(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 2 901 372 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
20.06.2018 Bulletin 2018/25

(21) Application number: **13840744.0**

(22) Date of filing: **25.09.2013**

(51) Int Cl.:
G06F 17/30 *(2006.01)* ***H04L 9/32*** *(2006.01)*

(86) International application number:
PCT/US2013/061750

(87) International publication number:
WO 2014/052492 (03.04.2014 Gazette 2014/14)

(54) **USING DIGITAL FINGERPRINTS TO ASSOCIATE DATA WITH A WORK**

VERWENDUNG VON DIGITALEN FINGERABDRÜCKEN ZUR VERKNÜPFUNG VON DATEN MIT EINER ARBEIT

UTILISATION D'EMPREINTES DIGITALES NUMÉRIQUES, POUR ASSOCIER DES DONNÉES À UN MOT

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **25.09.2012 US 201213626709**

(43) Date of publication of application:
05.08.2015 Bulletin 2015/32

(73) Proprietor: **Audible Magic Corporation**
Los Gatos, CA 95032 (US)

(72) Inventors:
- **GARSIDE, Vic**
  Los Gatos, CA 95032 (US)
- **FRIEDMAN, Jay**
  Monte Sereno, CA 95030 (US)
- **SCHREMPP, James**
  Saratoga, CA 95070 (US)
- **WILLIAMS, Robert**
  San Jose, CA 95111 (US)

(74) Representative: **WP Thompson**
8th Floor
1 Mann Island
Liverpool L3 1BP (GB)

(56) References cited:
US-A1- 2007 271 248 US-A1- 2010 104 259
US-A1- 2010 281 042 US-A1- 2010 290 667
US-A1- 2011 066 489

- **KUNIO KASHINO ET AL: "Robust Search Methods for Music Signals Based on Simple Representation", 2007 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING 15-20 APRIL 2007 HONOLULU, HI, USA, 23 October 2007 (2007-10-23), pages IV-1421, XP055296602, Piscataway, NJ, USA DOI: 10.1109/ICASSP.2007.367346 ISBN: 978-1-4244-0727-9**
- **Jaap Haitsma ET AL: "A Highly Robust Audio Fingerprinting System With an Efficient Search Strategy", Journal of New Music Research, 17 October 2002 (2002-10-17), pages 211-221, XP055259740, Lisse DOI: 10.1076/jnmr.32.2.211.16746 Retrieved from the Internet: URL:http://www.cs.northwestern.edu/~pardo/courses/casa2009/casa_papers_2009/papers/audio fingerprint - haitsma.pdf [retrieved on 2016-03-18]**
- **CONG JIN ET AL: "Applications of digital fingerprinting and digital watermarking for E-commerce security mechanism", AUDIO, LANGUAGE AND IMAGE PROCESSING, 2008. ICALIP 2008. INTERNATIONAL CONFERENCE ON, IEEE, PISCATAWAY, NJ, USA, 7 July 2008 (2008-07-07), pages 536-540, XP031298539, ISBN: 978-1-4244-1723-0**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 2 901 372 B1

Description

BACKGROUND OF THE INVENTION

[0001] Individuals often want to create annotations attached to points or segments in a work (for example, a video, a music recording, or the like) that can then be shared and used by others, or that they can access themselves when the work is replayed at a later time or different location. Standard methods of associating data with a work at a particular point in the work rely on using time offsets into the work such that there is an inherent risk in presenting the associated data too early or too late, thus diminishing a user experience.

[0002] Prior art document XP05529G602 - KUNIO KASHINO ET AL: "Robust Search methods for Music Signals Based on Simple Representation" - IEEE International Conference on Acoustics, Speech, and Signal Processing, 23 October 2007, pages IV - 1421, discloses a technique for music information retrieval designed to cope with additive interfering sounds by spectral partitioning. The technique is designed to be robust under multiplicative noise or distortion based on binary area representation.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] The embodiments described herein will be understood more fully from the detailed description given below and from the accompanying drawings, which, however, should not be taken to limit the application to the specific embodiments, but are for explanation and understanding only.

Figure 1 is a block diagram of an exemplary network architecture, in accordance with one embodiment of the present invention.

Figure 2A is a block diagram of one embodiment of a client media manager.

Figure 2B is a block diagram of one embodiment of a digital fingerprint manager.

Figure 3 is a flow diagram showing an embodiment for a method of associating data with a digital fingerprint.

Figure 4 is a How diagram showing an embodiment for a method of replacing a digital fingerprint on a user device with a digital fingerprint of a master copy of a work.

Figure 5 is a flow diagram showing an embodiment for a method of replacing a digital fingerprint on a user device based on the quality of a work during playback.

Figure 6 is a flow diagram showing an embodiment for a method of subscribing to data associated with a work for display during playback by a client device.

Figure 7 is a flow diagram showing an embodiment for a method of processing digital fingerprints and associated data received by a server.

Figure 8 is a flow diagram showing an embodiment for a method of streaming media work to a client with associated data.

Figure 9 is a flow diagram showing an embodiment for a method of processing client requests for data subscriptions by a server.

Figure 10 illustrates a sequence diagram of transmission of digital fingerprints and data between a producer client, a server, and a consumer client, in accordance with one embodiment of the present invention.

Figure 11 is a block diagram illustrating an exemplary computer system, in accordance with one embodiment of the present invention.

DETAILED DESCRIPTION OF THE PRESENT INVENTION

[0004] Methods and systems for enabling a user device to receive user-entered data (or other user originated data) and associate that data with a fingerprint of a segment for a work are described. Methods and systems for enabling a user device or server to transmit data associated with a fingerprint for playback are also described. The user device may be any content rendering device that includes a modem or network interface for connecting the user device to a network. Examples of such user devices include cellular telephones, personal digital assistants (PDAs), portable media players, installed media players, tablet computers, laptop computers, netbooks, televisions, smart TVs (televisions with integrated internet capabilities) cable and satellite television set-top boxes, smart glasses, smart clothing, smart watches, smart appliances, and the like.

[0005] In one embodiment, a user device plays a work and receives data to be associated with a particular point in the work. Alternatively, the user device may use a microphone and/or camera to sample a work from a surrounding environment (e.g., to sample a performance of a work or a work being rendered by another device). The user device may receive data to be associated with a particular point of the sampled work. A work may be anything that is fixed in a tangible medium. Some examples of works include, but are not limited to, audio renderings, video renderings, images and video/audio renderings. An example of an audio rendering is a song or other audio track. Examples of video renderings include animation or video sequence. Examples of an image include photographs and paintings. Examples of audio/video renderings include movies, television shows, cartoons, and any audio/video combination that is fixed using MPEG, AVI, VHS or any of the analog or digital encoding technologies as are commonly known. A single work may be rendered in many different ways in a tangible medium. For example, one fixed audio segment may be fixed on an analog tape, in an MP3 audio stream, and as part of a soundtrack in an audio/video rendering, and so forth. It will be obvious to anyone fluent in the art that embodiments of the

present invention can apply to one entire work and to segments of a work. It will also be obvious that one work may contain segments of another work. Note that in some embodiments (e.g., in the case of sampling a live performance), the work being sampled may not initially be embodied in a tangible medium.

[0006] The user device may generate a user digital fingerprint of a segment of the work at the particular point in the work and associate the data to the fingerprint. As used herein, a user digital fingerprint may be a fingerprint of a work generated by a user device. The data and user digital fingerprint may be sent to a server to be saved for use during a later playback of the work. In one embodiment, the data and user fingerprint together form a tag. The server may replace the user digital fingerprint in the tag with a master digital fingerprint. As used herein, a master digital fingerprint may be a fingerprint of a high quality version of the known work. A second user device may receive the digital fingerprint and data at some point in time. The second user device may use the associated user fingerprint or master fingerprint to determine the point in the playback of the work that corresponds to the point that has data associated with it. The second user device may present the associated data to the user of the second user device. For example, a first user may use a mobile phone to play a video and enter text commentary at particular points in the video. A second user device, while playing back the same work, may identify the particular point in the video using the digital fingerprint, and may present the commentary entered by the first user at the particular point in the video.

[0007] In another embodiment a first device may be rendering a media stream. A second device may capture samples of the rendering by means of a microphone or camera on the second device. Data may be input by a user, and may be associated with a user fingerprint created from the captured samples of the media stream. For example the first device may be rendering a DVD or radio program. The second device may be sampling this media stream by way of a camera or microphone. The second device generates a user digital fingerprint of an immediate portion of the media stream being samples and associates the data with the user digital fingerprint for later use. In this example, the term immediate portion may mean a portion of the captured samples of the media stream (which may occur as much as two minutes before or two minutes after the point in the media stream where the data should be associated).

[0008] In another embodiment, both actions may occur on the same device instead of separate devices. A user device may play a work and receive data to be associated with a particular point in the work. The user device generates a user digital fingerprint of the work, or portion thereof, and associates the data to the user fingerprint. The data and user fingerprint (or master fingerprint) may be saved locally on the user device for use during a later playback of the work. During later playback of the work on the same device (which may be a same rendition of the work or a different rendition of the work), the user device may determine the point in the work that has associated data and present it to the user. For example, a user may download a video to a tablet computer, play it on the tablet computer, and enter text commentary at particular points in the video. The same tablet computer while playing back the video, may present the commentary entered by the first user at the particular point in the video.

[0009] In another embodiment the data associated with a point in the work using the fingerprint may instruct the client device to take some action. By means of example and not limitation, some actions include - communicating with another application on the device, such as a to-do list application; communicating with another server, such as one that would provide a purchase recommendation to the user; communicating with another server, such as one that would log the exposure of the device to the work at the point of the fingerprint; communicating with another device to cause it to behave in some way, such as vibrate, rotate, elevate, release a scent, change color, change temperature, emit sounds, and/or display some information; and so forth.

[0010] In another embodiment the user device may generate a user fingerprint from a media stream. The user device may use this fingerprint to identify the work being played. The user device may then download a master fingerprint of a known work which corresponds to the work being played. The user device then associates data with the master fingerprint of the known work. Note that in an alternative embodiment, the user device may extract a portion of the master fingerprint and associate the data with the extracted portion of the master fingerprint. For example, the master fingerprint may be a fingerprint of multiple segments. The user device may extract a portion of the master fingerprint that is associated with a particular segment or set of segments.

[0011] **Figure 1** is a block diagram of an exemplary network architecture 100 in which embodiments described herein may operate. The network architecture 100 may include a server system 140, a media server 150, and one or more user devices 105-110 capable of communicating with the server system 140 and/or the media server 150 via a network 135 (e.g., a public network such as the Internet or a private network such as a local area network (LAN)) and/or a wireless communication system 130. In one embodiment, a user device 105 connects to server system 140 and/or media server 150 via the wireless communications system 130 and network 135. In another embodiment a user device 110 connects to server system 140 and/or media server 150 via network 135. In another embodiment user device 105 is not connected to a network and performs the described functions in isolation.

[0012] The user devices 105-110 may be variously configured with different functionality to enable consumption of one or more types of media items and/or sampling of works from an environment in which the user devices

105-110 are placed. The media items may be any type of format of digital content, including, for example, digital audio (e.g., music, audible books, etc.), digital video (e.g., movies, television, short clips, etc.), and multi-media content. Each media item may include a specific work. The user devices 105-110 may include any type of content rendering devices such as mobile phones, laptop computers, portable media players, tablet computers, cameras, video cameras, netbooks, notebooks, desktop computers, televisions, cable and satellite television set-top boxes, audio or image sensing eyewear, and the like. In one embodiment, the user devices 105-110 are mobile devices. The user devices 105-110 may store content locally on storage media or in memory for later playback. Additionally, media server 150 may deliver media to the user devices 105-110, such as during streaming of music or videos to the user devices 105-110.

[0013] In one embodiment, user devices 105-110 may include a microphone 123 and/or a camera 124, which may be integrated into the device or attached externally. Microphone 123 may be used to sample audio of a surrounding environment of user devices 105-110. Similarly, Camera 124 may be used to sample audio, video, or images of a surrounding environment of the user device 105-110. For example, if a nearby media rendering device 175 (e.g., a radio having a radio receiver 178 or a television (TV) having a TV receiver 180, or a DVD player) is playing a work, the microphone may pick up segments of the work being played by the media rendering device 175.

[0014] In one embodiment, one or more of the user devices 105-110 include a client media manager 115. The client media manager 115 may manage the processing of media items on the user devices 105-110, the provisioning of sampled audio and/or video data that includes segments of works, as well as the generation of digital fingerprints of the work (or works) included in the media items. For example, the client media manager 115 may manage the generation of user digital fingerprints of audio, video, or images captured by user devices 105-100 via microphone 123 or camera 124 from the surrounding environment. The client media manager 115 may also manage the generation of user digital fingerprints directly from the media being rendered by playback module 122. The client media manager 115 may also manage the receipt of data, association of the data with the user digital fingerprints (e.g., to generate a tag), sending and receiving information (e.g., tags) to and from server system 140, and the presentation of data during playback of a work. The client media manager 115 may be comprised of a producer module 120 and a consumer module 125.

[0015] The producer module 120 may handle processing involved in the management of the data to be associated with a work played on or sampled by the user devices 105-110. In one embodiment, producer module 120 monitors playback of a work (e.g., a work included in a media item played on the user device or a work sampled from an environment by the user device), and generates user digital fingerprints of the work locally on the user devices 105-110. In one embodiment, a playback module 122 on the same user device 105-110 as the client media manager 115 plays the work that is fingerprinted. Alternatively, the user device 105-110 may not include a playback module 122, and may rely on sampling of a work from an environment of the user device 105-110. The work may be in streamed media being played by a player on the user device, such as a streamed movie or song being provided by a media server. The work may also be in a stored media item that is being played on the user device, such as playback of a digital video disk (DVD), stored movie file, stored audio file, and so forth. The work may also be a broadcast of the work that is received by a tuner on the user device. For example, the work may be a radio broadcast, a television broadcast, and so forth. The work may also be sampled from an environment of the user device. For example, the work may be rendered from any source by another device, the rendering of which is then sensed by the user device.

[0016] In one embodiment, digital fingerprints are generated continuously during playback. Alternatively, a cache of recently played segments of a work may be maintained, and fingerprints may be generated from such cached segments, present segments, and/or subsequent segments responsive to a user command to add data to the work at a particular point in the work. Producer module 120 receives a command to associate data with the work at a particular point in the work to be presented at the same point during playback of the work at a later time. For example, the data may be text commentary, a digital image, a uniform resource locator (URL), a video, a sound effect or a command.

[0017] Producer module 120 associates the received data with the user digital fingerprint that corresponds to the segment of the work at the particular point in the work and saves it for later use. In one embodiment, producer module 120 generates a tag that includes the user digital fingerprint and the data. A segment of a work may be less than an entirety of the work and may include at least one of audio or video or image media content. In one embodiment, a segment is a clip (e.g., a sequence of frames) of the work. For example, a segment may be between a single frame and a 10 second clip of a video or song. For example, a segment may be between 5% and 100% of an image. For example, a segment may contain the entire visual image and some irrelevant other data around the image, as when one takes a video of a television screen but the television screen does not consume the entire image space. In one embodiment, the data and associated user digital fingerprint are saved locally on user devices 105-110 for later use on the same user device. In another embodiment, the data and associated user digital fingerprint are transmitted to server system 140 to be saved in data storage 160 and fingerprint storage 155 for later use on the same or other user devices. The server system 140 may additionally replace

the user digital fingerprint with a higher quality master digital fingerprint (or a portion of a higher quality master digital fingerprint). Such replacement may be particularly useful when the user fingerprint was generated from a segment of a work sampled from a microphone or camera.

[0018] The consumer module 125 may handle processing involved in presenting the data on the same or different user devices 105-110 during later playback of the same work. Said playback may occur on the same user device 105-110. Alternatively said playback rendering may occur on some other device and be sampled by microphone 123 or camera 124. During playback, consumer module 125 may generate fingerprints of segments of the work and compare the digital fingerprint of the segment of the work being played back to the user digital fingerprints or master fingerprints stored locally on the user devices 105-110. Alternatively, consumer module 125 may generate fingerprints and send these fingerprints to server system 140 for comparison to user fingerprints or master fingerprints. In such an embodiment, the server system may respond with data associated with a master fingerprint or user fingerprint matching the fingerprint that was sent to the server system. In one embodiment, segments may overlap. For example, a first segment may include seconds 1-5 in a video, a second segment may include seconds 2-6, a third segment may include seconds 3-7, and so on. Accordingly, consumer module 125 may generate a continuous stream of fingerprints from overlapping segments during playback. In one embodiment, fingerprints of segments are generated before those segments are played. This enables a match to be made to the stored fingerprint associated with the data before the segment associated with the data is played. Note that a match may be an approximate match. In one embodiment, two fingerprints are determined to match if they differ by less than a threshold amount (e.g., 10%, 20%, etc.).

[0019] Note that playback of the work may begin before the user device 105-110 has received any data or associated master fingerprints (or user fingerprints). In such an embodiment, consumer module 125 may query server system 140 for data and associated fingerprints after playback of the work by playback module 122 has begun, or before the rendering device 175 has begun rendering the work.

[0020] When consumer module 125 encounters a match between a generated digital fingerprint and a saved digital fingerprint that is associated with data, the data will be used. One use of the data is that it be presented to the user on the user devices 105-110. Another use is that the data be presented on another device such as a TV display or radio. Additionally, other actions may be performed, such as communicating with another device or server, communicating with another application on the user device, loading a web page, posting data to a blog or social media site, and so forth. One can imagine a user device posting to a user's Facebook page a message that the user has just watched some particular part of a movie or television show.

[0021] In one embodiment, a playback time range is associated with the digital fingerprint. The playback time range may indicate a point during playback at which digital fingerprints shall be generated for comparison to the saved fingerprint. This may save processing power by minimizing the number of fingerprints that are generated during playback. The playback time range may include a time in the work associated with the segment used to generate the fingerprint associated with the data plus or minus an error factor (e.g., 10 minutes). The error factor may account for changes that are made to the work (e.g., cutting the play time of the work to add commercials, changing a starting time of the work, and so forth).

[0022] In another embodiment, consumer module 125 may send a request to server 140 to subscribe to any data saved on the server system 140 that is entered by a particular user on any user device or any data saved for a particular work. During playback of a work, consumer module 125 may generate a digital fingerprint of the segment of the work being played back and send a request to server system 140 to determine whether there is any saved data associated with a master fingerprint or user fingerprint that matches the fingerprint generated on the user device. If server system 140 finds any saved data associated with a matching master or user fingerprint, it may then send that saved data to consumer module 125. When consumer module 125 encounters the particular point in the work where the fingerprint matches that of the master or user fingerprint associated with the data, the data will be used in any of the ways previously described.

[0023] Alternatively, consumer module 125 may send a request to server system 140 to receive all fingerprints with associated data for a particular work that have been generated by one or more entities to which the user subscribes. Server system 140 may then send the fingerprints and the associated data to the user device. Consumer module 125 may match the fingerprints with the generated fingerprints during playback and when the particular points in the work are encountered, use the associated data as appropriate.

[0024] In one embodiment, the data may be presented by displaying text associated with the particular point in the work. Additionally, the user devices 105-110 may display a digital image, launch a web browser to display a URL, play a video clip or sound effect, or execute a command on the user devices 105-110. Other types of data may also be associated with the fingerprint to perform other actions.

[0025] User devices 105-110 may connect to the server system 140 via a wireless communication system 130. The wireless communication system 130 may provide a wireless infrastructure that allows users to use the user devices 105-110 to consume items provided by the server system 140 without being tethered via hardwired links. The wireless communications systems 130 may be wire-

less fidelity (WiFi) hotspots connected with the network 135. The wireless communication systems 130 may alternatively be a wireless carrier system (e.g., as provided by Verizon®, AT&T®, T-Mobile®, etc.) that can be implemented using various data processing equipment, communication towers, etc.

[0026] The server system 140 may include one or more machines (e.g., one or more server computer systems, routers, gateways, etc.). In one embodiment, the server system 140 includes one or more cloud based servers (such as server systems made accessible to network connectivity via the internet or wireless carrier systems), which may be hosted, for example, by cloud based hosting services such as Amazon's® Elastic Compute Cloud® (EC2).

[0027] The server system 140 may have greater resources than the user devices 105-110. Accordingly, the server system 140 may implement resource intensive algorithms for matching the fingerprint of the work during playback to fingerprints associated with data for presentation on the user devices 105-110. Thus, in some instances improved throughput and/or improved fingerprint matching may be achieved by having the server system 140 perform fingerprint matching for the user devices 105-110.

[0028] In one embodiment, server system 140 may include a digital fingerprint manager 145, a fingerprint storage 155, and a data storage 160. Digital fingerprint manager 145 is configured to match user digital fingerprints generated by user devices 105-110 to master digital fingerprints of known works stored and to previously generated user digital fingerprints stored on the server system 140 in fingerprint storage 155. Additionally, digital fingerprint manager 145 manages the receipt, storage, and distribution of data associated with master digital fingerprints of known works and data associated with user digital fingerprints. Data received by digital fingerprint manager 145 is stored in data storage 160 for later distribution. Alternatively, master and/or user digital fingerprints and data may be stored together in a single file, or may be stored in a relational database. Further, digital fingerprint manager 145 processes subscription requests from user devices 105-110, utilizing fingerprint storage 155 and data storage 160.

[0029] For example, producer module 120 on user devices 105-110 may create a user fingerprint and associated data combination, then send the combination to server system 140 via network 135. Digital fingerprint manager 145 on server 140 will receive the user fingerprint and data combination, and compare the user fingerprint generated by the user device to the master fingerprints stored in fingerprint storage 155. Once a corresponding master fingerprint is located in fingerprint storage 155, digital fingerprint manager 145 may store the data in data storage 160 and associate that data with the master fingerprint of the known work stored in fingerprint storage 155. Thus, the data may be stored on the server system 140 such that it is associated with a high quality master digital fingerprint of the work. This may improve an ability to match future user fingerprints generated from the segments of a low quality of a work being sampled to the digital fingerprint associated with the data.

[0030] In one embodiment, if no match to a master fingerprint can be made, then digital fingerprint manager 145 may store the data in data storage 160, store the user digital fingerprint in fingerprint storage 155, and associate the two. Thus, it is possible that some elements of data in data storage 160 are only associated with user fingerprints and not with any master fingerprint. Alternatively, even if a match is made between a user fingerprint and a master fingerprint, the digital fingerprint manager 145 may choose to store the user digital fingerprint in fingerprint storage 155 and associate the data with both fingerprints. Thus, the digital fingerprint manager 145 may have several fingerprints associated with data. This multiplicity of fingerprints may be useful to improve the ability to match future user fingerprints.

[0031] In another embodiment, consumer module 125 on user devices 105-110 sends a subscription request for a media work that is saved on server system 140. The work may be downloaded and stored locally on user devices 105-110 or it may be streamed to the user devices 105-110 via media server 150, or it may be streamed to an ancillary media rendering device and captured via microphone 123 or camera 124. For works stored locally on user devices, digital fingerprint manager 145 may transmit some or all digital fingerprints of the work stored in fingerprint storage 155 along with associated data from data storage 160 to the user device so the user device can manage playback and presentation of the data locally.

[0032] For works streamed via media server 150, digital fingerprint manager 145 may transmit the fingerprints and associated data to media server 150 so that it may embed these in the media stream such that the user device need not compare locally generated digital fingerprints to those on the server. In this embodiment, media server 150 streams the work with the associated data already integrated at the appropriate place in the work. In one embodiment, media server 150 includes a fingerprint comparison module 185 and a data insertion module 188. The fingerprint comparison module 185 may generate fingerprints of a work being streamed, and may compare those fingerprints to a fingerprint received with associated data from server system 140. Once a match is identified, data insertion module 188 may insert the data into the work at a position in the work indicated by the matching fingerprints. Thus, media server 150 may stream both the data and the work to the user device 105, 110. Then the user device would not need to perform any fingerprint generation or comparison. The user device may simply act on the data when it is encountered in the stream.

[0033] In another embodiment the digital fingerprint manager 145 may transmit a master digital fingerprint of a known work from fingerprint storage 155 to the user

devices where producer module 120 will make the comparison between the producer generated fingerprint and the master fingerprint of the known work. When the producer module 120 is instructed to create a tag it may associate the data for the tag with a master fingerprint taken from a portion of the fingerprint of the known work.

[0034] **Figure 2A** is a block diagram of one embodiment of a client media manager 200, which may correspond to the client media manager 115 of **Figure 1.** The client media manager 200 may include one or more of a producer module 205 and a consumer module 210. Additionally, client media manager 200 may receive an incoming media work 240, incoming digital fingerprints 241 and incoming data 242, as well as transmit outgoing digital fingerprints 243 and outgoing data 244. The incoming work 240 may be from digital stream, a playback module 225, a microphone capture 250, a camera 255, or other sensor capable of fixing the work in a digital form. The incoming work 240 may be a sampling of a work being played on a separate device.

[0035] Producer module 205 is configured to enable a user device to associate data to a segment of a work at a particular point in the work. In one embodiment, producer module 205 includes a fingerprint generator 215, a data association module 220 and a data action module 230. In one embodiment, producer module 205 may correspond to the producer module 120 of **Figure 1.**

[0036] Fingerprint generator 215 is configured to generate a digital fingerprint from a segment of a work, such as audio and/or video content. In one embodiment, fingerprint generator 215 may receive the audio and/or video content via incoming work 240. Fingerprint generator 215 may create the digital fingerprint using conventional digital signal processing known in the art. For example, a digital fingerprint of a sound stream may be generated according to acoustic and/or perceptual features over time. In one embodiment, digital fingerprints may be generated from the media content of a work according to a predetermined or user-defined interval. In another embodiment, a digital fingerprint may be generated when a user attempts to associate data with a segment of a work at a particular point in the work.

[0037] In one embodiment, digital fingerprints are generated by generating a feature vector of a segment of the work. The feature vector may be generated from an audio portion of the work, an image portion of the work, or both. For example, in one embodiment, fingerprint generator 215 measures a variety of acoustical features of a segment. The acoustical features may include loudness, pitch, bass, brightness, bandwidth, Mel-frequency cepstral coefficients (MFCCs), and so forth. Fingerprint generator 215 may also compute first and/or second derivatives of some or all of these features to determine changes in the features and rates of changes in the features. Fingerprint generator 215 may compute statistical measurements such as mean and standard deviations of each of these features over the segment as well. Some or all of these values may be included in the feature vector. For example, in one embodiment, fingerprint generator 215 measures a variety of visual features of a frame. The visual features may include luminosity, pixel colors, shading, etc. For example in one embodiment, fingerprint generator 215 measures the variation in visual features over the segment of the work. The time varying features may include the motion of pixels or pixel regions, the change in luminosity, the gross difference in pixel regions, and so forth.

[0038] An optional playback module 225 configured to render a work on the user device or otherwise play the work. In one embodiment, an audio file, video file, video stream, audio stream, etc. may be played by playback module 225. Playback module 225 may provide played (or to be played) segments of the work to client media manager 200. Alternatively, a microphone 250 and/or camera 255 may capture segments of a work from a surrounding environment, and may provide these segments of the work to client media manager 200.

[0039] Data association module 220 is configured to associate data with a segment of a work at a particular point in the work. In one embodiment, data association module 220 may be invoked by producer module 205 when a user issues a command to associate data with a work being played or sampled. For example, a user may watch a video and issue a command to enter text commentary at a particular point in the video. In another embodiment, data association module 220 may be invoked by producer module 205 when a user issues a command to associate data with a work being captured by a microphone 250. Data association module 220 may receive the user-entered data via a keyboard and associate the data with that point in the video. In one embodiment this data may be user generated commentary. In one embodiment, this association may be accomplished by utilizing the user fingerprint generated by fingerprint generator 215 for that segment of the work. In one embodiment data association module 220 matches the user fingerprint from fingerprint generator 215 to a master fingerprint (or a portion of a master fingerprint) of a known work 241 and associates the data with the master fingerprint (or portion thereof) of the known work. Data association module 220 may save the data and associated fingerprint locally on the user device, either combined or separately in a file or database. Alternatively, the data and digital fingerprint may be sent to a server system via outgoing data 244 and outgoing digital fingerprints 243 respectively.

[0040] Data action module 230 may be configured to display data entered by a user during the process of associating that data with a work. In one embodiment, if a user issues a command to associate text commentary to a video, data action module 230 may present a window on the screen of the user device to allow the user to observe the text being entered. In an alternative embodiment, if a user issues a command to associate a sound file to a work, data action module 230 may present a graphical representation of a recording device to permit

the user to start and stop the sound file recording. Thus, a user may review how generated data will be presented along with the work. In another embodiment, the user may perform a new work that is captured via a microphone or camera or both, and this new work may become the data associated with the user fingerprint.

[0041] Consumer module 210 is configured to enable a user device to receive and display data associated with segments of a work at particular points in the work. Consumer module 210 may also perform other actions at particular points in the work based on the data. In one embodiment, consumer module 210 includes a fingerprint generator 215, a fingerprint matching module 235, and a data action module 230. In one embodiment, consumer module 210 may correspond to the consumer module 125 of **Figure 1.**

[0042] Fingerprint generator 215 is configured to generate a user digital fingerprint from a segment of a work, as described above. In one embodiment, fingerprint generator 215 may also generate user digital fingerprints during playback of a work in order to determine if and when associated data is to be presented to the user (or other actions are to be performed). For example, if a user plays a video after having subscribed to data and fingerprints entered by an entity for that particular video, fingerprint generator 215 may generate digital user fingerprints for the video during playback that will be used to compare to tag fingerprints (fingerprints associated with or included in a tag) to determine the particular point in the video to present the data. In one embodiment, the entity may be a user who enters data associated with a media work. For example, a user may associate text commentary data with particular points in a video work. In another embodiment, the entity may be an automated transcription service that associated data with a media work from a data store. For example, an automated service may add statistical information at particular points in a broadcast of an athletic event, such as historical batting averages of baseball players during the broadcast of a baseball game. Another example is an automated object recognition system associating data about a recognized object, such as a particular automobile, actor, location, or product, to particular points in the work.

[0043] Fingerprint matching module 235 is configured to match digital fingerprints of a work with stored user or master fingerprints associated with data in order to determine points in the work to present the data. In one embodiment, fingerprint matching module 235 receives digital fingerprints generated by fingerprint generator 215 in consumer module 210 from incoming work 240. In one embodiment, a user device may receive a collection of digital fingerprints with associated data for a particular video from a server and store them locally. This collection of fingerprints may contain user fingerprints from other users, user fingerprints from this user, and/or master fingerprints. During playback of that video, fingerprint matching module 235 may compare the fingerprints received from the server (e.g., fingerprints with associated data) to those fingerprints generated by fingerprint generator 215 (e.g., newly generated user fingerprints) in order to use the associated data at the appropriate point in the video. In another embodiment, a user device may receive a collection of fingerprints with associated data for a number of works and store them locally. In another embodiment, a user device may receive fingerprints with associated data for a number of works, all the data originating from a particular source.

[0044] Data action module 231 is configured to display data associated with a work at a particular point in the work and/or to perform other actions at the particular point in the work. In one embodiment, data action module 231 may receive a notification during playback of a work that there is data associated with a particular point in a segment of that work. Data action module 231 may then present the data on the user device at the appropriate point during playback. For example, during playback of a video, data action module 231 may receive notification from fingerprint matching module 235 that there is text commentary associated with a particular point in the video. Data action module 231 may then present a window on the screen of the user device to display the text commentary on the screen. Alternatively, data action module may perform other actions at the particular point in the video, such as loading a web page, loading an application, communicating with a remote device or server, and so forth.

[0045] **Figure 2B** is a block diagram of one embodiment of a digital fingerprint manager 250, which may correspond to the digital fingerprint manager 145 of **Figure 1.** The digital fingerprint manager 250 may include one or more of a fingerprint matching engine 255, a data management engine 260, a media quality analyzer 265, and a subscription manager 270. Additionally, digital fingerprint manager 250 may receive incoming digital fingerprints 275 and incoming data 280, as well as transmit outgoing digital fingerprints 285 and outgoing data 290.

[0046] Fingerprint matching engine 255 is configured to match user digital fingerprints generated by a user device with high quality master digital fingerprints of known works stored on a server system. For example, a user device that captures user digital fingerprints of a work in order to associate data with particular points in that work may send the user digital fingerprints to fingerprint matching engine 255 as part of the process to store the associated data on the server

[0047] Data management engine 260 is configured to store and distribute data entered by a user device associated with a work at a particular point in the work. Data may be stored in a storage facility such as data storage 160 from **Figure 1.** In one embodiment, data management engine 260 receives incoming data 280 in combination with incoming user digital fingerprints 275. Once the incoming user digital fingerprints 275 are matched with master digital fingerprints of known works via fingerprint matching engine 255, the incoming data 280 may be stored on the server and associated with the master

digital fingerprints of the known work. In another embodiment, when a user device subscribes to data of a known work, data management engine 260 is invoked to query the data storage and transmit outgoing data 290 to the user device. In another embodiment the user device requests a master fingerprint from the data management engine, does the matching on the user device, potentially extracts an appropriate portion of the master fingerprint, and returns that to the data management engine with the data.

[0048] Media quality analyzer 265 is configured to analyze the quality of media being played back on the user device. Such an analysis may be made based on the incoming user digital fingerprints 275 received from the user device. In one embodiment, media quality may be analyzed using objective audio and video techniques known in the art such as signal-to-noise ratio (SNR) or peak signal-to-noise ratio (PSNR). Media quality analyzer 265 may also examine an audio or video work for the presence of network transmission artifacts during playback when compared to a master copy to determine whether the transmitted signal has been degraded. In one embodiment, a received fingerprint includes an associated media quality value that may have been generated by a producer module that generated the fingerprint. The producer module may have generated the media quality value using the above described techniques.

[0049] Alternatively, media quality analyzer 265 may determine the media quality from the digital fingerprint itself. For example the producer module may determine that the user fingerprint matches poorly with a master fingerprint thus indicating that the user fingerprint was generated from a poor quality rendering of the original work or that substantial noise was included in the sampling of the work from a microphone or camera. The producer module may determine that the digital user fingerprint matches in some segments but not in others. The producer module may infer that the user fingerprint was generated from a rendering that included pauses, or that failed to render some sections of the original work.

[0050] The incoming user fingerprints may be replaced in the outgoing digital fingerprint with the master fingerprints stored on the server for the particular segment of the work. Media quality analyzer 265 may thus lower network traffic and limit processing when the work being played back is known to be a higher quality sample such as a movie stored locally on the device or being streamed via a high speed broadband connection. Additionally, replacing fingerprints determined to be degraded may improve performance when matching against the same work in future playbacks. Alternatively, the fingerprint may be replaced with the corresponding fingerprint of the master copy of the work regardless of the media quality.

[0051] Subscription manager 270 is configured to process subscription requests from user devices. In one embodiment, a user device sends a request to the server to receive all data and associated fingerprints entered for a particular work or by a particular entity. In one embodiment, the entity may be a user who enters data associated with a media work. For example, a user may associate commentary data with particular points in a video work. In another embodiment, the entity may be an automated transcription service that associated data with a media work from a data store. For example, an automated service may add statistical information at particular points in a broadcast of an athletic event, such as historical batting averages of baseball players during the broadcast of a baseball game. As another example, an automated service may add purchase information for items that are shown in a particular scene of a television show. As another example, an automated service may add travel information for a location mentioned in a song. Subscription manager 270 receives this request, and operates in conjunction with fingerprint matching engine 255 and data management engine 260 to transmit the data and fingerprints associated with a work to user devices of subscribing users.

[0052] **Figure 3** is a flow diagram of an embodiment for a method 300 of associating data with a digital fingerprint. Method 300 is performed by processing logic that may comprise hardware (circuitry, dedicated logic, etc.), software (such as is run on a general purpose computer system or a dedicated machine), or a combination of both. In one embodiment, the method 300 is performed by a user device 105-110 of **Figure 1.** For example, the method of **Figure 3** may be performed by a client media manager of a user device.

[0053] At block 305 of method 300, processing logic begins sampling of a work on a computing device. In one embodiment, the sampling may be of a media file or stream playing back on the device itself. In another embodiment, the sampling may be of a work being played on another device directly coupled to the computing device. In another embodiment, the sampling may be an audio and/or video sampling of the environment of the user device using a microphone or camera coupled to or contained within the user device. The sampling may capture segments of the work. In one embodiment, the computing device may be a personal computer. In an alternative embodiment, the computing device may be a tablet computer, smartphone, PDA, smart TV, smart eyewear, or the like on which media can be played or sampled. In an example, a television or radio or Internet broadcast may be played by a TV or radio, and a microphone or video camera coupled to the computing device may pick up the played TV or radio broadcast.

[0054] At block 310, processing logic receives a command to associate data with the work being played at a particular point in the work. In one embodiment the point in the work is specified by the user indicating that the data should be associated with the point in the work that has just been rendered. In one embodiment, the data may represent text commentary manually entered by the user via a keyboard. In an alternative embodiment, the data may be a uniform resource locator (URL), a video, a sound effect, or a command. In another embodiment

the data may be a new work created by the user, such as a video, a drawing, or a bit of audio commentary. Other types of data are also envisioned. At block 315, processing logic generates a user digital fingerprint of the segment of the work associated with the particular point in the work identified at block 310. At block 320, processing logic associates the data created at block 310 with the user digital fingerprint generated at block 315. This may include generating a single file that contains both the data and fingerprint. Alternatively, the data and fingerprint may be generated and stored separately and connected logically using a reference link or digital pointer (such as a hyperlink).

[0055] At block 325, processing logic transmits the user digital fingerprint and associated data to a server system, which may be, for example, a cloud based server and/or a server provided by a wireless carrier. The server may then store the user digital fingerprint and associated data on the server system. In one embodiment, the server may store the digital fingerprint and associated data together in a single record of a relational database. In another embodiment, the server system may store the digital fingerprint and associated data separately, such as in fingerprint storage 155 and data storage 160 from **Figure 1** respectively.

[0056] **Figure 4** is a flow diagram showing an embodiment for a method 400 of replacing a user digital fingerprint on a user device with a master digital fingerprint of a master copy of a work. Method 400 is performed by processing logic that may comprise hardware (circuitry, dedicated logic, etc.), software (such as is run on a general purpose computer system or a dedicated machine), or a combination of both. In one embodiment, the method 400 is performed by a user device 105-110 of **Figure 1.** For example, the method of **Figure 4** may be performed by a client media manager of a user device.

[0057] At block 405 of method 400, processing logic generates a user digital fingerprint of a segment of a work that corresponds to a particular point in the work and associates data with the user fingerprint. At block 410, processing logic queries a server with the user digital fingerprint generated in block 405. The server system may analyze a quality of the user digital fingerprint and if found not to meet a predetermined threshold, send a replacement master digital fingerprint to the user device. In one embodiment, the server may do so from media quality analyzer 265 in **Figure 2B.** At block 415, processing logic receives a master digital fingerprint from the server. At block 420, processing logic replaces the user digital fingerprint generated in block 405 with the new master digital fingerprint received from the server in block 415.

[0058] In one embodiment, a user may play or sample a work on a user device, and enter data associated with that work to be replayed at a later time on that same user device or a different user device. In this embodiment, the user device may be streaming a video from an online source where playback quality is low, thereby resulting in a low quality user digital fingerprint generated on the user device. In another embodiment, the user device may be sampling the rendered work via a microphone or camera, thereby resulting in a sample of media contaminated by environmental noise. The user device may query a server to obtain a higher quality master digital fingerprint of the work from a database of fingerprints of known works to replace the generated user fingerprint in order to prevent possible problems during later playback.

[0059] **Figure 5** is a flow diagram showing an embodiment for a method 500 of replacing a user digital fingerprint on a user device based on the quality of a work during playback or sampling. Method 500 is performed by processing logic that may comprise hardware (circuitry, dedicated logic, etc.), software (such as is run on a general purpose computer system or a dedicated machine), or a combination of both. In one embodiment, the method 500 is performed by a user device 105-110 of **Figure 1.** For example, the method of **Figure 5** may be performed by a client media manager of a user device.

[0060] At block 505 of method 500, processing logic begins sampling or playback of a work by a user device. At block 510, processing logic generates a user digital fingerprint of a segment of the work at a particular point in the work and associates data with the fingerprint. At block 515, processing logic analyzes the quality of the segment of the work processed in block 510. At block 520, processing logic determines whether the quality of the segment of the work analyzed at block 515 satisfies a predetermined threshold. For example, a threshold may be set for video quality limiting the number of frames that can be dropped due to online bandwidth limitations. Additionally, a threshold may be set for an audio track that limits the amount of audio noise encountered during the playback of the segment.

[0061] If at block 520 processing logic determines that the segment quality meets the threshold, the method continues to block 535, and processing logic stores the user digital fingerprint. If not, the method continues to block 525, and processing logic queries a server for a replacement fingerprint. The query may include a copy of the generated user digital fingerprint. At block 530, processing logic receives a replacement fingerprint generated from a higher quality segment stored on the server system. At block 530 the received replacement fingerprint may be a master fingerprint of a known work or a higher quality user fingerprint.

[0062] At block 535, processing logic stores the digital fingerprint on the user device. If a replacement fingerprint was received by the server based on the threshold analysis conducted at block 520, the replacement is stored. Otherwise, the original user fingerprint is stored. The method may repeat continuously so long as a work is being played and data is being entered on the user device.

[0063] **Figure 6** is a flow diagram showing an embodiment for a method 600 of subscribing to data associated with a work for display during playback by a client device.

Method 600 is performed by processing logic that may comprise hardware (circuitry, dedicated logic, etc.), software (such as is run on a general purpose computer system or a dedicated machine), or a combination of both. In one embodiment, the method 600 is performed by a user device 105-110 of **Figure 1.** For example, the method of **Figure 6** may be performed by a client media manager of a user device.

[0064] At block 605 of method 600, processing logic subscribes to data provided by another entity. In one embodiment, the entity may be a user who enters data associated with a media work. For example, a user may associate text commentary data with particular points in a video work. In another embodiment, the entity may be an automated transcription service that associated data with a media work from a data store. For example a user may subscribe to all data entered by some particular user. For example a user may subscribe to all data entered by some particular transcription service. For example a user may subscribe to all data entered by any entity for a particular work.

[0065] In another embodiment, tags may be categorized, and processing logic may request all tags of a particular category that is associated with any work. For example, a user may subscribe to tags that that have been categorized as being associated with a particular celebrity, animal, science, politics, or other subject matter. These categories may be indicated by a user who generates the data. Alternatively, the data may be categorized using an automated categorization process or through a human mediated editorial process. Automated categorization may be performed by parsing data and analyzing it using one or more categorization rules. Each categorization rule may search for specific terms, words, phrases, etc. If such information is found, the rule may cause the data to be categorized with a category associated with that rule.

[0066] At block 610, processing logic receives data and associated fingerprints from a server. In one embodiment, a user device may receive all data and fingerprints for a particular work generated by the entity to which the user device is subscribed so that they may be stored on the user device prior to playback. For example, a user that intends to play a particular movie on a device may download all data and associated fingerprints from the server prior to playback in order to improve operating efficiency and reduce the risk of issues arising during transmission during playback. In another embodiment the device uses a user fingerprint to identify the work that is being sampled and downloads tags associated with the work.

[0067] At block 615, processing logic begins sampling of the work on the user device. At block 620, processing logic identifies a point in the work using the digital fingerprint that is associated with data received from the server. This may be performed by generating a sequence of overlapping digital fingerprints of the work during playback of the work. These digital fingerprints may be compared to the digital fingerprint associated with the data. If a matching digital fingerprint is found, the data may be used when the segment that was used to generate the matching fingerprint is played. For example, during the playback of a video on a device, processing logic may identify a particular point in the video where a prior user had associated text commentary. In another embodiment, during the viewing of a television show, processing logic may identify a particular point in the show where a URL should be displayed. At block 625, processing logic uses the data during playback of the particular point in the work. In one embodiment, if a prior user had entered text commentary regarding the particular point in the work, processing logic would display that commentary on the user device such that it is viewable by the user. In another embodiment, if a prior user had entered a video regarding the particular point in the work, processing logic would display that video on the user device such that it is viewable by the user. This presentation may consume the entire display surface. Alternatively this presentation may take place in a window consuming only a portion of the display surface. Alternatively the featured presentation may be reduced to a window consuming only a portion of the display surface and the presentation may consume the rest of the surface. Alternatively the presentation may be audio and play without consuming any of the visual display surface. Alternatively the featured presentation may pause while the prior user's video and audio is played. Alternatively the presentation may be made on another device, operatively coupled to the user device. Alternatively the user device may command another device to begin the presentation. As discussed above, other actions may also be performed based on the data at the particular point in the work.

[0068] **Figure 7** is a flow diagram showing an embodiment for a method 700 of processing digital fingerprints and associated data received by a server. Method 700 is performed by processing logic that may comprise hardware (circuitry, dedicated logic, etc.), software (such as is run on a general purpose computer system or a dedicated machine), or a combination of both. In one embodiment, the method 700 is performed by a server system 140 of **Figure 1.** For example, the method of **Figure 7** may be performed by a digital fingerprint manager of a server system.

[0069] At block 705 of method 700, processing logic receives a digital fingerprint and associated data from a user device. At block 710, processing logic compares the user digital fingerprint received to a collection of master digital known work fingerprints stored on the server system such as in fingerprint storage 155 of **Figure 1.** At block 715, processing logic identifies an additional fingerprint from the collection of master fingerprints of known works or some portion thereof that approximately matches the user digital fingerprint received at block 705. A person skilled in the art will recognize that there are a plethora of digital fingerprint techniques available to perform this matching. In one embodiment, the method may

identify fingerprints matching known works such as movies, songs, episodic video, user generated video, speeches, paintings, photographs, or the like.

[0070] At block 720, processing logic associates the data received with the matching master digital fingerprint stored on the server. In one embodiment, the associated data may be located in the same data store record as the stored digital fingerprint. In another embodiment, the associated data may be referenced by the stored digital fingerprint via a pointer or other linking mechanism known in the art such that the associated data and digital fingerprint need not be in the same storage media. At block 725, processing logic stores the associated data for later presentation during media playback. In one embodiment, the associated data may be stored on the server system in data storage 160 of **Figure 1.**

[0071] **Figure 8** is a flow diagram showing an embodiment for an optional method 800 of streaming media work to a client with associated data. Method 800 is performed by processing logic that may comprise hardware (circuitry, dedicated logic, etc.), software (such as is run on a general purpose computer system or a dedicated machine), or a combination of both. In one embodiment, the method 800 is performed by a server system 140 of **Figure 1.** For example, the method of **Figure 8** may be performed by a media server of a server system.

[0072] At block 805 of method 800, processing logic receives a request from a client to stream a known work. In one embodiment, a client may be a playback module 122 that executes on a user device 105-110 of **Figure 1.** For example, the client may be a personal computer, tablet computer, smart phone, smart eyewear, or other computing device. Since server systems generally possess more processing power than client devices, the processing used to determine the particular point in a work to present associated data may be conducted by the server and streamed to the client such that the client need only display the contents of the stream without further intervention. This method allows for client devices with limited processing resources to receive associated data at the proper point in the playback of a work.

[0073] At block 810, processing logic streams the known work to the client. It is understood that the method may utilize streaming technologies and known network protocols known in the art to transmit the media to the client across a network. In one embodiment, the method may access a media server such as media server 150 of **Figure 1** to stream the media to the client via network 135 of **Figure 1.**

[0074] At block 815, processing logic generates a user fingerprint of the streamed work. At block 820, processing logic compares the user digital fingerprint generated at block 815 to fingerprints associated with stored data. At block 825, processing logic determines whether the generated user fingerprint matches a stored fingerprint associated with the data. If so, the method continues to block 830. Otherwise, the method proceeds to block 840,

[0075] At block 830, processing logic determines the point in the known work to insert data associated with the matching fingerprint. At block 835, processing logic sends the associated data to the client for presentation at the point in the known work. In one embodiment, the method may send the associated data embedded in the work as part of a single stream of data. In another embodiment, the method may establish two connections, ports, or "threads" with the client device sending the work and the associated data separately to be received by the client application. The client may simply present the work and the data without performing any additional processing to determine where to insert the data. At block 840, processing logic determines whether the work is still being streamed. If so, the method returns to block 815. Otherwise, the method ends.

[0076] **Figure 9** is a flow diagram showing an embodiment for a method 900 of processing client requests for data subscriptions by a server. Method 900 is performed by processing logic that may comprise hardware (circuitry, dedicated logic, etc.), software (such as is run on a general purpose computer system or a dedicated machine), or a combination of both. In one embodiment, the method 900 is performed by a server system 140 of **Figure 1.** For example, the method of **Figure 9** may be performed by a subscription manager 220 of **Figure 2B.**

[0077] At block 905 of method 900, processing logic receives a request from a client to subscribe to data provided by an entity. Alternatively, the request may be to subscribe to data having a particular categorization. At block 910, processing logic identifies a work to be played by the client. In one embodiment, processing logic may receive a digital fingerprint from a client to identify the work. In another embodiment, processing logic may receive the identity of the work (e.g., as part of the subscription request received at block 905). For example, processing logic may receive a request from a client to subscribe to all commentary for a specific video work.

[0078] At block 915, processing logic identifies data associated with the work. In one embodiment, the method may identify this information by utilizing data storage 160 of **Figure 1.** At block 920, processing logic transmits the data and an associated fingerprint to the requesting client for use. In one embodiment, the method may transmit all data and fingerprints for the work to the client such that the client can store the information and use the data when the work is played back or sampled at some later time. In an alternative embodiment, the method may transmit the data as the work is being played back on the client. In another embodiment, the method may transmit all data and fingerprints from a particular user to the subscribing client such that the client can store the information and make use of it whenever the work is played or sampled.

[0079] **Figure 10** illustrates a sequence diagram of an example transmission of digital fingerprints and data between a producer client 1005, a server 1010, and a consumer client 1015, in accordance with one embodiment of the present invention. In one embodiment, producer

client 1005 and/or consumer client 1015 may be user devices 105-110 of **Figure 1.** In one embodiment, producer client 1005 and consumer client 1015 may logically represent the same physical device. In one embodiment, server 1010 may be server system 140 of **Figure 1.**

[0080] At block 1020, the producer client 1005 generates a user fingerprint of a segment of a work and associates data with the fingerprint. Producer client 1005 then executes processing logic to send the user fingerprint and data to server 1025. At block 1030, the server associates the data and user fingerprint received with a known work fingerprint of a known work. The data is then stored on the server system. In one embodiment, the data is stored in data storage 160 of **Figure 1.**

[0081] At block 1035, the server identifies a data subscription to the data stored at block 1030. The server then executes processing logic to send fingerprint and data to client 1040. Consumer client 1015 receives the fingerprint and data and begins sampling a media stream (or playback of a work) at block 1045. At block 1050, consumer client 1015 determines a point in the work to present the associated data. At block 1055, consumer client 1015 presents the data at the particular point in the work.

[0082] Note that the above embodiments have been described with reference to recorded works. However, embodiments of the present invention may additionally be applied to live video recordings. For example, a user may include a video camera that records a scene at a particular location (e.g., at Alcatraz island). The video camera may additionally include a display that presents the image being recorded. In one embodiment, the user may add data to the generated video recording, which may be sent to a server along with an associated fingerprint. When another individual takes an image of the same scene, his video camera may receive the data, and may present the data previously entered by the first user when the scene is detected. In one embodiment, a geo-location is associated with the data and fingerprint. When a user device determines that it is at that geo-location, it may retrieve data that has been recorded by other users at that geo-location.

[0083] **Figure 11** is a block diagram illustrating an exemplary computer system 1100 configured to perform any one or more of the methodologies performed herein. In one embodiment, the computer system 1100 corresponds to a user device 105-110 of **Figure 1.** For example, computer system 1100 may be any type of computing device such as a PDA, a mobile phone, a laptop computer, a portable media player, a tablet computer, a camera, a video camera, a netbook, a desktop computer, a gaming console, a DVD player, a computing pad, a media center, and the like. Computer system 1100 may also correspond to one or more devices of the server system 140 of **Figure 1.** For example, computer system 100 may be a rackmount server, a desktop computer, a network router, switch or bridge, or any other computing device. The computer system 1100 may operate in the capacity of a server or a client machine in client-server network environment, or as a peer machine in a peer-to-peer (or distributed) network environment. Further, while only a single machine is illustrated, the computer system 1100 shall also be taken to include any collection of machines that individually or jointly execute a set (or multiple sets) of instructions to perform any one or more of the methodologies discussed herein.

[0084] The computer system 1100 includes one or more processing devices 1135, which may include general-purpose processing devices such as central processing units (CPUs), microcontrollers, microprocessors, systems on a chip (SoC), or the like. The processing devices 1135 may further include field programmable gate arrays, dedicated chipsets, application specific integrated circuits (ASIC), a field programmable gate arrays (FPGA), digital signal processors (DSP), network processors, or the like. The computer system 1100 also includes system memory 1105, which may correspond to any combination of volatile and/or non-volatile storage mechanisms. The system memory 1105 stores information which may provide an operating system component 1110, various program modules 1115 such as digital fingerprint manager 1185, program data 1120, and/or other components. The computer system 1100 may perform functions by using the processing device(s) 1135 to execute instructions provided by the system memory 1105. Such instructions may be provided as software or firmware. Alternatively, or additionally, the processing device(s) 1135 may include hardwired instruction sets (e.g., for performing functionality of the digital fingerprint manager 1185). The processing device 1135, system memory 1105 and additional components may communicate via a bus 1180.

[0085] The computer system 1100 also includes a data storage device 1125 that may be composed of one or more types of removable storage and/or one or more types of non-removable storage. The data storage device 1125 includes a computer-readable storage medium 1130 on which is stored one or more sets of instructions embodying any one or more of the methodologies or functions described herein. As shown, instructions for the digital fingerprint manager 1185 may reside, completely or at least partially, within the computer readable storage medium 1130, system memory 1105 and/or within the processing device(s) 1135 during execution thereof by the computer system 1100, the system memory 1105 and the processing device(s) 1135 also constituting computer-readable media. While the computer-readable storage medium 1130 is shown in an exemplary embodiment to be a single medium, the term "computer-readable storage medium" should be taken to include a single medium or multiple media (e.g., a centralized or distributed database, and/or associated caches and servers) that store the one or more sets of instructions. The term "computer-readable storage medium" shall also be taken to include any medium that is capable of storing or encoding a set of instructions for execution by the machine

and that cause the machine to perform any one or more of the methodologies of the present disclosure. The term "computer-readable storage medium" shall accordingly be taken to include, but not be limited to, solid-state memories, optical media, and magnetic media.

[0086] The computer system 1100 may also include one or more input devices 1140 (keyboard, mouse device, specialized selection keys, etc.) and one or more output devices 1145 (displays, printers, audio output mechanisms, etc.). In one embodiment, the computer system 1100 is a user device that includes one or more microphones 1150 and one or more speakers 1155.

[0087] The computer system may additionally include a wireless modem 1160 to allow the computer system 1100 to communicate via a wireless network (e.g., such as provided by a wireless communication system) with other computing devices, such as remote user devices, a server system, and so forth. The wireless modem 1160 allows the computer system 1100 to handle both voice and non-voice communications (such as communications for text messages, multimedia messages, media downloads, web browsing, etc.) with a wireless communication system. The wireless modem 1160 may provide network connectivity using any type of mobile network technology including, for example, cellular digital packet data (CDPD), general packet radio service (GPRS), enhanced data rates for GSM evolution (EDGE), universal mobile telecommunications system (UMTS), 1 times radio transmission technology (1xRTT), evaluation data optimized (EVDO), high-speed down-link packet access (HSDPA), WiFi, long term evolution (LTE), worldwide interoperability for microwave access (WiMAX), etc.

[0088] Computer system 1100 may additionally include a network interface device 1175 such as a network interface card (NIC) to connect to a network.

[0089] In the above description, numerous details are set forth. It will be apparent, however, to one of ordinary skill in the art having the benefit of this disclosure, that embodiments of the invention may be practiced without these specific details. In some instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the description.

[0090] Some portions of the detailed description are presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

[0091] It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the above discussion, it is appreciated that throughout the description, discussions utilizing terms such as "receiving", "generating", "associating", "querying", "analyzing", "replacing" or the like, refer to the actions and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (e.g., electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

[0092] Some portions of the detailed description are presented in terms of methods. These methods may be performed by processing logic that may comprise hardware (circuitry, dedicated logic, etc.), software (such as is run on a general purpose computer system or a dedicated machine), or a combination of both. In certain embodiments, the methods are performed by a user device, such as user devices 105-110 of **Figure 1.** In other embodiments, the methods are performed by server devices, such as server system 140 of **Figure 1.**

[0093] Embodiments of the invention also relate to an apparatus for performing the operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, and magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic instructions.

[0094] It is to be understood that the above description is intended to be illustrative, and not restrictive. Many other embodiments will be apparent to those of skill in the art upon reading and understanding the above description. The scope of the invention should, therefore, be determined with reference to the appended claims, along with the full scope of equivalents to which such claims are entitled.

Claims

1. A method comprising:

during sampling or a first playback of a work, receiving, by a user device, a command to associate data with the work at a segment of the

work occuring at a particular point in time in the work wherein the work is one of an audio rendering, a video rendering, or a video/audio rendering;

generating, by the user device, a user digital fingerprint of a segment of the work, wherein the segment comprises at least one of a sequence of frames of the work or adjacent in time samples of the work that begin at the particular point in time in the work, and wherein the user digital fingerprint comprises a feature vector of the segment of the work generated by the user device; and

associating the data with the user digital fingerprint, wherein the user digital fingerprint associated with the data will cause the user device or an additional user device to identify the segment of the work at a point in a time range that corresponds to the particular point in time in the work plus or minus an error factor during a subsequent playback of the work and to present the data during the subsequent playback of the segment at the point in the time range.

2. The method of claim 1, wherein said data comprises at least one of a text commentary, a digital image, a uniform resource locator URL, a video, a sound effect, or a command.

3. The method of claim 1, further comprising:

querying a server, the query comprising the user digital fingerprint;

receiving a master digital fingerprint of a segment of a copy of the work from the server, the segment of the copy corresponding to the segment of the work, wherein the master digital fingerprint is a digital fingerprint from a high quality version of the work that comprises a feature vector of the high quality version of the work, wherein the high quality version of the work has at least one of a high signal to noise ratio SNR or a high peak signal to noise ratio PSNR ; and

replacing the user digital fingerprint with the master digital fingerprint.

4. The method of claim 3, further comprising:

analyzing a quality of the segment of the work based on at least one of characteristics of the segment or characteristics of the user digital fingerprint; and

querying the server in response to determining, based on the analysis, that the quality of the segment of the work fails to satisfy a quality threshold.

5. The method of claim 1, further comprising:

sending a request to a server to subscribe to additional data provided by an entity that generated the additional data;

receiving the additional data and an additional fingerprint associated with the additional data, the additional fingerprint having been generated from an additional segment of the work ;

during the subsequent playback of the work on the user device, identifying the additional segment of the work using the additional fingerprint; and

presenting, by the user device, the additional data during the playback of the additional segment of the work.

6. The method of claim 5, wherein identifying the additional segment of the work comprises:

generating a plurality of fingerprints for a plurality of segments of the work; and

comparing the additional fingerprint to the plurality of fingerprints to find one of the plurality of fingerprints that approximately matches the additional fingerprint.

7. The method of claim 1, wherein during the subsequent playback of the work by the user device or the additional user device, the user device or the additional user device is to:

generate a plurality of additional digital fingerprints of the work;

identify an approximate match between one of the plurality of additional digital fingerprints and the user digital fingerprint or a master digital fingerprint corresponding to the user digital fingerprint;

determine the point in the time range of the work associated with the additional digital fingerprint that matches the user digital fingerprint or the master digital fingerprint; and

present the data at the point in the time range, wherein the master digital fingerprint is a digital fingerprint generated from a high quality version of the work, wherein the high quality version of the work has at least one of a high signal to noise ratio SNR or a high peak signal to noise ratio PSNR.

8. A computer readable storage medium having instructions that, when executed by the user device, cause the user device to perform the method of any of claims 1-6.

9. A system comprising:

a first user device comprising a first memory and a first processing device operatively coupled to the first memory, wherein the first user device is to:

during sampling or a first playback of a work, receiving, by a user device, a command to associate data with the work at a segment of the work occurring at a particular point in time in the work, wherein the work is one of an audio rendering, a video rendering, or a video/audio rendering;
generating, by the user device, a user digital fingerprint of a segment of the work, wherein the segment comprises at least one of a sequence of frames of the work or adjacent in time samples of the work that begin at the particular point in time in the work, and wherein the user digital fingerprint comprises a feature vector of the segment of the work generated by the user device; and
associating the data with the user digital fingerprint, wherein the user digital fingerprint associated with the data will cause the user device or an additional user device to identify the segment of the work at a point of time range that corresponds to the particular point in time in the work plus or minus an error factor during a subsequent playback of the work and to present the data during the subsequent playback of the segment at the point in the time range.

10. The system of claim 9, further comprising: a second user device comprising a second memory and a second processing device operatively coupled to the second memory, wherein the second user device is to:

receive the data and the user digital fingerprint or a master digital fingerprint corresponding to the user digital fingerprint;
initiate a subsequent playback of the work; and perform the following during the subsequent playback of the work:
generate a plurality of additional digital fingerprints of the work;
identify an approximate match between one of the plurality of additional digital fingerprints and the user digital fingerprint or the master digital fingerprint corresponding to the user digital fingerprint;
determine the point in the time range in the work associated with the additional digital fingerprint that matches the user digital fingerprint or the master digital fingerprint; and
present the data at the point in the time range in the work, wherein the master digital fingerprint is a digital fingerprint generated from a high quality version of the work,
wherein the high quality version of the work has at least one of a high signal to noise ratio SNR or a high peak signal to noise ratio PSNR.

11. The system of claim 9, wherein the first user device is further to:

initiate a subsequent playback of the work; and perform the following during the subsequent playback of the work:

generate a plurality of additional digital fingerprints of the work;
identify an approximate match between one of the plurality of additional digital fingerprints and the user digital fingerprint or a master digital fingerprint corresponding to the user digital fingerprint;
determine the point in the time range in the work associated with the additional digital fingerprint that matches the user digital fingerprint or the master digital fingerprint; and
present the data at the point in the time range in the work, wherein the master digital fingerprint is a digital fingerprint generated from a high quality version of the work, wherein the high quality version of the work has at least one of a high signal to noise ratio SNR or a high peak signal to noise ratio PSNR.

12. The system of claim 9, wherein said data comprises at least one of a text commentary, a digital image, a uniform resource locator URL, a video, a sound effect, or a command.

13. The system of claim 9, wherein the first user device is further to:

query a server, the query comprising the user digital fingerprint;
receive a master digital fingerprint of a segment of a copy of the work from the server, the segment of the copy corresponding to the segment of the work, wherein the master digital fingerprint is a digital fingerprint from a high quality version of the work that comprises a feature vector of the high quality version of the work, wherein the high quality version of the work has at least one of a high signal to noise ratio SNR or a high peak signal to noise ratio PSNR, and
replace the user digital fingerprint with the master digital fingerprint.

14. The system of claim 13, wherein the first user device is further to:

analyze a quality of the segment of the work based on at least one of characteristics of the segment or characteristics of the user digital fingerprint; and
query the server in response to determining,

based on the analysis, that the quality of the segment of the work fails to satisfy a quality threshold.

15. The system of claim 9, wherein the first user device is further to:

send a request to a server to subscribe to additional data provided by an entity that generates the additional data;
receive the additional data and an additional fingerprint associated with the additional data, the additional fingerprint having been generated from an additional segment of the work;
during the subsequent playback of the work on the user device, identify the additional segment of the work using the additional fingerprint, wherein identifying the additional segment of the work comprises:

generating a plurality of fingerprints for a plurality of segments of the work; and
comparing the additional fingerprint to the plurality of fingerprints to find one of the plurality of fingerprints that approximately matches the additional fingerprint; and

present the additional data during the playback of the additional segment.

Patentansprüche

1. Verfahren, das Folgendes beinhaltet:

Empfangen, während des Abtastens oder einer ersten Wiedergabe einer Arbeit, durch ein Benutzergerät, eines Befehls zum Assoziieren von Daten mit der Arbeit an einem Segment der Arbeit, das an einem bestimmten Zeitpunkt in der Arbeit auftritt, wobei die Arbeit eines aus Audio-Rendering, Video-Rendering und Video/Audio-Rendering ist;
Erzeugen, durch das Benutzergerät, eines digitalen Benutzerfingerabdrucks eines Segments der Arbeit, wobei das Segment wenigstens eines aus einer Sequenz von Frames der Arbeit oder zeitlich benachbarten Samples der Arbeit umfasst, die an dem bestimmten Zeitpunkt in der Arbeit beginnen, und wobei der digitale Benutzerfingerabdruck einen Merkmalsvektor des Segments der von dem Benutzergerät erzeugten Arbeit umfasst; und
Assoziieren der Daten mit dem digitalen Benutzerfingerabdruck, wobei der mit den Daten assoziierte digitale Benutzerfingerabdruck bewirkt, dass das Benutzergerät oder ein zusätzliches Benutzergerät das Segment der Arbeit an einem Punkt in einem Zeitbereich identifiziert, der dem bestimmten Zeitpunkt in der Arbeit plus oder minus einem Fehlerfaktor während einer nachfolgenden Wiedergabe der Arbeit entspricht, und um die Daten während der nachfolgenden Wiedergabe des Segments an dem Punkt in dem Zeitbereich zu präsentieren.

2. Verfahren nach Anspruch 1, wobei die genannten Daten wenigstens eines aus einem Textkommentar, einem digitalen Bild, einem URL (Uniform Resource Locator), einem Video, einem Toneffekt und einem Befehl umfasst.

3. Verfahren nach Anspruch 1, das ferner Folgendes beinhaltet:

Abfragen eines Servers, wobei die Abfrage den digitalen Benutzerfingerabdruck umfasst;
Empfangen eines digitalen Masterfingerabdrucks eines Segments einer Kopie der Arbeit von dem Server, wobei das Segment der Kopie dem Segment der Arbeit entspricht, wobei der digitale Masterfingerabdruck ein digitaler Fingerabdruck von einer Hochqualitätsversion der Arbeit ist, die einen Merkmalsvektor der Hochqualitätsversion der Arbeit umfasst, wobei die Hochqualitätsversion der Arbeit wenigstens eines aus einem hohen Signal-Rausch-Verhältnis SNR oder einem hohen Spitzen-Signal-Rausch-Verhältnis PSNR hat; und
Ersetzen des digitalen Benutzerfingerabdrucks durch den digitalen Masterfingerabdruck.

4. Verfahren nach Anspruch 3, das ferner Folgendes beinhaltet:

Analysieren einer Qualität des Segments der Arbeit wenigstens auf der Basis von Charakteristiken des Segments oder Charakteristiken des digitalen Benutzerfingerabdrucks; und
Abfragen des Servers als Reaktion auf die Feststellung, auf der Basis der Analyse, dass die Qualität des Segments der Arbeit eine Qualitätsschwelle nicht erfüllt.

5. Verfahren nach Anspruch 1, das ferner Folgendes beinhaltet:

Senden einer Anforderung zu einem Server zum Abonnieren von zusätzlichen Daten, die von einer Entität bereitgestellt werden, die die zusätzlichen Daten erzeugt hat;
Empfangen der zusätzlichen Daten und eines mit den zusätzlichen Daten assoziierten zusätzlichen Fingerabdrucks, wobei der zusätzliche Fingerabdruck von einem zusätzlichen Segment der Arbeit erzeugt wurde;

Identifizieren, während der nachfolgenden Wiedergabe der Arbeit auf dem Benutzergerät, des zusätzlichen Segments der Arbeit mit dem zusätzlichen Fingerabdruck; und
Präsentieren, durch das Benutzergerät, der zusätzlichen Daten während der Wiedergabe des zusätzlichen Segments der Arbeit.

6. Verfahren nach Anspruch 5, wobei das Identifizieren des zusätzlichen Segments der Arbeit Folgendes beinhaltet:

Erzeugen von mehreren Fingerabdrücken für mehrere Segmente der Arbeit; und
Vergleichen des zusätzlichen Fingerabdrucks mit den mehreren Fingerabdrücken, um einen der mehreren Fingerabdrücke zu finden, der etwa mit dem zusätzlichen Fingerabdruck übereinstimmt.

7. Verfahren nach Anspruch 1, wobei während der nachfolgenden Wiedergabe der Arbeit durch das Benutzergerät oder das zusätzliche Benutzergerät das Benutzergerät oder das zusätzliche Benutzergerät dient zum:

Erzeugen von mehreren zusätzlichen digitalen Fingerabdrücken der Arbeit;
Identifizieren einer ungefähren Übereinstimmung zwischen einem der mehreren zusätzlichen digitalen Fingerabdrücke und dem digitalen Benutzerfingerabdruck oder einem digitalen Masterfingerabdruck entsprechend dem digitalen Benutzerfingerabdruck;
Bestimmen des Punkts in dem Zeitbereich der mit dem zusätzlichen digitalen Fingerabdruck assoziierten Arbeit, der mit dem digitalen Benutzerfingerabdruck oder dem digitalen Masterfingerabdruck übereinstimmt; und
Präsentieren der Daten an dem Punkt in dem Zeitbereich, wobei der digitale Masterfingerabdruck ein von einer Hochqualitätsversion der Arbeit erzeugter digitaler Fingerabdruck ist, wobei die Hochqualitätsversion der Arbeit wenigstens eines aus einem hohen Signal-Rausch-Verhältnis SNR oder einem hohen Spitzen-Signal-Rausch-Verhältnis PSNR ist.

8. Computerlesbares Speichermedium mit Befehlen, die bei Ausführung durch das Benutzergerät bewirken, dass das Benutzergerät das Verfahren nach einem der Ansprüche 1-6 durchführt.

9. System, das Folgendes umfasst:
ein erstes Benutzergerät, das einen ersten Speicher und ein erstes Verarbeitungsgerät umfasst, das operativ mit dem ersten Speicher gekoppelt ist, wobei das erste Benutzergerät dient zum:

Empfangen, während des Abtastens oder einer ersten Wiedergabe einer Arbeit, durch ein Benutzergerät, eines Befehls zum Assoziieren von Daten mit der Arbeit an einem Segment der Arbeit, das an einem bestimmten Zeitpunkt in der Arbeit auftritt, wobei die Arbeit eines aus Audio-Rendering, Video-Rendering und Video/Audio-Rendering ist;
Erzeugen, durch das Benutzergerät, eines digitalen Benutzerfingerabdrucks eines Segments der Arbeit, wobei das Segment wenigstens eines aus einer Sequenz von Frames der Arbeit oder zeitlich benachbarten Samples der Arbeit umfasst, die an dem bestimmten Zeitpunkt in der Arbeit beginnen, und wobei der digitale Benutzerfingerabdruck einen Merkmalsvektor des Segments der von dem Benutzergerät erzeugten Arbeit umfasst; und
Assoziieren der Daten mit dem digitalen Benutzerfingerabdruck, wobei der mit den Daten assoziierte digitale Benutzerfingerabdruck bewirkt, dass das Benutzergerät oder ein zusätzliches Benutzergerät das Segment der Arbeit an einem Punkt des Zeitbereichs identifiziert, der dem bestimmten Zeitpunkt in der Arbeit plus oder minus einem Fehlerfaktor während einer nachfolgenden Wiedergabe der Arbeit entspricht, und zum Präsentieren der Daten während der nachfolgenden Wiedergabe des Segments an dem Punkt in dem Zeitbereich.

10. System nach Anspruch 9, das ferner Folgendes umfasst:
ein zweites Benutzergerät, das einen zweiten Speicher und ein zweites Verarbeitungsgerät umfasst, das operativ mit dem zweiten Speicher gekoppelt ist, wobei das zweite Benutzergerät dient zum:

Empfangen der Daten und des digitalen Benutzerfingerabdrucks oder eines digitalen Masterfingerabdrucks entsprechend dem digitalen Benutzerfingerabdruck;
Einleiten einer nachfolgenden Wiedergabe der Arbeit; und
Durchführen des Folgenden während der nachfolgenden Wiedergabe der Arbeit:

Erzeugen von mehreren zusätzlichen digitalen Fingerabdrücken der Arbeit;
Identifizieren einer ungefähren Übereinstimmung zwischen einem der mehreren zusätzlichen digitalen Fingerabdrücke und dem digitalen Benutzerfingerabdruck oder dem digitalen Masterfingerabdruck entsprechend dem digitalen Benutzerfingerabdruck;
Bestimmen des Punkts in dem Zeitbereich in der mit dem zusätzlichen digitalen Fin-

gerabdruck assoziierten Arbeit, der mit dem digitalen Benutzerfingerabdruck oder dem digitalen Masterfingerabdruck übereinstimmt; und

Präsentieren der Daten an dem Punkt in dem Zeitbereich in der Arbeit, wobei der digitale Masterfingerabdruck ein von einer Hochqualitätsversion der Arbeit erzeugter digitaler Fingerabdruck ist, wobei die Hochqualitätsversion der Arbeit wenigstens eines aus einem hohen Signal-Rausch-Verhältnis SNR oder einem hohen Spitzen-Signal-Rausch-Verhältnis PSNR hat.

11. System nach Anspruch 9, wobei das erste Benutzergerät ferner dient zum:

Einleiten einer nachfolgenden Wiedergabe der Arbeit; und

Durchführen des Folgenden während der nachfolgenden Wiedergabe der Arbeit:

Erzeugen von mehreren zusätzlichen digitalen Fingerabdrücken der Arbeit;

Identifizieren einer ungefähren Übereinstimmung zwischen einem der mehreren zusätzlichen digitalen Fingerabdrücke und dem digitalen Benutzerfingerabdruck oder einem digitalen Masterfingerabdruck entsprechend dem digitalen Benutzerfingerabdruck;

Bestimmen des Punkts in dem Zeitbereich in der mit dem zusätzlichen digitalen Fingerabdruck assoziierten Arbeit, der mit dem digitalen Benutzerfingerabdruck oder dem digitalen Masterfingerabdruck übereinstimmt; und

Präsentieren der Daten an dem Punkt in dem Zeitbereich in der Arbeit, wobei der digitale Masterfingerabdruck ein von einer Hochqualitätsversion der Arbeit erzeugter digitaler Fingerabdruck ist, wobei die Hochqualitätsversion der Arbeit wenigstens eines aus einem hohen Signal-Rausch-Verhältnis SNR oder einem hohen Spitzen-Signal-Rausch-Verhältnis PSNR hat.

12. System nach Anspruch 9, wobei die genannten Daten wenigstens eines aus einem Textkommentar, einem digitalen Bild, einem URL (Uniform Resource Locator), einem Video, einem Toneffekt und einem Befehl umfasst.

13. System nach Anspruch 9, wobei das erste Benutzergerät ferner dient zum:

Abfragen eines Servers, wobei die Abfrage den digitalen Benutzerfingerabdruck umfasst;

Empfangen eines digitalen Masterfingerabdrucks eines Segments einer Kopie der Arbeit von dem Server, wobei das Segment der Kopie dem Segment der Arbeit entspricht, wobei der digitale Masterfingerabdruck ein digitaler Fingerabdruck von einer Hochqualitätsversion der Arbeit ist, der einen Merkmalsvektor der Hochqualitätsversion der Arbeit umfasst, wobei die Hochqualitätsversion der Arbeit wenigstens eines aus einem hohen Signal-Rausch-Verhältnis SNR oder einem hohen Spitzen-Signal-Rausch-Verhältnis PSNR hat; und

Ersetzen des digitalen Benutzerfingerabdrucks durch den digitalen Masterfingerabdruck.

14. System nach Anspruch 13, wobei das erste Benutzergerät ferner dient zum:

Analysieren einer Qualität des Segments der Arbeit wenigstens auf der Basis von Charakteristiken des Segments oder von Charakteristiken des digitalen Benutzerfingerabdrucks; und

Abfragen des Servers als Reaktion auf die Feststellung, auf der Basis der Analyse, dass die Qualität des Segments der Arbeit eine Qualitätsschwelle nicht erfüllt.

15. System nach Anspruch 9, wobei das erste Benutzergerät ferner dient zum:

Senden einer Anforderung zu einem Server zum Abonnieren von zusätzlichen Daten, die von einer Entität bereitgestellt werden, die die zusätzlichen Daten erzeugt;

Empfangen der zusätzlichen Daten und eines mit den zusätzlichen Daten assoziierten zusätzlichen Fingerabdrucks, wobei der zusätzliche Fingerabdruck von einem zusätzlichen Segment der Arbeit erzeugt wurde;

Identifizieren, während der nachfolgenden Wiedergabe der Arbeit auf dem Benutzergerät, des zusätzlichen Segments der Arbeit anhand des zusätzlichen Fingerabdrucks, wobei das Identifizieren des zusätzlichen Segments der Arbeit Folgendes beinhaltet:

Erzeugen von mehreren Fingerabdrücken für mehrere Segmente der Arbeit; und

Vergleichen des zusätzlichen Fingerabdrucks mit den mehreren Fingerabdrücken, um einen der mehreren Fingerabdrücke zu finden, der ungefähr mit dem zusätzlichen Fingerabdruck übereinstimmt; und

Präsentieren der zusätzlichen Daten während der Wiedergabe des zusätzlichen Segments.

Revendications

1. Procédé comprenant les étapes consistant à :

pendant un échantillonnage ou une première lecture d'un travail, recevoir, par un dispositif d'utilisateur, une commande pour associer des données au travail au niveau d'un segment du travail se produisant à un point particulier du travail dans le temps, le travail étant une restitution audio, une restitution vidéo ou une restitution audio/vidéo ;
générer, par le dispositif d'utilisateur, une empreinte numérique d'utilisateur d'un segment du travail, le segment comprenant au moins une trame d'une séquence de trames du travail ou adjacentes dans des échantillons de temps du travail qui commencent au point particulier du travail dans le temps, et l'empreinte numérique d'utilisateur comprenant un vecteur d'attribut du segment du travail généré par le dispositif d'utilisateur ; et
associer les données à l'empreinte numérique d'utilisateur, l'empreinte numérique d'utilisateur associée aux données amenant le dispositif d'utilisateur ou un dispositif d'utilisateur supplémentaire à identifier le segment du travail à un point d'une plage de temps qui correspond au point particulier du travail dans le temps plus ou moins un facteur d'erreur pendant une lecture ultérieure du travail, et à présenter les données pendant la lecture ultérieure du segment au point de la plage de temps.

2. Procédé selon la revendication 1, dans lequel lesdites données comprennent au moins un élément parmi un commentaire de texte, une image numérique, un localisateur uniforme de ressources, URL, une vidéo, un effet sonore ou une commande.

3. Procédé selon la revendication 1, comprenant en outre les étapes consistant à :

interroger un serveur, l'interrogation comprenant l'empreinte numérique d'utilisateur ;
recevoir une empreinte numérique maîtresse d'un segment d'une copie du travail en provenance du serveur, le segment de la copie correspondant au segment du travail, l'empreinte numérique maîtresse étant une empreinte numérique d'une version de haute qualité du travail, qui comprend un vecteur d'attribut de la version de haute qualité du travail, la version de haute qualité du travail ayant au moins un rapport parmi un rapport signal-bruit, SNR, élevé ou une valeur de crête de rapport signal-bruit, PSNR, élevée ; et
remplacer l'empreinte numérique d'utilisateur par l'empreinte numérique maîtresse.

4. Procédé selon la revendication 3, comprenant en outre les étapes consistant à :

analyser une qualité du segment du travail sur la base d'au moins une des caractéristiques du segment ou des caractéristiques de l'empreinte numérique d'utilisateur ; et
interroger le serveur en réponse à la détermination, sur la base de l'analyse, que la qualité du segment du travail ne satisfait pas à un seuil de qualité.

5. Procédé selon la revendication 1, comprenant en outre les étapes consistant à :

envoyer une demande à un serveur pour s'abonner à des données supplémentaires fournies par une entité qui a généré les données supplémentaires ;
recevoir les données supplémentaires et une empreinte supplémentaire associée aux données supplémentaires, l'empreinte supplémentaire ayant été générée à partir d'un segment supplémentaire du travail ;
pendant la lecture ultérieure du travail sur le dispositif d'utilisateur, identifier le segment supplémentaire du travail au moyen de l'empreinte supplémentaire ; et
présenter, par le dispositif d'utilisateur, les données supplémentaires pendant la lecture du segment supplémentaire du travail.

6. Procédé selon la revendication 5, dans lequel l'identification du segment supplémentaire du travail comprend les étapes consistant à :

générer une pluralité d'empreintes pour une pluralité de segments du travail ; et
comparer l'empreinte supplémentaire à la pluralité d'empreintes pour trouver parmi la pluralité d'empreintes une empreinte qui concorde approximativement avec l'empreinte supplémentaire.

7. Procédé selon la revendication 1, dans lequel, pendant la lecture ultérieure du travail par le dispositif d'utilisateur ou le dispositif d'utilisateur supplémentaire, le dispositif d'utilisateur ou le dispositif d'utilisateur supplémentaire est destiné à :

générer une pluralité d'empreintes numériques supplémentaires du travail ;
identifier une concordance approximative entre une empreinte de la pluralité d'empreintes numériques supplémentaires et l'empreinte numérique d'utilisateur ou une empreinte numérique

maîtresse correspondant à l'empreinte numérique d'utilisateur ;
déterminer le point de la plage de temps du travail associé à l'empreinte numérique supplémentaire qui concorde avec l'empreinte numérique d'utilisateur ou l'empreinte numérique maîtresse ; et
présenter les données au point de la plage de temps, l'empreinte numérique maîtresse étant une empreinte numérique générée à partir d'une version de haute qualité du travail, la version de haute qualité du travail ayant au moins un rapport parmi un rapport signal-bruit, SNR, élevé ou une valeur de crête de rapport signal-bruit, PSNR, élevée.

8. Support de stockage lisible par ordinateur, comprenant des instructions qui, lorsqu'elles sont exécutées par le dispositif d'utilisateur, amènent le dispositif d'utilisateur à réaliser le procédé selon l'une quelconque des revendications 1 à 6.

9. Système comprenant :
un premier dispositif d'utilisateur comprenant une première mémoire et un premier dispositif de traitement couplé de manière fonctionnelle à la première mémoire, le premier dispositif d'utilisateur étant destiné à :

pendant un échantillonnage ou une première lecture d'un travail, recevoir, par un dispositif d'utilisateur, une commande pour associer des données au travail au niveau d'un segment du travail se produisant à un point particulier du travail dans le temps, le travail étant une restitution audio, une restitution vidéo ou une restitution audio/vidéo ;
générer, par le dispositif d'utilisateur, une empreinte numérique d'utilisateur d'un segment du travail, le segment comprenant au moins une trame d'une séquence de trames du travail ou adjacentes dans des échantillons de temps du travail qui commencent au point particulier du travail dans le temps, et l'empreinte numérique d'utilisateur comprenant un vecteur d'attribut du segment du travail généré par le dispositif d'utilisateur ; et
associer les données à l'empreinte numérique d'utilisateur, l'empreinte numérique d'utilisateur associée aux données amenant le dispositif d'utilisateur ou un dispositif d'utilisateur supplémentaire à identifier le segment du travail à un point d'une plage de temps qui correspond au point particulier du travail dans le temps plus ou moins un facteur d'erreur pendant une lecture ultérieure du travail, et à présenter les données pendant la lecture ultérieure du segment au point de la plage de temps.

10. Système selon la revendication 9, comprenant en outre :
un second dispositif d'utilisateur comprenant une seconde mémoire et un second dispositif de traitement couplé de manière fonctionnelle à la seconde mémoire, le second dispositif d'utilisateur étant destiné à :

recevoir les données et l'empreinte numérique d'utilisateur ou une empreinte numérique maîtresse correspondant à l'empreinte numérique d'utilisateur ;
amorcer une lecture ultérieure du travail ; et
réaliser les étapes suivantes pendant la lecture ultérieure du travail :

générer une pluralité d'empreintes numériques supplémentaires du travail ;
identifier une concordance approximative entre une empreinte de la pluralité d'empreintes numériques supplémentaires et l'empreinte numérique d'utilisateur ou l'empreinte numérique maîtresse correspondant à l'empreinte numérique d'utilisateur ;
déterminer le point de la plage de temps du travail associé à l'empreinte numérique supplémentaire qui concorde avec l'empreinte numérique d'utilisateur ou l'empreinte numérique maîtresse ; et
présenter les données au point de la plage de temps du travail, l'empreinte numérique maîtresse étant une empreinte numérique générée à partir d'une version de haute qualité du travail, la version de haute qualité du travail ayant au moins un rapport parmi un rapport signal-bruit, SNR, élevé ou une valeur de crête de rapport signal-bruit, PSNR, élevée.

11. Système selon la revendication 9, dans lequel le premier dispositif d'utilisateur est en outre destiné à :

amorcer une lecture ultérieure du travail ; et
réaliser les étapes suivantes pendant la lecture ultérieure du travail :

générer une pluralité d'empreintes numériques supplémentaires du travail ;
identifier une concordance approximative entre une empreinte de la pluralité d'empreintes numériques supplémentaires et l'empreinte numérique d'utilisateur ou une empreinte numérique maîtresse correspondant à l'empreinte numérique d'utilisateur ;
déterminer le point de la plage de temps du travail associé à l'empreinte numérique supplémentaire qui concorde avec l'empreinte numérique d'utilisateur ou l'em-

preinte numérique maîtresse ; et
présenter les données au point de la plage de temps du travail, l'empreinte numérique maîtresse étant une empreinte numérique générée à partir d'une version de haute qualité du travail, la version de haute qualité du travail ayant au moins un rapport parmi un rapport signal-bruit, SNR, élevé ou une valeur de crête de rapport signal-bruit, PSNR, élevée.

12. Système selon la revendication 9, dans lequel lesdites données comprennent au moins un élément parmi un commentaire de texte, une image numérique, un localisateur uniforme de ressources, URL, une vidéo, un effet sonore ou une commande.

13. Système selon la revendication 9, dans lequel le premier dispositif d'utilisateur est en outre destiné à :

interroger un serveur, l'interrogation comprenant l'empreinte numérique d'utilisateur ;
recevoir une empreinte numérique maîtresse d'un segment d'une copie du travail en provenance du serveur, le segment de la copie correspondant au segment du travail, l'empreinte numérique maîtresse étant une empreinte numérique d'une version de haute qualité du travail, qui comprend un vecteur d'attribut de la version de haute qualité du travail, la version de haute qualité du travail ayant au moins un rapport parmi un rapport signal-bruit, SNR, élevé ou une valeur de crête de rapport signal-bruit, PSNR, élevée ; et
remplacer l'empreinte numérique d'utilisateur par l'empreinte numérique maîtresse.

14. Système selon la revendication 13, dans lequel le premier dispositif d'utilisateur est en outre destiné à :

analyser une qualité du segment du travail sur la base d'au moins une des caractéristiques du segment ou des caractéristiques de l'empreinte numérique d'utilisateur ; et
interroger le serveur en réponse à la détermination, sur la base de l'analyse, que la qualité du segment du travail ne satisfait pas à un seuil de qualité.

15. Système selon la revendication 9, dans lequel le premier dispositif d'utilisateur est en outre destiné à :

envoyer une demande à un serveur pour s'abonner à des données supplémentaires fournies par une entité qui génère les données supplémentaires ;
recevoir les données supplémentaires et une empreinte supplémentaire associée aux données supplémentaires, l'empreinte supplémentaire ayant été générée à partir d'un segment supplémentaire du travail ;
pendant la lecture ultérieure du travail sur le dispositif d'utilisateur, identifier le segment supplémentaire du travail au moyen de l'empreinte supplémentaire, l'identification du segment supplémentaire du travail comprenant les étapes consistant à :

générer une pluralité d'empreintes pour une pluralité de segments du travail ; et
comparer l'empreinte supplémentaire à la pluralité d'empreintes pour trouver parmi la pluralité d'empreintes une empreinte qui correspond approximativement à l'empreinte supplémentaire ; et

présenter les données supplémentaires pendant la lecture du segment supplémentaire.

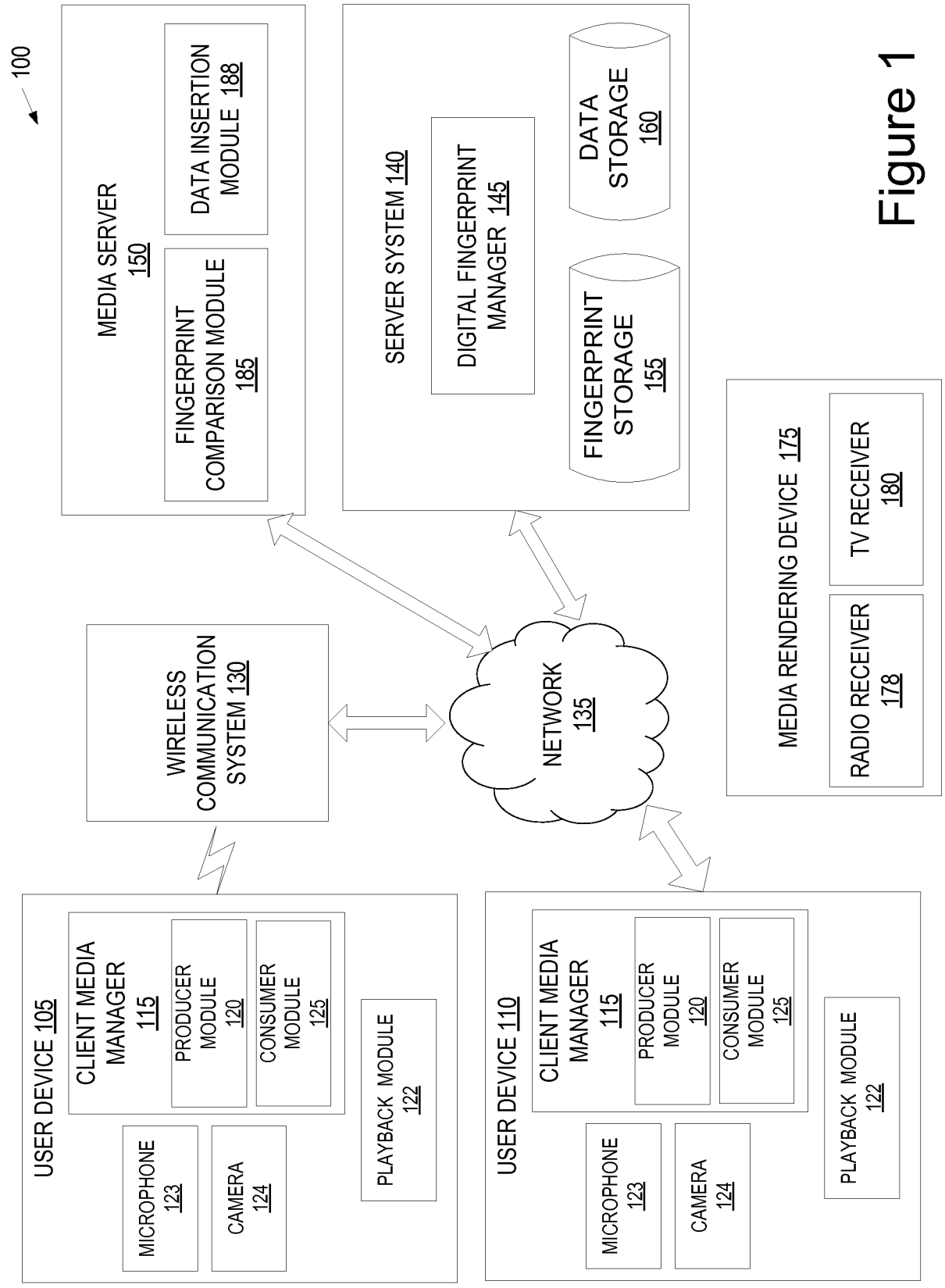
100
MEDIA SERVER 150
DATA INSERTION MODULE 188
FINGERPRINT COMPARISON MODULE 185
SERVER SYSTEM 140
DIGITAL FINGERPRINT MANAGER 145
DATA STORAGE 160
FINGERPRINT STORAGE 155
MEDIA RENDERING DEVICE 175
TV RECEIVER 180
RADIO RECEIVER 178
WIRELESS COMMUNICATION SYSTEM 130
NETWORK 135
USER DEVICE 105
CLIENT MEDIA MANAGER 115
PRODUCER MODULE 120
CONSUMER MODULE 125
PLAYBACK MODULE 122
MICROPHONE 123
CAMERA 124
USER DEVICE 110
CLIENT MEDIA MANAGER 115
PRODUCER MODULE 120
CONSUMER MODULE 125
PLAYBACK MODULE 122
MICROPHONE 123
CAMERA 124

Figure 1

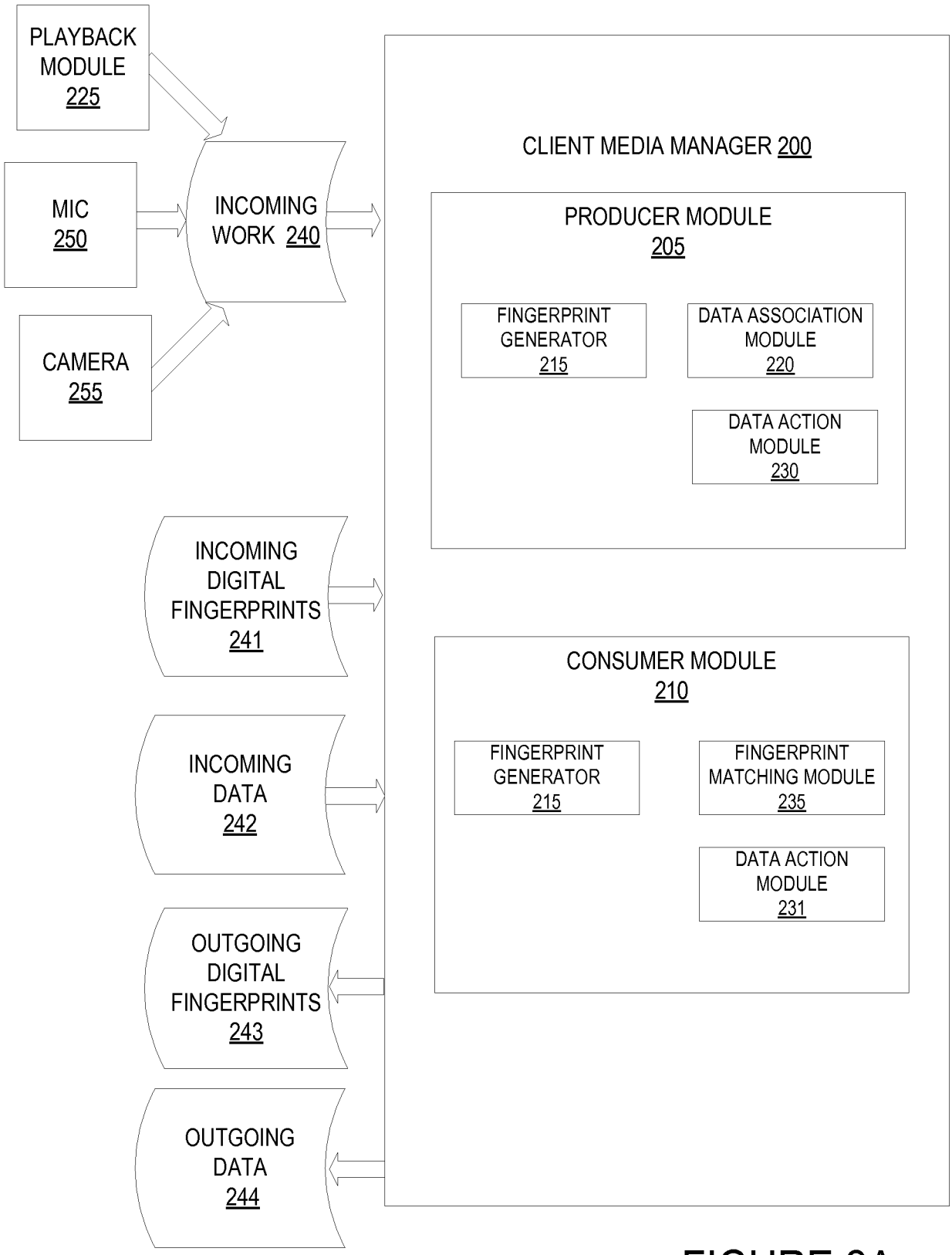
PLAYBACK MODULE 225
MIC 250
CAMERA 255
INCOMING WORK 240
CLIENT MEDIA MANAGER 200
PRODUCER MODULE 205
FINGERPRINT GENERATOR 215
DATA ASSOCIATION MODULE 220
DATA ACTION MODULE 230
INCOMING DIGITAL FINGERPRINTS 241
CONSUMER MODULE 210
FINGERPRINT GENERATOR 215
FINGERPRINT MATCHING MODULE 235
DATA ACTION MODULE 231
INCOMING DATA 242
OUTGOING DIGITAL FINGERPRINTS 243
OUTGOING DATA 244

FIGURE 2A

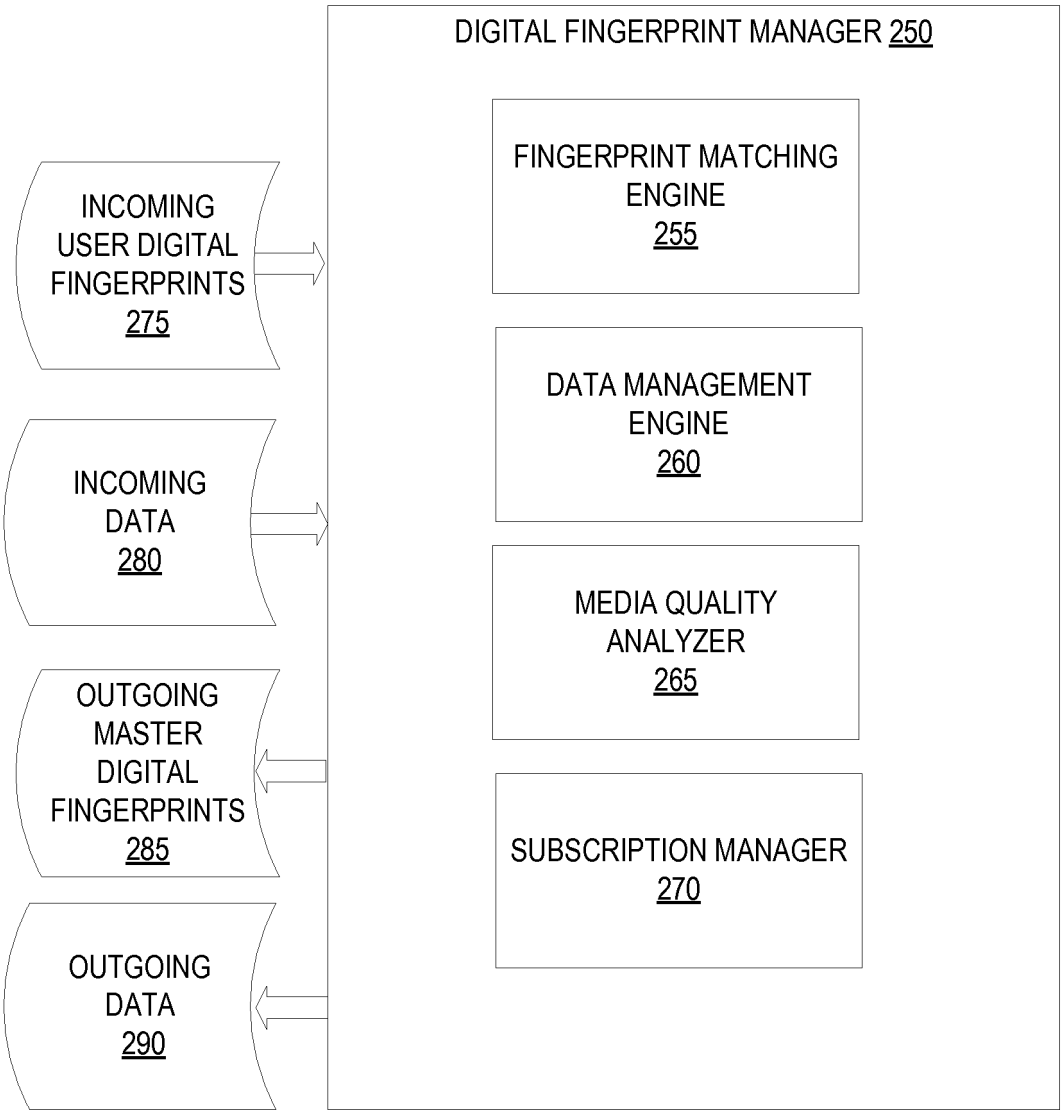
DIGITAL FINGERPRINT MANAGER 250
FINGERPRINT MATCHING ENGINE 255
DATA MANAGEMENT ENGINE 260
MEDIA QUALITY ANALYZER 265
SUBSCRIPTION MANAGER 270
INCOMING USER DIGITAL FINGERPRINTS 275
INCOMING DATA 280
OUTGOING MASTER DIGITAL FINGERPRINTS 285
OUTGOING DATA 290

FIGURE 2B

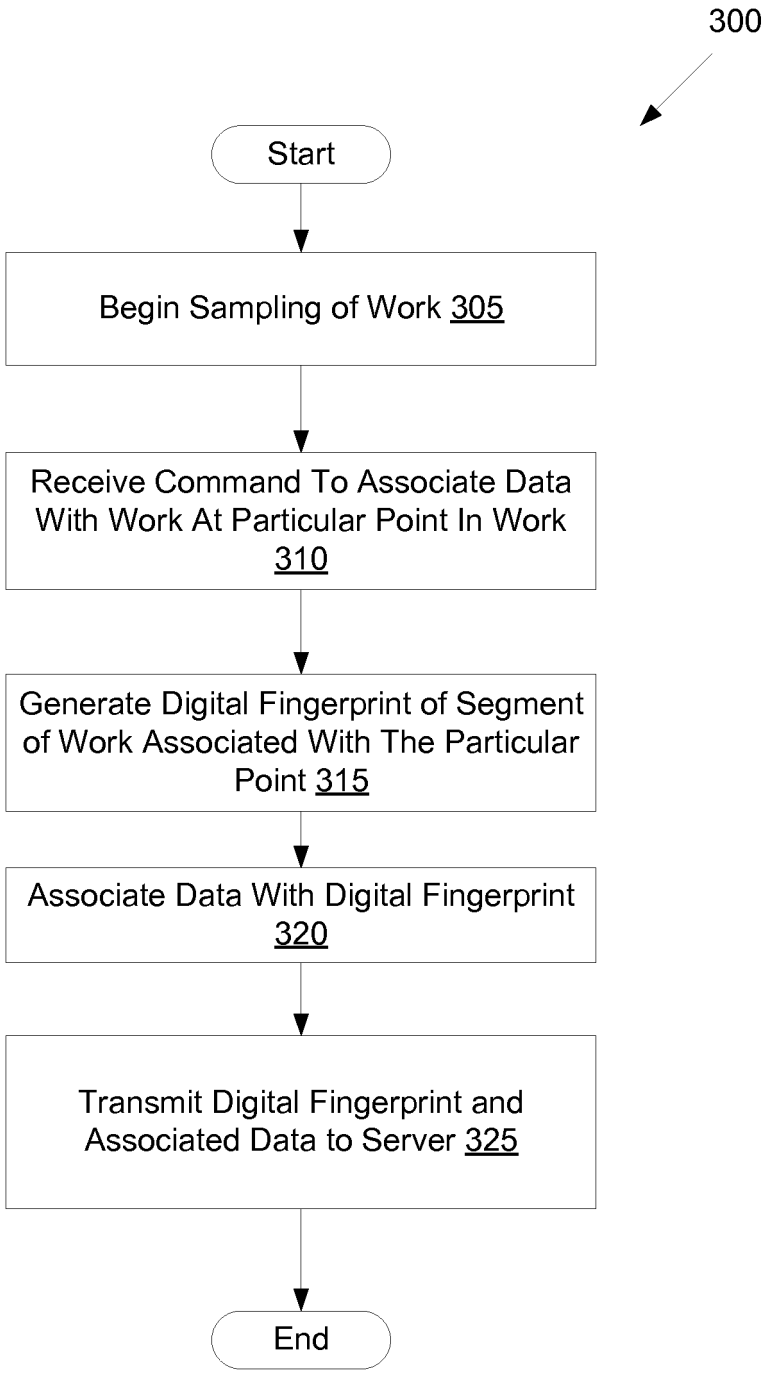
300
Start
Begin Sampling of Work 305
Receive Command To Associate Data With Work At Particular Point In Work 310
Generate Digital Fingerprint of Segment of Work Associated With The Particular Point 315
Associate Data With Digital Fingerprint 320
Transmit Digital Fingerprint and Associated Data to Server 325
End

Figure 3

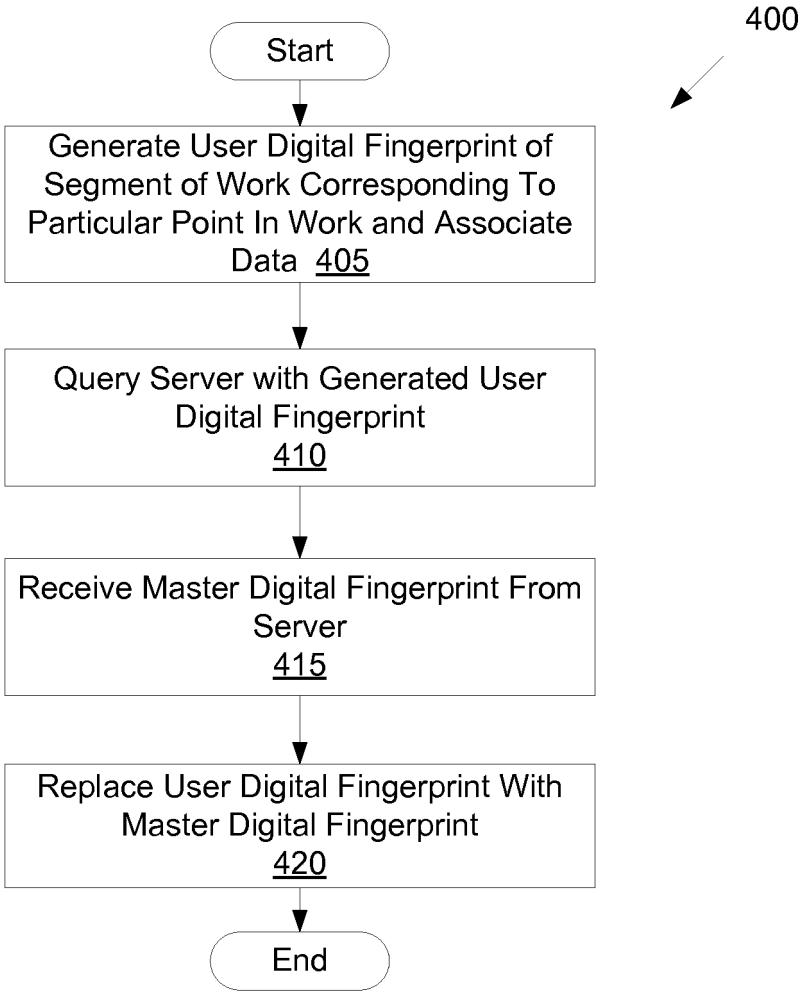
400
Start
Generate User Digital Fingerprint of Segment of Work Corresponding To Particular Point In Work and Associate Data 405
Query Server with Generated User Digital Fingerprint 410
Receive Master Digital Fingerprint From Server 415
Replace User Digital Fingerprint With Master Digital Fingerprint 420
End

Figure 4

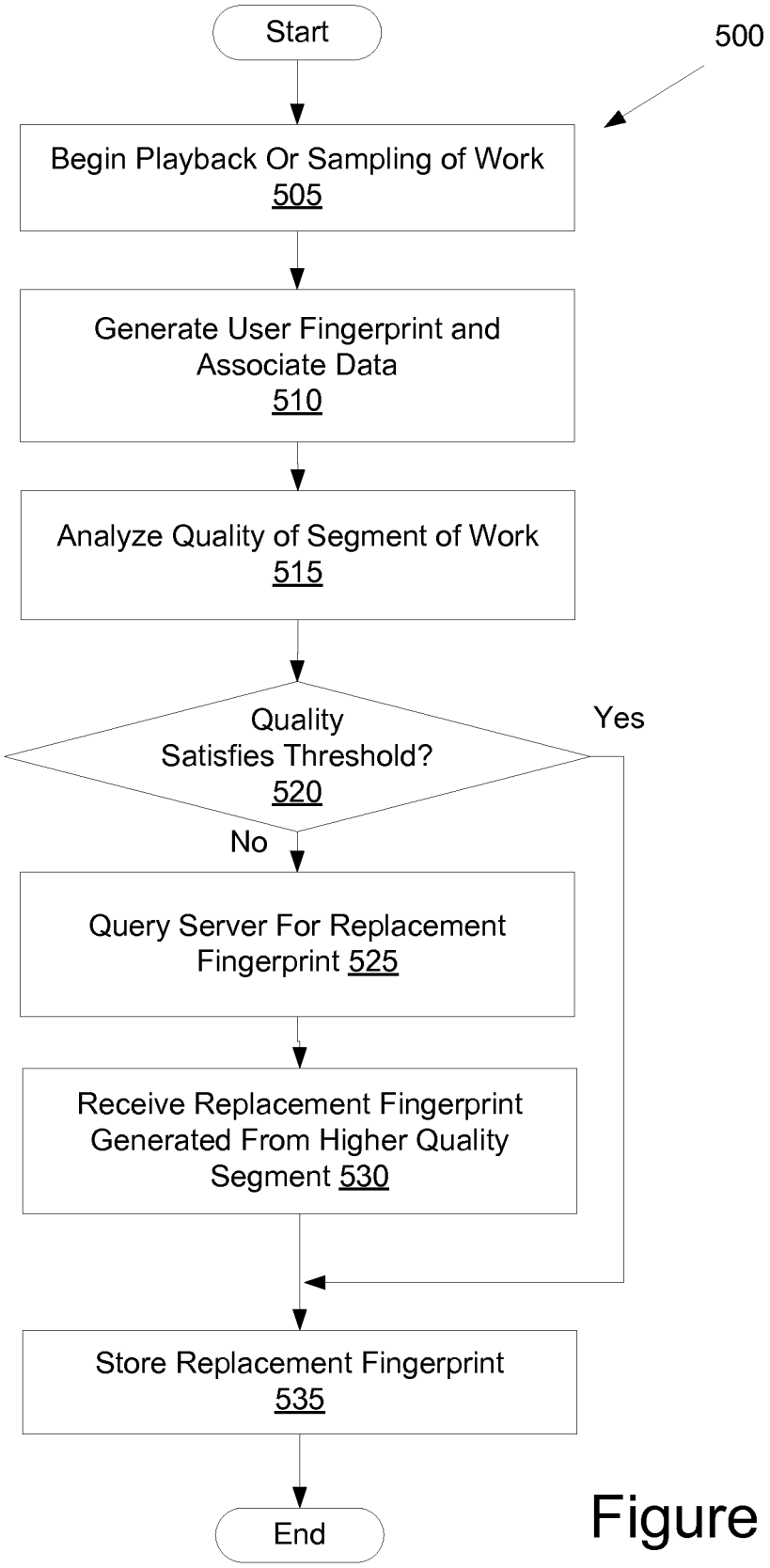
Start
500
Begin Playback Or Sampling of Work
505
Generate User Fingerprint and
Associate Data
510
Analyze Quality of Segment of Work
515
Quality
Satisfies Threshold?
520
Yes
No
Query Server For Replacement
Fingerprint 525
Receive Replacement Fingerprint
Generated From Higher Quality
Segment 530
Store Replacement Fingerprint
535
End

Figure 5

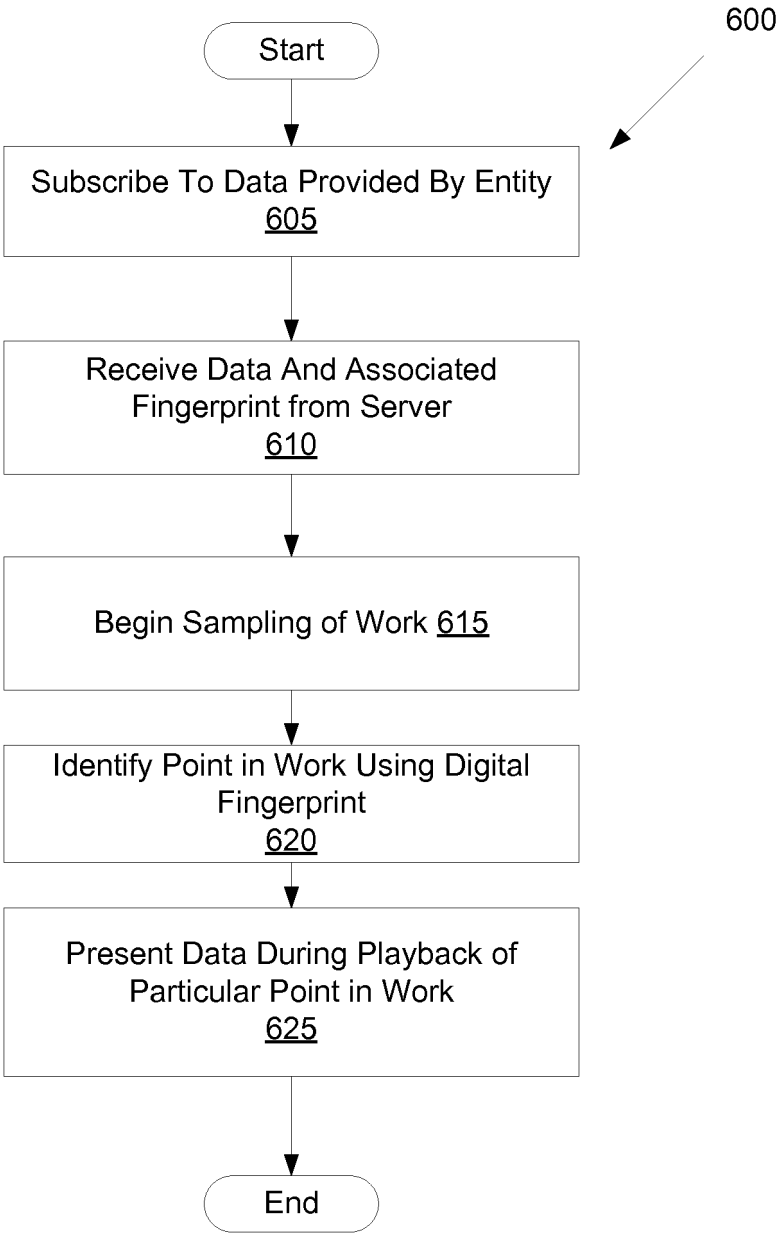
600
Start
Subscribe To Data Provided By Entity
605
Receive Data And Associated Fingerprint from Server
610
Begin Sampling of Work 615
Identify Point in Work Using Digital Fingerprint
620
Present Data During Playback of Particular Point in Work
625
End

Figure 6

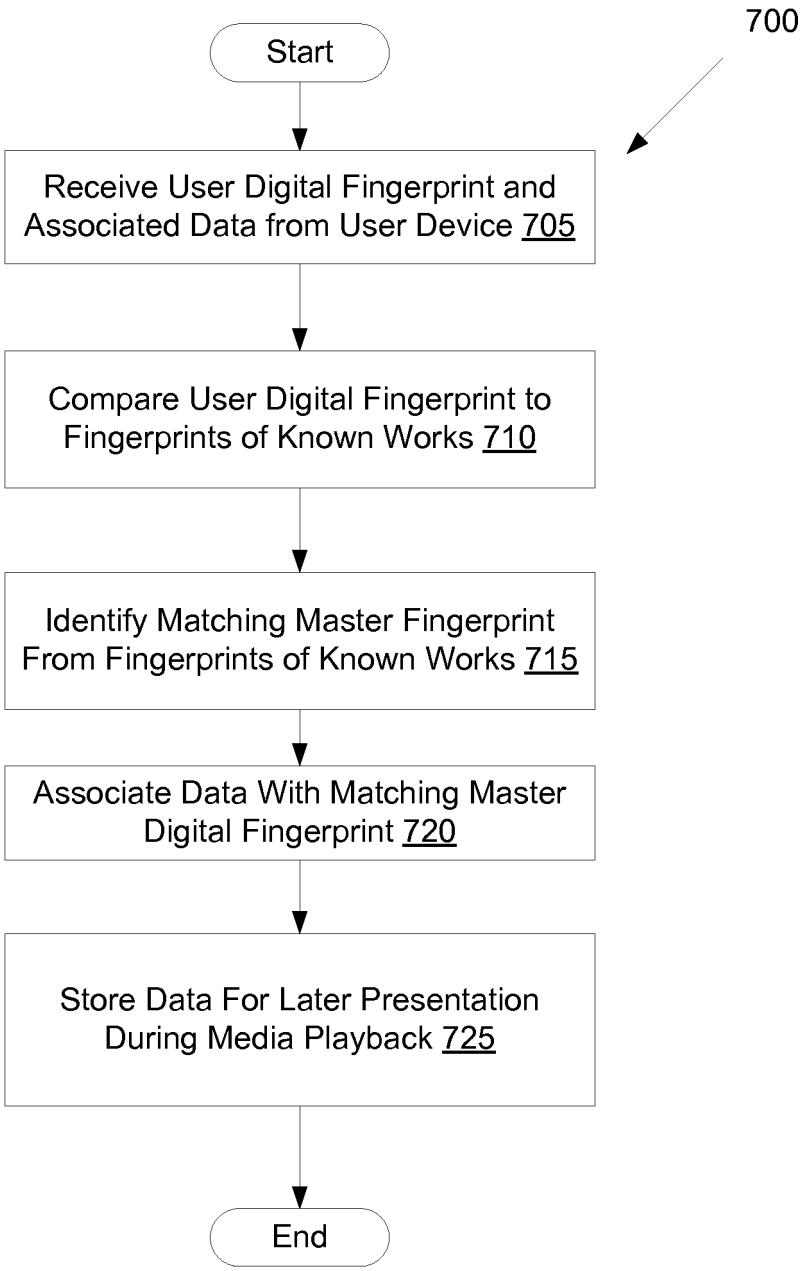
700
Start
Receive User Digital Fingerprint and Associated Data from User Device 705
Compare User Digital Fingerprint to Fingerprints of Known Works 710
Identify Matching Master Fingerprint From Fingerprints of Known Works 715
Associate Data With Matching Master Digital Fingerprint 720
Store Data For Later Presentation During Media Playback 725
End

Figure 7

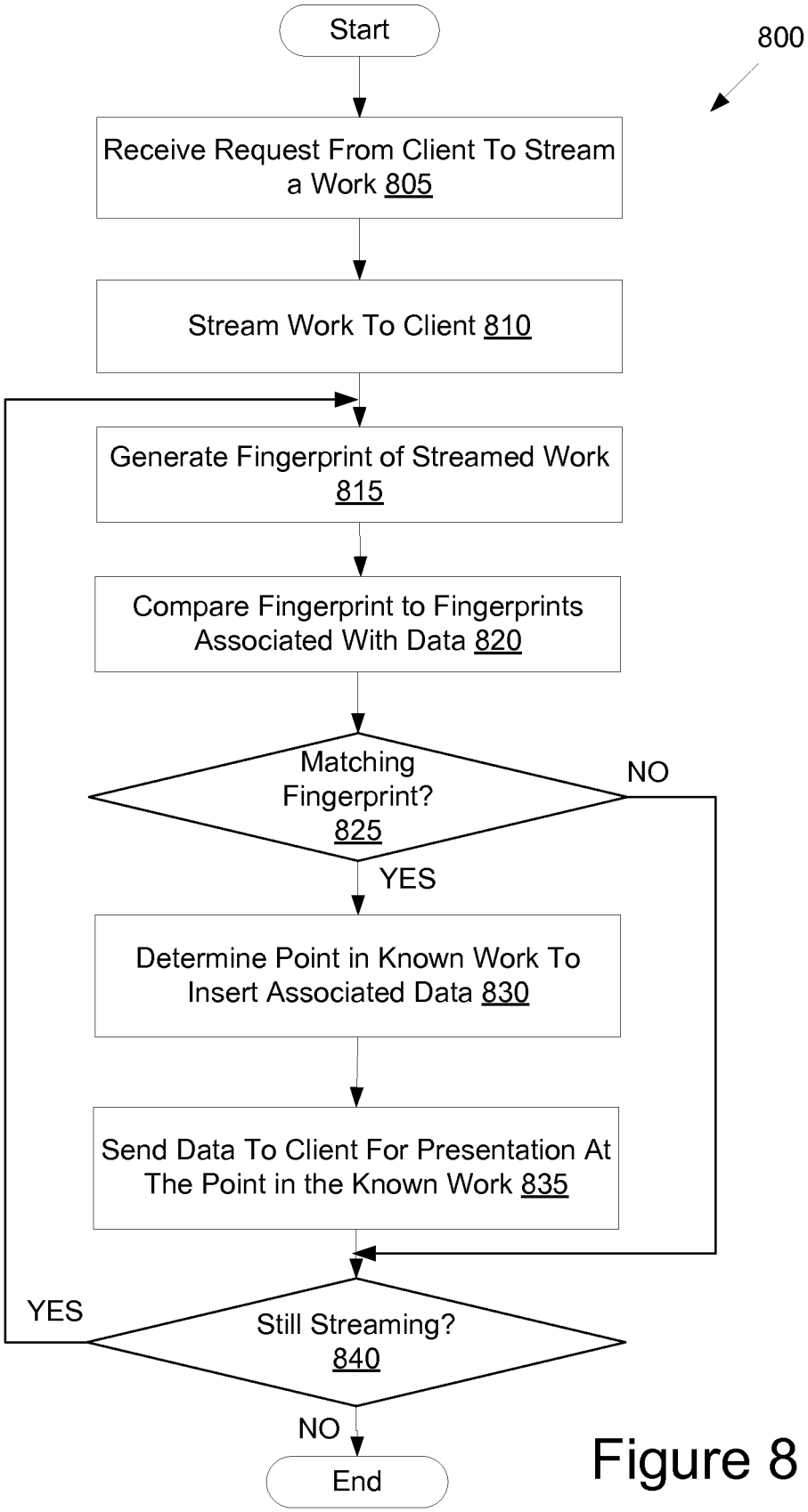
Start
800
Receive Request From Client To Stream a Work 805
Stream Work To Client 810
Generate Fingerprint of Streamed Work 815
Compare Fingerprint to Fingerprints Associated With Data 820
Matching Fingerprint? 825
NO
YES
Determine Point in Known Work To Insert Associated Data 830
Send Data To Client For Presentation At The Point in the Known Work 835
YES
Still Streaming? 840
NO
End

Figure 8

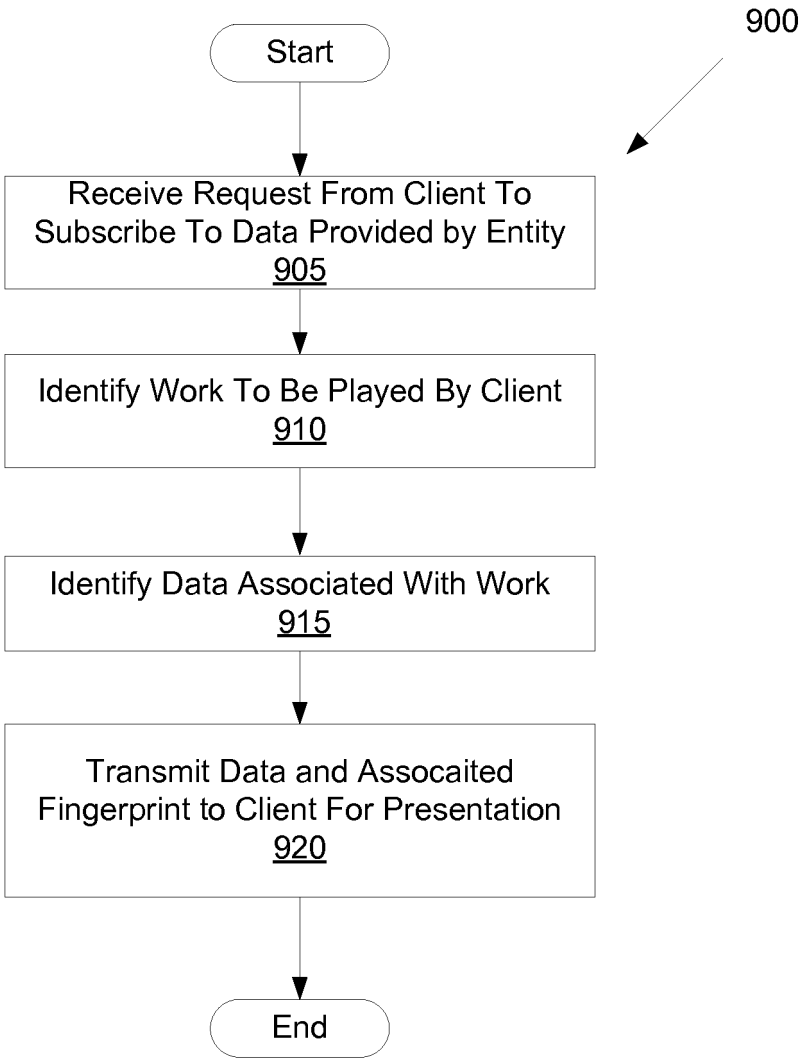
900
Start
Receive Request From Client To Subscribe To Data Provided by Entity
905
Identify Work To Be Played By Client
910
Identify Data Associated With Work
915
Transmit Data and Assocaited Fingerprint to Client For Presentation
920
End

Figure 9

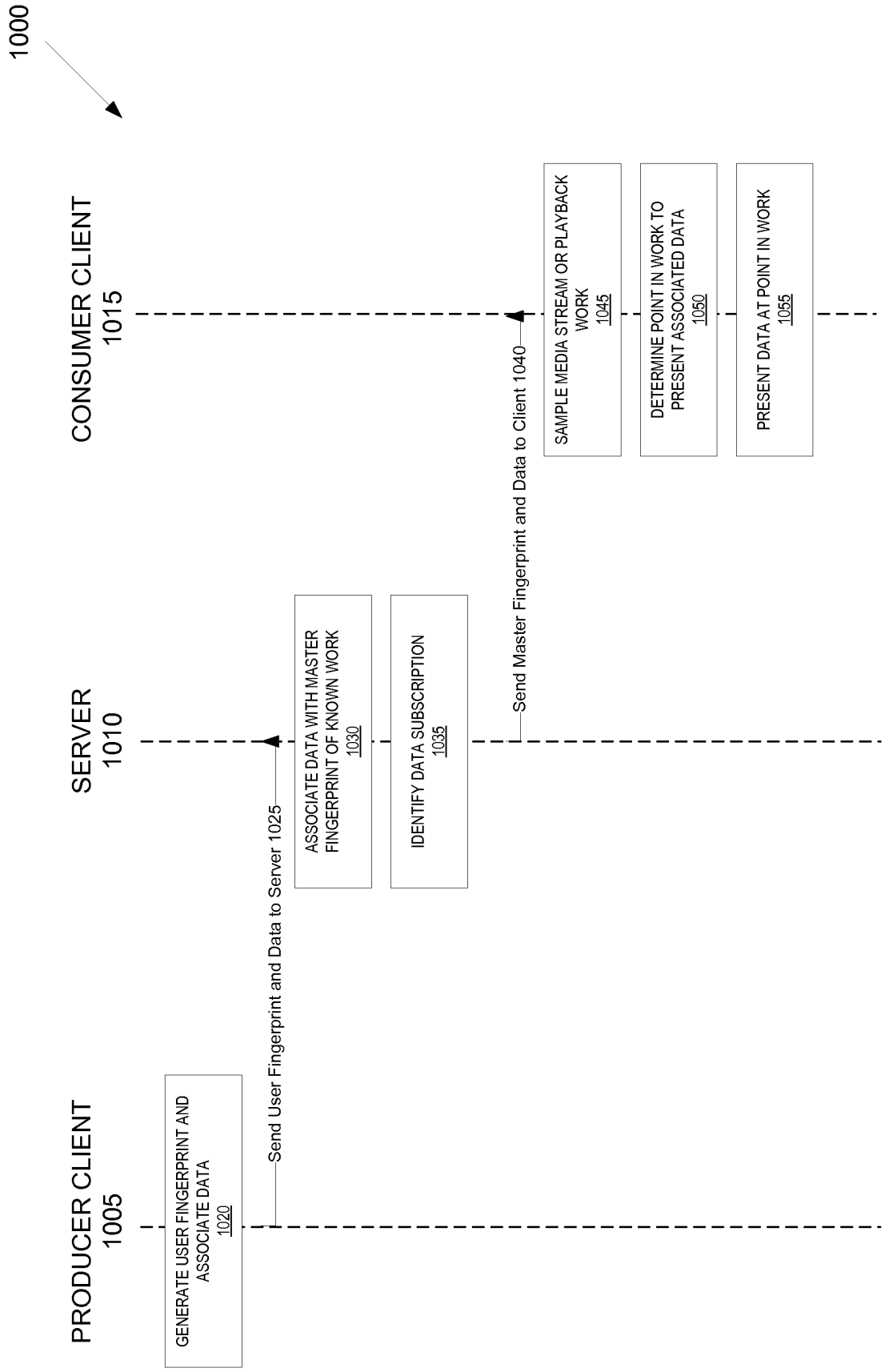
1000
PRODUCER CLIENT
1005
SERVER
1010
CONSUMER CLIENT
1015
GENERATE USER FINGERPRINT AND ASSOCIATE DATA
1020
Send User Fingerprint and Data to Server 1025
ASSOCIATE DATA WITH MASTER FINGERPRINT OF KNOWN WORK
1030
IDENTIFY DATA SUBSCRIPTION
1035
Send Master Fingerprint and Data to Client 1040
SAMPLE MEDIA STREAM OR PLAYBACK WORK
1045
DETERMINE POINT IN WORK TO PRESENT ASSOCIATED DATA
1050
PRESENT DATA AT POINT IN WORK
1055

Figure 10

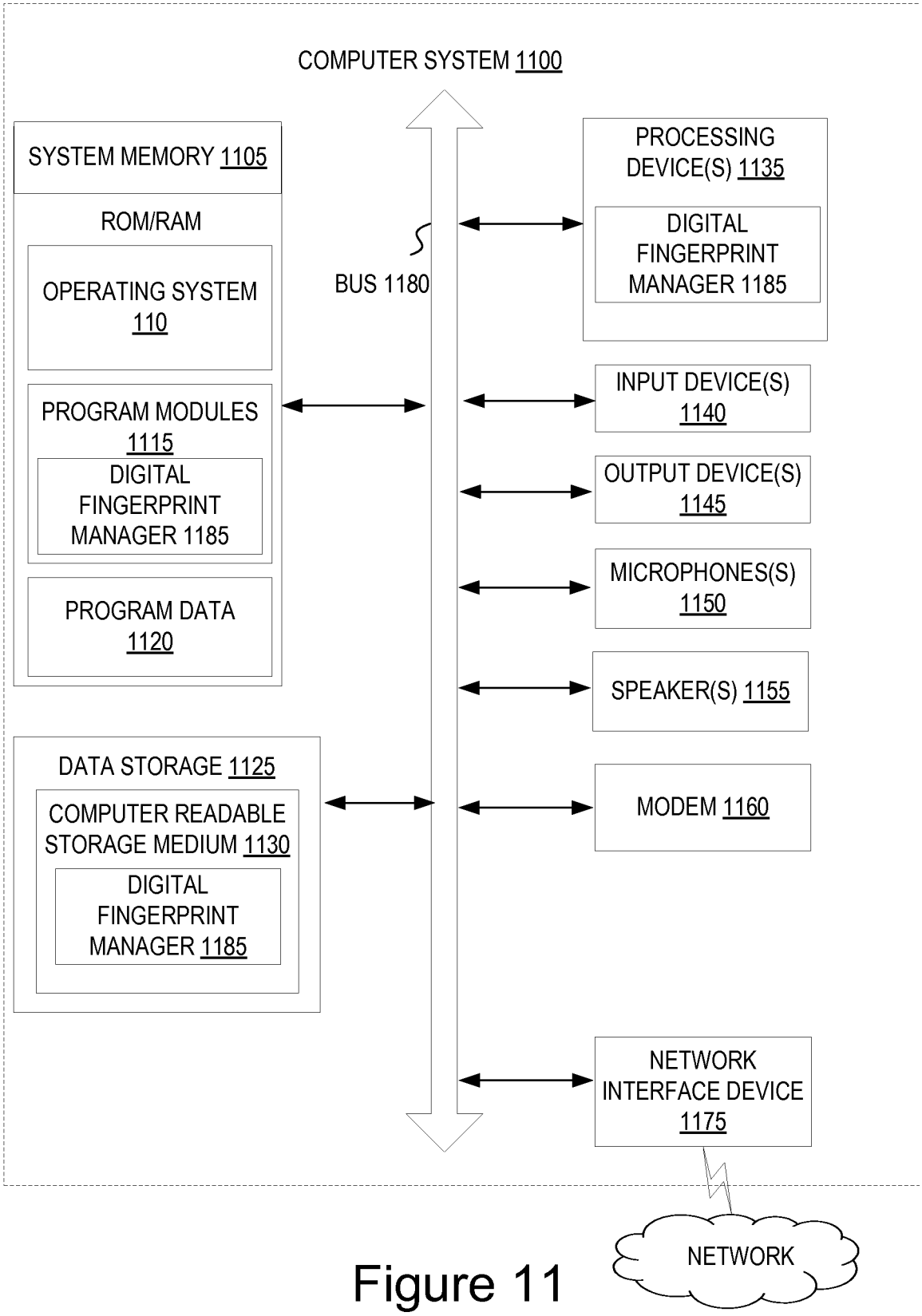
COMPUTER SYSTEM 1100
SYSTEM MEMORY 1105
ROM/RAM
OPERATING SYSTEM 110
PROGRAM MODULES 1115
DIGITAL FINGERPRINT MANAGER 1185
PROGRAM DATA 1120
BUS 1180
PROCESSING DEVICE(S) 1135
DIGITAL FINGERPRINT MANAGER 1185
INPUT DEVICE(S) 1140
OUTPUT DEVICE(S) 1145
MICROPHONES(S) 1150
SPEAKER(S) 1155
MODEM 1160
DATA STORAGE 1125
COMPUTER READABLE STORAGE MEDIUM 1130
DIGITAL FINGERPRINT MANAGER 1185
NETWORK INTERFACE DEVICE 1175
NETWORK

Figure 11

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Non-patent literature cited in the description

- **KUNIO KASHINO et al.** Robust Search methods for Music Signals Based on Simple Representation. *IEEE International Conference on Acoustics, Speech, and Signal Processing,* 23 October 2007, IV - 1421 **[0002]**