June 2023

AI-Assisted Video Editing and Sharing Framework

Priyam Chatterjee

Aayush Bhutani

Roderick Hodgson

Yury Gitman

AI-Assisted Video Editing And Sharing framework

TECHNICAL FIELD

[1] This disclosure generally relates to video editing and sharing.

BRIEF DESCRIPTION OF THE DRAWINGS

[2] FIG. 1 illustrates an example workflow for AI-Assisted Video Editing.

[3] FIG. 2 illustrates an example of a video editor interface showing a suggested segment in a video.

[4] FIG. 3 illustrates an example of the video editor interface showing the location of the current playback in both the timeline in the left pane and the transcript in the right pane.

[5] FIG. 4 illustrates a user selecting a video segment by selecting a portion of the transcript.

[6] FIG. 5 illustrates an example method for editing and publishing a video using particular embodiments described herein.

[7] FIG. 6 illustrates an example network environment associated with a social-networking system.

[8] FIG. 7 illustrates an example computer system.

DESCRIPTION OF EXAMPLE EMBODIMENTS

[9] Publishers upload a substantial number of videos onto social media platforms on a daily basis. The uploaded content is then consumed by viewers via a variety of “surfaces,” which is used herein to refer to the graphical user interface of any device, software, or platform through which the content could be displayed. Such surfaces may differ from each other in the aspect ratio of the viewport or interface, as well as the minimum and maximum duration of the content that may be deemed to be optimum or even eligible. Examples of such surfaces are Facebook Watch, Reels, Stories, etc. Oftentimes, publishers need to edit these videos and clip out interesting segments. In addition, if they want to target different surfaces (example, Watch, Stories, Reels) where different formats are needed, they would need to edit their content offline and upload separate videos for each surface, without any way to link between the different posts. Such

publishers want simple and easy-to-use tools to help them edit videos and publish to different surfaces to connect with wider audiences.

[10] Embodiments described herein, collectively referred to as AI-Assisted Video Editing, allow creators and publishers (referred to as Users) to upload a single video and share it across multiple surfaces (e.g., social media webpage, mobile app, watch, virtual reality, augmented reality, etc.), with target-specific transformations (such as trimming and reframing) added as part of the single workflow.

[11] We built these tools to democratize the use of video editing to creators. These creators may not have the expertise to edit videos on their own using expensive, sophisticated tools or hire publishing houses to do it for them. Additionally, we can also employ AI tools to suggest best practices for the surface that the user wants to export their content to.

[12] FIG. 1 illustrates an example workflow for AI-Assisted Video Editing. The user interface for the feature may be presented via a webpage or an app, for example. At step 110, a user uploads a video clip to be published to a server of a media platform. At step 120, the user sees a list (e.g., one or more) of clips that are pre-selected by an AI system of the media platform. The list of clips may be shared on one or more surfaces (e.g., website, app, etc.) or edited before sharing. The one or more clips presented to the user may include clips that were previously uploaded by the user, clips shared with the user, etc. The user previews and selects one clip, and the user can choose to either share this clip directly or edit it. At step 130, the user may choose to open the editor to edit the selected clip. At step 140, the media platform may present the user with a text editor and suggest interesting segments within the selected clip. In particular embodiments, the suggested interesting segments within the clip may be detected using an AI model. At step 150, the user may review the automatically selected segments and determine whether they are satisfactory. If not, at step 160, the user may proceed with editing the clip. For example, the user may change the suggested segments by dragging and dropping the text to highlight words that highlight the corresponding portion in the video playback and timeline. The user may clip the highlighted segment and the corresponding part of the video will be clipped, leading to an edited video. Additional AI systems are applied to this edited video to reframe/crop the clips for different surfaces (e.g., landscape videos are transformed to portrait mode for app surfaces, small clippings are published to watch surfaces, etc.). In particular embodiments, the reframe/cropping process may be done automatically depending on the surfaces to which the user wishes to publish. At step

170, the user may preview the clip before publishing. In particular embodiments, the user may be presented with previews of the clips as they would be viewed on different surfaces (e.g., a simulated watch may show how the clip would appear on a watch, a simulated phone to show how the clip would appear on a mobile app, etc.). At step 180, once the user is satisfied with the previews, the user may proceed with publishing the clip. The platform may selectively publish clips with the appropriate aspect ratios/dimensions to the appropriate surfaces. For example, a small clip of the edited video segments may be published to watch surfaces, a portrait-mode clipping of the edited video segments may be published to mobile apps, a landscape-mode clipping of the edited video segments may be published on websites, etc.

[13] FIG. 2 illustrates an example of a video editor interface 200 showing a suggested segment in a video. In particular embodiments, any suitable video-highlights detection technology may be used to analyze the video clip selected by the user to pre-create one or more interesting regions that can be shared on one or more surfaces. In FIG. 2, the automatically-selected segment 210 is highlighted in the timeline under the video player in the left pane. The right pane shows a text editor, which includes a transcript of utterances detected in the video, listed with their corresponding timestamps. The transcript may be automatically detected using Automatic Speech Recognition or Speech-to-Text technology. The highlighted text 220 corresponds to the transcript of the utterance detected within the automatically-selected segment 210.

[14] FIG. 3 illustrates an example of the video editor interface 200 showing the location of the current playback in both the timeline in the left pane and the transcript in the right pane. The cursor 310 shown in the timeline corresponds to the current location of the video playback. The word uttered at the time corresponding to cursor 310 is shown as highlighted text 320 in the text editor. This feature allows for a seamless editing experience and gives the impression of editing text on paper.

[15] The user can click on a word outside the highlighted segment (e.g., text 320 shown in FIG. 3) to move the cursor (playing head) to that position both in the text editor and the timeline. The video player would also jump to the video frame corresponding to that point in time.

[16] If and when users choose to modify the AI system's selection, they can use the text-based editing console to select parts of the transcript that they want to clip. The corresponding selected segments in the video are highlighted (e.g., in light blue), and the user can choose to clip these through the "create clip from selection" button, which pops up when the user leaves/ drops

the mouse after dragging. In particular embodiments, this selection may include regions that are not highlighted in text to create seamless start and stop regions of the clip. For example, FIG. 4 illustrates a user selecting a portion 420 of the transcript. This results in an additional portion of the video segment 410 to be highlighted. This portion 410 may initially be highlighted in dotted lines to visually indicate that portion that would be selected once the user confirms the selection. Once the user confirms, the dotted portion 410 would be converted into solid lines (or any other suitable visual indicator) to inform the user that it is part of the current clipping.

[17] Users can see the duration of the currently selected segments as they are dragging and highlighting, which is helpful if they want the final output to be under a specific duration to better suit different consumption surfaces like Watch, Stories, Reels, etc.

[18] If the surface that the user chooses to share the clip would benefit from additional modifications (for example, aspect ratio), the AI system of the media platform could support such edits as well. For example, the system may support automatic reframing to create the best portrait videos from the clips if they have different formats (e.g., landscape or portrait mode). Additional automated features may include, e.g., baking in the title or catchy descriptions in a video without being intrusive, the addition of audio assets in a video such as laughter tracks, sad music, etc.

[19] After the user is done editing, a final preview of the clipped and reframed video may be shown to the user. In particular embodiments, the preview may be shown in the format of the desired surface through which the video clip would be shown. For example, if the video clip is to be shared via a mobile app, a preview of the video would be displayed within a virtual representation of a mobile app to show the user how the video would look. Once the user is ready to publish, the video will be shared via the selected surfaces.

FIG. 5 illustrates an example method 500 for editing and publishing a video using particular embodiments described herein. The method may begin at step 510, where a computing system (e.g., a server, a computer, etc.) receives a selection of a video from a user. At step 520, the system may determine a recommended segment of the video for publication using a machine-learning model. At step 530, the system may display a video editor comprising a video playback region, a video timeline, and a text editor including a transcript of utterances made within the video. At step 540, the system may highlight a first portion of the video timeline that corresponds to the recommended segment of the video and highlight a first portion of the transcript that corresponds to utterances in the recommended segment of the video. At step 550, the system may

receive, from the user, instructions to highlight a second portion of the transcript, the second portion of the transcript corresponding to a user-selected segment of the video. At step 560, the system may highlight a second portion of the video timeline corresponding to the user-selected segment of the video. At step 570, the system may generate an output video based on the recommended segment of the video and the user-selected segment of the video. At step 580, the system may reformat the output video based on one or more surfaces through which the output video is to be published. At step 590, the system may publish the reformatted output video through the one or more surfaces. Although this disclosure describes and illustrates particular steps of the method of FIG. 5 as occurring in a particular order, this disclosure contemplates any suitable steps of the method of FIG. 5 occurring in any suitable order and may include all or some of the steps of the method of FIG. 5, where appropriate.

[20] **FIG. 6** illustrates an example network environment 600 associated with a social-networking system. Network environment 600 includes a user 601, a client system 630, a social-networking system 660, and a third-party system 670 connected to each other by a network 610. Although FIG. 6 illustrates a particular arrangement of user 601, client system 630, social-networking system 660, third-party system 670, and network 610, this disclosure contemplates any suitable arrangement of user 601, client system 630, social-networking system 660, third-party system 670, and network 610. As an example and not by way of limitation, two or more of client system 630, social-networking system 660, and third-party system 670 may be connected to each other directly, bypassing network 610. As another example, two or more of client system 630, social-networking system 660, and third-party system 670 may be physically or logically co-located with each other in whole or in part. Moreover, although FIG. 6 illustrates a particular number of users 601, client systems 630, social-networking systems 660, third-party systems 670, and networks 610, this disclosure contemplates any suitable number of users 601, client systems 630, social-networking systems 660, third-party systems 670, and networks 610. As an example and not by way of limitation, network environment 600 may include multiple users 601, client system 630, social-networking systems 660, third-party systems 670, and networks 610.

[21] In particular embodiments, user 601 may be an individual (human user), an entity (e.g., an enterprise, business, or third-party application), or a group (e.g., of individuals or entities) that interacts or communicates with or over social-networking system 660. In particular embodiments, social-networking system 660 may be a network-addressable computing system

hosting an online social network. Social-networking system 660 may generate, store, receive, and send social-networking data, such as, for example, user-profile data, concept-profile data, social-graph information, or other suitable data related to the online social network. Social-networking system 660 may be accessed by the other components of network environment 600 either directly or via network 610. In particular embodiments, social-networking system 660 may include an authorization server (or other suitable component(s)) that allows users 601 to opt in to or opt out of having their actions logged by social-networking system 660 or shared with other systems (e.g., third-party systems 670), for example, by setting appropriate privacy settings. A privacy setting of a user may determine what information associated with the user may be logged, how information associated with the user may be logged, when information associated with the user may be logged, who may log information associated with the user, whom information associated with the user may be shared with, and for what purposes information associated with the user may be logged or shared. Authorization servers may be used to enforce one or more privacy settings of the users of social-networking system 30 through blocking, data hashing, anonymization, or other suitable techniques as appropriate. In particular embodiments, third-party system 670 may be a network-addressable computing system that can host any media publishing feature. Third-party system 670 may be accessed by the other components of network environment 600 either directly or via network 610. In particular embodiments, one or more users 601 may use one or more client systems 630 to access, send data to, and receive data from social-networking system 660 or third-party system 670. Client system 630 may access social-networking system 660 or third-party system 670 directly, via network 610, or via a third-party system. As an example and not by way of limitation, client system 630 may access third-party system 670 via social-networking system 660. Client system 630 may be any suitable computing device, such as, for example, a personal computer, a laptop computer, a cellular telephone, a smartphone, a tablet computer, or an augmented/virtual reality device.

[22] This disclosure contemplates any suitable network 610. As an example and not by way of limitation, one or more portions of network 610 may include an ad hoc network, an intranet, an extranet, a virtual private network (VPN), a local area network (LAN), a wireless LAN (WLAN), a wide area network (WAN), a wireless WAN (WWAN), a metropolitan area network (MAN), a portion of the Internet, a portion of the Public Switched Telephone Network (PSTN), a

cellular telephone network, or a combination of two or more of these. Network 610 may include one or more networks 610.

[23] Links 650 may connect client system 630, social-networking system 660, and third-party system 670 to communication network 610 or to each other. This disclosure contemplates any suitable links 650. In particular embodiments, one or more links 650 include one or more wireline (such as for example Digital Subscriber Line (DSL) or Data Over Cable Service Interface Specification (DOCSIS)), wireless (such as for example Wi-Fi or Worldwide Interoperability for Microwave Access (WiMAX)), or optical (such as for example Synchronous Optical Network (SONET) or Synchronous Digital Hierarchy (SDH)) links. In particular embodiments, one or more links 650 each include an ad hoc network, an intranet, an extranet, a VPN, a LAN, a WLAN, a WAN, a WWAN, a MAN, a portion of the Internet, a portion of the PSTN, a cellular technology-based network, a satellite communications technology-based network, another link 650, or a combination of two or more such links 650. Links 650 need not necessarily be the same throughout network environment 600. One or more first links 650 may differ in one or more respects from one or more second links 650.

[24] **FIG. 7** illustrates an example computer system 700. In particular embodiments, one or more computer systems 700 perform one or more steps of one or more methods described or illustrated herein. In particular embodiments, one or more computer systems 700 provide functionality described or illustrated herein. In particular embodiments, software running on one or more computer systems 700 performs one or more steps of one or more methods described or illustrated herein or provides functionality described or illustrated herein. Particular embodiments include one or more portions of one or more computer systems 700. Herein, reference to a computer system may encompass a computing device, and vice versa, where appropriate. Moreover, reference to a computer system may encompass one or more computer systems, where appropriate.

[25] This disclosure contemplates any suitable number of computer systems 700. This disclosure contemplates computer system 700 taking any suitable physical form. As example and not by way of limitation, computer system 700 may be an embedded computer system, a system-on-chip (SOC), a single-board computer system (SBC) (such as, for example, a computer-on-module (COM) or system-on-module (SOM)), a desktop computer system, a laptop or notebook computer system, an interactive kiosk, a mainframe, a mesh of computer systems, a mobile

telephone, a personal digital assistant (PDA), a server, a tablet computer system, an augmented/virtual reality device, or a combination of two or more of these. Where appropriate, computer system 700 may include one or more computer systems 700; be unitary or distributed; span multiple locations; span multiple machines; span multiple data centers; or reside in a cloud, which may include one or more cloud components in one or more networks. Where appropriate, one or more computer systems 700 may perform without substantial spatial or temporal limitation one or more steps of one or more methods described or illustrated herein. As an example and not by way of limitation, one or more computer systems 700 may perform in real time or in batch mode one or more steps of one or more methods described or illustrated herein. One or more computer systems 700 may perform at different times or at different locations one or more steps of one or more methods described or illustrated herein, where appropriate.

[26] In particular embodiments, computer system 700 includes a processor 702, memory 704, storage 706, an input/output (I/O) interface 708, a communication interface 710, and a bus 712. Although this disclosure describes and illustrates a particular computer system having a particular number of particular components in a particular arrangement, this disclosure contemplates any suitable computer system having any suitable number of any suitable components in any suitable arrangement.

[27] In particular embodiments, processor 702 includes hardware for executing instructions, such as those making up a computer program. As an example and not by way of limitation, to execute instructions, processor 702 may retrieve (or fetch) the instructions from an internal register, an internal cache, memory 704, or storage 706; decode and execute them; and then write one or more results to an internal register, an internal cache, memory 704, or storage 706. In particular embodiments, processor 702 may include one or more internal caches for data, instructions, or addresses. This disclosure contemplates processor 702 including any suitable number of any suitable internal caches, where appropriate. As an example and not by way of limitation, processor 702 may include one or more instruction caches, one or more data caches, and one or more translation lookaside buffers (TLBs). Instructions in the instruction caches may be copies of instructions in memory 704 or storage 706, and the instruction caches may speed up retrieval of those instructions by processor 702. Data in the data caches may be copies of data in memory 704 or storage 706 for instructions executing at processor 702 to operate on; the results of previous instructions executed at processor 702 for access by subsequent instructions executing

at processor 702 or for writing to memory 704 or storage 706; or other suitable data. The data caches may speed up read or write operations by processor 702. The TLBs may speed up virtual-address translation for processor 702. In particular embodiments, processor 702 may include one or more internal registers for data, instructions, or addresses. This disclosure contemplates processor 702 including any suitable number of any suitable internal registers, where appropriate. Where appropriate, processor 702 may include one or more arithmetic logic units (ALUs); be a multi-core processor; or include one or more processors 702. Although this disclosure describes and illustrates a particular processor, this disclosure contemplates any suitable processor.

[28] In particular embodiments, memory 704 includes main memory for storing instructions for processor 702 to execute or data for processor 702 to operate on. As an example and not by way of limitation, computer system 700 may load instructions from storage 706 or another source (such as, for example, another computer system 700) to memory 704. Processor 702 may then load the instructions from memory 704 to an internal register or internal cache. To execute the instructions, processor 702 may retrieve the instructions from the internal register or internal cache and decode them. During or after execution of the instructions, processor 702 may write one or more results (which may be intermediate or final results) to the internal register or internal cache. Processor 702 may then write one or more of those results to memory 704. In particular embodiments, processor 702 executes only instructions in one or more internal registers or internal caches or in memory 704 (as opposed to storage 706 or elsewhere) and operates only on data in one or more internal registers or internal caches or in memory 704 (as opposed to storage 706 or elsewhere). One or more memory buses (which may each include an address bus and a data bus) may couple processor 702 to memory 704. Bus 712 may include one or more memory buses, as described below. In particular embodiments, one or more memory management units (MMUs) reside between processor 702 and memory 704 and facilitate accesses to memory 704 requested by processor 702. In particular embodiments, memory 704 includes random access memory (RAM). This RAM may be volatile memory, where appropriate. Where appropriate, this RAM may be dynamic RAM (DRAM) or static RAM (SRAM). Moreover, where appropriate, this RAM may be single-ported or multi-ported RAM. This disclosure contemplates any suitable RAM. Memory 704 may include one or more memories 704, where appropriate. Although this disclosure describes and illustrates particular memory, this disclosure contemplates any suitable memory.

[29] In particular embodiments, storage 706 includes mass storage for data or instructions. As an example and not by way of limitation, storage 706 may include a hard disk drive (HDD), a floppy disk drive, flash memory, an optical disc, a magneto-optical disc, magnetic tape, or a Universal Serial Bus (USB) drive or a combination of two or more of these. Storage 706 may include removable or non-removable (or fixed) media, where appropriate. Storage 706 may be internal or external to computer system 700, where appropriate. In particular embodiments, storage 706 is non-volatile, solid-state memory. In particular embodiments, storage 706 includes read-only memory (ROM). Where appropriate, this ROM may be mask-programmed ROM, programmable ROM (PROM), erasable PROM (EPROM), electrically erasable PROM (EEPROM), electrically alterable ROM (EAROM), or flash memory or a combination of two or more of these. This disclosure contemplates mass storage 706 taking any suitable physical form. Storage 706 may include one or more storage control units facilitating communication between processor 702 and storage 706, where appropriate. Where appropriate, storage 706 may include one or more storages 706. Although this disclosure describes and illustrates particular storage, this disclosure contemplates any suitable storage.

[30] In particular embodiments, I/O interface 708 includes hardware, software, or both, providing one or more interfaces for communication between computer system 700 and one or more I/O devices. Computer system 700 may include one or more of these I/O devices, where appropriate. One or more of these I/O devices may enable communication between a person and computer system 700. As an example and not by way of limitation, an I/O device may include a keyboard, keypad, microphone, monitor, mouse, printer, scanner, speaker, still camera, stylus, tablet, touch screen, trackball, video camera, another suitable I/O device or a combination of two or more of these. An I/O device may include one or more sensors. This disclosure contemplates any suitable I/O devices and any suitable I/O interfaces 708 for them. Where appropriate, I/O interface 708 may include one or more device or software drivers enabling processor 702 to drive one or more of these I/O devices. I/O interface 708 may include one or more I/O interfaces 708, where appropriate. Although this disclosure describes and illustrates a particular I/O interface, this disclosure contemplates any suitable I/O interface.

[31] In particular embodiments, communication interface 710 includes hardware, software, or both providing one or more interfaces for communication (such as, for example, packet-based communication) between computer system 700 and one or more other computer

systems 700 or one or more networks. As an example and not by way of limitation, communication interface 710 may include a network interface controller (NIC) or network adapter for communicating with an Ethernet or other wire-based network or a wireless NIC (WNIC) or wireless adapter for communicating with a wireless network, such as a WI-FI network. This disclosure contemplates any suitable network and any suitable communication interface 710 for it. As an example and not by way of limitation, computer system 700 may communicate with an ad hoc network, a personal area network (PAN), a local area network (LAN), a wide area network (WAN), a metropolitan area network (MAN), or one or more portions of the Internet or a combination of two or more of these. One or more portions of one or more of these networks may be wired or wireless. As an example, computer system 700 may communicate with a wireless PAN (WPAN) (such as, for example, a BLUETOOTH WPAN), a WI-FI network, a WI-MAX network, a cellular telephone network (such as, for example, a Global System for Mobile Communications (GSM) network), or other suitable wireless network or a combination of two or more of these. Computer system 700 may include any suitable communication interface 710 for any of these networks, where appropriate. Communication interface 710 may include one or more communication interfaces 710, where appropriate. Although this disclosure describes and illustrates a particular communication interface, this disclosure contemplates any suitable communication interface.

[32] In particular embodiments, bus 712 includes hardware, software, or both coupling components of computer system 700 to each other. As an example and not by way of limitation, bus 712 may include an Accelerated Graphics Port (AGP) or other graphics bus, an Enhanced Industry Standard Architecture (EISA) bus, a front-side bus (FSB), a HYPERTRANSPORT (HT) interconnect, an Industry Standard Architecture (ISA) bus, an INFINIBAND interconnect, a low-pin-count (LPC) bus, a memory bus, a Micro Channel Architecture (MCA) bus, a Peripheral Component Interconnect (PCI) bus, a PCI-Express (PCIe) bus, a serial advanced technology attachment (SATA) bus, a Video Electronics Standards Association local (VLB) bus, or another suitable bus or a combination of two or more of these. Bus 712 may include one or more buses 712, where appropriate. Although this disclosure describes and illustrates a particular bus, this disclosure contemplates any suitable bus or interconnect.

[33] Herein, a computer-readable non-transitory storage medium or media may include one or more semiconductor-based or other integrated circuits (ICs) (such, as for example, field-

programmable gate arrays (FPGAs) or application-specific ICs (ASICs)), hard disk drives (HDDs), hybrid hard drives (HHDs), optical discs, optical disc drives (ODDs), magneto-optical discs, magneto-optical drives, floppy diskettes, floppy disk drives (FDDs), magnetic tapes, solid-state drives (SSDs), RAM-drives, SECURE DIGITAL cards or drives, any other suitable computer-readable non-transitory storage media, or any suitable combination of two or more of these, where appropriate. A computer-readable non-transitory storage medium may be volatile, non-volatile, or a combination of volatile and non-volatile, where appropriate.

[34] Herein, "or" is inclusive and not exclusive, unless expressly indicated otherwise or indicated otherwise by context. Therefore, herein, "A or B" means "A, B, or both," unless expressly indicated otherwise or indicated otherwise by context. Moreover, "and" is both joint and several, unless expressly indicated otherwise or indicated otherwise by context. Therefore, herein, "A and B" means "A and B, jointly or severally," unless expressly indicated otherwise or indicated otherwise by context.

[35] The scope of this disclosure encompasses all changes, substitutions, variations, alterations, and modifications to the example embodiments described or illustrated herein that a person having ordinary skill in the art would comprehend. The scope of this disclosure is not limited to the example embodiments described or illustrated herein. Moreover, although this disclosure describes and illustrates respective embodiments herein as including particular components, elements, feature, functions, operations, or steps, any of these embodiments may include any combination or permutation of any of the components, elements, features, functions, operations, or steps described or illustrated anywhere herein that a person having ordinary skill in the art would comprehend. Furthermore, reference in the appended claims to an apparatus or system or a component of an apparatus or system being adapted to, arranged to, capable of, configured to, enabled to, operable to, or operative to perform a particular function encompasses that apparatus, system, component, whether or not it or that particular function is activated, turned on, or unlocked, as long as that apparatus, system, or component is so adapted, arranged, capable, configured, enabled, operable, or operative. Additionally, although this disclosure describes or illustrates particular embodiments as providing particular advantages, particular embodiments may provide none, some, or all of these advantages.

CLAIMS

What is claimed is:

1. A method comprising, by a computing system:

receiving a selection of a video from a user;

determining a recommended segment of the video for publication using a machine-learning model;

displaying a video editor comprising a video playback region, a video timeline, and a text editor including a transcript of utterances made within the video;

highlighting a first portion of the video timeline that corresponds to the recommended segment of the video and highlighting a first portion of the transcript that corresponds to utterances in the recommended segment of the video;

receiving, from the user, instructions to highlight a second portion of the transcript, the second portion of the transcript corresponding to a user-selected segment of the video;

highlighting a second portion of the video timeline corresponding to the user-selected segment of the video;

generating an output video based on the recommended segment of the video and the user-selected segment of the video;

reformatting the output video based on one or more surfaces through which the output video is to be published; and

publishing the reformatted output video through the one or more surfaces.

ABSTRACT

An AI-assisted video editing and sharing framework that provides video clipping recommendations and automatically reformats the final video based on surfaces through which the final video is to be published.

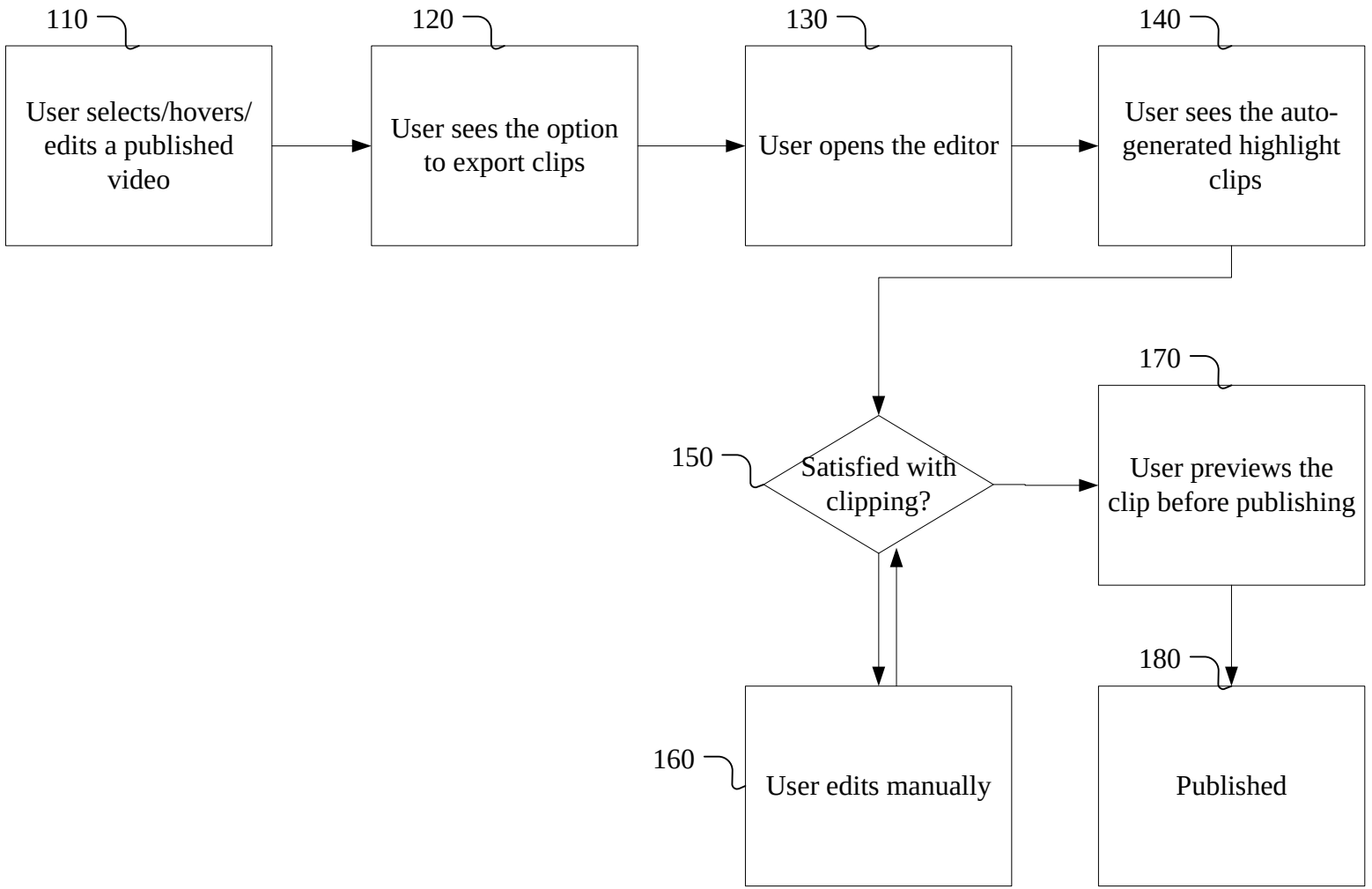
110
User selects/hovers/ edits a published video
120
User sees the option to export clips
130
User opens the editor
140
User sees the auto-generated highlight clips
150
Satisfied with clipping?
170
User previews the clip before publishing
160
User edits manually
180
Published

FIG. 1

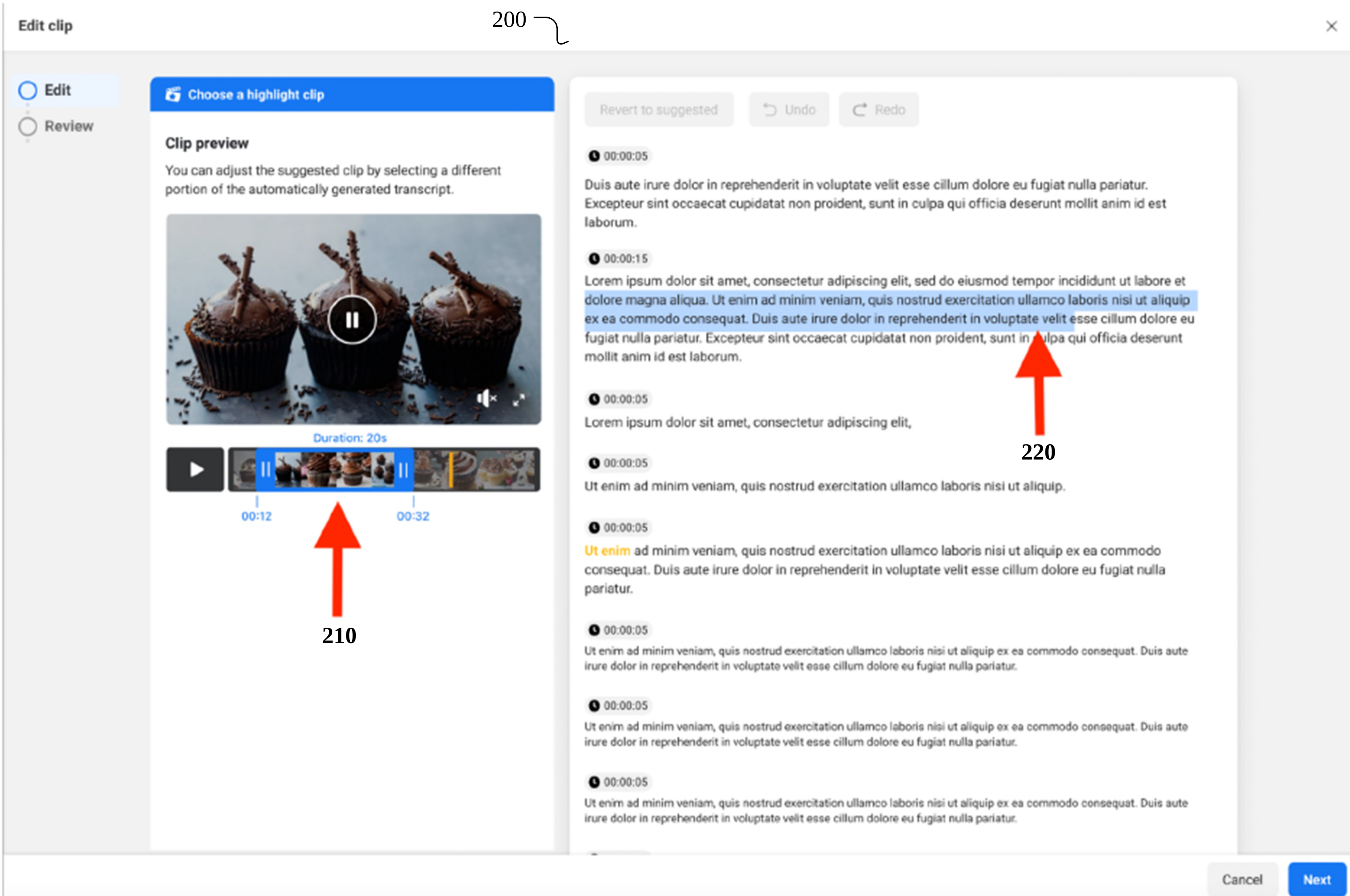

FIG. 2

Sheet 2/7

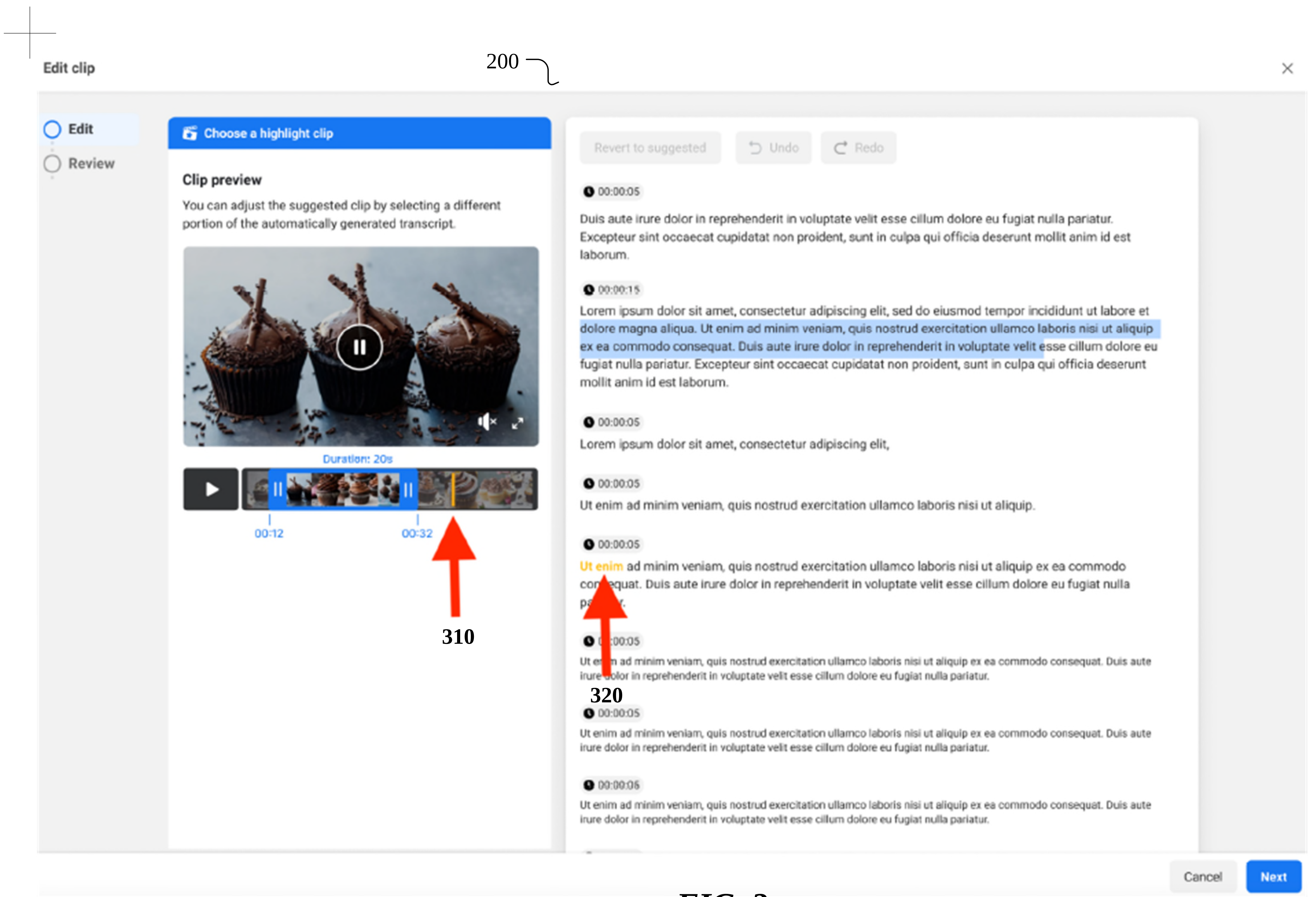

FIG. 3

Sheet 3/7

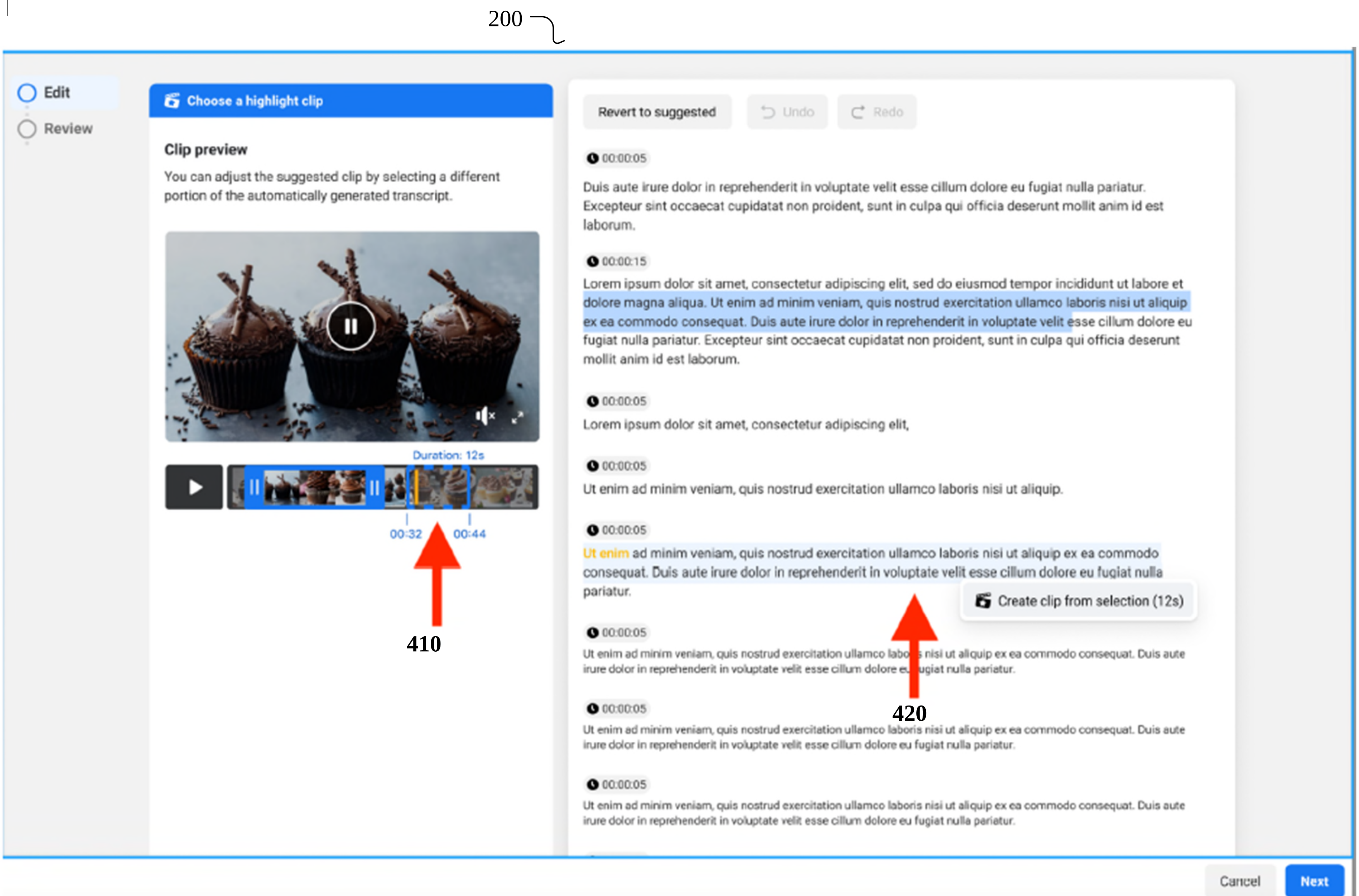

FIG. 4

Sheet 4/7

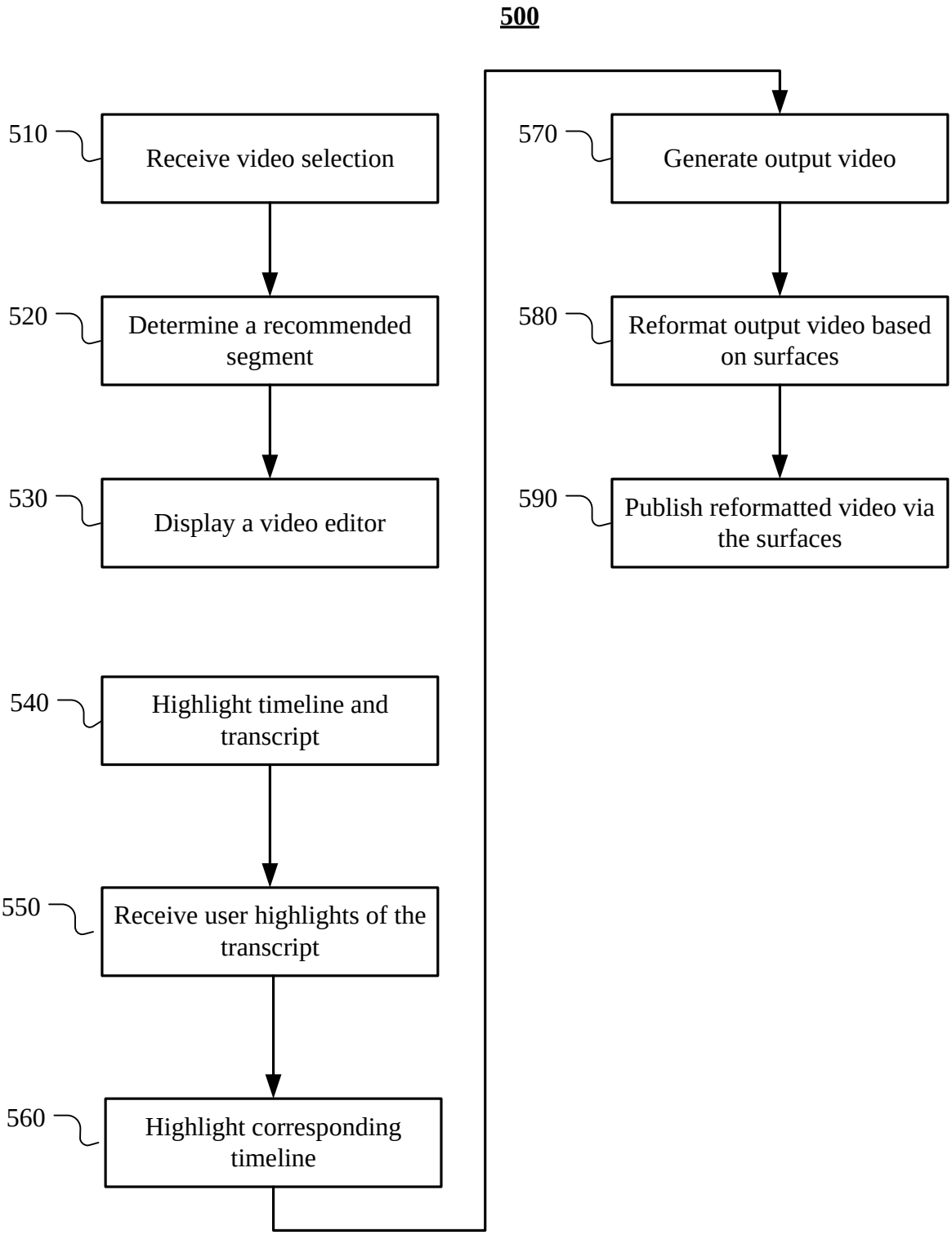
500
510
Receive video selection
520
Determine a recommended segment
530
Display a video editor
540
Highlight timeline and transcript
550
Receive user highlights of the transcript
560
Highlight corresponding timeline
570
Generate output video
580
Reformat output video based on surfaces
590
Publish reformatted video via the surfaces

FIG. 5

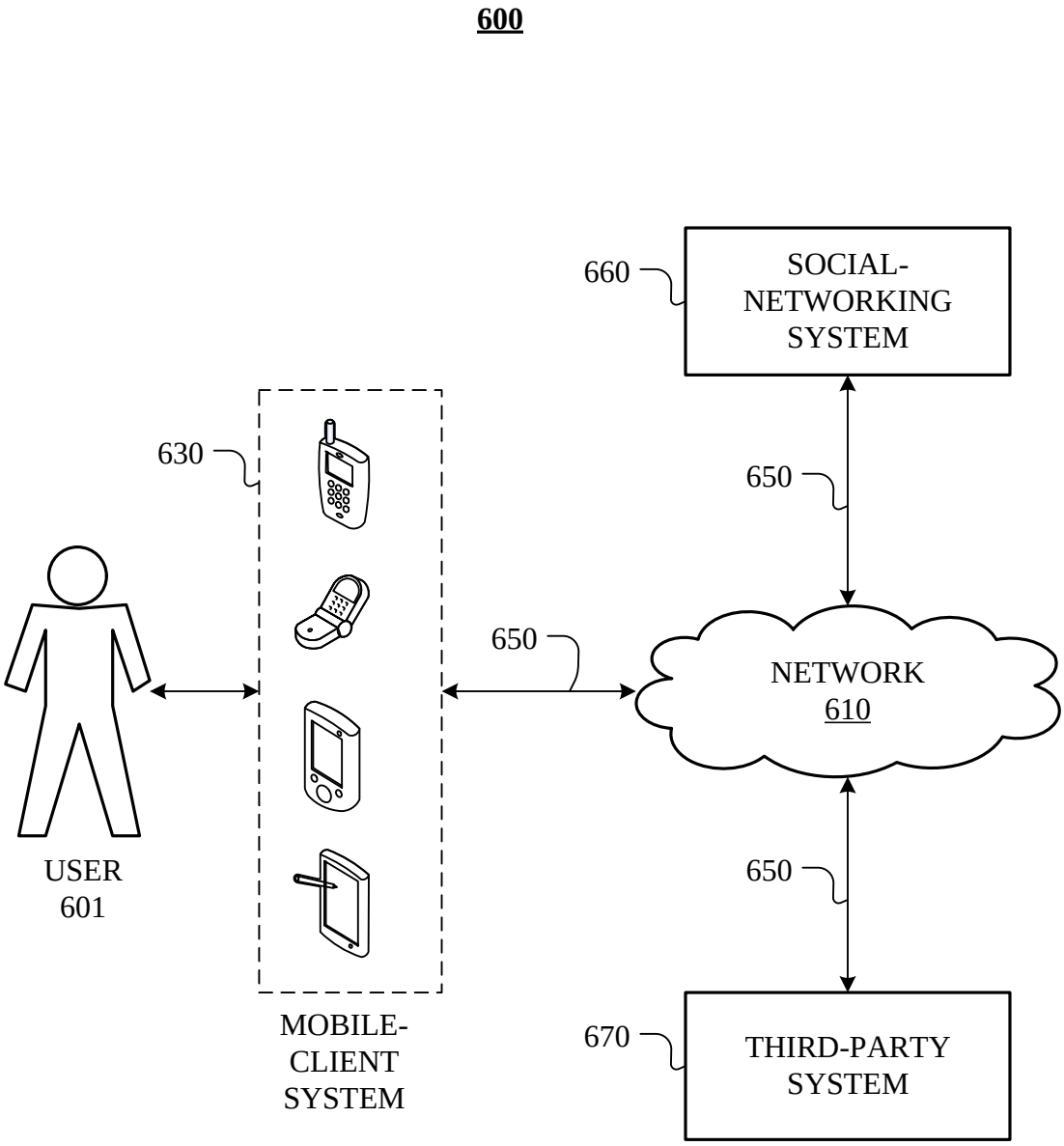
600
660
SOCIAL-NETWORKING SYSTEM
650
630
650
NETWORK 610
650
USER 601
MOBILE-CLIENT SYSTEM
670
THIRD-PARTY SYSTEM

FIG. 6

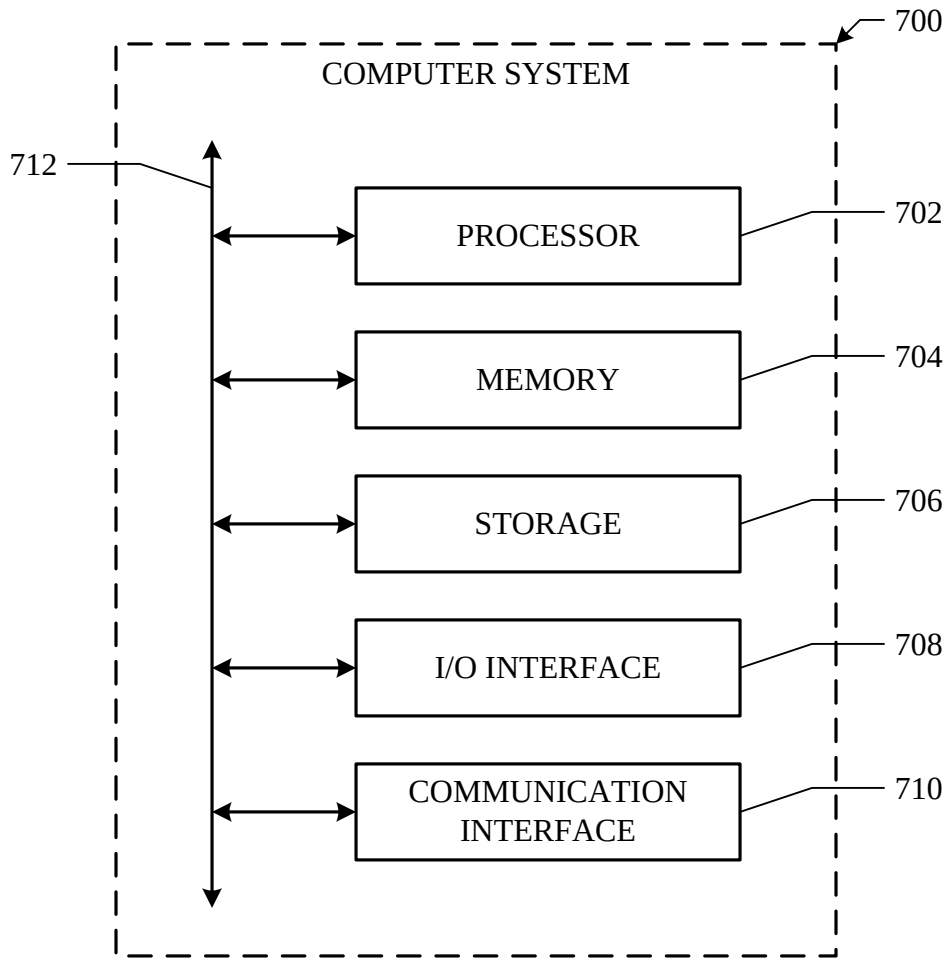
700
COMPUTER SYSTEM
712
PROCESSOR
702
MEMORY
704
STORAGE
706
I/O INTERFACE
708
COMMUNICATION INTERFACE
710

FIG. 7