer channel selection is required. In step 502 the viewer changes the desired field of view or zoom level for an event in progress. The viewer's device transmits new information to communications interface 350 which is transmitted to

UPP 320 (step 504). The UPP 320 then changes selection according to the user request and begins transmitting the new view (step 506).

FIG. 6 depicts the various views than can be constructed from the composite image.

An important advantage of our invention is the ease with which views of zoom in and zoom out can be supplied. In the preferred embodiment, the composite picture is at higher resolution than the user receiving device (TV) handles. For example, the cameras may be HDTV format and the users TV may be NTSC. The cameras may also be NTSC, but many of them are used, each with a close-up, to create the resolution composite. Zoom is achieved by interpreting the users commands to define a rectangle within the composite image which will be image processed to fill the field of view of the receiving device. This rectangle may be from explicit specification from the user device, or may be derived dynamically so as to include objects of interest to the user. For example, in a football game, the zoom automatically changes to include the viewers favorite player and the football. If the selected resolution is lower than that of the composite image, then bits are eliminated by well known averaging techniques to produce a lower resolution image from a higher one. If the selected resolution is higher than that of the composite image, then bits are introduced by well known up-conversion techniques to produce a higher resolution image from a lower one. Examples of this art are embodied in the Snell-Wilcox chips which are capable of video conversion.

Tradeoffs in placement at the private channel

If the private channel, described in the following sections, is materialized closer to the viewer, less bandwidth in total is needed on the distribution network to serve all its viewers. A relatively small number of camera angles are sent to the video compositing devices. Users will compose their own pictures from this source material. This is shown pictorially in FIG. 6. Three camera outputs have been

combined into one composited picture (as explained earlier in the disclosure in conjunction with FIG. 3). Viewers D, E, and F are each satisfied with one camera view, camera 1, 2, and 3 respectively. However, viewers A, B, and C each prefer a more customized view of the composite picture. They have each selected a different "virtual camera" represented by the field of view propagated by the parenthesis set by D, E, and F. To minimize distribution bandwidth, embedding the VPSS function in the subscriber's unit (a TV, set-top box, or the like) is the optimum solution.

If the private channel is implemented closer to the event, more bandwidth in total is needed in the distribution network to serve its viewers; and in this case each private view must traverse more of the distribution network, in this case, a relatively large number of "virtual camera" views, are composed closer to the event, and more bandwidth (channels) are needed in the distribution network to accommodate the large number of viewers, each of whom has a specific permutation of the view they desire. This method may allow some economies of scale in the VPSS units, but in either case parts of the VPSS unit are on a "per viewer" or possibly a "per view" basis, specifically the User Proxy Processor (explained later). In one embodiment of the invention, a maximum number of virtual camera views is established, from a smaller number of actual camera views which provide source material to the compositors. In this case a fixed number of View Servers is established. All viewers sharing the same virtual camera view, share the same view server. When a viewer acts to change his view, he or she is fed the output from the new View Server representing that view. To accomplish this, a video distributing function, commonly known in the art, would be added to FIG. 3. There may be fewer accessible views (each represented by a View Server) than is possible given the number of pixels in the receiving station. In this case, the viewer will "snap to" the closest matching View, much as the snap to function in drawing programs aligns the objects drawn to a predetermined grid. Note that in this implementation of the

invention, zoom functions cannot be performed at the VPPS. They may be performed at the set top box, or may be absent entirely.

A third option is to split the entire video processing operation into a distributed, hierarchical design where some of the processing is done at the VPPS in the centralized locations, and the rest at a point closer to the termination point (TV or set-top unit), or at an intermediate VPPS (IVPPS). In a cable environment for example, the IVPPS could be at the cable headend. Such a split could be accomplished by generating and storing the composited view close to the event, and sending the composited view through the distribution network to the IVPPS. The bandwidth required between the VPPS and IVPPS must be enough to carry the composite. This is less than the number of channels which would be required for each camera, but certainly more than one. Out of the IVPPS, the bandwidth requirements for private channel remain. The IVPPS finishes the Video processing, using the aforementioned User Proxy Processors to allow for the customized view. In terms of FIG. 3, the function of UPPs 320 through 325 would be embedded in the IVPPS.

In a fourth option, the set top box serves as the IVPPS. The composite is sent to the terminal units (set top box, TV, PC etc.) This requires the use of multiple channel bandwidth to send the composite to each terminal unit, but the bandwidth required is less than the amount of custom channels which may be created from it. The terminal unit completes the video processing, with each unit containing one UPP, in terms of FIG. 3, blocks 320 through 325 are embedded in the terminal units. Each of these embodiments are operative and each has tradeoffs in bandwidth and cost and all are contemplated by our invention.

Additional Features and Functions

This invention also allows for the use of a database (not shown) as part of the VPSS facility, or separately as part of the service providers network. This database

is used to maintain viewer profiles and viewer preference history data so that when a given event occurs, the "default" image shown is the user specific default. This could be determined by direct user input (e.g., setting preferences via the Internet or via a telephone system connection to the database), or by heuristic data gathered based an the VPSS selections. Depending on the capabilities of the viewing equipment this database could also select default parameters for enhancing or reducing the quality of the image as stated earlier in the disclosure.

This invention also allows the distribution of the total composite image to a plurality of end users. The total composite image is formed by combining the outputs of several at the cameras as is shown in FIGs. 2 and 3. The transmission of the composite: image will require more bandwidth than an individual private view, but less bandwidth than the sum of the individual camera outputs. The total composite image is transmitted to each end user and is processed within his terminal device, to generate his personal view. The transmission of the total composite image can be achieved by well known transmission techniques already referenced. For example, the total composite image may require 10 times the bandwidth of a personal channel. Current digital compression techniques will allow the transmission of five digital channels within one 6 MHZ NTSC channel. Two channels would be used in conjunction in order to transmit the total composite picture. This would require the terminal equipment to have a multi-channel tuner, multi-channel demodulation, and associated digital processing capability. An example of this in analog would be the "picture-in-picture" tuners available today.

This invention also allows the use of The Second Audio Program (SAP) channel to deliver different audio streams that may be associated with different camera views. Each camera sound system would use its audio facility (spokesman, microphone in the camera, etc.) to feed into the VPSS. The VPSS separates the audio stream as in today's TV signal processing, and associates it with either a discrete camera view, or a "portion" of the composited view, in today's TV designs,

only two audio channels are available so the VPSS would "switch" the audio presentation when a viewer "panned" from one part of the composite view to another. If future TV designs support multiple audio streams, then this technique would be used on a more granular boundary. Alternatively, the audio inputs could be mixed, with emphasis placed on certain inputs relative to a selected area of view.

Viewer Interface and Selection Methodology

One of the goals of this invention is to make the viewer interface as uncomplicated as possible, given the rich selection of function described as a result of the VPSS capability. There are several envisioned additional capabilities that would be supported by either the set-top box or the TV itself to accomplish this. All of those are easily accomplished using any two way communication capability as currently supported by cable and satellite networks. Broadcast networks do not now have two way capability, but could accomplish the same effect using a telephone call-in system. In this case, caller-ID or an equivalent function would identify the caller and allow the camera selection to occur.

In any case, the VPSS creates selection menus for the available processing features and camera angles, which are limited by the number of VPSS, UPPs and the number of cameras and OOIs that the cameras can be assigned to follow. This list of limited selections could be presented as a "drop down box" or a "dialog box" as commonly used in computer program graphical user interfaces, where a selection list is presented when the box is opened. The intelligence in the terminal unit controls the appearance of the box and its contents. The user may select the desired item(s) via remote control using the channel buttons on today's remotes, or future function/selection keys as newer remote controls are designed to take advantage of emerging capabilities in set-tops and TVS.

In an alternative implementation, the VPSS could segment the video into selectable areas, based on either a field of view or an OOI, and present a limited number of these to the viewer. By dividing the screen into a number of "blocks," the VPSS could orient the camera angle to the selected block.

In another implementation, using a new remote control with "zoom" and "pan" functions, the user could dynamically select the area of interest. Again, the return channel in a two way system would relay this information to the VPSS where the UPP would execute the requested commands. These functions would operate in a way similar to the game controls on today's computer games.

The response times will vary depending on whether the VPSS function is located locally in the set-top or TV, remotely at a centralized location, or split into hierarchical sets as explained earlier.

New Business Model

Personalized view provides a new business model for broadcasters. This invention enables the sale of enhanced entertainment, enhanced education, etc. as a premium service in conjunction with traditional broadcast.

This invention also allows for the carriers to develop a new business model where they work in co-operation in delivering the broadcast rather than competition. In today's environment, a network such as ABC may have the exclusive rights to broadcast a sporting event. In most cases, the network will not share this right with other networks, one notable exception being arrangements where "pools" are required. The network will make the broadcast available to all its owned and affiliated stations. These local stations often compete for viewers in a given area (e.g., channels 10 and 25 in South Florida). With the aforementioned need to carry more channels, and the ability to use different channels to concurrently carry different portions of the picture, different affiliates could each carry a portion of the composited picture in their standard broadcast spectrum (8 MHZ in today's analog

network). They could each get a portion of the revenue derived from the enhanced, private channel and virtual camera services which neither of them could earn if they could carry only a single view of the event. The broadcast media are not limited to telecommunication lines. The principles of the invention would work whether the channels were delivered via broadcast, CATV, or satellite, or other business model such as enhanced TV's and set top boxes. The cost is based on the bandwidth used (e.g. smart set-tops cost less to support than implementations that need the carrier to be the whole VPSS.

The scope of the invention is not to be restricted, therefore, to the specific embodiment, and it is intended that the appended claims cover any and all such applications, modifications, and embodiments within the scope of the present invention. Thus, even though the example above relates to the distribution of video content (preferably digital) with accompanying audio, the personalized perspective could also be a selection of various audio, text, or other content broadcasted over a network.

CLAIMS

1. A video processing personalization system (VPPS) coupled to at least one video source and to at least one end user unit, the VPPS comprising:

a receiver for receiving composite video signals representing more than one than one view of an event;

an end user input for receiving signals representing a selection of at least one of the more than one view of the video; and

a rendering device for creating a personalized view for presentation to an end user.

2. The VPPS of claim 1, wherein the video signals are streaming video signals.

3. The VPPS of claim 2, wherein the rendering device comprises at least one user proxy processor for rendering at least one selected view to the end user.

4. The VPPS of claim 3, wherein the end user input comprises a communication processor for receiving user requests for selected views and for transmitting said user requests to a source of the video signals.

5. The VPPS of claim 4, further comprising a composite memory for receiving and storing at least some frames of composited view.

6. The VPPS of claim 5, further comprising a plurality of compositing processors, each for receiving a plurality of overlapping frame images to produce at least one composited view.



7. The VPPS of claim 6, further comprising a video memory comprising a plurality of image pipelines, each pipeline for storing a plurality of the most recent composite frames.

8. The VPPS of claim 6, further comprising a plurality of video input processors, each for receiving images from a plurality of video cameras and for processing said images.

9. The VPPS of claim 1, wherein the video comprises sound and the VPPS further comprises a speaker.

10. The VPPS of claim 2, wherein the view comprises a zoomed view.

11. The VPPS of claim 2, wherein the view comprises a location-based view.

12. The VPPS of claim 2, wherein the view comprises a view of an object of interest that has been tagged.



13. A method for personalizing a video transmission for an end user, comprising:

receiving a video transmission comprising a plurality of views of a video event;

receiving a signal from the end user selecting at least one view for rendering thereof; and

rendering a selected view to the user.

14. The method of claim 13, further comprising the following step after receiving a video transmission:

receiving an end user signal selecting an event for viewing.

15. The method of claim 13, further comprising the following step preceding receiving a video transmission:

receiving an end user signal selecting a camera angle and focus;

16. The method of claim 13, further comprising:

associating a user proxy process with the end user and communicating initial end user choices to an apparatus for compositing the plurality of views.



17. A computer readable medium comprising program instructions for:

receiving a broadcast transmission comprising a plurality of views;

receiving signal from the user selecting at least one view for presentation to the user; and

rendering a selected view to the user.

18. The medium of claim 17 further comprising the following instruction after receiving a broadcast transmission:

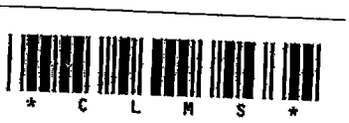
receiving an end user signal selecting an event for viewing.

19. The medium of claim 17 further comprising the following instruction preceding receiving a broadcast transmission:

receiving an end user signal selecting a camera angle and focus;

20. The medium of claim 17 further comprising:

associating a user proxy process with the end user and communicating initial end user choices to an apparatus for compositing the plurality of views.



21. A method for viewer selection of a desired field of view and zoom level comprising:

receiving a viewer selection to watch a selected event and a selection of the desired camera angle or location and focus (the selection information);

transmitting the selection information to a video processor;

associating a user proxy processor with the viewer and communicating initial viewer choices;

initiating a user proxy process to provide service to the viewer, and initializing said process with the appropriate information as to viewer camera angle and zoom selection;

aggregating video into a desired field of view and zoom level;

determining the tuning location for the personalized channel for the viewer;

creating the desired field of view, and informing the user's device where to receive the personalized channel;

and transmitting the location to the viewer device.

* C L M S *

22. An end user data processing unit comprising:

a plurality of video input processors, each for receiving video image data representing different views of an event;

a video memory, coupled to the video processors, and comprising at least some of the most recently received frames of the video image data;

a plurality of compositing processors, each coupled to the video memory, for receiving overlapping frame images from the video memory, and producing one composite view;

a composite memory for storing a plurality of composite views, each comprising a plurality of storage areas, each storage area for storing a frame of the composite view; and

a user communication processor for receiving viewing selections from the user and for interacting with the composite memory to render a selected view to the end user.



23. A television head end unit comprising:

a plurality of video input processors, each for receiving video image data representing different views of an event;

a video memory, coupled to the video processors, and comprising at least some of the most recently received frames Of the video image data;

a plurality of compositing processors, each coupled to the video memory, for receiving overlapping frame images from the video memory, and producing one composite view;

a composite memory for storing a plurality of composite views, each comprising a plurality of storage areas, each storage area for storing a frame of the composite view; and

means for receiving selections of views from a user proxy processor.

