March 18, 2015

SYNCHRONIZING AUDIO AND VIDEO PLAYBACK ON MULTIPLE DEVICES

Julien Marchand

Alexander Taboriskiy

SYNCHRONIZING AUDIO AND VIDEO PLAYBACK ON MULTIPLE DEVICES

ABSTRACT

An audio and video playback synchronization system can be used to synchronize an audio and a video of a content playing on separate devices. The devices fetch the audio and/or video content independently. The playback synchronization can be implemented via two separate embodiments. In a first embodiment, the system synchronizes a video of a content playing on a first device with audio of the same content playing on a second device. The system detects, at the first device playing video of the content, the audio of the content from the second device. The system then compares the audio detected from the second device with the audio associated with the video content played at the first device. The system then calculates an audio delay between the audio detected from the second device and the audio associated with the video content played at the first device. Accordingly, the system adjusts the playback of the video content at the first device based on the calculated audio delay.

In a second embodiment, the system synchronizes a first clock signal of a first device with a second clock signal of a second device. The system then detects a message received at the second device from the first device, and the message is stamped with a playback start time of the first device. The system then compares the playback start time received from the first device with the current clock time of the second device to calculate a time difference. Accordingly, the system then adjusts the playback at the second device based on the calculated time difference.

PROBLEM STATEMENT

Users consume audiovisual content through a multitude of electronic devices. These devices range from specialized audio devices, such as surround sound speaker systems, to video screens, such as high definition televisions. In some instances, the user may use multiple devices in parallel to consume content. For example, a user may wish to watch the video of the content on his smartphone screen but listen to the audio of the same content via his surround sound speaker system. In another scenario, the user may wish to view video media content on a high definition screen and listen to its corresponding audio via over-the-ear headphones.

In such scenarios, a user may experience asynchronization of the video and audio streams for the content. This may occur due to different playback or buffering speeds of the devices simultaneously outputting the video and audio streams for the content. For example, even if the devices start their respective playbacks in synchronization with each other, they can still diverge over time due to their different playback speeds. Moreover, the devices may buffer at different speeds resulting in asynchronization. In order to overcome this problem, an audio and video playback synchronization system is disclosed that synchronizes the audio and video of a content playing on different devices.

AUDIO AND VIDEO PLAYBACK SYNCHRONIZATION SYSTEM

The systems and techniques described here relate to an audio and video synchronization system that can be used to synchronize audio and video of a content playing on separate devices. The system can be implemented for use in an Internet, an intranet, or another client and server

environment. The devices can be any electronic device, e.g., a laptop, mobile phone, desktop computer, tablet, electronic wearable device, PDA, over-the-ear headphones, surround sound speakers, home theatre, etc. The system can be locally installed on the electronic device or can be web-based. For example, it can be a functionality implemented at a cloud, a server, or a remote memory location.

FIG. 1 illustrates an example method for synchronizing video of a content being played at a first device with audio of the content being played at a second device. The first and second devices fetch the audio and/or video of the content independently. The method can be performed by a system such as an audio and video playback synchronization system.

The system detects, at the first device playing video of a content, audio of the content from a second device (Block 102). The second device may be any device capable of outputting audio, e.g., a surround sound speaker system, and the first device may be any device capable of playing and displaying video content and detecting audio, e.g., a smartphone. The system may detect the audio output by the second device via the first device's input interface, such as a microphone. For example, a surround sound speaker system plays the audio of a music video while a smartphone plays the video of the same music video while detecting the audio from the surround sound speaker system.

The devices may access the content in different manners. The first and second devices may access the content from the same or different sources. For example, the first device, i.e., the smartphone, may stream the video of the content over the Internet and the second device, i.e., the speakers, may play the audio of the content from local storage. The first and second device may

access the content at the same time or at different times. For example, the first device may access the content and cache the video stream on a local storage drive.

The system then compares the audio detected from the second device with the audio associated with the video content played at the first device (Block 104). Continuing with the above example, the user may choose to play only the video of the content on the smartphone and play the audio of the content through the speakers. The system compares the audio associated with the video playing at the first device with the audio detected from the surround sound speakers.

The system calculates an audio delay between the audio detected from the second device and the audio associated with the video content played at the first device (Block 106). To compute the audio delay, the system first determines whether the audio detected from the second device, e.g., the speaker system, has advanced or is trailing the current playback position of the first device, e.g.,such as the smartphone. If a difference in the playback positions is determined, the system calculates the difference between the audio streams, the audio delay. The system may also perform audio processing to remove background noise and external sounds from the detected audio from the second device before computing the audio delay.

Upon calculating the audio delay, the system adjusts the playback of the video content at the first device based on the calculated audio delay (Block 108). The system determines a desired video playback position at the first device based on the calculated audio delay. For example, the video can be rewinded or forwarded by an amount of time that corresponds to the calculated audio delay. The system seeks to the desired position in the video so as to cancel out the delay and synchronize the audio and video playback. Alternatively, or additionally, the system can

synchronize the video by playing the video on the first device faster or slower in order to compensate for the audio delay between the two devices.

However, if the audio for both the devices is in synchronization, i.e., the calculated audio delay is zero, the system continues to detect the audio from the second device and re-calculating the audio delay.

FIG. 2 illustrates an example implementation of a first embodiment of the audio and video synchronization system. FIG. 2 shows two electronic devices, a smartphone 204 and a speaker system 202. A user of the devices chooses to play video of a media content on the smartphone 204 and audio associated with the same media content on the speaker system 202. In this example, the smartphone 204 streams the media content over the Internet. The smartphone 204 while streaming the the media content, receives the audio as well as the associated video but only plays the video based on the user's selection. The speaker system 202 may retrieve the audio associated with the media content from a local storage device, for example, a CD ROM, digital storage device, pen drive, magnetic disc, etc., or over the Internet. The audio and video synchronization system receives an audio output 208 from the speaker system 202, via the smartphone's audio input interface, i.e., the microphone 210. The system then compares and calculates any delay between audio received from the speaker system and the audio associated with the streaming video 206. If a delay is present, the system determines a desired playback position of the video running on the smartphone and adjusts the video playback to the desired playback position.

FIG. 3 illustrates another example method for synchronizing audio of a content with video of the same content, both being played at separate devices. The method can be performed by a system such as, an audio and video playback synchronization system.

The audio and video synchronization system synchronizes a first clock signal of a first device with a second clock signal of a second device (Block 302). For example, the first device can be an electronic display device such as a liquid crystal display (LCD) and the second device can be over-the-ear headphones. The over-the-ear headphones may be playing audio associated with a content and the LCD may be playing video associated with the same content. The clock synchronization can be achieved between the two devices via various methods. One method of clock synchronization can be done by employing an External Network Time Protocol (NTP) server. A Network Time Protocol is a networking protocol for clock synchronization between computer systems over packet-switched, variable-latency data networks. In this method, the devices synchronize their clocks with a NTP server over the Internet. NTP synchronizes both the participating devices to within a few milliseconds of Coordinated Universal Time (UTC). Another method of clock synchronization can be to utilize a Local NTP server. The first and the second device synchronize their clocks with the same NTP server on a local network. In this method, one of the two devices can act as an NTP server and achieve a clock synchronization accuracy of less than 1 ms with the other device. Another method of clock synchronization can be manual synchronization. The first device for example, can periodically send its clock time to the second device. The second device can adjust its own clock in accordance with the received clock time of the first device. The method also takes into account any network delays which may arise while sending the clock signals. The network delays are measured by sending message

pings and measuring their round trip time between the devices. Depending on the capabilities of the devices, the pings can be sent either over the local network or through the cloud.

After clock synchronization is achieved between the devices, the system detects a message received at the second device from the first device, and the message is stamped with a playback start time of the first device (Block 304). Continuing with the above example, the audio and video synchronization system detects a message from the LCD device at the over- the-ear headphones. When the LCD device starts the video playback of the content, it sends a message to over-the-ear headphones stamped with the playback start time of the video content. The LCD device can also periodically send its current playback time stamped with its clock time (precise to the ms). The periodicity can be for example, every 10 seconds. As the devices are already in synchronization, network delays are not considered,.

The system then compares the playback start time received from the first device with the current clock time of the second device (Block 306) to calculate a time difference (Block 308). The system can compute this time difference locally on the second device or or at a cloud server.

The system then adjusts the playback at the second device based on the calculated time difference (Block 310). After the system calculates the time difference between the current clock signal of the second device and the playback start time sent by the first device, the system determines a desired playback position for the content. The system then directly seeks to the desired position in the audio so as to cancel out the time difference and put the audio and video in synchronization. Alternatively, or additionally, the system can synchronize the audio by playing the audio on the second device faster or slower in order to compensate for the time difference between the two devices. Alternatively, or additionally, the system can synchronize

the audio by rewinding or forwarding the audio on the second device in order to cancel out the delay.

The system may also take into account any delays that might occur due to network latency issues. The system computes the difference between the current clock of the second device and the timestamp associated with the received playback start time message received from the first device. The timestamp signifies the time when the message was sent from the first device. The system adds this difference to the playback time reported by the first device and then compares it with the clock of the second device.

FIG. 4 illustrates an example implementation of a second embodiment of the audio and video synchronization system. FIG. 4 shows two electronic devices, a high definition video screen 402 and over-the-ear headphones 404. A user of the devices selects to play video of a media content on the high definition video screen 402 and audio associated with the same media content at the over-the-ear headphones 404. The headphones are connected to a portable music player 406 through which the media content is accessed. The headphones may also be able to stream the media content over the Internet. The high definition video screen may stream the media content over the Internet or from a local storage.

Both the electronic devices have an internal playback clock. The system first synchronizes the clock of both the devices. The clock synchronization can be achieved via an external NTP server or a local NTP server in which the clocks of both the devices are synced (within tens of milliseconds). The clock of the devices can be synchronized manually as well as by exchanging messages stamped with their respective clocks periodically.

After the clocks are synchronized, the system detects a message sent by the high definition screen 402 to the over-the-ear headphones 404. This message is sent by the high definition video screen when it starts video playback. The message is stamped with the video playback start time of the high definition video screen 402.

The system then compares the received playback start time with the current clock of the over-the-ear headphones to calculate a time difference, if any. If a time difference is present, the system determines a desired playback position of the audio running on the over-the-ear headphones and adjusts the audio playback to the desired playback position.

FIG. 5 is a block diagram of an exemplary environment that shows components of a system for implementing the techniques described in this disclosure. The environment includes client devices 510, servers 530, and network 540. Network 540 connects client devices 510 to servers 530. Client device 510 is an electronic device. Client device 510 may be capable of requesting and receiving data/communications over network 540. Example client devices 510 are personal computers (e.g., laptops), mobile communication devices, (e.g. smartphones, tablet computing devices), set-top boxes, game-consoles, embedded systems, and other devices 510' that can send and receive data/communications over network 540. Client device 510 may execute an application, such as a web browser 512 or 514 or a native application 516. Web applications 513 and 515 may be displayed via a web browser 512 or 514. Server 530 may be a web server capable of sending, receiving and storing web pages 532. Web page(s) 532 may be stored on or accessible via server 530. Web page(s) 532 may be associated with web application 513 or 515 and accessed using a web browser, e.g., 512. When accessed, webpage(s) 532 may be transmitted and displayed on a client device, e.g., 510 or 510'. Resources 518 and 518' are

resources available to the client device 510 and/or applications thereon, or server(s) 530 and/or web pages(s) accessible therefrom, respectively. Resources 518' may be, for example, memory or storage resources; a text, image, video, audio, JavaScript, CSS, or other file or object; or other relevant resources. Network 540 may be any network or combination of networks that can carry data communication.

The subject matter described in this disclosure can be implemented in software and/or hardware (for example, computers, circuits, or processors). The subject matter can be implemented on a single device or across multiple devices (for example, a client device and a server device). Devices implementing the subject matter can be connected through a wired and/or wireless network. Such devices can receive inputs from a user (for example, from a mouse, keyboard, or touchscreen) and produce an output to a user (for example, through a display). Specific examples disclosed are provided for illustrative purposes and do not limit the scope of the disclosure.

FIGURES

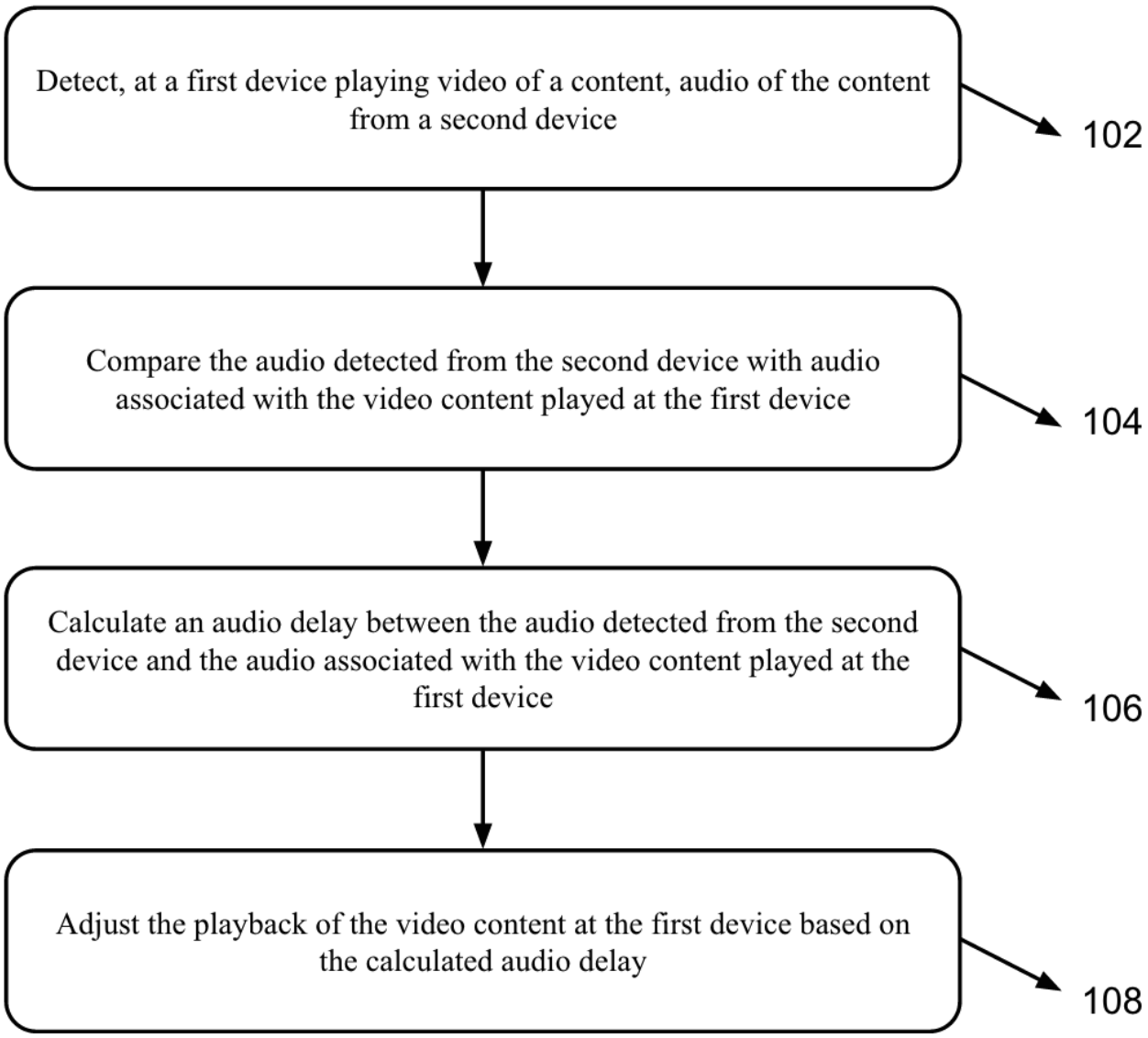

FIGURE 1

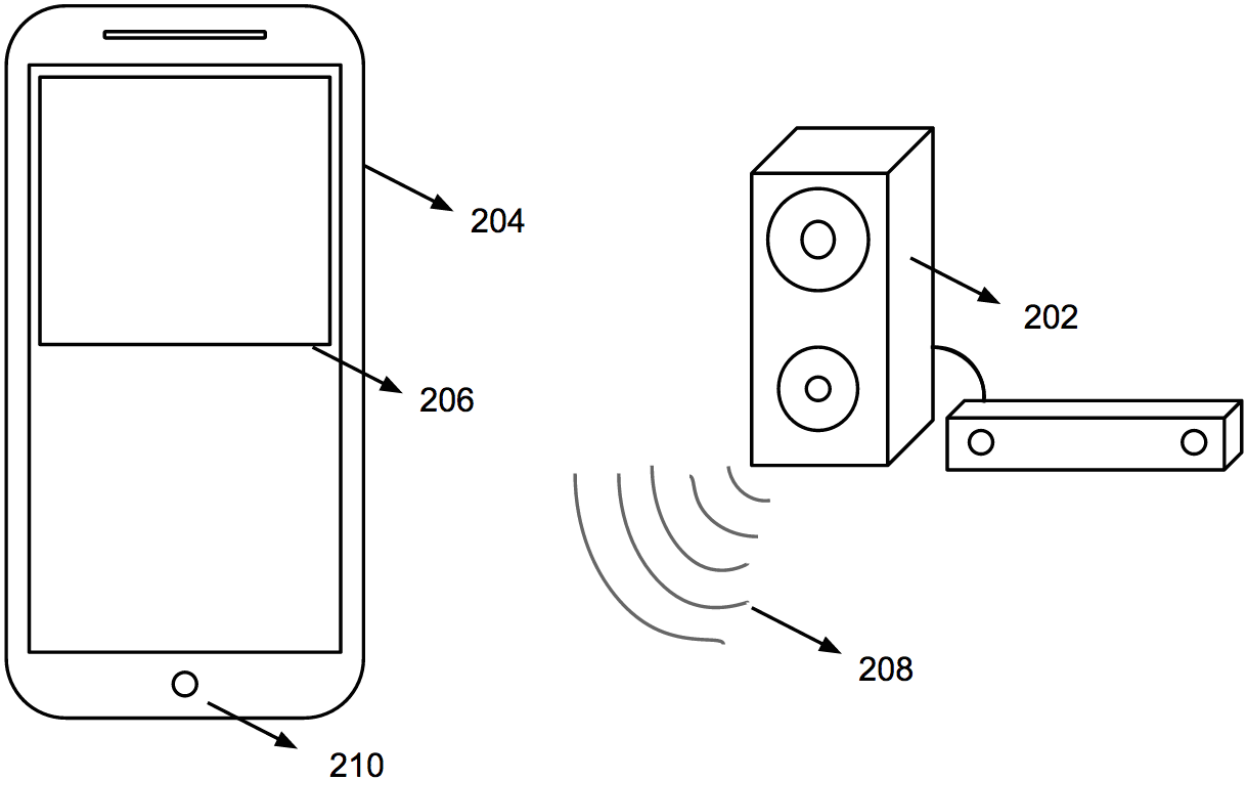

FIGURE 2

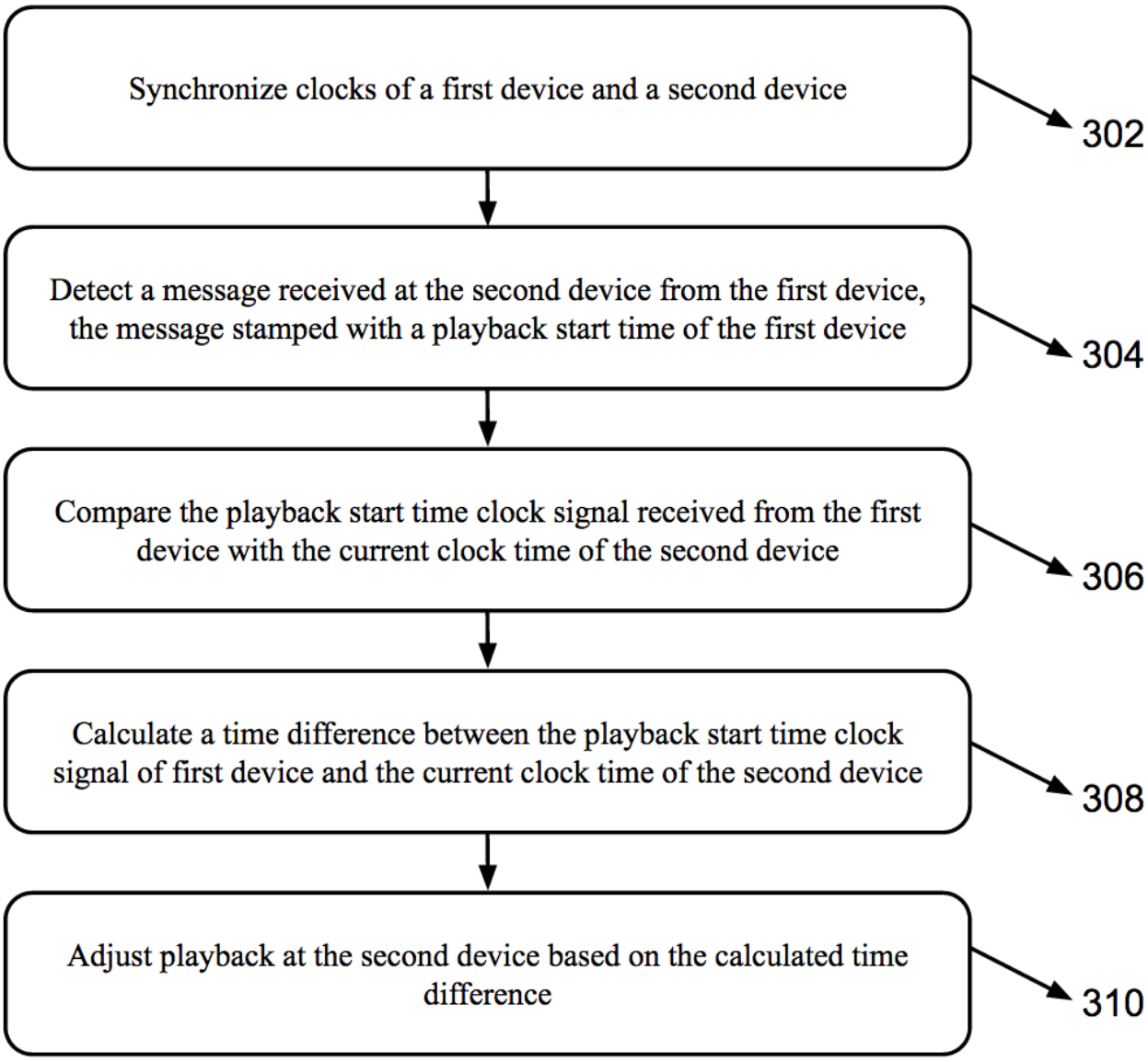
Synchronize clocks of a first device and a second device
302
Detect a message received at the second device from the first device, the message stamped with a playback start time of the first device
304
Compare the playback start time clock signal received from the first device with the current clock time of the second device
306
Calculate a time difference between the playback start time clock signal of first device and the current clock time of the second device
308
Adjust playback at the second device based on the calculated time difference
310

FIGURE 3

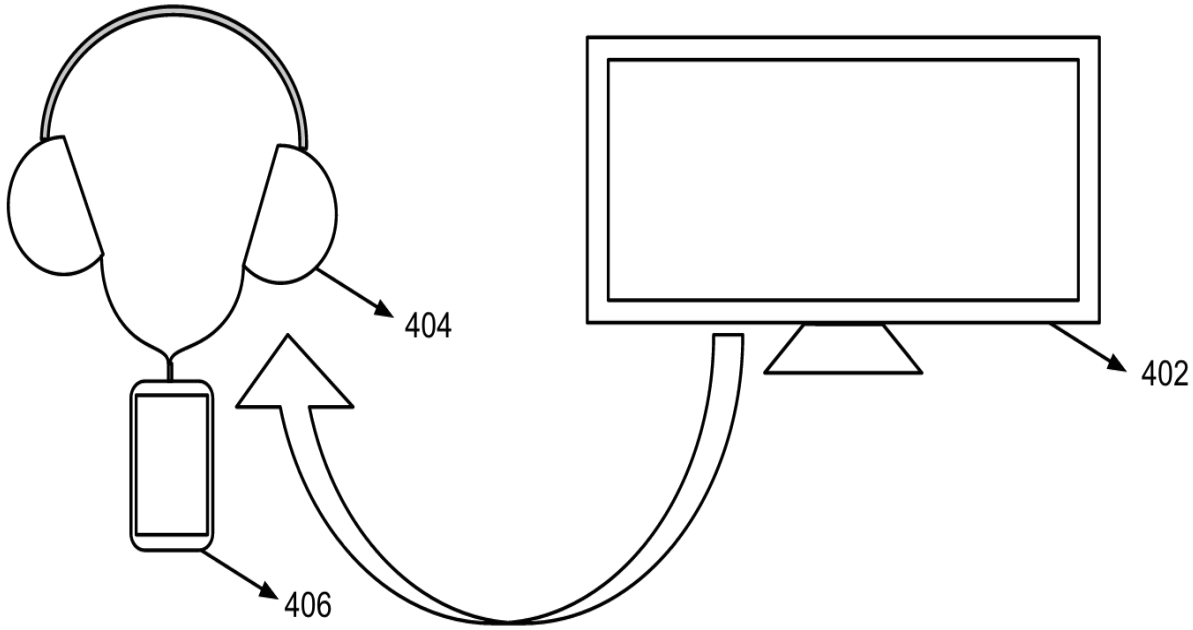

FIGURE 4

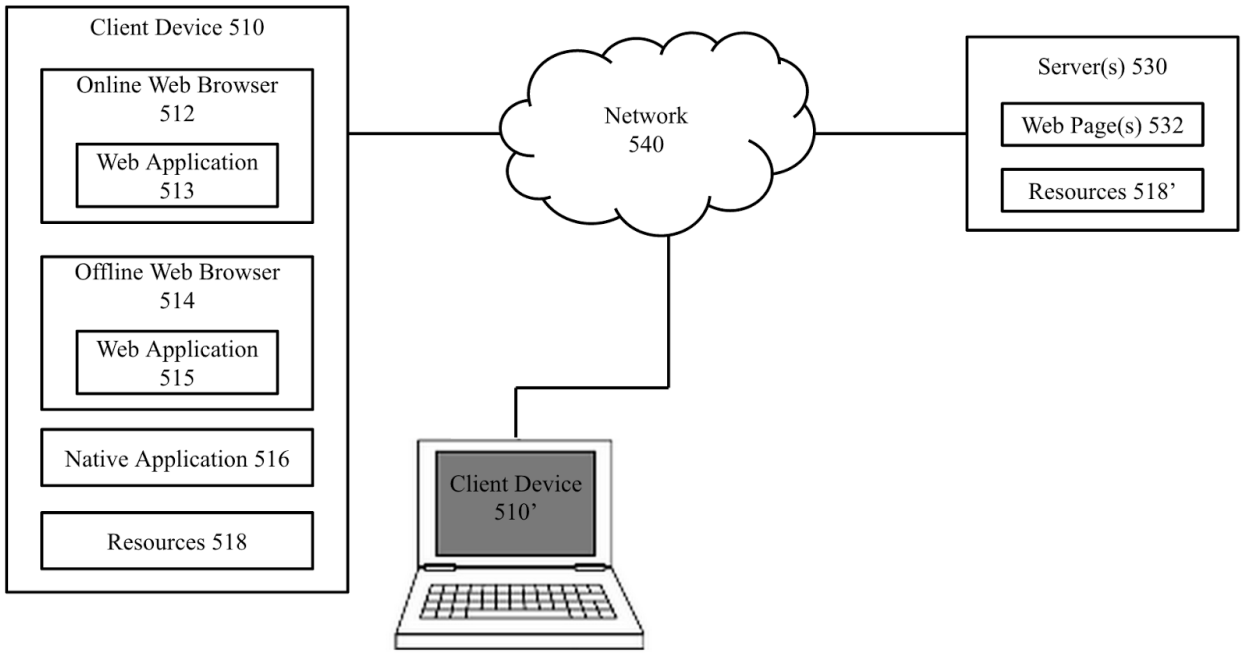

FIGURE 5