

METHOD AND APPARATUS FOR REPLACEMENT OF AUDIO DATA IN A RECORDED AUDIO/VIDEO STREAM

BACKGROUND

[0001] Since the advent of the video cassette recorder (VCR), viewers of transmitted video and audio programming have been able to privately record their favorite programs. As a result, such technology has allowed the viewer to “time-shift” television programs of interest to a subsequent, more convenient time. More recently, digital video recorder (DVR) units, often provided in satellite and cable television receivers, or “set-top boxes,” have supplied the viewer a more convenient means of time-shifting audio/video streams by allowing a larger number of programs to be recorded in a more automated fashion without the use of removable storage media.

[0002] Given this time-shifting capability and the amount of data storage available in such devices, users often do not view a recorded program or other audio/video stream until several days, or even weeks, after the stream was first recorded. As a result, some information provided in the audio/video stream, such as the subject matter of the program of interest, or even the informational content of a commercial, may be outdated by the time the user finally views the recorded stream.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] Fig. 1 is a flow diagram of a method according to an embodiment of the invention for replacing audio data within a recorded audio/video stream.

[0004] Fig. 2 is a block diagram of a system according to an embodiment of the invention configured to replace audio data within a contiguous block of audio/video data.

[0005] Fig. 3 is a graphical representation of an audio/video stream or contiguous block of audio/video data, and a version of the audio/video stream that is updated with replacement audio data, employing the system of Fig. 2.

[0006] Fig. 4 is a graphical representation of the use of text data of an audio/video stream to identify a location in the audio/video stream for replacement of audio data according to an embodiment of the invention.

[0007] Fig. 5 is a block diagram of a receiving device according to an embodiment of the invention for replacing audio data within a recorded audio/video stream.

[0008] Fig. 6 is a block diagram of a system in which multiple receiving devices are communicatively coupled to a communication network according to an embodiment of the invention.

DETAILED DESCRIPTION

[0009] Fig. 1 provides a flow diagram of a method 100 for replacing audio data within a recorded audio/video stream. Generally, an audio/video stream is a contiguous block of audio data and associated video data that may be transmitted to, and received by, an electronic device, such as a terrestrial (“over-the-air”) television receiver, a cable television receiver, a satellite television receiver, a computer, a portable electronic device, or the like. The video data of the audio/video stream may include a contiguous series of video frames, while the audio data of the audio/video stream may include consecutive samples of audio information. Further, the audio/video stream may delivered by any transmission method, such as broadcast, multicast, simulcast, closed circuit, pay-per-view, point-to-point (by “streaming,” file transfer, or other means), or other method. Additionally, the audio/video stream may be transmitted by way of any communication technology, such as by satellite, wire or optical cable, wireless, or other means. The audio/video stream may also be transferred over any type of communication network, such as the Internet or other wide area network, a local area network, a private network, a mobile communication system, a terrestrial television network, a cable television network, and a satellite television network.

[0010] In the method 100, a first audio/video stream that includes audio data, video data, and supplemental data is stored (operation 102). Also, an audio data segment is received (operation 104), as is location information (operation 106). The

location information references the supplemental data of the first audio/video stream to identify a location within the stream. Also, the location information is received in a data file separately from the first audio/video stream. A portion of the audio data of the first audio/video stream is replaced at the identified location with the audio data segment to produce a second audio/video stream (operation 108). At least a portion of the second audio/video stream is transferred for presentation (operation 110).

[0011] Fig. 2 presents an annotated block diagram of a system 200 for replacing audio data within a contiguous block of audio/video data according to another embodiment of the invention. The contiguous block is alternatively termed an audio/video stream, as employed above. The system 200 includes a communication network 202, a receiving device 210, and a display device 214.

[0012] The communication network 202 may be any communication network capable of transmitting an audio/video stream, such as a terrestrial over-the-air television network, a cable television network, a satellite television network (such as a direct broadcast satellite (DBS) system), a wide area network (WAN) (such as the Internet), a local area network (LAN), or the like. In another embodiment, the communication network 202 may be a combination of two or more network types listed above. Also, multiple components possibly included in the communication network 202, such as transmitters, repeaters, transponders, routers, and the like, are not shown in Fig. 2 to facilitate brevity in the following discussion.

[0013] The receiving device 210 of Fig. 2 may be any device capable of receiving an audio/video stream from the communication network 202. For example, in the case of the communication network 202 being a cable or satellite television network, the receiving device 210 may be a set-top box configured to communicate with the communication network 202. In another example, the receiving device 202 may be computer, a personal digital assistant (PDA), or similar device configured to communicate with the Internet or comparable communication network 202.

[0014] The display device 214 may be any device configured to receive an audio/video stream from the receiving device 210. Examples of the display device 214 include a television, a video monitor, or similar device capable of presenting audio and video information to a user. In some implementations, the display device

214 may be integrated within the receiving device. 210. For example, each of a computer, a PDA, and a mobile communication device may serve as both the receiving device 210 and the display device 214 by providing the capability of receiving audio/video streams from the communication network 202 and presenting the received audio/video streams to a user. In another implementation, a cable-ready television may include a converter device for receiving audio/video streams from the communication network 202 and displaying the streams to a user.

[0015] In the system 200, the communication network 202 transmits each of a first audio/video stream 204, an audio data segment 206, and location information 208 to the receiving device 210. The first audio/video stream 204 includes audio data and video data. In one embodiment, the video data includes a series of digital frames, or single images to be presented in a serial fashion to a user. Similarly, the audio data may be composed of a series of audio samples to be presented simultaneously with the video data to the user. In one example, the audio data and the video data may be formatted according to one of the Motion Picture Experts Group (MPEG) encoding standards, such as MPEG-2 or MPEG-4, as may be used in DBS systems and terrestrial Advanced Television Systems Committee (ATSC) systems. However, different audio and video data formats may be utilized in other implementations.

[0016] Also included in the first audio/video stream 204 is supplemental data providing information relevant to the audio data, the video data, or both, of the first audio/video stream 204. In one implementation, the supplemental data includes text data, such as closed captioning data, available for visual presentation to the user during the presentation of the associated audio and video data of the audio/video data stream 204. In one example, the text data may conform to any text data or closed captioning standard, such as the Electronic Industries Alliance 708 (EIA-708) standard employed in ATSC transmissions. When the text data is available to the display device 214, the user may configure the display device 214 to present the text data to the user in conjunction with the video data.

[0017] Each of a number of portions of the text data may be associated with a corresponding portion of the audio data or video data also included in the audio/video stream 204. For example, one or more frames of the video data of the audio/video stream 204 may be specifically identified with a portion of the text data included in

the stream 204. As a result, multiple temporal locations within the audio/video stream 204 may be identified by way of an associated portion of the text data. For example, a particular text string or phrase within the text data may be associated with one or more specific frames of the video data within the same audio/video stream 204 so that the text string is presented to the user simultaneously with its associated video data frames. Therefore, the particular text string or phrase may provide an indication of a location of these video data frames, as well as the portion of the audio data synchronized or associated with the frames.

[0018] In another embodiment, the supplemental data may include timestamp data, wherein some or all of the video frames or audio data are associated with one of the timestamps. In one embodiment, some or all of the video frames are “marked” or labeled serially in presentation order. In this case, the supplemental data may also include a unique program identifier (UPID) for each separate program or event. Each audio/video stream may include one or more such programs and, thus, one or more UPIDs. Therefore, timestamp data, possibly in conjunction with a UPID, may also be utilized to identify a particular location within the first audio/video stream 204. The use of supplemental data to specify a location within the first audio/video stream 204 is discussed in greater detail below.

[0019] The communication network 202 also transmits an audio data segment 206 and location information 208 to the receiving device 210. Generally, the audio data segment 206 is provided to replace or supplant some or all of the audio data of the first audio/video stream 204 at a location within the stream 204 designated by the location information 208. Fig. 3 provides a graphical representation of the first audio/video stream 204 received by the receiving device 210, and a second audio/video stream 212, as indicated in the system 200 of Fig. 2. The first audio/video stream 204 includes video data 302 and audio data 304. Also indicated is an identified location 306 within the first audio/video stream 204, which is indicated by way of the location information 208. The supplemental data of the audio/video stream 204 is not shown in Fig. 3 to simplify the diagram.

[0020] In the specific example of Fig. 3, the identified location 306 is the starting point at which the audio data segment 206 is to replace a portion of the audio data 304. In Fig. 3, the portion of the audio data 304 to be replaced is a second

portion 304B, located after a first audio data portion 304A and before a third audio data portion 304C. While Fig. 3 employs the use of an starting point for the audio data segment 206 as the identified location 306 within the first audio/video stream 204, other locations, such as an ending point for the audio data segment 206, or both a starting point and an ending point, may be used in other implementations.

[0021] As a result of this audio replacement, a second audio/video stream 212 is produced, in which the audio data segment 206 replaces the second audio data portion 304B from the first audio/video stream 204. In one implementation, the audio data segment 206 is formatted according to the same formatting or encoding standard represented by the second audio data portion 304B. The remainder of the second audio/video stream 212 includes the video data 302, the first audio data portion 304A, and the third audio data portion 304C from the first audio/video stream 204.

[0022] To specify the identifying location 306, the location information 208 of Fig. 2 references a portion of the supplemental data of the first audio/video stream 204, such as the text data and/or timestamp data described above. In the case of text data, the text data of the first audio/video stream 204 may be perused to select a substantially unique portion or string within the text data that may be unambiguously detected at the receiving device 210. The text data may consist of a single character, several characters, an entire word, multiple consecutive words, or the like. In one embodiment, the selected text string may be located within the first audio/video stream 204 close to the location at which the audio data segment 206 is to reside. In one implementation, a human operator bears responsibility for selecting the text string. In other examples, the text string selection occurs automatically under computer control, or by way of human-computer interaction.

[0023] A node within the communication network 202 may then transmit the selected text string to the receiving device 210 as the location information 208. Further, if the selected text string is located a significant distance from the location to be indicated within the first audio/video stream 204, the location information 208 may include an offset. Fig. 4 depicts a situation in which a selected text string 412 within text data 408 of an audio/video stream 400 is located at string location 412, which is located in the audio/video stream 400 after the identified location 406 where replacement of a portion of the audio data 404 is to occur. An offset 414 may thus

indicate the distance within the stream 204 between the location 412 of the selected text string 410 and the location 406 at which an audio data segment is to be placed. In this specific example, the direction of the offset 414 is toward the starting end of the audio/video stream 400, which may be indicated by way of a negative offset value. Oppositely, situations in which the selected text string is located in the associated audio/video stream prior to the identified location at which audio replacement is to occur may be indicated via a positive offset value.

[0024] In one example, the offset of the selected text string from the identified location may be stated relative to the video data of the audio/video stream, as the text data of the stream is often presented visually. However, the location of the selected text string may be stated relative to the audio data of the stream, or the stream in general, in other implementations. The offset may be expressed in any convenient unit capable of indicating a distance along the audio/video stream, such as a number of video frames, a length of time, or other value.

[0025] Returning to Figs. 2 and 3, instead of employing text data, timestamp data within the supplemental data of the first audio/video stream 204 may be employed to specify the identified location 306 at which the audio data segment 206 is to be placed in the audio data 304 of the first audio/video stream 204. In one embodiment, each video frame of the first audio/video stream 204, or some subset of the video frames, may be associated with a unique timestamp. Further, the supplemental data may include a UPID identifying a particular program or broadcast event constituting at least a portion of the first audio/video stream 204. Thus, the UPID and the timestamp data may be employed in tandem to indicate the identified location 306 as specified in the location information 208.

[0026] Depending on the resiliency and other characteristics of the supplemental data, the node of the communication network 202 generating and transmitting the location information 208 may issue more than one instance of the location information 208 to the receiving device 210. For example, text data, such as closed captioning data, is often error-prone due to limited error correction capability protecting the text data. As a result, the receiving device 210 may not be able to detect some of the text data, including the text data selected for specifying the identified location 306. To address this issue, multiple unique text strings may be

selected from the text data of the first audio/video stream 204 to indicate the identified location 306. The use of multiple text strings (each possibly accompanied with its own offset) may thus result in multiple sets of location information 208 transmitted over the communication network 202 to the receiving device, each of which is associated with the same audio data segment 206. Each set of location information 208 may be issued separately, or may be transmitted with one more other sets.

[0027] The audio data segment 206 and the location information 208 may be logically associated with one another to prevent incorrect association of the location information 208 with other audio data segments 206 being received at the receiving device 210. To this end, the audio data segment 206 may include an identifier or other indication associating the audio data segment 206 with its appropriate location information 208. Conversely, the location information 208 may include such an identifier, or both the audio data segment 206 and the location information 208 may do so. Use of an identifier may be appropriate in the case the audio data segment 206 and the location information 208 are transmitted separately, such as in separate data files. In another embodiment, the audio data segment 206 and the location information 208 may be packaged within the same transmission or data file to the receiving device 210 so that the receiving device 210 may identify the location information 208 with the audio data segment 206 on that basis.

[0028] Further, both the audio data segment 206 and the location information 208 may be associated with the first audio/video stream 204 to prevent any incorrect association of these data with another audio/video stream. Thus, an identifier, such as that discussed above, may be included with the first audio/video stream 204 to relate the stream 204 to its audio data segment 206 and location information 208. In one particular example, the identifier may be a UPID, mentioned earlier. Use of an identifier in this context addresses situations in which the audio data segment 206 and the location information 208 are created after the first audio-video stream 204 has been transmitted over the communication network 202 to the receiving device 210. In another scenario, the audio data segment 206 and the location information 208 may be available for transmission by the time the first audio/video stream 204 is transmitted. In this case, the communication network 202 may transmit the audio data segment 206 and the location information 208 with the first audio/video stream 204, thus

associating all three sets of data 204, 206, 208. In one example, the audio data segment 206 and the location information 208 may be included as supplemental data within the first audio/video stream 204.

[0029] Once the receiving device 210 has received the first audio/video stream 204, the audio data segment 206, and the location information 208, the receiving device 210 may process this information to produce a second audio/video stream 212 for presentation to the display device 214. More specifically shown in Fig. 3, the second audio/video stream 212 includes the video data 302 of the first audio/video data stream 204, as well as most of the audio data 304 for the first stream 204. The portion of the audio data 304 located at the identified location 306 specified in the location information 308 is then replaced in the second audio/video stream 212 with the audio data segment 206. In the specific example of Fig. 3, the replaced portion of the audio data 304 is the second portion 304B, as described earlier. In other situations, other portions of the audio data 304, including all of the audio data 304, may be replaced. In other implementations, multiple segments of the audio data 304 may be supplanted, possibly by different audio data segments 206, and indicated by separate sets of location information 208.

[0030] A more explicit view of a receiving device 510 according to one embodiment is portrayed in Fig. 5. The receiving device 510 includes a communication interface 502, a storage unit 516, an audio/video interface 518, and control logic 520. In some implementations, a user interface 522 may also be employed in the receiving device 510. Other components possibly included in the receiving device 510, such as demodulation circuitry, decoding logic, and the like, are not shown explicitly in Fig. 5 to facilitate the following discussion.

[0031] The communication interface 502 may include circuitry to receive a first audio/video stream 504, an audio data segment 506, and location information 508. For example, if the receiving device 510 is a satellite set-top box, the communication interface 502 may be configured to receive satellite programming, such as the first audio/video stream 402, via an antenna from a satellite transponder. If, instead, the receiving device 510 is a cable set-top box, the communication interface 502 may be operable to receive cable television signals and the like over a coaxial cable. In either case, the communication interface 502 may receive the audio

data segment 506 and the location information 508 by employing the same technology used to receive the first audio/video stream 402. In another implementation, the communication interface 502 may receive the audio data segment 506 and the location information 508 by way of another communication technology, such as the Internet, a standard telephone network, or other means. Thus, the communication network 502 may employ one or more different communication technologies, including wired and wireless communication technologies, to communicate with a communication network, such as the communication network 202 of Fig. 2.

[0032] Coupled with the communication interface 502 is a storage unit 516, which is configured to store both the first audio/video stream 504 and the audio data segment 506. The storage unit 516 may include any storage component configured to store one or more such audio/video streams. Examples include, but are not limited to, a hard disk drive, an optical disk drive, and flash semiconductor memory. Further, the storage unit 516 may include either or both volatile and nonvolatile memory.

[0033] Communicatively coupled with the storage unit 516 is an audio/video interface 518, which is configured to transfer audio/video streams from the receiving unit 510 to a display device 514 for presentation to a user. The audio/video interface 518 may incorporate circuitry to transfer the audio/video streams in any format recognizable by the display device 514, including composite video, component video, the Digital Visual Interface (DVI), and the High-Definition Multimedia Interface (HDMI). The audio/video interface 518 may also incorporate circuitry to support multiple types of these or other audio/video formats. In one example, the display device 514, such as a television monitor or similar display component, may be incorporated within the receiving device 510, as indicated earlier.

[0034] In communication with the communication interface 502, the storage unit 516, and the audio/video interface 518 is control logic 520 configured to control the operation of each of these three components 502, 516, 518. In one implementation, the control logic 520 includes a processor, such as a microprocessor, microcontroller, digital signal processor (DSP), or the like for execution of software configured to perform the various control functions described herein. In another embodiment, the control logic 520 may include hardware logic circuitry in lieu of, or

in addition to, a processor and related software to allow the control logic 520 to control the other components of the receiving device 510.

[0035] Optionally, the control logic 520 may communication with a user interface 522 configured to receive user input 523 directing the operation of the receiving device 510. The user input 523 may be generated by way of a remote control device 524, which may transmit the user input 523 to the user interface 522 by the use of, for example, infrared (IR) or ultra-high frequency (UHF) signals. In another embodiment, the user input 523 may be received more directly by the user interface 522 by way of a touchpad or other manual interface incorporated into the receiving device 510.

[0036] The receiving device 510, by way of the control logic 520, is configured to receive the first audio/video stream 504 by way of the communication interface 502, and store the audio/video stream 504 in the storage unit 516. The receiving device 510 is also configured to receive the audio data segment 506 over the communication interface 502, possibly storing the audio data segment 506 in the storage unit 516 as well. The location information 508 is also received at the communication interface 502, which may pass the location information 508 to the control logic 520 for processing. In another embodiment, the location information 508 may be stored in the storage unit 516 for subsequent retrieval and processing by the control logic 520.

[0037] At some point after the location information 508 is processed, the control logic 520 generates and transmits a second audio/video stream 512 over the audio/video interface 518 to the display device 514. In one embodiment, the control logic 520 generates and transmits the second audio/video stream 512 in response to the user input 523. For example, the user input 523 may command the receiving device 510 to transfer the first audio/video stream 504 to the display device 514 for presentation. In response, the control logic 520 instead generates and transmits the second audio/video stream 512. As described above in conjunction with Fig. 2, the second audio/video stream 512 includes the audio and video data of the first audio/video stream 504, but with the audio data segment 506 replacing the original audio data of the first audio/video stream 504 at the location specified in the location

information 508, as described in detail above in conjunction with Fig. 3 with respect to the first audio/video stream 204 of Fig. 2.

[0038] Depending on the implementation, the second audio/video stream 512 may or may not be stored as a separate data structure in the storage unit 516. In one example, the control logic 520 generates and stores the entire second audio/video stream 512 in the storage unit 516. The control logic 520 may further overwrite the first audio/video stream 504 with the second audio/video stream 512 to save storage space within the storage unit 516. Otherwise, both the first audio/video stream 504 and the second audio/video stream 512 may reside within the storage unit 516.

[0039] In another implementation, the second audio/video stream 512 may not be stored separately within the storage unit 516. For example, the control logic 520 may instead generate the second audio/video stream 512 “on the fly” by transferring the audio data and the video data of the first audio/video stream 504 in presentation order from the storage unit 516 to the audio/video interface 518. At the point at which the audio data indicated by the location information 508 is to be transferred, the control logic 520 may then cause the audio data segment 506 to be transmitted from the storage unit 516, thus replacing the corresponding audio data from the first stream 504 at the audio/video interface 518. Once the last of the audio data segment 506 has been transferred from the storage unit 516, the control logic 520 may cause the remainder of the original audio data of the first stream 504 not replaced by the audio data segment 506 to be transferred to the audio/video interface 518 for presentation to the display device 514.

[0040] In one implementation, a user may select by way of the user input 523 whether the first audio/video stream 504 or the second audio/video stream 512 is transferred to the display device 514 by way of the audio/video interface 518. In another embodiment, a content provider of the first audio/video stream 504 may prevent the user from maintaining such control by way of additional information delivered to the receiving device 510.

[0041] If more than one audio data segment 506 is available in the storage unit 516 to replace a specified portion of the audio data of the first audio/video stream 504, the user may select via the user input 523 which of the audio data segments 506 are to

replace the corresponding portion of the audio data of the first stream 504 upon transmission to the display device 514. Such a selection may be made in a menu system incorporated in the user interface 522 and presented to the user via the display device 514.

[0042] In a broadcast environment, such as that depicted in the system 600 of Fig. 6, multiple receiving devices 610 may be coupled with a communication network 602 to receive audio/video streams, any of which may be recorded, in whole or in part, by any of the receiving devices 610. In conjunction with any number of these audio/video streams, audio data segments serving as partial or total replacements for audio data in the streams, as well as the location information for each of the audio data segments, may be transferred to the multiple receiving units 610. In response to receiving the audio/video streams, each of the receiving units 610 may record any number of the audio/video streams received. For any audio data segments and associated location information that are transmitted over the communication network 602, each receiving device 610 may then review whether the received audio data segments and location information is associated with an audio/video stream currently stored in the device 610. If the associated stream is not stored therein, the receiving device 610 may delete or ignore the related audio data segment and location information received.

[0043] In another embodiment, instead of broadcasting each possible audio data segment and related location information, the transfer of an audio/video stream stored within the receiving unit 614 to an associated display device 614 may cause the receiving unit 614 to query the communication network 602 for any outstanding replacement audio data segments that apply to the stream to be presented. As a result, the broadcasting of each audio data segment and related location information would not be required, thus potentially reducing the amount of consumed bandwidth over the communication network 602.

[0044] Various embodiments as described herein may provide any of a number of benefits. Generally, the methods and systems disclosed above allow replacement of all or a portion of the audio content of an audio/video stream, such as a program or commercial message, after the audio/video stream has been recorded by the user. For example, a commercial message announcing a “this weekend only” sale that is

originally broadcast during a recorded program may be obsolete by the time the user views the recorded program the following week. To address this problem, the audio portion of the commercial may be replaced with updated audio data announcing a future sale date or other more recent information. The same may be true of information presented in news, weather, and sports programs, as well as other types of audio/video streams.

[0045] Under another scenario, some programs may contain language that some users deem offensive or objectionable. To render the program palatable to a wider range of viewers, the content provider may make alternative audio segments of the audio portion of the program available. A user who has recorded the program may then select a milder form of the audio portion for viewing.

[0046] Similarly, some or all of the audio portion of an audio/video stream or program may be available in alternative languages. Again, a user may then indicate a language preference to the receiving device, which may then provide the appropriate audio data segments for that language in place of the original audio data upon presentation of the recorded audio/video stream to the user.

[0047] In each of these examples, the replacement audio data segment may be made available to the receiving device after the audio/video stream has been recorded at the device, thus providing a significant level of flexibility as to when the replacement audio data is provided.

[0048] While several embodiments of the invention have been discussed herein, other embodiments encompassed by the scope of the invention are possible. For example, while various embodiments have been described primarily within the context of satellite, cable, and terrestrial set-top boxes, other receiving devices capable of replacing portions of audio data of a recorded audio/video stream, such as computers, personal digital assistants (PDAs), and mobile communication devices, may be utilized in other embodiments. In addition, aspects of one embodiment disclosed herein may be combined with those of alternative embodiments to create further implementations of the present invention. Thus, while the present invention has been described in the context of specific embodiments, such descriptions are

provided for illustration and not limitation. Accordingly, the proper scope of the present invention is delimited only by the following claims and their equivalents.

CLAIMS

What is claimed is:

1. A method for replacing audio data within a recorded audio/video stream, the method comprising:
 - storing a first audio/video stream, wherein the first audio/video stream comprises audio data, video data, and supplemental data;
 - receiving an audio data segment;
 - receiving location information in a data file separately from the first audio/video stream, wherein the location information references the supplemental data of the first audio/video stream to identify a location within the first audio/video stream;
 - replacing at least a portion of the audio data of the first audio/video stream at the identified location with the audio data segment to produce a second audio/video stream; and
 - transferring at least a portion of the second audio/video stream for presentation.
2. The method of claim 1, further comprising:
 - before transferring the portion of the second audio/video stream, storing the second audio/video stream.
3. The method of claim 1, wherein the supplemental data comprises text data corresponding with the first audio/video stream.
4. The method of claim 1, wherein the data file is received after completing reception of the first audio/video stream.
5. The method of claim 1, wherein the location information comprises an indication of a starting point for the audio data segment within the second audio/video stream, wherein the indication of the starting point references the supplemental data.
6. The method of claim 1, wherein:
 - the supplemental data comprises text data; and

the location information comprises a reference to a portion of the text data to specify the identified location.

7. The method of claim 6, wherein:

the location information further comprises a value indicating an offset between the portion of the text data and the identified location.

8. The method of claim 1, wherein:

the supplemental data comprises timestamp data; and

the location information comprises a reference to a portion of the timestamp data associated with the identified location.

9. The method of claim 1, further comprising:

receiving a second audio data segment; and

receiving a selection of one of the first audio data segment and the second audio data segment;

wherein replacing the portion of the audio data of the first audio/video stream is performed with either the first audio data segment or the second audio data segment according to the selection.

10. The method of claim 1, further comprising:

receiving second location information referencing the supplemental data to identify the location;

wherein the identified location may be determined using at least one of the first location information and the second location information.

11. A method for replacing audio data within a contiguous block of audio/video data, comprising:

transmitting a contiguous block of audio/video data over a communication network to a receiving device, wherein the contiguous block comprises audio data, video data, and supplemental data;

storing the contiguous block at the receiving device;

transmitting an audio data segment over the communication network to the receiving device;

transmitting location information over the communication network to the receiving device in a data file separately from the contiguous block, wherein the location information references the supplemental data to indicate a location within the contiguous block;

replacing at least a portion of the audio data of the contiguous block at the indicated location with the audio data segment to yield a second contiguous block; and

transferring at least a portion of the second contiguous block from the receiving device to a display device.

12. The method of claim 11, wherein the receiving device comprises one of a terrestrial television receiver, a cable television receiver, a satellite television receiver, and a computer.

13. The method of claim 11, wherein the communication network comprises at least one of a terrestrial television network, a cable television network, a satellite television network, a wide area network, and a local area network.

14. A receiving device for replacing audio data within a recorded audio/video stream, the receiving device comprising:

a communication interface configured to receive a first audio/video stream, an audio data segment, and location information, wherein the first audio/video stream comprises audio data, video data, and supplemental data, and wherein the location information is received in a data file separate from the first audio/video stream and references the supplemental data to identify a location within the first audio/video stream;

a storage unit configured to store the first audio/video stream and the audio data segment;

an audio/video interface configured to transmit audio/video streams to a display device; and

control logic configured to replace at least a portion of the audio data of the first audio/video stream at the identified location with the audio data segment to produce a second audio/video stream, and transfer at least a portion of the second audio/video stream to the audio/video interface.

15. The receiving device of claim 14, wherein the control logic is configured to store the second audio/video stream in the storage unit.
16. The receiving device of claim 14, further comprising:
a user interface configured to receive a user input;
wherein the control logic is configured to transfer the second audio/video stream to the audio/video interface based on the user input.
17. The receiving device of claim 14, further comprising:
a user interface configured to receive a user input;
wherein the communication interface is configured to receive a second audio data segment;
wherein the storage unit is configured to store the second audio data segment;
and
wherein the control logic is configured to replace the portion of the audio data of the first audio/video stream at the identified location with either the first audio data segment or the second audio data segment based on the user input.
18. The receiving device of claim 14, wherein the communication interface is configured to receive the data file after completing reception of the first audio/video stream, wherein the data file comprises at least one of the location information and the audio data segment.
19. The receiving device of claim 14, wherein the supplemental data comprises text data corresponding with the first audio/video stream.
20. The receiving device of claim 14, further comprising the display device.