September 2022

# Artificial Reality Enhanced Learning

Anonymous

# Artificial Reality Enhanced Learning


## ABSTRACT

This disclosure describes techniques for performing a medical operation in an artificial reality environment. The described techniques include providing, via a user device configured to render the artificial reality environment, first artificial reality content that simulates performance of the medical operation. Anatomical features of a subject being treated by the medical operation are labeled. An audio signal corresponding to a video frame of the first artificial reality content is determined and the video frame is annotated based on the audio signal. The techniques further include reproducing the audio signal based on determining second artificial reality content corresponding to the performance of the medical operation. A portion of the first artificial reality content matching a portion of the second artificial reality content based on the anatomical features is provided in the artificial reality environment.



## KEYWORDS

- Artificial reality (AR)
- Mixed reality (MR)
- Virtual reality (VR)
- Extra reality (XR)
- Medical operation
- Surgery
- Video annotation
- Anatomical feature
- Anatomical component
- Object annotation
- Object detection
- AR overlay
- Pace learning
- Head-mounted display (HMD)

**DESCRIPTION**

The present disclosure generally relates to artificial reality enhanced pace learning. For example, the disclosure to performing a medical operation in an artificial reality environment. The present disclosure provides techniques, including systems (e.g., computing devices such as servers or clients), methods (e.g., implemented via computer software and/or in computer hardware), and computer-readable media for artificial reality enhanced pace learning. In an aspect, the present disclosure relates to techniques for performing a medical operation in an artificial reality environment. The techniques further include providing, via a user device configured to render the artificial reality environment, first artificial reality content that simulates performance of the medical operation. The techniques further include labeling anatomical features of a subject being treated by the medical operation. The techniques further include determining an audio signal corresponding to a video frame of the first artificial reality content. The techniques further include annotating the video frame based on the audio signal. The techniques further include reproducing, based on determining second artificial reality content corresponding to the performance of the medical operation, the audio signal. The techniques further include providing, in the artificial reality environment, a portion of the first artificial reality content matching a portion of the second artificial reality content based on the anatomical features.

The disclosure includes several figures in which the most significant digit or digits in a reference number refers to the figure number in which that element is first introduced. In various implementations, not all of the depicted components in each figure may be required, and one or more implementations may include additional components not shown in a figure. Variations in the arrangement and type of the components may be made without departing from the scope of the subject disclosure. Additional components, different components, or fewer components may be utilized within the scope of the subject disclosure.

In the following detailed description, numerous specific details are set forth to provide a full understanding of the present disclosure. It will be apparent, however, to one ordinarily skilled in the art that the embodiments of the present disclosure may be practiced without some of these specific details. In other instances, well-known structures and techniques have not been shown in detail so as not to obscure the disclosure.

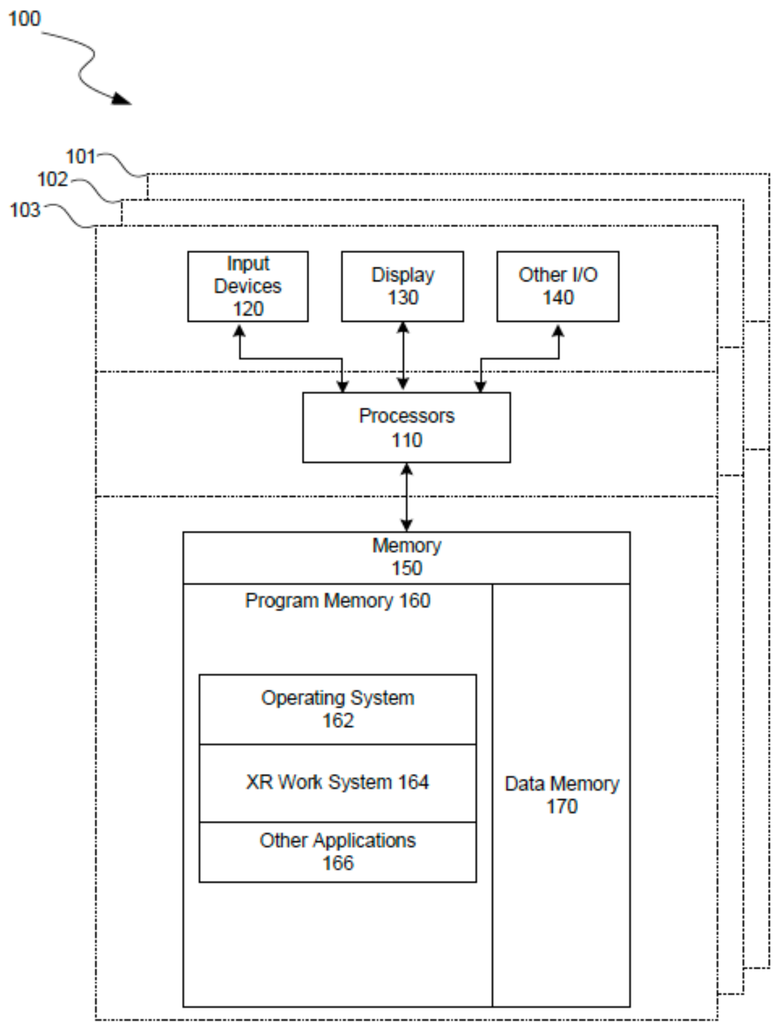


**FIG. 1: Device overview**

FIG. 1 is a block diagram illustrating an overview of devices on which some implementations of the disclosed technology can operate. The devices can comprise hardware components of a computing system 100 that can create, administer, and provide interaction modes for an artificial reality collaborative working environment. In various implementations, computing system 100 can include a single computing device 103 or multiple computing devices (e.g., computing device 101, computing device 102, and computing device 103) that communicate over wired or wireless channels to distribute processing and share input data.

In some implementations, computing system 100 can include a stand-alone headset capable of providing a computer created or augmented experience for a user without the need for external processing or sensors. In other implementations, computing system 100 can include multiple computing devices such as a headset and a core processing component (such as a console, mobile device, or server system) where some processing operations are performed on the headset and others are offloaded to the core processing component. Example headsets are described below in relation to FIGS. 2A and 2B. In some implementations, position and environment data can be gathered only by sensors incorporated in the headset device, while in other implementations one or more of the non-headset computing devices can include sensor components that can track environment or position data.

Computing system 100 can include one or more processor(s) 110 (e.g., central processing units (CPUs), graphical processing units (GPUs), holographic processing units (HPUs), etc.) Processors 110 can be a single processing unit or multiple processing units in a device or distributed across multiple devices (e.g., distributed across two or more of computing devices 101-103).

Computing system 100 can include one or more input devices 120 that provide input to the processors 110, notifying them of actions. The actions can be mediated by a hardware controller that interprets the signals received from the input device and communicates the information to the processors 110 using a communication protocol. Each input device 120 can include, for example, a mouse, a keyboard, a touchscreen, a touchpad, a wearable input device (e.g., a haptics glove, a bracelet, a ring, an earring, a necklace, a watch, etc.), a camera (or other light-based input device, e.g., an infrared sensor), a microphone, or other user input devices.

Processors 110 can be coupled to other hardware devices, for example, with the use of an internal or external bus, such as a PCI bus, SCSI bus, or wireless connection. The processors 110 can communicate with a hardware controller for devices, such as for a display 130. Display 130 can be used to display text and graphics. In some implementations, display 130 includes the input device as part of the display, such as when the input device is a touchscreen or is equipped with an eye direction monitoring system. In some implementations, the display is separate from the input device. Examples of display devices are: an LCD display screen, an LED display screen, a projected, holographic, or augmented reality display (such as a heads-up display device or a head-mounted device), and so on. Other I/O devices 140 can also be coupled to the processor, such as a network chip or card, video chip or card, audio chip or card, USB, firewire or other external device, camera, printer, speakers, CD-ROM drive, DVD drive, disk drive, etc.

Computing system 100 can include a communication device capable of communicating wirelessly or wire-based with other local computing devices or a network node. The communication device can communicate with another device or a server through a network using, for example, TCP/IP protocols. Computing system 100 can utilize the communication device to distribute operations across multiple network devices.

The processors 110 can have access to a memory 150, which can be contained on one of the computing devices of computing system 100 or can be distributed across of the multiple computing devices of computing system 100 or other external devices. A memory includes one or more hardware devices for volatile or non-volatile storage and can include both read-only and writable memory. For example, a memory can include one or more of random-access memory (RAM), various caches, CPU registers, read-only memory (ROM), and writable non-volatile memory, such as flash memory, hard drives, floppy disks, CDs, DVDs, magnetic storage devices, tape drives, and so forth. A memory is not a propagating signal divorced from underlying hardware; a memory is thus non-transitory. Memory 150 can include program memory 160 that stores programs and software, such as an operating system 162, XR work system 164, and other application programs 166. Memory 150 can also include data memory 170 that can include information to be provided to the program memory 160 or any element of the computing system 100.

Some implementations can be operational with numerous other computing system environments or configurations. Examples of computing systems, environments, and/or configurations that may be suitable for use with the technology include, but are not limited to, XR headsets, personal computers, server computers, handheld or laptop devices, cellular telephones, wearable electronics, gaming consoles, tablet devices, multiprocessor systems, microprocessor-based systems, set-top boxes, programmable consumer electronics, network PCs, minicomputers, mainframe computers, distributed computing environments that include any of the above systems or devices, or the like.

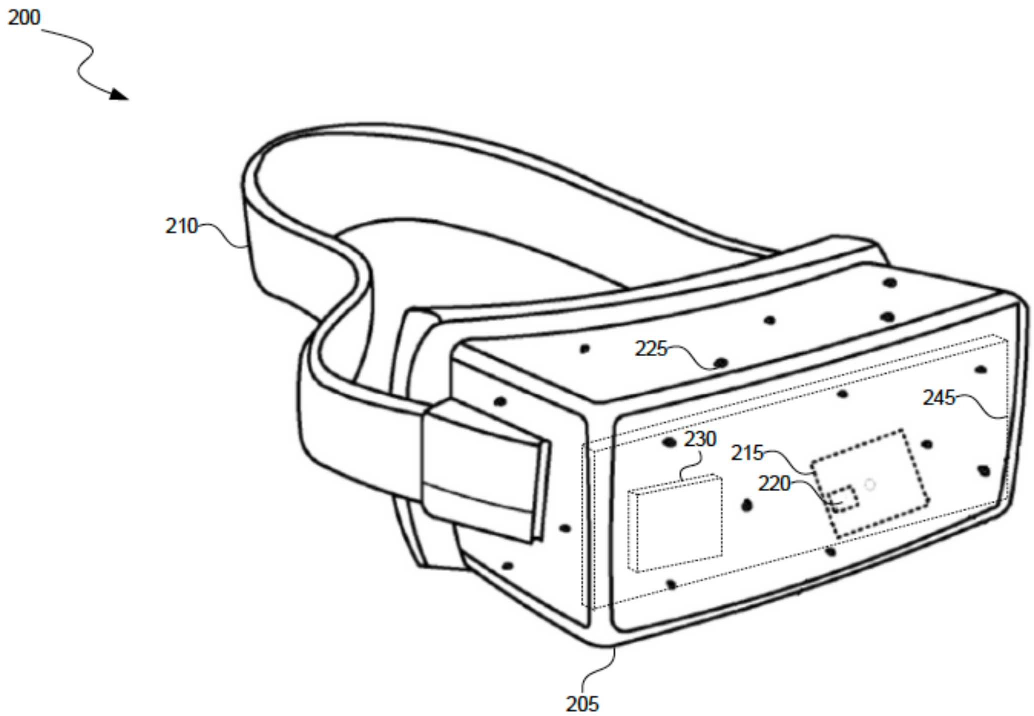


**FIG. 2A: Wire diagram of a virtual reality head-mounted display (HMD)**

FIG. 2A is a wire diagram of an example virtual reality head-mounted display (HMD) 200. The HMD 200 includes a front rigid body 205 and a band 210. The front rigid body 205 includes one or more electronic display elements of an electronic display 245, an inertial motion unit (IMU) 215, one or more position sensors 220, locators 225, and one or more compute units 230. The position sensors 220, the IMU 215, and compute units 230 may be internal to the HMD 200 and may not be visible to the user.

The IMU 215, position sensors 220, and locators 225 can track movement and location of the HMD 200 in the real-world and in a virtual environment in three degrees of freedom (3DoF) or six degrees of freedom (6DoF). For example, the locators 225 can emit infrared light beams which create light points on real objects around the HMD 200. As another example, the IMU 215 can include e.g., one or more accelerometers, gyroscopes, magnetometers, other non-camera-based position, force, or orientation sensors, or combinations thereof. One or more cameras (not shown) integrated with the HMD 200 can detect the light points. Compute units 230 in the HMD 200 can use the detected light points to extrapolate position and movement of the HMD 200 as well as to identify the shape and position of the real objects surrounding the HMD 200.

The electronic display 245 can be integrated with the front rigid body 205 and can provide image light to a user as dictated by the compute units 230. In various embodiments, the electronic display 245 can be a single electronic display or multiple electronic displays (e.g., a display for each user eye). Examples of the electronic display 245 include: a liquid crystal display (LCD), an organic light-emitting diode (OLED) display, an active-matrix organic light-emitting diode display (AMOLED), a display including one or more quantum dot light-emitting diode (QOLED) sub-pixels, a projector unit (e.g., microLED, LASER, etc.), some other display, or some combination thereof.

In some implementations, the HMD 200 can be coupled to a core processing component such as a personal computer (PC) (not shown) and/or one or more external sensors (not shown). The external sensors can monitor the HMD 200 (e.g., via light emitted from the HMD 200) which the PC can use, in combination with output from the IMU 215 and position sensors 220, to determine the location and movement of the HMD 200.

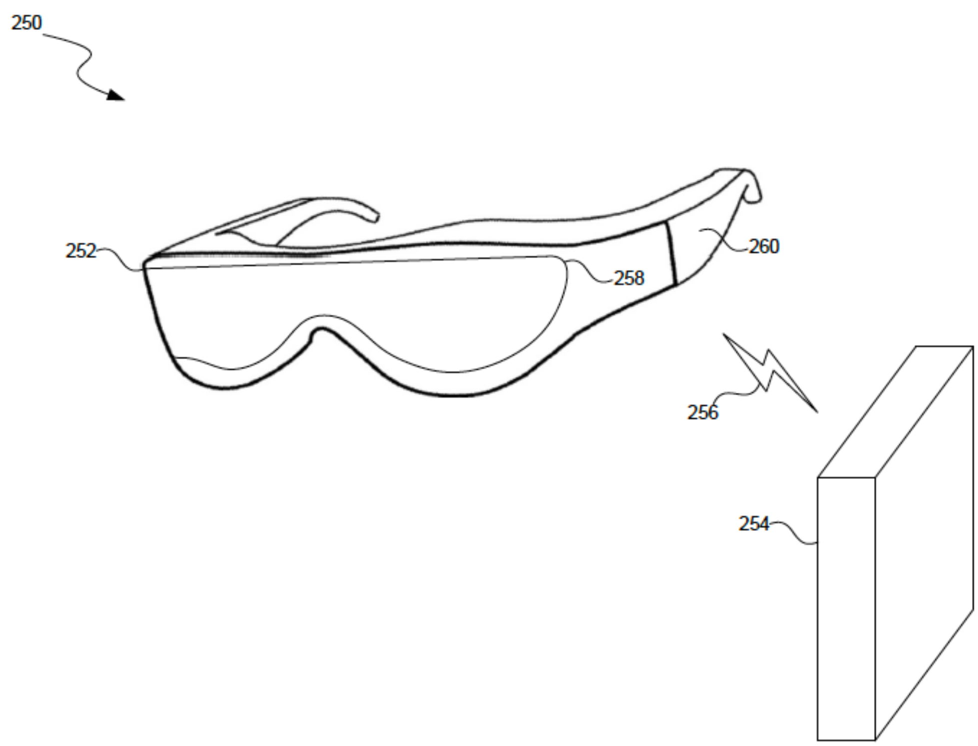


**FIG. 2B: Wire diagram of a mixed reality HMD system which includes a mixed reality HMD and a core processing component**

FIG. 2B is a wire diagram of a mixed reality HMD system 250 which includes a mixed reality HMD 252 and a core processing component 254. The mixed reality HMD 252 and the core processing component 254 can communicate via a wireless connection (e.g., a 60 GHz link) as indicated by link 256. In some implementations, the mixed reality system 250 includes a headset only, without an external compute device or includes other wired or wireless connections between the mixed reality HMD 252 and the core processing component 254. The mixed reality HMD 252 includes a pass-through display 258 and a frame 260. The frame 260 can house various electronic components (not shown) such as light projectors (e.g., LASERs, LEDs, etc.), cameras, eye-tracking sensors, MEMS components, networking components, etc.

The projectors can be coupled to the pass-through display 258, e.g., via optical elements, to display media to a user. The optical elements can include one or more waveguide assemblies, reflectors, lenses, mirrors, collimators, gratings, etc., for directing light from the projectors to a user's eye. Image data can be transmitted from the core processing component 254 via link 256 to HMD 252. Controllers in the HMD 252 can convert the image data into light pulses from the projectors, which can be transmitted via the optical elements as output light to the user's eye. The output light can mix with light that passes through the display 258, allowing the output light to present virtual objects that appear as if they exist in the real-world.

Similar to the HMD 200, the HMD system 250 can also include motion and position tracking units, cameras, light sources, etc., which allow the HMD system 250 to, e.g., track itself in 3DoF or 6DoF, track portions of the user (e.g., hands, feet, head, or other body parts), map virtual objects to appear as stationary as the HMD 252 moves, and have virtual objects react to gestures and other real-world objects.

Conventionally, education or further training is often promised to be the next big breakthrough in AR applications. Examples include a system that allows a student to learn at its pace and in an AR-enhanced fashion. To illustrate, an example is used of a surgeon teaching its students how to perform a given surgical procedure; however, other applications are contemplated and within the scope of this disclosure.

The teacher may perform the surgical procedure while wearing AR glasses. While the teacher speaks, for example, saying “and here, we move the spleen on the bit before inserting the camera there,” a highly specialized object detection model (in this case recognizing organs, bones, etc.) may annotate the video frames. For example, the audio stream may be matched with the recognized objects. As an example, the audio stream may be transformed to text with a speech to text and the salient terms of that transcript are matched against part of the video frames.

Once it is the student’s turn, they may be also wearing AR glasses and as they start performing the procedure, the audio of the teacher can be replayed. Instead of playing like in a regular audio player, an example smarter player may wait until the student finishes something. For example, if it took the teacher 3 seconds to move the spleen and insert the camera but 30 seconds for the student, the audio will not move forward until the student is done. The student can ask the system to re-show (with an overlay of the teacher's hands for instance) what exactly needs to be done. For example, the current step at which the student is may be matched with the corresponding one in the recorded teacher video; and with the object annotations (e.g., where the spleen is and where the camera was inserted with respect to that recognized object), show a reconstruction of the hands of the teacher moving on what the student sees.

The disclosed system(s) address a problem in traditional content modification techniques tied to computer technology, namely, the technical problem of synchronizing AR teaching content with actions of a person being trained using the AR teaching content, and/or other technical problems. The disclosed system(s) solve this technical problem by providing a solution also rooted in computer technology, namely, by providing for artificial reality enhanced pace learning. The disclosed subject technology further provides improvements to the functioning of the computer itself because it improves processing and efficiency in artificial reality enhanced pace learning.

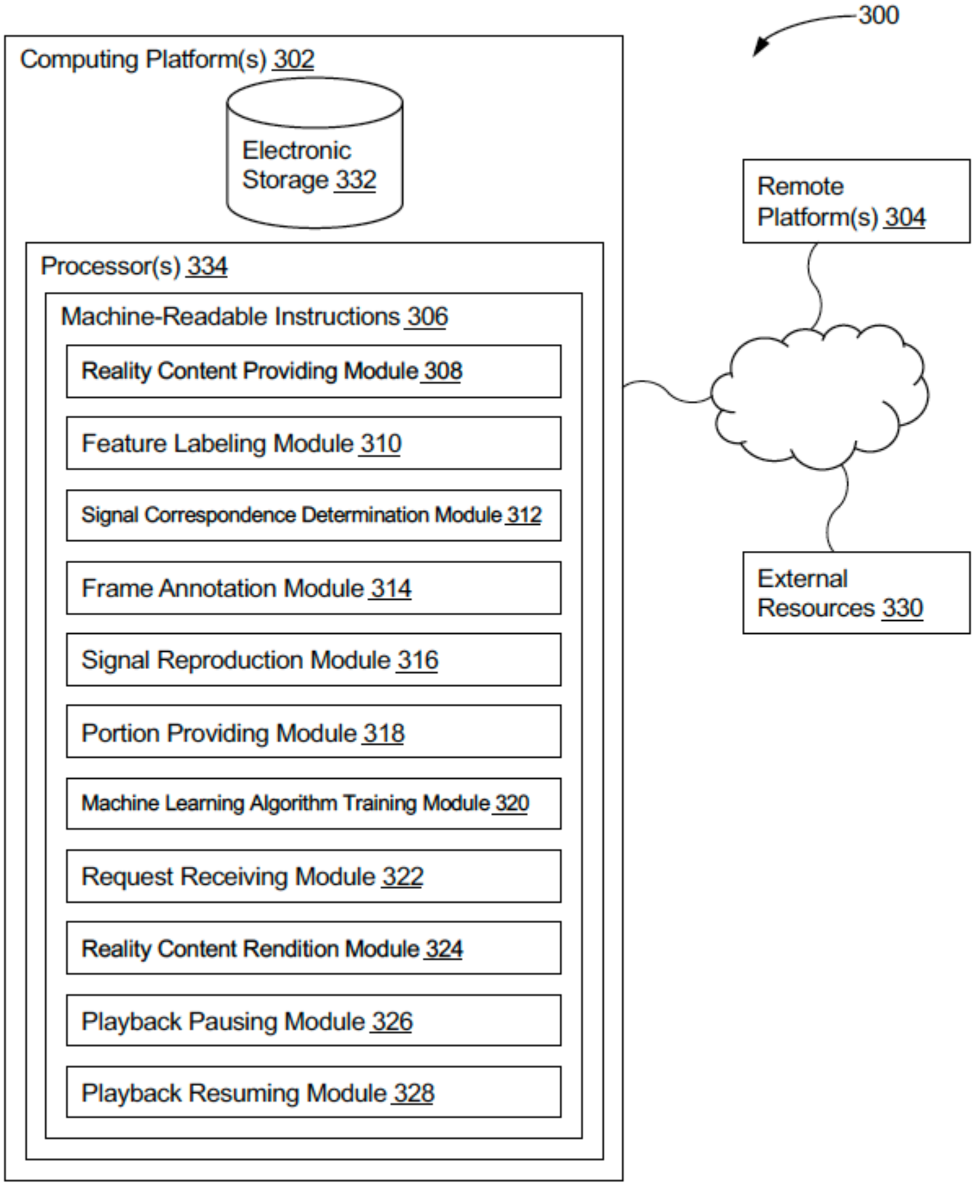


**FIG. 3: System configured for performing a medical operation in an artificial reality environment**

FIG. 3 illustrates an example system 300 configured for performing a medical operation in an artificial reality environment. System 300 includes one or more computing platforms 302. Computing platform(s) 302 are configured to communicate with one or more remote platforms 304 according to a client/server architecture, a peer-to-peer architecture, and/or other architectures. Remote platform(s) 304 are configured to communicate with other remote platforms via computing platform(s) 302 and/or according to a client/server architecture, a peer-to-peer architecture, and/or other architectures. Users can access system 300 via remote platform(s) 304.

Computing platform(s) 302 are configured by machine-readable instructions 306. Machine-readable instructions 306 include one or more instruction modules that include computer program modules. For example, the instruction modules can include one or more of reality content providing module 308, feature labeling module 310, signal correspondence determination module 312, frame annotation module 314, signal reproduction module 316, portion providing module 318, machine learning algorithm training module 320, request receiving module 322, reality content rendition module 324, playback pausing module 326, playback resuming module 328, and/or other instruction modules.

Reality content providing module 308 is configured to provide, via a user device configured to render the artificial reality environment, first artificial reality content that simulates performance of the medical operation. For example, the medical operation may include a surgical operation and the user device includes at least one of artificial reality glasses, artificial reality headset, virtual reality glasses, virtual reality headset, artificial reality compatible device, or virtual reality compatible device.

Reality content providing module 308 is configured to provide the second artificial reality content while playing back the audio signal.

Feature labeling module 310 is configured to label anatomical features of a subject being treated by the medical operation. By way of non-limiting example, labeling the anatomical features of the subject may include determining, via an object detection model, at least one of an organ, a bone, an artery, a vessel, a tissue, or an anatomical component of the subject.

Signal correspondence determination module 312 is configured to determine an audio signal corresponding to a video frame of the first artificial reality content. Annotating the video frame can include converting the audio signal to a textual transcript. Annotating the video frame can include annotating, based on matching portions of the textual transcript to portions of the video frame, the video frame with corresponding textual terms of the textual transcript.

Frame annotation module 314 is configured to annotate the video frame based on the audio signal.

Signal reproduction module 316 is configured to reproduce, based on determining second artificial reality content corresponding to the performance of the medical operation, the audio signal. Reproducing the audio signal can include playing back the audio signal corresponding to a teacher simulating performance of the medical operation while rendering the second artificial reality content corresponding to a student simulating performance of the medical operation in the artificial reality environment.

Portion providing module 318 is configured to provide, in the artificial reality environment, a portion of the first artificial reality content matching a portion of the second artificial reality content based on the anatomical features.

Machine learning algorithm training module 320 is configured to train, according to a machine learning algorithm.

Request receiving module 322 is configured to receive a request from the user device to show simulated performance of at least a portion of the medical procedure.

Reality content rendition module 324 is configured to render at least a portion of the first artificial reality content based on matching a portion of first artificial reality content to a step of the medical operation associated with the second artificial reality content.

Playback pausing module 326 is configured to pause playback of the audio signal based on the second artificial reality content being indicative of simulated performance of the medical operation by a student exceeding a temporal threshold.

Playback resuming module 328 is configured to resume playback of the audio signal based on the second artificial reality content being indicative of simulated completed performance of a portion of the medical operation by the student. Providing the portion of the first artificial reality content can include providing a reconstruction of hands of a teacher simulating performance of the medical operation in the artificial reality environment. Providing the reconstruction can include providing object annotations based on the anatomical features. Providing the portion of the first artificial reality content can include rendering an overlay of a reconstruction of simulated performance of the medical procedure overlaid onto the portion of the second artificial reality content.

In some examples, determining the audio signal includes determining audio from a teacher that is indicative of how to perform a surgical step of the medical operation.

Computing platform(s) 302, remote platform(s) 304, and/or external resources 330 may be operatively linked via one or more electronic communication links. For example, such electronic communication links may be established, at least in part, via a network such as the Internet and/or other networks or via some other communication media.

A given remote platform 304 can include one or more processors configured to execute computer program modules. The computer program modules re configured to enable an expert or user associated with the given remote platform 304 to interface with system 300 and/or external resources 330, and/or provide other functionality attributed herein to remote platform(s) 304. A given remote platform 304 and/or a given computing platform 302 can include one or more of a server, a desktop computer, a laptop computer, a handheld computer, a tablet computing platform, a smartphone, a gaming console, and/or other computing platforms.

External resources 330 can include sources of information outside of system 300, external entities participating with system 300, and/or other resources. Alternatively, some or all of the functionality attributed herein to external resources 330 may be provided by resources included in system 300.

Computing platform(s) 302 include electronic storage 332, one or more processors 334, and/or other components. Computing platform(s) 302 can also include communication lines, or ports to enable the exchange of information with a network and/or other computing platforms. Computing platform(s) 302 include a plurality of hardware, software, and/or firmware components operating together to provide the functionality attributed herein to computing platform(s) 302. For example, computing platform(s) 302 can be implemented by a cloud of computing platforms operating together as computing platform(s) 302.

Electronic storage 332 can comprise non-transitory storage media that electronically stores information. The electronic storage media of electronic storage 332 may include one or both of system storage that is provided integrally (i.e., substantially non-removable) with computing platform(s) 302 and/or removable storage that is removably connectable to computing platform(s) 302 via, for example, a port (e.g., a USB port, a firewire port, etc.) or a drive (e.g., a disk drive, etc.). Electronic storage 332 may include one or more of optically readable storage media (e.g., optical disks, etc.), magnetically readable storage media (e.g., magnetic tape, magnetic hard drive, floppy drive, etc.), electrical charge-based storage media (e.g., EEPROM, RAM, etc.), solid-state storage media (e.g., flash drive, etc.), and/or other electronically readable storage media. Electronic storage 332 may include one or more virtual storage resources (e.g., cloud storage, a virtual private network, and/or other virtual storage resources). Electronic storage 332 may store software algorithms, information determined by processor(s) 334, information received from computing platform(s) 302, information received from remote platform(s) 304, and/or other information that enables computing platform(s) 302 to function as described herein.

Processor(s) 334 are configured to provide information processing capabilities in computing platform(s) 302. As such, processor(s) 334 may include one or more of a digital processor, an analog processor, a digital circuit designed to process information, an analog circuit designed to process information, a state machine, and/or other mechanisms for electronically processing information. In various examples, processor(s) 334 may include a single processor or a plurality of processing units. The processing units may be physically located within the same device, or processor(s) 334 may represent processing functionality of a plurality of devices operating in coordination.

Processor(s) 334 are configured to execute modules 308, 310, 312, 314, 316, 318, 320, 322, 324, 326, and/or 328, and/or other modules. Processor(s) 334 may be configured to execute modules 308, 310, 312, 314, 316, 318, 320, 322, 324, 326, and/or 328, and/or other modules by software; hardware; firmware; some combination of software, hardware, and/or firmware; and/or other mechanisms for configuring processing capabilities on processor(s) 334. As used herein, the term "module" refers to any component or set of components that perform the functionality attributed to the module. This can include one or more physical processors during execution of processor readable instructions, the processor readable instructions, circuitry, hardware, storage media, or any other components.

While modules 308, 310, 312, 314, 316, 318, 320, 322, 324, 326, and/or 328 are illustrated in FIG. 3 as being implemented within a single processing unit, in examples in which processor(s) 334 includes multiple processing units, one or more of modules 308, 310, 312, 314, 316, 318, 320, 322, 324, 326, and/or 328 can be implemented remotely from the other modules. The description of the functionality provided by the different modules 308, 310, 312, 314, 316, 318, 320, 322, 324, 326, and/or 328 described below is for illustrative purposes, and is not intended to be limiting, as any of modules 308, 310, 312, 314, 316, 318, 320, 322, 324, 326, and/or 328 may provide more or less functionality than is described.

The techniques described herein can be implemented as method(s) that are performed by physical computing device(s); as one or more non-transitory computer-readable storage media storing instructions which, when executed by computing device(s), cause performance of the method(s); or, as physical computing device(s) that are specially configured with a combination of hardware and software that causes performance of the method(s).

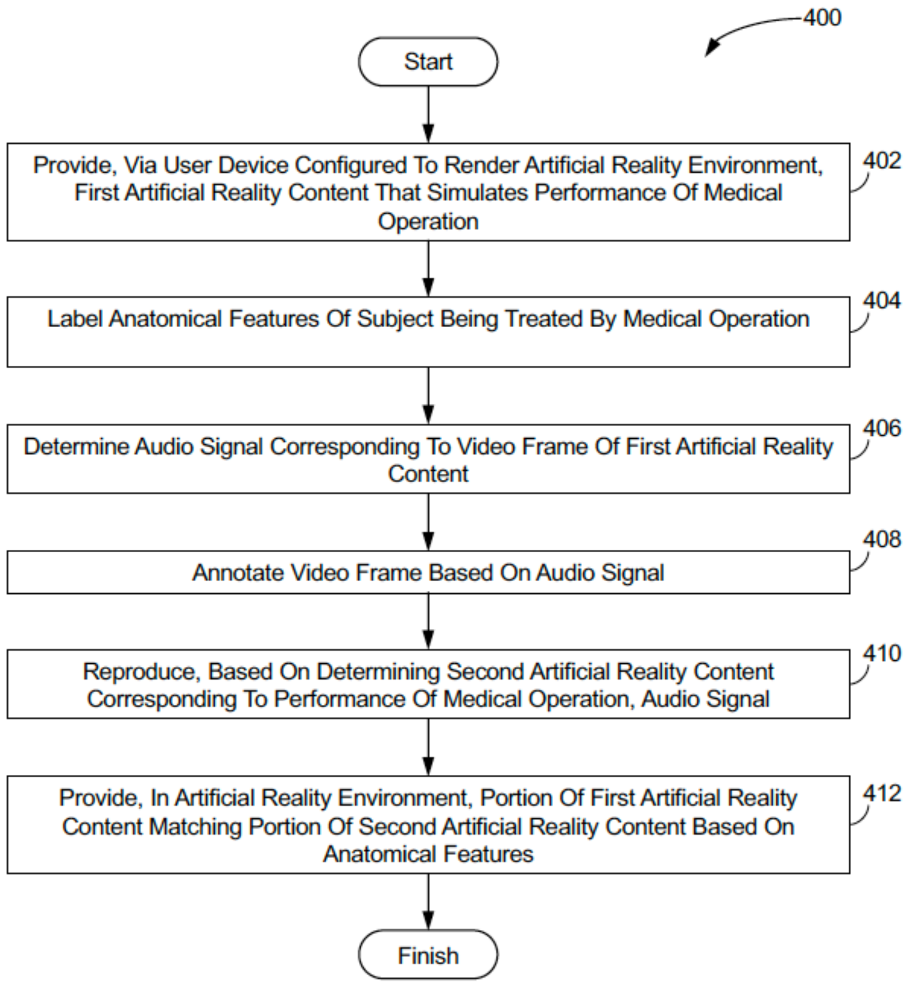


**FIG. 4: An example flow diagram for performing a medical operation in an artificial reality environment**

FIG. 4 an example flow diagram (e.g., process 400) for performing a medical operation in an artificial reality environment as an example of artificial reality enhanced pace learning, according to certain aspects of the disclosure. For explanatory purposes, the example process 400 is described herein with reference to FIG. 3. Further for explanatory purposes, the steps of the example process 400 are described herein as occurring in serial, or linearly. However, multiple instances of the example process 400 may occur in parallel. For purposes of explanation of the subject technology, the process 400 is discussed in reference to FIG. 3.

At step 402, the process 400 may include providing, via a user device configured to render the artificial reality environment, first artificial reality content that simulates performance of the medical operation. At step 404, the process 400 may include labeling anatomical features of a subject being treated by the medical operation. At step 406, the process 400 may include determining an audio signal corresponding to a video frame of the first artificial reality content. At step 408, the process 400 may include annotating the video frame based on the audio signal. At step 410, the process 400 may include reproducing, based on determining second artificial reality content corresponding to the performance of the medical operation, the audio signal. At step 412, the process 400 may include providing, in the artificial reality environment, a portion of the first artificial reality content matching a portion of the second artificial reality content based on the anatomical features. Step 412 may be performed by one or more hardware processors configured by machine-readable instructions including a module that is the same as or similar to portion providing module 318, according to certain aspects of the disclosure.

For example, as described above in relation to FIG. 3, at step 402, the process 400 may include providing, via a user device configured to render the artificial reality environment, first artificial reality content that simulates performance of the medical operation, through reality content providing module 308. At step 404, the process 400 may include labeling anatomical features of a subject being treated by the medical operation, through feature labeling module 310. At step 406, the process 400 may include determining an audio signal corresponding to a video frame of the first artificial reality content, through signal correspondence determination module 312. At step 408, the process 400 may include annotating the video frame based on the audio signal, through frame annotation module 314. At step 410, the process 400 may include reproducing, based on determining second artificial reality content corresponding to the performance of the medical operation, the audio signal, through signal reproduction module 316. At step 412, the process 400 may include providing, in the artificial reality environment, a portion of the first artificial reality content matching a portion of the second artificial reality content based on the anatomical features, through portion providing module 318.

According to an aspect, labeling the anatomical features of the subject may include determining, via an object detection model, at least one of: an organ, a bone, an artery, a vessel, a tissue, or an anatomical component of the subject.

According to an aspect, the process 400 may include training, according to a machine learning algorithm, the object detection model.

According to an aspect, determining the audio signal may include determining audio from a teacher that is indicative of how to perform a surgical step of the medical operation.

According to an aspect, annotating the video frame may include converting the audio signal to a textual transcript. According to an aspect, annotating the video frame comprises annotating, based on matching portions of the textual transcript to portions of the video frame, the video frame with corresponding textual terms of the textual transcript.

According to an aspect, reproducing the audio signal may include playing back the audio signal corresponding to a teacher simulating performance of the medical operation while rendering the second artificial reality content corresponding to a student simulating performance of the medical operation in the artificial reality environment.

According to an aspect, providing the portion of the first artificial reality content may include providing a reconstruction of hands of a teacher simulating performance of the medical operation in the artificial reality environment.

According to an aspect, providing the reconstruction may include providing object annotations based on the anatomical features.

According to an aspect, providing the portion of the first artificial reality content may include rendering an overlay of a reconstruction of simulated performance of the medical procedure overlaid onto the portion of the second artificial reality content.

According to an aspect, the process 400 may include receiving a request from the user device to show simulated performance of at least a portion of the medical procedure. According to an aspect, the process 400 may include rendering at least a portion of the first artificial reality content based on matching a portion of first artificial reality content to a step of the medical operation associated with the second artificial reality content.

According to an aspect, the process 400 may include providing the second artificial reality content while playing back the audio signal. According to an aspect, the process 400 may include pausing playback of the audio signal based on the second artificial reality content being indicative of simulated performance of the medical operation by a student exceeding a temporal threshold. According to an aspect, the process 400 may include resuming playback of the audio signal based on the second artificial reality content being indicative of simulated completed performance of a portion of the medical operation by the student.

According to an aspect, the medical operation may include a surgical operation and the user device comprises at least one of: artificial reality glasses, artificial reality headset, virtual reality glasses, virtual reality headset, artificial reality compatible device, or virtual reality compatible device.

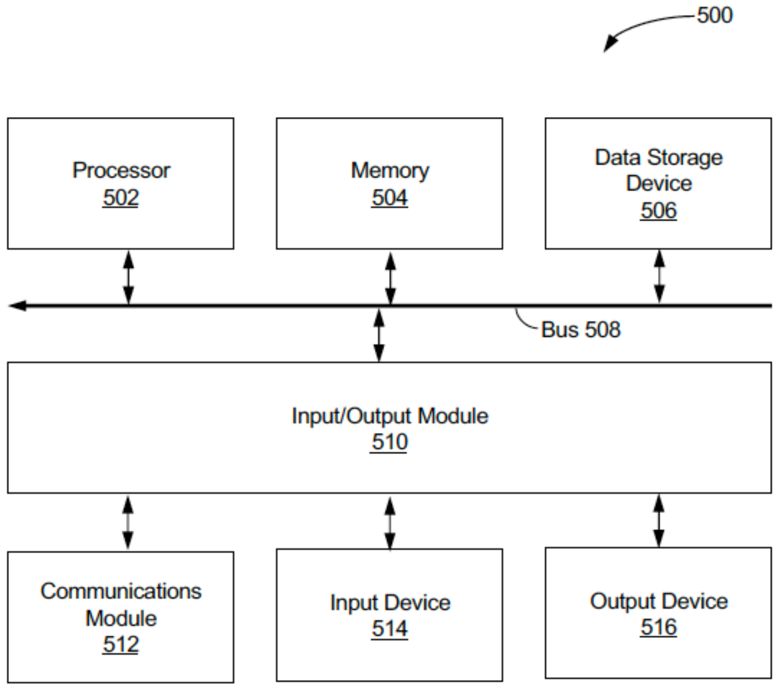


**FIG. 5: Example computer system (e.g., representing both client and server)**

FIG. 5 is a block diagram illustrating an exemplary computer system 500 with which aspects of the subject technology can be implemented. In certain aspects, the computer system 500 may be implemented using hardware or a combination of software and hardware, either in a dedicated server, integrated into another entity, or distributed across multiple entities.

Computer system 500 (e.g., server and/or client) includes a bus 508 or other communication mechanism for communicating information, and a processor 502 coupled with bus 508 for processing information. By way of example, the computer system 500 may be implemented with one or more processors 502. Processor 502 may be a general-purpose microprocessor, a microcontroller, a Digital Signal Processor (DSP), an Application Specific Integrated Circuit (ASIC), a Field Programmable Gate Array (FPGA), a Programmable Logic Device (PLD), a controller, a state machine, gated logic, discrete hardware components, or any other suitable entity that can perform calculations or other manipulations of information.

Computer system 500 can include, in addition to hardware, code that creates an execution environment for the computer program in question, e.g., code that constitutes processor firmware, a protocol stack, a database management system, an operating system, or a combination of one or more of them stored in an included memory 504, such as a Random Access Memory (RAM), a flash memory, a Read Only Memory (ROM), a Programmable Read-Only Memory (PROM), an Erasable PROM (EPROM), registers, a hard disk, a removable disk, a CD-ROM, a DVD, or any other suitable storage device, coupled to bus 508 for storing information and instructions to be executed by processor 502. The processor 502 and the memory 504 can be supplemented by, or incorporated in, special purpose logic circuitry.

The instructions may be stored in the memory 504 and implemented in one or more computer program products, i.e., one or more modules of computer program instructions encoded on a computer readable medium for execution by, or to control the operation of, the computer system 500, and according to any method well-known to those of skill in the art, including, but not limited to, computer languages such as data-oriented languages (e.g., SQL, dBase), system languages (e.g., C, Objective-C, C++, Assembly), architectural languages (e.g., Java, .NET), and application languages (e.g., PHP, Ruby, Perl, Python). Instructions may also be implemented in computer languages such as array languages, aspect-oriented languages, assembly languages, authoring languages, command line interface languages, compiled languages, concurrent languages, curly-bracket languages, dataflow languages, data-structured languages, declarative languages, esoteric languages, extension languages, fourth-generation languages, functional languages, interactive mode languages, interpreted languages, iterative languages, list-based languages, little languages, logic-based languages, machine languages, macro languages, metaprogramming languages, multiparadigm languages, numerical analysis, non-English-based languages, object-oriented class-based languages, object-oriented prototype-based languages, off-side rule languages, procedural languages, reflective languages, rule-based languages, scripting languages, stack-based languages, synchronous languages, syntax handling languages, visual languages, Wirth languages, and xml-based languages. Memory 504 may also be used for storing temporary variable or other intermediate information during execution of instructions to be executed by processor 502.

A computer program as discussed herein does not necessarily correspond to a file in a file system. A program can be stored in a portion of a file that holds other programs or data (e.g., one or more scripts stored in a markup language document), in a single file dedicated to the program in question, or in multiple coordinated files (e.g., files that store one or more modules, subprograms, or portions of code). A computer program can be deployed to be executed on one computer or on multiple computers that are located at one site or distributed across multiple sites and interconnected by a communication network. The processes and logic flows described in this specification can be performed by one or more programmable processors executing one or more computer programs to perform functions by operating on input data and generating output.

Computer system 500 further includes a data storage device 506 such as a magnetic disk or optical disk, coupled to bus 508 for storing information and instructions. Computer system 500 may be coupled via input/output module 510 to various devices. The input/output module 510 can be any input/output module. Exemplary input/output modules 510 include data ports such as USB ports. The input/output module 510 is configured to connect to a communications module 512. Exemplary communications modules 512 include networking interface cards, such as Ethernet cards and modems. In certain aspects, the input/output module 510 is configured to connect to a plurality of devices, such as an input device 514 and/or an output device 516. Exemplary input devices 514 include a keyboard and a pointing device, e.g., a mouse or a trackball, by which a user can provide input to the computer system 500. Other kinds of input devices 514 can be used to provide for interaction with a user as well, such as a tactile input device, visual input device, audio input device, or brain-computer interface device. For example, feedback provided to the user can be any form of sensory feedback, e.g., visual feedback, auditory feedback, or tactile feedback, and input from the user can be received in any form, including acoustic, speech, tactile, or brain wave input. Exemplary output devices 516 include display devices such as an LCD (liquid crystal display) monitor, for displaying information to the user.

According to one aspect of the present disclosure, the above-described gaming systems can be implemented using a computer system 500 in response to processor 502 executing one or more sequences of one or more instructions contained in memory 504. Such instructions may be read into memory 504 from another machine-readable medium, such as data storage device 506. Execution of the sequences of instructions contained in the main memory 504 causes processor 502 to perform the process steps described herein. One or more processors in a multi-processing arrangement may also be employed to execute the sequences of instructions contained in memory 504. In alternative aspects, hard-wired circuitry may be used in place of or in combination with software instructions to implement various aspects of the present disclosure. Thus, aspects of the present disclosure are not limited to any specific combination of hardware circuitry and software.

Various aspects of the subject matter described in this specification can be implemented in a computing system that includes a back end component, e.g., such as a data server, or that includes a middleware component, e.g., an application server, or that includes a front end component, e.g., a client computer having a graphical user interface or a Web browser through which a user can interact with an implementation of the subject matter described in this specification, or any combination of one or more such back end, middleware, or front end components. The components of the system can be interconnected by any form or medium of digital data communication, e.g., a communication network. The communication network can include, for example, any one or more of a LAN, a WAN, the Internet, and the like. Further, the communication network can include, but is not limited to, for example, any one or more of the following network topologies, including a bus network, a star network, a ring network, a mesh network, a star-bus network, tree or hierarchical network, or the like. The communications modules can be, for example, modems or Ethernet cards.

Computer system 500 can include clients and servers. A client and server are generally remote from each other and typically interact through a communication network. The relationship of client and server arises by virtue of computer programs running on the respective computers and having a client-server relationship to each other. Computer system 500 can be, for example, and without limitation, a desktop computer, laptop computer, or tablet computer. Computer system 500 can also be embedded in another device, for example, and without limitation, a mobile telephone, a PDA, a mobile audio player, a Global Positioning System (GPS) receiver, a video game console, and/or a television set top box.

The term “machine-readable storage medium” or “computer readable medium” as used herein refers to any medium or media that participates in providing instructions to processor 502 for execution. Such a medium may take many forms, including, but not limited to, non-volatile media, volatile media, and transmission media. Non-volatile media include, for example, optical or magnetic disks, such as data storage device 506. Volatile media include dynamic memory, such as memory 504. Transmission media include coaxial cables, copper wire, and fiber optics, including the wires that comprise bus 508. Common forms of machine-readable media include, for example, floppy disk, a flexible disk, hard disk, magnetic tape, any other magnetic medium, a CD-ROM, DVD, any other optical medium, punch cards, paper tape, any other physical medium with patterns of holes, a RAM, a PROM, an EPROM, a FLASH EPROM, any other memory chip or cartridge, or any other medium from which a computer can read. The machine-readable storage medium can be a machine-readable storage device, a machine-readable storage substrate, a memory device, a composition of matter effecting a machine-readable propagated signal, or a combination of one or more of them.

As the user computing system 500 reads game data and provides a game, information may be read from the game data and stored in a memory device, such as the memory 504. Additionally, data from the memory 504 servers accessed via a network the bus 508, or the data storage 506 may be read and loaded into the memory 504. Although data is described as being found in the memory 504, it will be understood that data does not have to be stored in the memory 504 and may be stored in other memory accessible to the processor 502 or distributed among several media, such as the data storage 506.

Embodiments of the disclosed technology may include or be implemented in conjunction with an artificial reality system. Artificial reality, extended reality, or extra reality (collectively "XR") is a form of reality that has been adjusted in some manner before presentation to a user, which may include, e.g., virtual reality (VR), augmented reality (AR), mixed reality (MR), hybrid reality, or some combination and/or derivatives thereof.

Artificial reality content may include completely generated content or generated content combined with captured content (e.g., real-world photographs). The artificial reality content may include video, audio, haptic feedback, or some combination thereof, any of which may be presented in a single channel or in multiple channels (such as stereo video that produces a three-dimensional effect to the viewer). Additionally, in some embodiments, artificial reality may be associated with applications, products, accessories, services, or some combination thereof, that are, e.g., used to create content in an artificial reality and/or used in (e.g., perform activities in) an artificial reality. The artificial reality system that provides the artificial reality content may be implemented on various platforms, including a head-mounted display (HMD) connected to a host computer system, a standalone HMD, a mobile device or computing system, a "cave" environment or other projection system, or any other hardware platform capable of providing artificial reality content to one or more viewers.

"Virtual reality" or "VR," as used herein, refers to an immersive experience where a user's visual input is controlled by a computing system. "Augmented reality" or "AR" refers to systems where a user views images of the real-world after they have passed through a computing system. For example, a tablet with a camera on the back can capture images of the real-world and then display the images on the screen on the opposite side of the tablet from the camera. The tablet can process and adjust or "augment" the images as they passthrough the system, such as by adding virtual objects. "Mixed reality" or "MR" refers to systems where light entering a user's eye is partially generated by a computing system and partially composes light reflected off objects in the real-world. For example, a MR headset could be shaped as a pair of glasses with a pass-through display, which allows light from the real-world to passthrough a waveguide that simultaneously emits light from a projector in the MR headset, allowing the MR headset to present virtual objects intermixed with the real objects the user can see. "Artificial reality," "extra reality," or "XR," as used herein, refers to any of VR, AR, MR, or any combination or hybrid thereof.

As used herein, the phrase “at least one of” preceding a series of items, with the terms “and” or “or” to separate any of the items, modifies the list as a whole, rather than each member of the list (i.e., each item). The phrase “at least one of” does not require selection of at least one item; rather, the phrase allows a meaning that includes at least one of any one of the items, and/or at least one of any combination of the items, and/or at least one of each of the items. By way of example, the phrases “at least one of A, B, and C” or “at least one of A, B, or C” each refer to only A, only B, or only C; any combination of A, B, and C; and/or at least one of each of A, B, and C.

To the extent that the terms “include”, “have”, or the like is used in the description or the claims, such term is intended to be inclusive in a manner similar to the term “comprise” as “comprise” is interpreted when employed as a transitional word in a claim. The word “exemplary” is used herein to mean “serving as an example, instance, or illustration”. Any embodiment described herein as “exemplary” is not necessarily to be construed as preferred or advantageous over other embodiments.

A reference to an element in the singular is not intended to mean “one and only one” unless specifically stated, but rather “one or more”. All structural and functional equivalents to the elements of the various configurations described throughout this disclosure that are known or later come to be known to those of ordinary skill in the art are expressly incorporated herein by reference and intended to be encompassed by the subject technology. Moreover, nothing disclosed herein is intended to be dedicated to the public regardless of whether such disclosure is explicitly recited in the above description.

While this disclosure contains many specifics, these should not be construed as limitations on the scope of what may be claimed, but rather as descriptions of particular implementations of the subject matter. Certain features that are described in this disclosure in the context of separate embodiments can also be implemented in combination in a single embodiment. Conversely, various features that are described in the context of a single embodiment can also be implemented in multiple embodiments separately or in any suitable subcombination. Moreover, although features may be described above as acting in certain combinations and even initially claimed as such, one or more features from a claimed combination can in some cases be excised from the combination, and the claimed combination may be directed to a subcombination or variation of a subcombination.

The subject matter of this disclosure has been described in terms of particular aspects, but other aspects can be implemented and are within the scope of the following claims. For example, while operations are depicted in the drawings in a particular order, this should not be understood as requiring that such operations be performed in the particular order shown or in sequential order, or that all illustrated operations be performed to achieve desirable results. The actions recited in the claims can be performed in a different order and still achieve desirable results. As one example, the processes depicted in the accompanying FIG. 4 do not necessarily require the particular order shown, or sequential order, to achieve desirable results. In certain circumstances, multitasking and parallel processing may be advantageous. Moreover, the separation of various system components in the aspects described above should not be understood as requiring such separation in all aspects, and it should be understood that the described program components and systems can generally be integrated together in a single software product or packaged into multiple software products. Other variations are within the scope of the following claims.

## CONCLUSION

This disclosure describes techniques for performing a medical operation in an artificial reality environment. The described techniques include providing, via a user device configured to render the artificial reality environment, first artificial reality content that simulates performance of the medical operation. Anatomical features of a subject being treated by the medical operation are labeled. An audio signal corresponding to a video frame of the first artificial reality content is determined and the video frame is annotated based on the audio signal. The techniques further include reproducing the audio signal based on determining second artificial reality content corresponding to the performance of the medical operation. A portion of the first artificial reality content matching a portion of the second artificial reality content based on the anatomical features is provided in the artificial reality environment.