APPARATUS, SYSTEMS AND METHODS FOR ENHANCED VIEWING EXPERIENCE USING AN AVATAR

Systems and methods are operable to enhance a user's experience during presentation of media content events. An exemplary embodiment outputs a media content event from a media device to a media presentation system, the media content event comprising a video portion and an audio portion, wherein the media presentation system is configured to present the media content event to at least one user; determines an emotional response of the at least one user to a currently presented portion of the media content event; animates an avatar to present an emotion, wherein the avatar emotion corresponds to the determined emotional response of the at least one user; and outputs the animated avatar to the media presentation system such that the animated avatar is presented concurrently with at least the video portion of the media content event.
APPARATUS, SYSTEMS AND METHODS FOR ENHANCED VIEWING EXPERIENCE USING AN AVATAR

BACKGROUND

[0001] Media devices, such as a set top box, a stereo, a television, a computer system, a game system, or the like, are configured to present a media content event, such as a program, movie, sporting event, game, or the like, to a user.

[0002] At times, the viewing experience may be enhanced when there are multiple users watching the media content event at a single location. For example, presentation of a sporting event may be more interesting and enjoyable when there is a shared viewing experience among a plurality of viewers. In such situations, viewer responses are experienced by the other viewers. Thus, when a favorite team scores during the presented sporting event, the cheers from some or all of the viewers may enhance the viewing experience of the viewers.

[0003] In some situations, the user may be alone when viewing the presented media content event. In such situations, there is no opportunity to enhance the viewing experience of the single user based on responses from a plurality of viewers (since no other viewers are present).

[0004] Accordingly, there is a need in the arts to facilitate an enhanced experience for at least a single user during presentation of a media content event.

SUMMARY

[0005] Systems and methods of enhancing a user's experience during presentation of media content events are disclosed. An exemplary embodiment is configured to output a media content event from a media device to a media presentation system, the media content event comprising a video portion and an audio portion, wherein the media presentation system is configured to present the media content event to at least one user; determine an emotional response of the at least one user to a currently presented portion of the media content event; animate an avatar to present an emotion, wherein the avatar emotion corresponds to the determined emotional response of the at least one user; and output the animated avatar to the media presentation system such that the animated avatar is presented concurrently with at least the video portion of the media content event.

BRIEF DESCRIPTION OF THE DRAWINGS
Preferred and alternative embodiments are described in detail below with reference to the following drawings:

FIGURE 1 is a block diagram of an embodiment of an avatar presentation system implemented in a media device; and 

FIGURE 2 is a block diagram illustrating greater detail of an embodiment of an example media device.

DETAILED DESCRIPTION

FIGURE 1 is a block diagram of an embodiment of an avatar presentation system 100 implemented in a media device 102, such as, but not limited to, a set top box (STB). Embodiments of the avatar presentation system 100 may be implemented in other media devices, such as, but not limited to, a television (TV), a digital video disc (DVD) player, a digital video recorder (DVR), a game playing device, a personal computer (PC), a notepad, a smart phone, or other media device that is configured to present a video-based media content event that is received in a media content stream 104.

The exemplary media device 102 is communicatively coupled to a media presentation system 106 that includes a visual display device 114, such as a television (hereafter, genetically a TV), and an audio presentation device 116, such as a surround sound receiver (hereafter, generically, a speaker). Other types of output devices may also be coupled to the media device 102, including those providing any sort of stimuli sensible by a human being, such as temperature, vibration and the like. The video portion 118 of the media content event is displayed on the display 120 and the audio portion of the streamed program content is reproduced as sounds by one or more speakers 122. In some embodiments, the media device 102 and the media presentation system 106 may integrated into a single electronic device.

Based on determining an emotional response of the at least one user to a currently presented portion of the media content event, embodiments of the avatar presentation system 100 present one or more avatars 124 on the display 120 concurrently with the video portion 118 of the media content event. The exemplary avatar 124 is presented in a manner such that the user (not shown) is able to perceive computer generated emotions presented by the avatar 124, such as facial expressions, gestures and/or audio comments. Presented avatar emotions may be based upon a variety of factors, including characteristics of identified responses of the user, characteristics of the currently presented portion of the media content event, and/or by avatar response specifications.
contained in the media content event data. Thus, an animated avatar 124 may smile, frown, wink, or the like. Further, the presented avatar 124 may perform various gestures or the like, such as, but not limited to, nodding their head, wave their arms, or even jumping up from a seat.

[0012] Some embodiments may be configured to present audio information that is perceived by the user as coming from, and/or being associated with, the avatar 124. For example, the avatar 124 may appear to provide auditory comments pertaining to the currently presented portion of the media content event. Some embodiments may present other verbal sounds apparently emitted by the avatar 124, such as, but not limited to, laughing, crying, cheering, booing, screaming, or the like. Alternatively, or additionally, the avatar 124 may appear to originate other non-verbal sounds, such as clapping, finger snapping, or the like.

[0013] Depending upon the nature of the presented media content event, a plurality of avatars 124 may be presented on the display 120. The plurality of presented avatars 124 may be made based upon a specification by the user. Alternatively, or additionally, information received in the media content event data may cause presentation of a plurality of avatars 124. For example, if the media content event typically associated with a group of people, such as at a sporting event, a movie, a concert, or the like, several different avatars 124 may be presented to simulate an audience. As another example, if the media content event pertains to a game, other participating game player avatars 124 may be represented.

[0014] The non-limiting exemplary media device 102 comprises a processor system 126, a memory 128, and an optional image capture device 130. The memory 128 comprises portions for storing the media device logic 132, the optional avatar management logic 134, the user image analysis logic 136, and the avatar model database 138. In some embodiments, the media device logic 132, the avatar management logic 134, and/or the user image analysis logic 136 may be integrated together, and/or may be integrated with other logic. In other embodiments, some or all of these memory and other data manipulation functions may be provided by and using a remote server or other electronic devices suitably connected via the Internet or otherwise to a client device. Other media devices 102 may include some, or may omit some, of the above-described media processing components. Further, additional components not described herein may be included in alternative embodiments.

[0015] The exemplary media device 102 is configured to receive commands from the user via a remote control 140. The remote control 140 includes one or more controllers 142 on its
surface. The user, by actuating one or more of the controllers 142, causes the remote control 140 to generate and transmit commands, via a wireless signal 144, to the media device 102. The commands control the media device 102 and/or control the components of the media presentation system 106. The wireless signal 144 may be an infrared (IR) signal or a radio frequency (RF) signal.

[0016] The processes performed by the media device 102 relating to the processing of the received media content stream 104 and communication of a presentable media content event to the components of the media presentation system 106 are generally implemented by the processor system 126 while executing the media device logic 132. Thus, the media device 102 may perform a variety of functions related to the processing and presentation of one or more media content events received in the media content stream 104.

[0017] The functions performed by the media device 102 relating to the processing of the avatar 124, and/or presentation of the avatar 124 to the components of the media presentation system 106, are generally implemented by the processor system 126 while executing the avatar management logic 134. In operation, an exemplary embodiment of the media device 102 presents the media content event without the exemplary avatar 124 by default. When the media device 102 is operation is such a default state, a user input is required to reconfigure the media device 102 so that the one or more avatars 124 are presented.

[0018] In an exemplary embodiment, the avatar 124 is presented in response to a user input. The user input may be provided in different manners. For example, but not limited to, user actuation of one or more of the controllers 142 causes the remote control 140 to generate and transmit an avatar presentation signal to cause presentation of the one or more avatars 124. An exemplary embodiment has a dedicated controller 142 that the user may conveniently activate to cause generation and transmission of the avatar presentation signal. Alternatively, or additionally, the user may operate the remote control 140 or another suitable user interface to present a graphical user interface (GUI), such as a selection menu or the like (not shown), on the display 120. The menu is configured to permit the user to select presentation of the one or more avatars 124. The user navigates about the presented menu to select presentation of the one or more avatars 124, and/or specify one or more characteristics of the presented one or more avatars 124.

[0019] Information used to generate a presented avatar 124 resides in the avatar model database 138. To present the avatar 124, an avatar model is selected and retrieved from the avatar model database 138. Information corresponding to the avatar 124 is then generated by the processor
system 126 under the management of the avatar management logic 134. The generated avatar information is then communicated from the media device 102 to the components of the media presentation system 106. The video portion of the computer generated avatar 124 is presented on the display 120. Audio information that is perceived to be associated with the computer generated avatar 124 is reproduced as sounds on the speakers 122 (or other speakers, such as speakers included in the exemplary visual display device 114).

[0020] The presented avatar 124, in an exemplary embodiment, is a "virtual person" or the like. The avatar 124 may be generated using any suitable avatar animation computer graphics logic that is configured to generate an image of the avatar 124. Further, the animation of the presented avatar 124 includes facial expressions, gestures, and/or other movements of the presented avatar 124 which are perceivable by the user. Alternatively, or additionally, the computer graphics logic may include voice synthesis logic and/or other sound synthesis logic configured to generate the audio information that is perceived to be associated with the presented avatar 124.

[0021] A single avatar model, or a plurality of different avatar models, may be used depending upon the embodiment and/or user preferences. For example, the exemplary avatar 124 presented on the display 120 is illustrated as a relatively young male human. Any suitable type of avatar 124 may be used by the various embodiments. Avatar models may be used to generate female avatars, cartoon avatars, older aged avatars, baby avatars, or the like. Non-human avatars 124 may be modeled, such as aliens, monsters, animals, or the like. Some avatar models may be based on an image of a person, such as a well known personality, an actor, a relative, and/or a friend. Some avatar images may be drawn or may be graphic-based images. Other avatar images may be based on a photograph and/or video of the person represented by the avatar 124.

[0022] In some embodiments, the particular type of avatar model that is used to generate and present the avatar 124 may be user selectable. For example, many different avatar models may be stored in the avatar model database 138. Embodiments may present a GUI type menu or the like providing thumbnail images of available avatars 124 that may be selected by the user. Alternatively, or additionally, the presented avatar 124 may be based on avatar model information received in the media content event data.

[0023] The presented avatar 124 is configured to enhance the user's viewing experience. Accordingly, in an exemplary embodiment, animated facial expressions of the presented avatar 124 will vary to reflect the currently presented portion of the media content event. In an exemplary
embodiment, the animated expressions of the avatar 124 will generally correspond to the emotions of the user. However, in some situations, the animated expressions of the avatar 124 may be different from the user's emotions so as to generate controversy.

[0024] For example, if the presented media content event is a sporting event, a missed scoring opportunity may be disappointing to the user. Information in the media content event data may indicate the missed scoring opportunity. Alternatively, or additionally, a user response to the missed score may be detected. Accordingly, the avatar 124 may become animated to display a frown or the like that is perceivable by the user to be compatible with their emotion relating to the missed score. Here, the avatar 124 and the user are "apparently" sharing the common experience of being disappointed by the missed score. Thus, the user's experience in enhanced.

[0025] In some embodiments, audio information corresponding to a groan or the like may be concurrently presented with the animated frown of the avatar 124. Alternatively, or additionally, audio commentary may be generated. For example, the avatar 124 may be perceived as saying a phrase, such as "look at the missed scoring opportunity!" Any suitable audio commentary may be generated.

[0026] A listing of suitable audio comments and/or avatar-based sounds may be stored in the avatar model database 138. Once the nature of the user's response has been determined and/or predicted, the nature of the user's response may be correlated with one or more response emotions and/or response behaviors. Based on the determined response emotion and/or response behavior of the user, one or more suitable audible comments may be selected and presented such that the animated avatar 124 appears to be speaking when the audible comments are reproduced as sounds by the components of the media presentation system 106. Concurrently, the avatar 124 may be animated so as to appear to be speaking. Alternatively, or additionally, particular audio comments may be included as part of the media content event data. Any suitable audio synthesis logic may be employed by the various embodiments to generate audio commentary that is apparently originating from a presented avatar 124.

[0027] At other times during the presented sporting event, the user's favored team may score, much to the pleasure of the user. Information in the media content event data may indicate the score. Alternatively, or additionally, the user's response to the score may be detected. Accordingly, the avatar 124 may become animated to display a smile or the like that is perceivable by the user. Further, the avatar 124 may be configured to apparently emit a cheer or other suitable commentary.
For example, the avatar 124 may be perceived as saying a phrase, such as "look at that amazing score!" Here, the avatar 124 and the user are "apparently" sharing the common experience of being excited by the score. Thus, the user's experience is enhanced.

[0028] Some embodiments may be configured to generate and present apparent gestures made by the avatar 124. For example, the animated avatar 124 may be shown as clapping or the like. Further, non-verbal sounds associated with the clapping or the like may be presented. In some situations, the presented avatar 124 might even stand up and cheer in response to the team's score.

[0029] Responses of the user to the presented media content event may be determined in a variety of manners. An exemplary embodiment includes the image capture device 130. The image capture device 130 is configured to capture images of the user as they are viewing the presented media content event. The image capture device 130 may capture one or more still images of the user. Alternatively, or additionally, video images of the user may be captured.

[0030] To capture images of the user, the image capture device 130 is oriented outwardly from the media device 102 so as to have a image capture field that corresponds to the user's location. In an exemplary embodiment, a set up process is used to set up the orientation of the image capture device 130 prior to presentation of the media content event. For example, but not limited to, images captured by the image capture device 130 may be presented on the display 120 so that the user may selectively orient the image capture device 130 in a desired direction. Some embodiments may permit the user to adjust the focus and/or aspect ratio of the captured images so that a higher quality user image is captured.

[0031] In some embodiments, the image capture device 130 is an integrated component that is built into the media device 102. Alternatively, or additionally, the image capture device 130 may be a separate device that is communicatively coupled to the media device 102. A separate image capture device 130 may be placed in any desired location in proximity to the user. The external image capture device 130 may communicate the captured images to the media device 102 via a suitable communication medium, such as a wire-based medium, an IR medium, and/or an RF medium.

[0032] The captured images of the user (or a plurality of users) are analyzed by the processor system 126, under management of the user image analysis logic 136, to determine or identify a physical response of the user to the currently presented portion of the media content event. For example, a user's physical response may include facial expressions, may include gestures,
and/or may include making sounds, such as shouts or the like. Exemplary embodiments are configured to identify the user's face in the captured image, and then to analyze facial features of the user. Based on the analyzed changes in the detected facial features of the user, a corresponding user emotion and/or user response may be determined. Then a corresponding emotion of the user may be determined such that a corresponding response (facial expression, responding gesture, and/or verbal commentary) for the animation of the avatar 124 may be determined. Any suitable facial analysis algorithm may be used by the various embodiments.

[0033] As another example, an exemplary embodiment may be configured to identify and/or analyze the user's gestures in the captured image, such as, but not limited to, clapping, hand waving, jumping, nodding or the like. Any suitable gesture analysis algorithm may be used by the various embodiments. Based on the identified and/or analyzed gestures of the user, a corresponding response (facial expression, responding gesture, and/or verbal commentary) for animation of the avatar 124 may be determined. In some embodiments, the avatar 124 may be generated so as to perform a similar gesture-type response.

[0034] For example, if the media content event is a sporting event and the favorite team of the user scores, the user may exhibit facial expressions that indicate that they are pleased with the score, such as a smile or raised eyebrow. Thus, the processor system 126 may identify a corresponding response (facial expression, responding gesture, and/or verbal commentary) of the avatar 124. Thus, based on the determined response of the user, the avatar 124 may then be animated based on the emotions determined from the verbal comments and/or sound characteristics. Thus, the animated avatar 124 may then appear to provide a corresponding (or alternatively, a contrary) facial expression, responding gesture, and/or verbal comment. For example, but not limited to, the avatar 124 may smile, clap in applause of the score, and/or shout "look at that great score!" Alternatively, the avatar presentation system 100 may be operating in a mode wherein the avatar 124 response is contrary to the determined user response. For example, the avatar 124, or even a different presented avatar 124, may frown, pull at their hair, and/or shout "no, not another score!"

[0035] Alternatively, or additionally, the media device 102 may include, or may be communicatively coupled to, an optional microphone 146. The microphone 146 detects sounds in the vicinity of the media device 102. The microphone 146 may be an integrated component of the
media device 102, or may be an external sound detecting device that is communicatively coupled to the media device 102.

[0036] Audio analysis logic may be included so that words spoken by the users in the sounds detected by the microphone 146 are identified and analyzed. In the various embodiments, the audio analysis logic is configured to differentiate sounds from the presented audio portion of the media content event and words spoken by the user. Speech recognition logic may be used to determine the words of the user's commentary, wherein the detected words may then be used to determine a verbal comment made by the user. The words and/or the verbal comment may then be associated with one or more corresponding user emotions. Alternatively, or additionally, emotions may be determined from other detected sound characteristics coming from the user. For example, but not limited to, levels of excitation or happiness of the user may be determined from the user's voice patterns and/or audible characteristics of the user's speech.

[0037] Based on the determined words and/or verbal comments from the user, and thus based on the user emotions determined there from, one or more audio phrases may be generated. The avatar 124 may then be animated based on the generated audio phrase. Thus, the animated avatar 124 may then appear to speak a verbal comment corresponding to the audio phrase. The audio phrase is combined with, or presented concurrently with, the presented audio portion of the media content event. Accordingly, the animated avatar 124 will appear to be making verbal commentary about the presented media content event that corresponds to (or alternatively, that is contrary to) the user's emotions. Additionally, or alternately, the avatar 124 may be animated to provide a corresponding (or alternatively, a contrary) facial expression and/or a responding gesture.

[0038] Alternatively, or additionally, the user may exhibit gestures that indicate that they are pleased with the score, such as clapping or the like. Thus, the processor system 126 may identify a corresponding response (facial expression, responding gesture, and/or verbal commentary) of the avatar 124. For example, if the user claps or jumps up from their chair when their favorite team scores, the avatar 124 may smile, clap in applause of the score, and/or shout "look at that great score!"

[0039] Some embodiments may be configured to base the response of the avatar 124 on characteristics of the currently presented portion of the media content event. An exemplary embodiment includes logic that analyzes audio and/or graphic characteristics of the media content event. For example, a score at a sporting event may be identified based on crowd cheering (from
audio analysis of the audio portion of the media content event) and/or may be based on a presented scoreboard that shows a change in score (from image analysis of the video portion of the media content event). Some embodiments may have special metadata included in the media content event. For example, but not limited to, the score may be included in the media content event as metadata.

[0040] In such embodiments, the user may specify certain conditions or events in the media content event that would be expected to elicit certain types of user responses. For example, if the media content event is a sporting event, the user may specify their favorite team. When their favorite team scores, the media content stream is analyzed to identify the scoring by the user's selected team, and then the avatar 124 is animated to exhibit characteristics that indicate that the avatar 124 is pleased with the score. Alternatively, when the opposing team scores, the avatar 124 may be animated to exhibit characteristics that indicate that the avatar 124 is disappointed with the score.

[0041] In some embodiments, the avatar presentation system 100 may receive specific avatar response information with the media content event. Animation of the avatar 124 is then based on predefined animation information in the avatar response information that is configured to animate the avatar 124. An exemplary embodiment receives the avatar response information in the metadata of the media content event using integrated technologies, such as, but not limited to, Hybrid broadcast broadband television (HbbTV) technologies. Alternatively, or additionally, the avatar response information may be communicated in other portions of the media content event data.

[0042] Alternatively, or additionally, the avatar response information may be separately communicated to the media device 102. For example, the media device 102 may be coupled to an external device 148 such that the avatar response information is provided by an Internet site or the like. The external device 148 may be synchronized with presentation of the media content event. Accordingly, the external device 148 may provide information that is used to generate a responsive avatar 124.

[0043] In some embodiments, the avatar response information received from the external device 148 may include sufficient information to fully generate the animated avatar 124. Accordingly, an avatar model would not need to be stored and/or selected from the avatar model database 138. In such embodiments, the avatar model database 138 may be optionally eliminated.

[0044] In some embodiments, the user may select the "nature" of the responses exhibited by the animated avatar 124. For example, the responses exhibited by the animated avatar 124 may
be subdued, exaggerated, etc. In some embodiments, the responses of the animated avatar 124 may correspond with the degree of response exhibited by the user. For example, a high level of enthusiasm detected in the user would cause the avatar presentation system 100 to animate the avatar 124 with a commensurate level of enthusiasm.

[0045] Avatar models in the avatar model database 138 may be updated from time to time in an exemplary embodiment. Avatar models may be received in the media content stream 104 and saved into the avatar model database 138. Alternatively, or additionally, the user may operate the media device 102 to receive avatar model updates from a remote source, such as an internet site or the like. Alternatively, or additionally, an avatar model may be automatically downloaded to the media device 102.

[0046] Further, specific characteristics of avatar models may be selected and/or may be communicated to the media device 102 based on characteristics of the presented media content event. The avatar characteristics may include clothing apparently worn by the avatar 124 and/or may include other objects apparently manipulated by the avatar 124. For example, if the media content event is a sporting event, then the avatar 124 may be clothed in sportswear associated with a team selected by the user. Alternatively, or additionally, the avatar 124 may be animated so as to wave a team flag or the like.

[0047] In some situations, a plurality of users may be at a common location viewing the presented media content event. Some embodiments may be configured to monitor a selected one of the users such that the avatar 124 is animated based on emotional responses of the monitored user. For example, the image capture device 130 may be oriented and focused on a particular chair where the monitored user is sitting. In another embodiment, user recognition logic may be used to identify the monitored user.

[0048] Alternatively, multiple users may be monitored such that the avatar 124 is animated based on one or more common identified emotional responses of the monitored users. For example, if all of, or a majority of users are excited by a particular play of a presented sporting event, the presented avatar 124 may be animated to become excited.

[0049] Alternatively, or additionally, two or more users may be separately monitored. Accordingly, a unique avatar 124 for each monitored user is animated based on the identified emotional response of the associated monitored user.
[0050] In some situations, personal information pertaining to a user may be available. The information may be received in the media content stream 104, received from an external device 148, and/or may be stored in the memory 128 of the media device. For example, but not limited to, the personal information may be the user's name and/or nickname. Such information may be available from account information associated with the user and/or the media device 102. In some embodiments, the personal information may be user-entered and stored into the memory 128. Accordingly, an audio phrase generated and presented (spoken by the avatar 124) may include the user's name or nickname, or other related personal information. For example, the user's name may be Igor. During a presented sporting event, the avatar 124 may be animated to appear to say that "hey Igor, did you see that amazing score?" As another non-limiting example, the avatar 124 might be animated to appear to say "your cousin Bob will surely dislike that score!" As another non-limiting example, the user may have a favorite team or the like such that the animation of the avatar 124 is based on corresponding team sportswear or the like.

[0051] In some embodiments, the animated avatar 124 may be configured to change their position and/or posture. For example, if the avatar 124 is apparently speaking, the avatar 124 may temporarily turn around and face the user while the speaking of the avatar 124 is simulated. As another example, the avatar 124 may be animated so as to appear to be following the scenes of the presented media content event. For example, if the media content event is a sporting event, the presented avatar 124 may be generated so as to appear to turn their heads to follow the currently presented play.

[0052] Some media presentation systems 106 are configured with three dimensional (3-D) displays 120 that are configured to present 3-D media (along with standard media). With such embodiments, 3-D avatar models may be used to generate and present one or more 3-D avatars 124 on the 3-D display 120.

[0053] Some media presentation systems 106 may be provisioned with holographic capabilities. In such embodiments, the animated avatars 124 may be presented using the holographic technologies such that the avatar 124 is more realistically presented to the user. In some embodiments, a separate holographic projector may be used to present the avatar 124 separately from the presented media content event. In such embodiments, the visual display device 114 may be used to present the media content event, while the holographic projector is used to present one or more holographic avatars 124 at selected locations in the media presentation room.
[0054] FIGURE 2 is a block diagram illustrating greater detail of an embodiment of an example media device 102. The non-limiting exemplary media device 102 comprises a media content stream interface 202, a processor system 126, a memory 128, a program buffer 204, an optional digital video recorder (DVR) 206, a presentation device interface 208, a remote interface 210, an optional communication system interface 212, an optional microphone 146, and an optional image capture device 130. The memory 128 comprises portions for storing the media device logic 132, the optional avatar management logic 134, the user image analysis logic 136, and the avatar model database 138. In some embodiments, the media device logic 132, the avatar management logic 134, and/or the user image analysis logic 136 may be integrated together, and/or may be integrated with other logic. In other embodiments, some or all of these memory and other data manipulation functions may be provided by and using remote server or other electronic devices suitably connected via the Internet or otherwise to a client device. Other media devices 102 may include some, or may omit some, of the above-described media processing components. Further, additional components not described herein may be included in alternative embodiments.

[0055] In embodiments without the image capture device 130, an image capture device interface (not shown) may be provided to couple the media device 102 to an external image capture device 130. Similarly, in embodiments without the microphone 146, a microphone interface (not shown) may be provided to couple the media device 102 to an external image microphone 146. The image capture device interface and/or the microphone interface may be any suitable interface configured to communicatively couple the media device and the image capture device 130 and/or the microphone 146. Any communication medium may be used, such as a wire-based communication medium, an RF communication medium, and/or an IR communication medium. In some embodiments, the interfaces may be configured to communicate with multiple and/or different devices and/or systems.

[0056] The functionality of the media device 102, here a set top box, is now broadly described. A media content provider provides media content that is received in one or more multiple media content streams 104 multiplexed together in one or more transport channels. The transport channels with the media content streams 104 are communicated to the media device 102 from a media system sourced from a remote head end facility (not shown) operated by the media content provider. Non-limiting examples of such media systems include satellite systems, cable system, and the Internet. For example, if the media content provider provides programming via a satellite-based
communication system, the media device 102 is configured to receive one or more broadcasted satellite signals detected by an antenna (not shown). Alternatively, or additionally, the media content stream 104 can be received from one or more different sources, such as, but not limited to, a cable system, a radio frequency (RF) communication system, or the Internet, or even an external device, such as a digital video disc (DVD) player, a video cassette recorder (VCR) or other memory medium devices that are configured to provide the media content stream 104.

[0057] The one or more media content streams 104 are received by the media content stream interface 202. One or more tuners 214 in the media content stream interface 202 selectively tune to one of the media content streams 104 in accordance with instructions received from the processor system 126. The processor system 126, executing the media device logic 132 and based upon a request for a media content event of interest specified by a user, parses out media content associated with the media content event of interest. The media content event of interest is then assembled into a stream of video and/or audio information which may be stored by the program buffer 204 such that the media content can be streamed out to the components of the media presentation system 106, such as the visual display device 114 and/or the audio presentation device 116, via the presentation device interface 208. Alternatively, or additionally, the parsed out media content may be saved into the DVR 206 for later presentation. The DVR 206 may be directly provided in, locally connected to, or remotely connected to, the media device 102.

[0058] The user not shown may view and listen to a selected media content event when presented on the exemplary display device 114 and/or the audio presentation device 116. That is, based upon the user commands, typically generated at and transmitted from the remote control 140 as a wireless signal 144 that is received by the remote interface 210, the media device 102 can then control itself and/or other various media devices that it is communicatively coupled to. Accordingly, a selected media content event is presented in accordance with the generated user commands. Further, one or more avatars 124 may be presented on the display device 114, and/or sounds associated with the one or more of the presented avatars 124 may be reproduced as sounds by the audio presentation device 116.

[0059] The above processes performed by the media device 102 are generally implemented by the processor system 126 while executing the media device logic 132. Thus, the media device 102 may perform a variety of functions related to the processing and presentation of one or more media content events received in the media content stream 104.
In an exemplary embodiment, information used to generate and present the one or more avatars 14 may be received from one or more external devices 148 to which the communication system interface 212 is coupled, via a communication system 216. Examples of the external devices 148 include, but are not limited to, a portable media device, a computer, a smart phone, an Internet site, or the like. The external devices 148 comprises a memory medium 218 that is configured to store models for the avatars 124 and/or other logic that controls the generation and management of the avatars 124. Thus, a completely animated avatar 124 that is presented concurrently with a media content event may be received by the media device 102, thereby reducing the computational effort required of the processor system 126 of the media device 102 for presentation of animated avatars 124.

The communication system 216 is illustrated as a generic communication system. In one embodiment, the communication system 216 comprises a cellular telephone system, such as a radio frequency (RF) wireless system. Accordingly, the media device 102 includes a suitable transceiver. Alternatively, the communication system 216 may be a telephony system, the Internet, a Wi-fi system, a microwave communication system, a fiber optics system, an intranet system, a local access network (LAN) system, an Ethernet system, a cable system, a radio frequency system, a cellular system, an infrared system, a satellite system, or a hybrid system comprised of multiple types of communication media. Additionally, embodiments of the media device 102 may be implemented to communicate using other types of communication technologies, such as but not limited to, digital subscriber loop (DSL), X.25, Internet Protocol (IP), Ethernet, Integrated Services Digital Network (ISDN) and asynchronous transfer mode (ATM). Also, embodiments of the media device 102 may be configured to communicate over combination systems having a plurality of segments which employ different formats for each segment that employ different technologies on each segment.

It should be emphasized that the above-described embodiments of the avatar presentation system 100 are merely possible examples of implementations of the invention. Many variations and modifications may be made to the above-described embodiments. All such modifications and variations are intended to be included herein within the scope of this disclosure and protected by the following claims.
Claims:

1. A method for enhancing a user's experience during presentation of media content events, the method comprising:
   - outputting a media content event from a media device to a media presentation system, the media content event comprising a video portion and an audio portion, and wherein the media presentation system is configured to present the media content event to at least one user;
   - determining at the media device an emotional response of the at least one user to a currently presented portion of the media content event;
   - animating an avatar to present an emotion corresponding to the determined emotional response of the at least one user; and
   - outputting the animated avatar to the media presentation system such that the animated avatar is presented concurrently with at least the video portion of the media content event.

2. The method of Claim 1, further comprising:
   - combining a video portion of the animated avatar with the video portion of the media content event at the media device, wherein the combined video portions of the animated avatar and the media content event are communicated to the media presentation system.

3. The method of Claim 1, wherein determining the emotional response of the at least one user to the presented media content event comprises,
   - capturing at least one image of the at least one user with an image capture device;
   - identifying a physical response of the at least one user from the at least one captured image at the media device; and
   - determining the emotional response of the at least one user based on the identified physical response of the at least one user.

4. The method of Claim 3, wherein determining the physical response of the at least one user and animating the avatar comprises:
identifying a facial expression of the at least one user from the at least one captured image; and
animating the avatar to have a facial expression corresponding to the identified facial expression of the user.

5. The method of Claim 3, wherein determining the physical response of the at least one user and animating the avatar comprises:
   identifying a gesture made by the at least one user from the at least one captured image; and
   animating the avatar to make a gesture corresponding to the determined emotional response of the at least one user.

6. The method of Claim 1, wherein determining the emotional response of the at least one user to the presented media content event comprises,
   detecting sounds with a microphone;
   identifying a words spoken by the at least one user from the detected sounds at the media device; and
   determining the emotional response of the at least one user based on the identified words.

7. The method of Claim 6, further comprising:
   generating an audio phrase corresponding to the identified words;
   combining the audio phrase with the audio portion of the media content event, wherein the combined audio portion of the media content event and phrase of the animated avatar is communicated to the media presentation system; and
   animating the avatar to appear to speak the audio phrase.

8. The method of Claim 7, generating the audio phrase further comprising:
   adding at least one word corresponding to a personal fact pertaining to the user into the generated audio phrase.

9. The method of Claim 7, further comprising:
   animating the avatar to make at least one facial expression corresponding to the identified words.
10. The method of Claim 7, further comprising:
animating the avatar to make a gesture corresponding to the identified words.

11. The method of Claim 1, wherein the avatar emotion corresponding to the determined emotional response of the at least one user is contrary to the determined emotional response of the at least one user.

12. The method of Claim 1, wherein the avatar is a first avatar presenting a first emotion, and further comprising:
animating a second avatar to present a second emotion, wherein the avatar emotion corresponds to the determined emotional response of the at least one user; and
outputting the animated second avatar to the media presentation system such that the animated second avatar is presented concurrently with the video portion of the media content event and the first avatar.

13. The method of Claim 1, further comprising:
receiving avatar response information in a media content stream, wherein the received avatar response information comprises predefined animation information configured to animate the avatar; and
animating the avatar to present a predefined animation information.

14. A media device, comprising:
a media content stream interface configured to receive a media content event;
a presentation device interface configured to communicate a video portion and an audio portion of the media content event to a media presentation system, wherein the media presentation system is configured to present the media content event to at least one user;
a memory configured to store at least one model of an avatar; and
a processor system communicatively coupled to the media content stream interface, the presentation device interface and the memory, wherein the processor system is configured to:
determine an emotional response of the at least one user to a currently presented portion of the media content event;
retrieve the avatar model from the memory;
animate the avatar model to present an emotion, wherein the avatar emotion corresponds to the determined emotional response of the at least one user; and output the animated avatar to the media presentation system such that the animated avatar is presented concurrently with at least the video portion of the media content event.

15. The media device of Claim 14, further comprising:
an image capture device communicatively coupled to the processor and configured to capture at least one image of the user, wherein the processor system is further configured to:
identify a facial expression of the at least one user from the at least one captured image; and
animate the avatar to have a facial expression corresponding to the identified facial expression of the user.

16. The media device of Claim 14, further comprising:
a microphone communicatively coupled to the processor and configured to detect sounds in proximity to the media device, wherein the processor system is further configured to:
identify words spoken by the at least one user from the detected sounds;
generate an audio phrase corresponding to the identified words; and
determine the emotional response of the at least one user based on the identified words.

17. The media device of Claim 14, further comprising:
a remote interface communicatively coupled to the processor system and configured to receive a wireless signal from a remote control, wherein the processor system is further configured to:
receive an avatar signal transmitted by the remote control, wherein the avatar model is retrieved from the memory for presentation in response to receiving the avatar presentation signal.
18. The media device of Claim 14, further comprising:
   a remote interface communicatively coupled to the processor system and configured to
   establish a communication link from the media device to an external device,
   wherein the processor system is further configured to:
   process another avatar model received from the external device.

19. A method for enhancing a user's experience during presentation of media content events,
    the method comprising:
    capturing an image of a user with an image capture device;
    determining an emotional response of the user to a currently presented portion of the
    media content event based on the captured image of the user;
    animating an avatar to simulate an emotion corresponding to the determined emotional
    response of the user; and
    outputting the animated avatar to the media presentation system such that the animated
    avatar is presented concurrently with a video portion of the media content event.

20. The method of Claim 19, further comprising:
    generating an audio phrase associated with the determined emotional response of the
    user;
    further animating the avatar such that the avatar appears to speak the phrase;
    outputting the audio phrase to the media presentation system such that sound
    corresponding to the audio phrase is presented concurrently with an audio portion of
    the media content event; and
    animating the avatar to appear to speak the audio phrase.
INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N21/422 H04N21/4223 H04N21/81 H04N21/431
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04N G06N G06K

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal , WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>paragraphs [0022] - [0025] , [0028] , [0031] ; figure 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>paragraphs [0240] - [0246] , [0254]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>----</td>
<td></td>
</tr>
</tbody>
</table>

Date of the actual completion of the international search: 6 July 2012

Date of mailing of the international search report: 16/07/2012

Name and mailing address of the ISA/Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016
Goy, Sebastien

* Special categories of cited documents:

*"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document relating to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

*"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"A" document member of the same patent family

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

Form PCT/ISA/210 (second sheet) (April 2003)
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>GORANKA ZORIC ET AL: &quot;On creating multimodal virtual humans real time speech driven facial gesture&quot;, MULTIMEDIA TOOLS AND APPLICATIONS, KLUWER ACADEMIC PUBLISHERS, BO, vol 54, no. 1, 29 April 1 2010 (2010-04-29), pages 165-179, XP019909826, ISSN: 1573-7721, DOI: 10.1007/S11042-010-0526-Y</td>
<td>7, 9, 10, 16, 20</td>
</tr>
<tr>
<td>Patent document cited in search report</td>
<td>Publication date</td>
<td>Patent family member(s)</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>-----------------</td>
<td>-------------------------</td>
</tr>
</tbody>
</table>