



US 20210004405A1

(19) **United States**

(12) **Patent Application Publication**
Solomon et al.

(10) **Pub. No.: US 2021/0004405 A1**

(43) **Pub. Date: Jan. 7, 2021**

(54) **ENHANCING TANGIBLE CONTENT ON
PHYSICAL ACTIVITY SURFACE**

G06K 9/00 (2006.01)

G06F 16/738 (2006.01)

G06F 16/783 (2006.01)

(71) Applicant: **Tangible Play, Inc.**, Palo Alto, CA
(US)

(52) **U.S. CL.**

CPC **G06F 16/7335** (2019.01); **G09B 5/06**

(2013.01); **G06F 16/783** (2019.01); **G06K**

9/00456 (2013.01); **G06F 16/738** (2019.01);

G09B 5/12 (2013.01)

(72) Inventors: **Mark Solomon**, San Jose, CA (US);
Jerome Scholler, San Francisco, CA
(US); **Pramod Kumar Sharma**, San
Jose, CA (US)

(21) Appl. No.: **16/880,882**

(22) Filed: **May 21, 2020**

Related U.S. Application Data

(60) Provisional application No. 62/871,195, filed on Jul.
7, 2019.

Publication Classification

(51) **Int. CL.**

G06F 16/732 (2006.01)

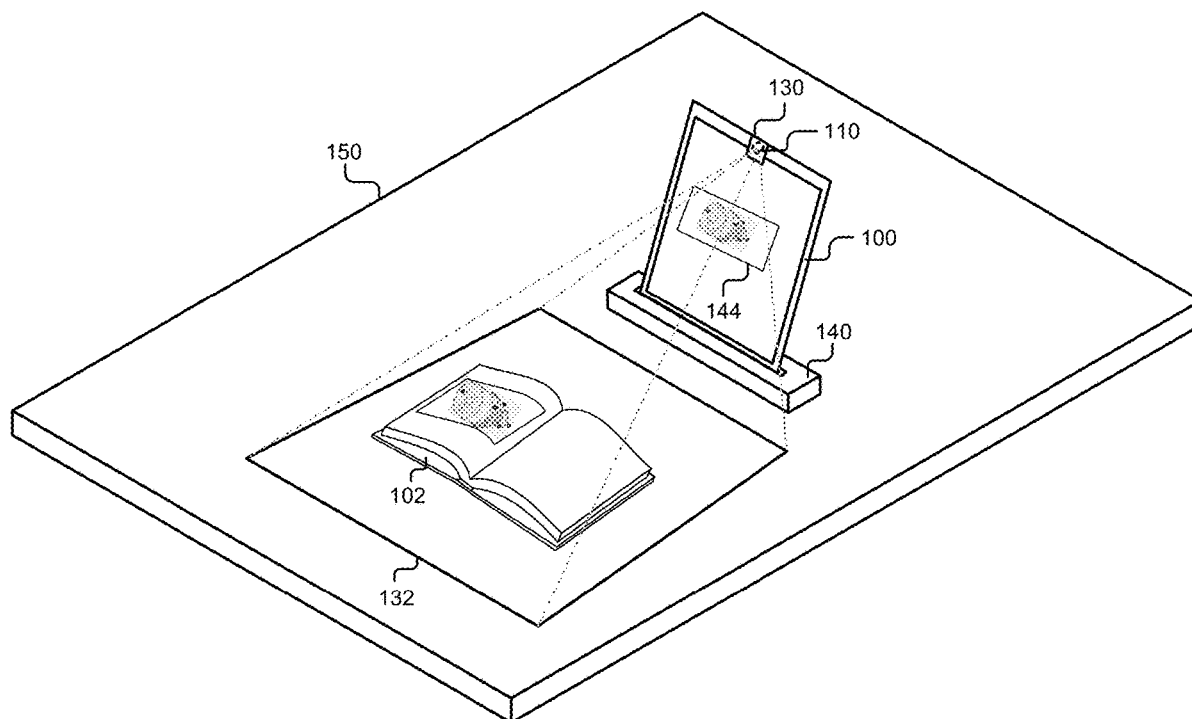
G09B 5/06 (2006.01)

G09B 5/12 (2006.01)

(57)

ABSTRACT

A method for enhancing tangible content on a physical activity surface is described. In an example implementation, the method includes capturing, using a video capture device of a computing device, a video stream that includes an activity scene of a physical activity surface; detecting in the video stream, using a detector executable on the computing device, a tangible content item on the physical activity surface; recognizing, from the video stream, one or more visually instructive elements in the tangible content item; determining a tangible identifier based on the one or more visually instructive elements in the tangible content item; retrieving a digital content item using the tangible identifier; and providing the digital content item on the computing device.



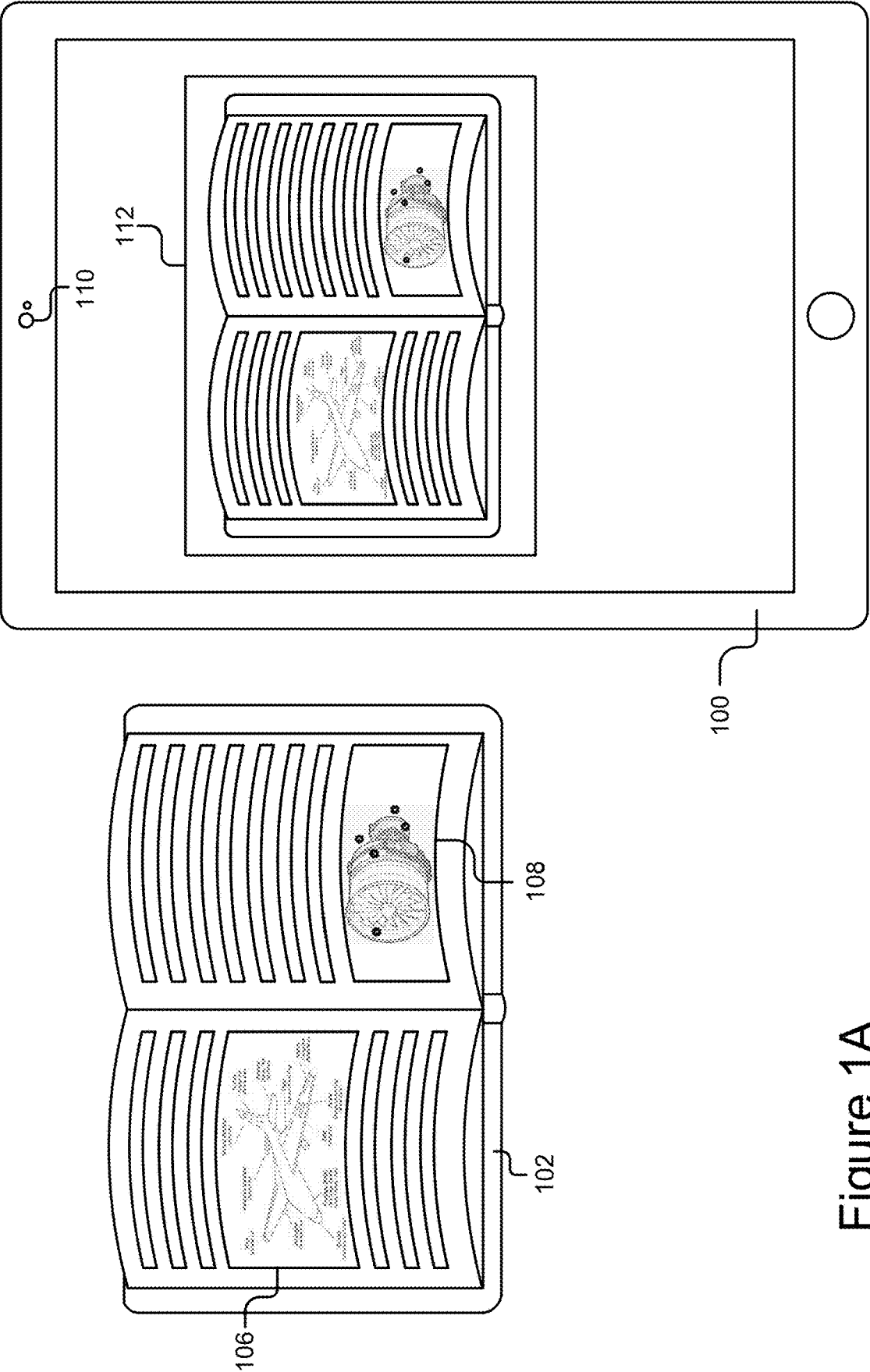


Figure 1A

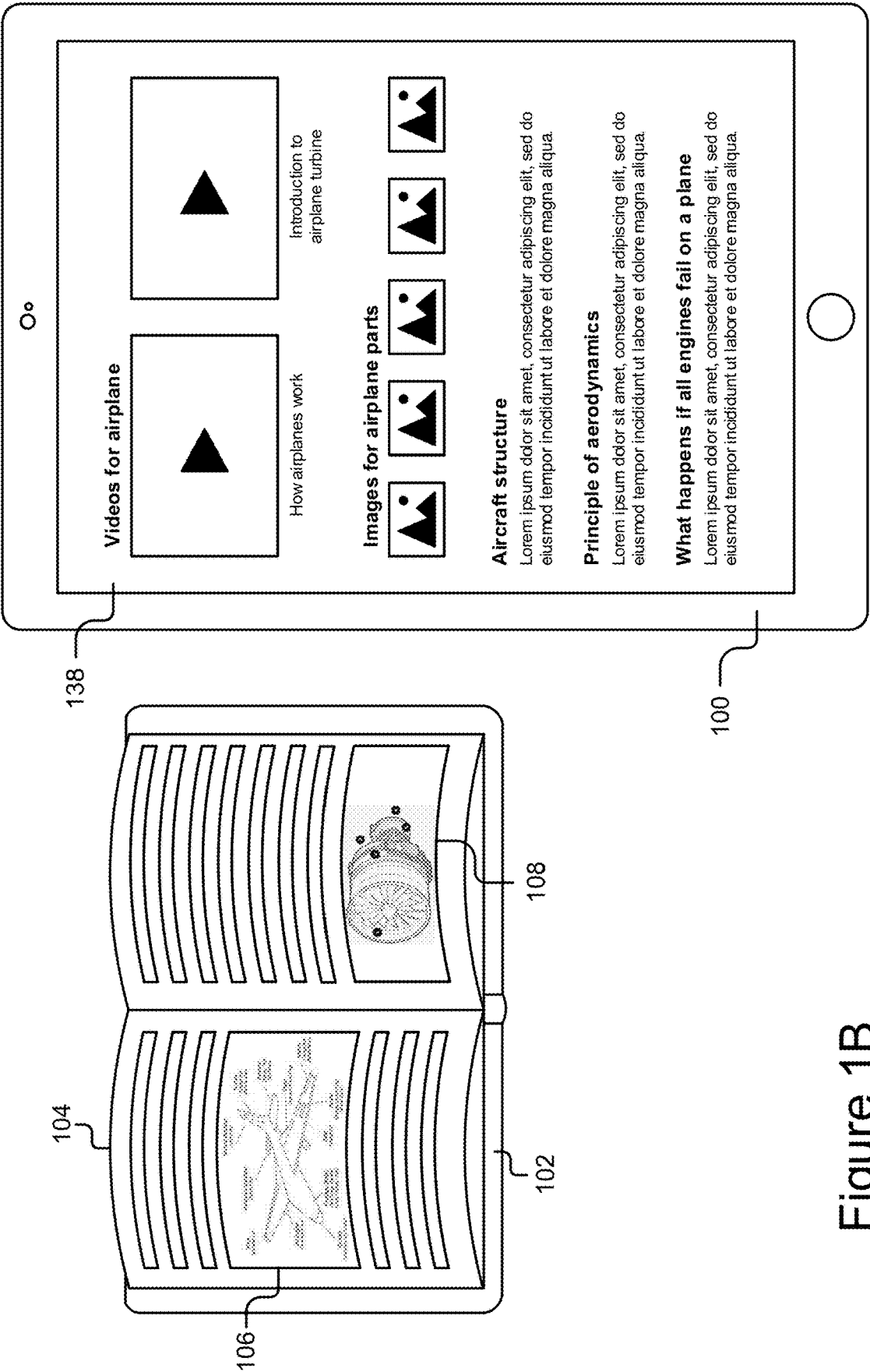


Figure 1B

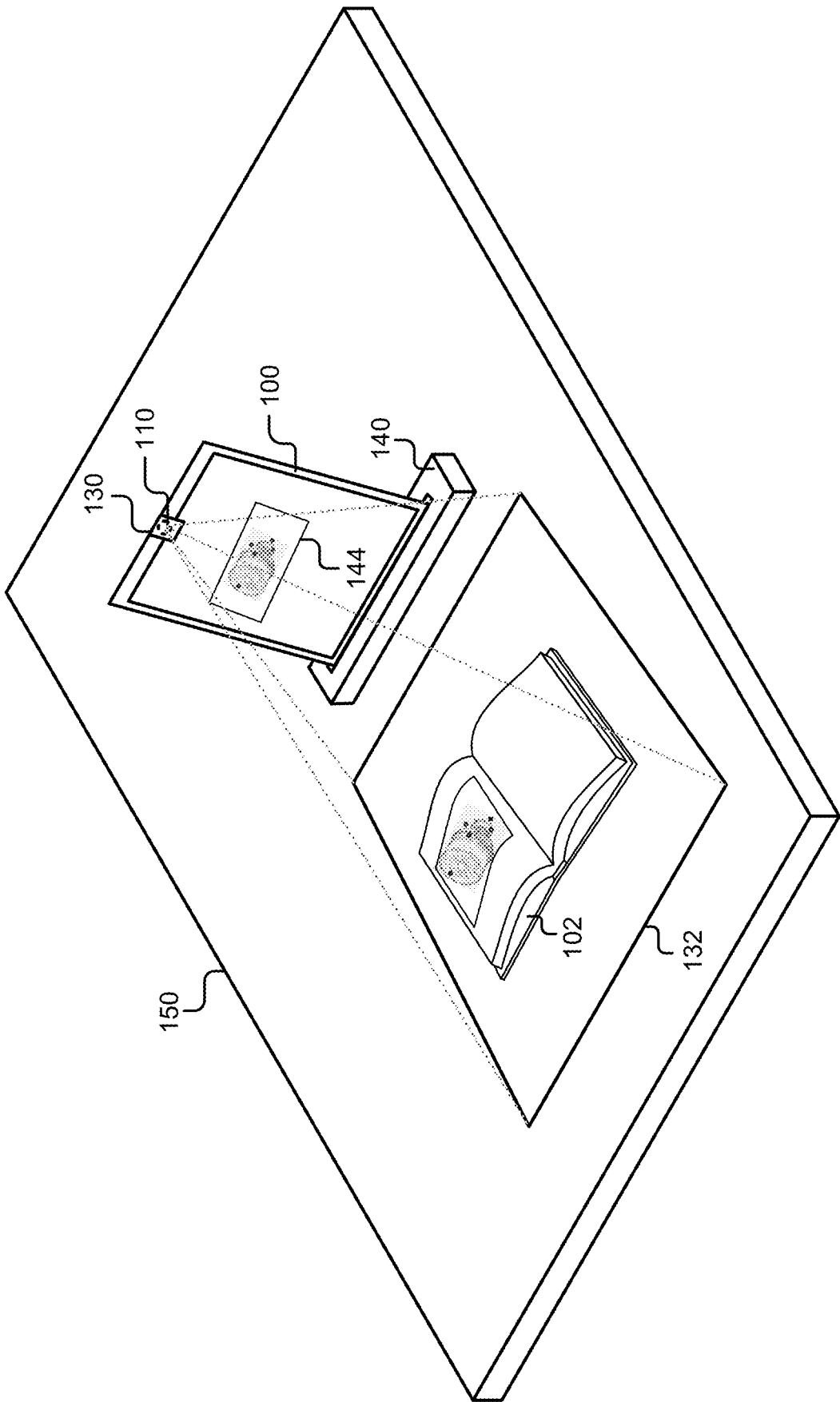


Figure 1C

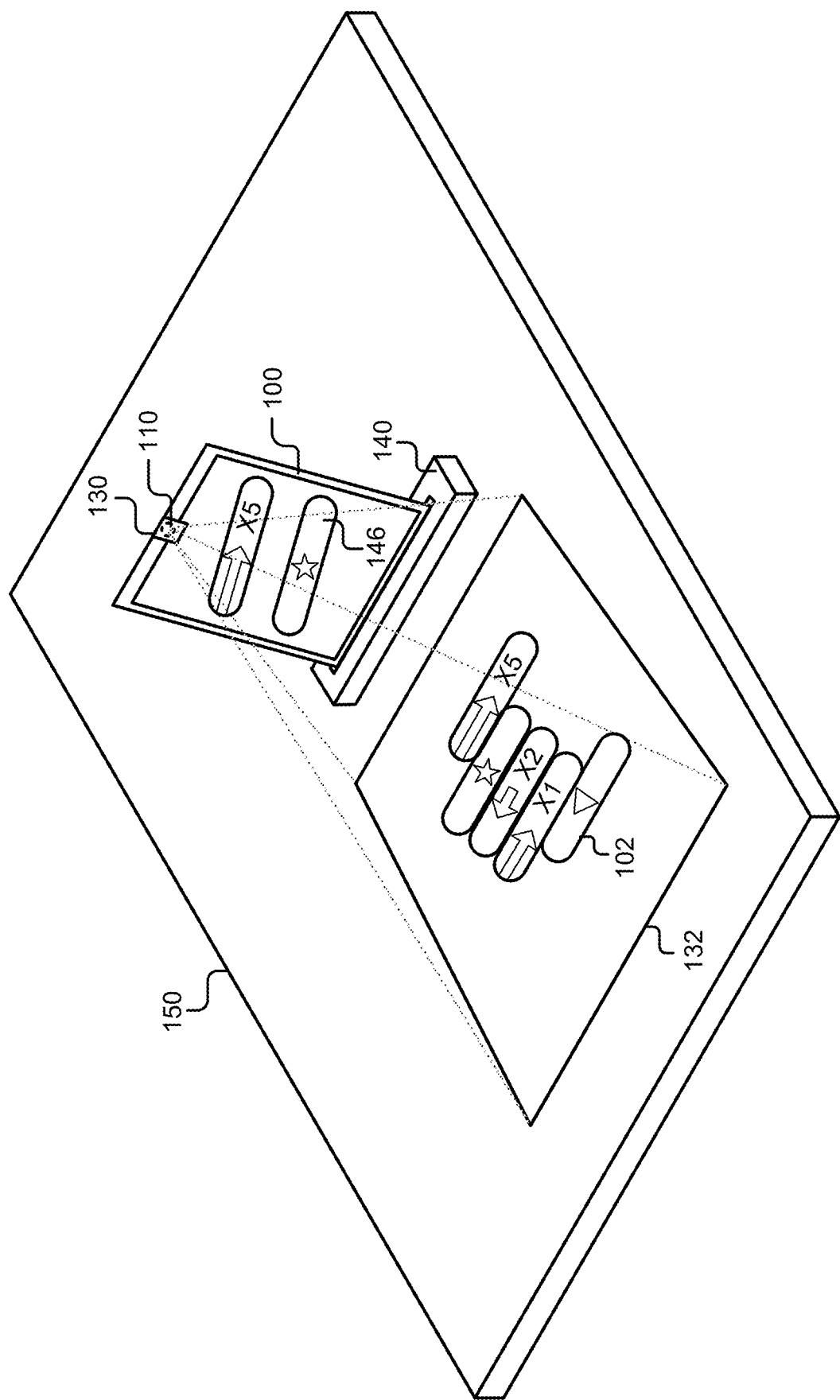


Figure 1D

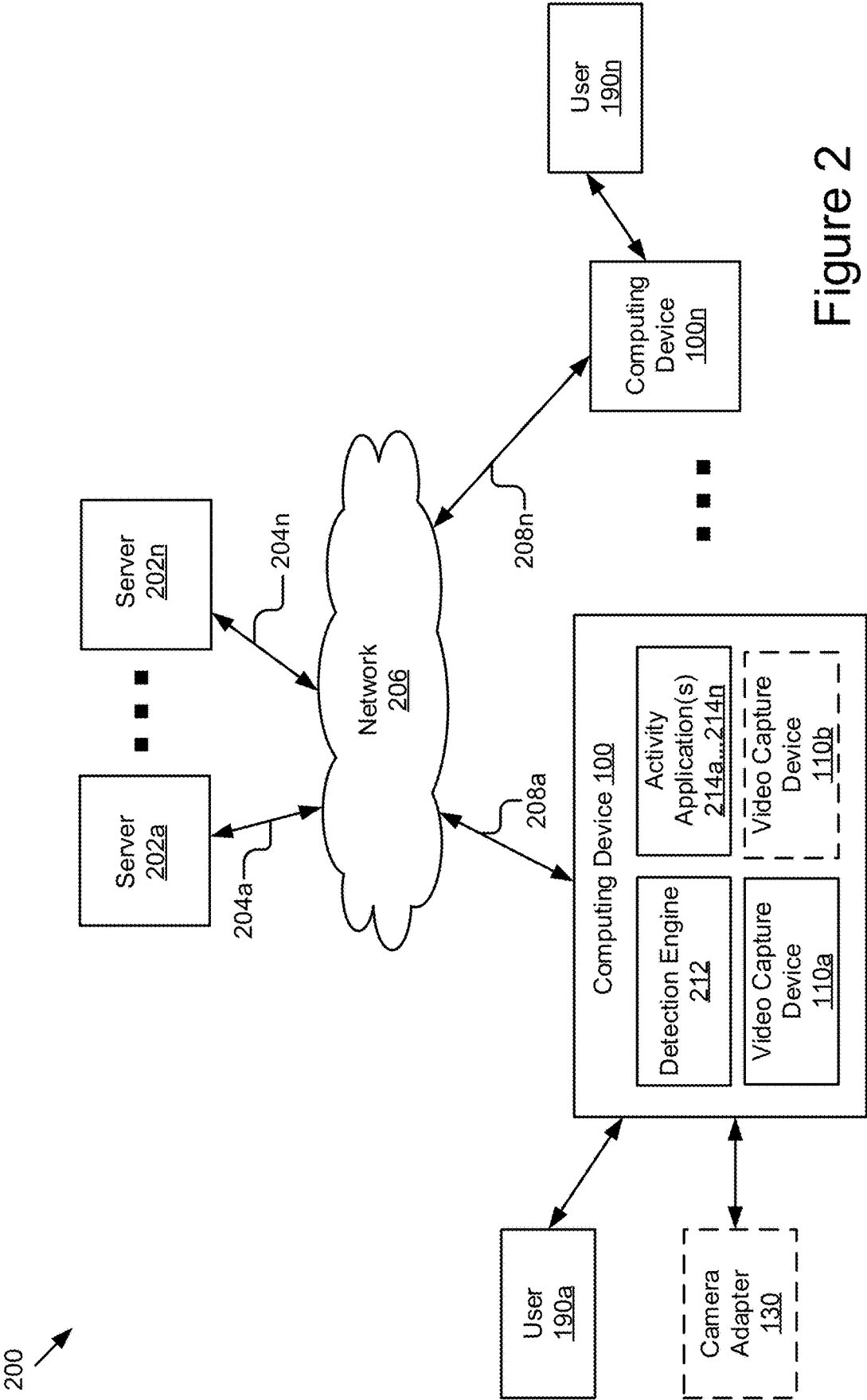


Figure 2

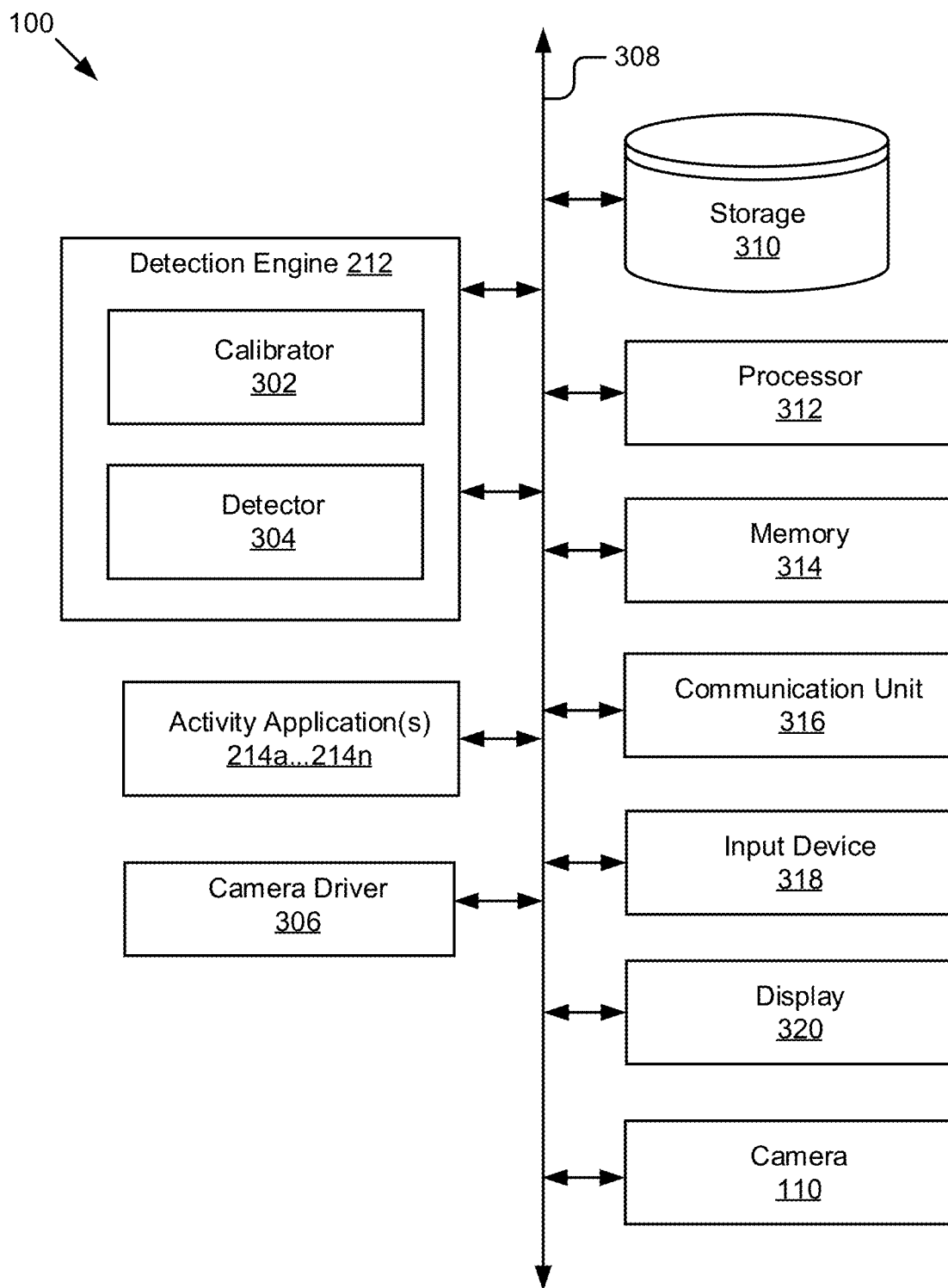


Figure 3

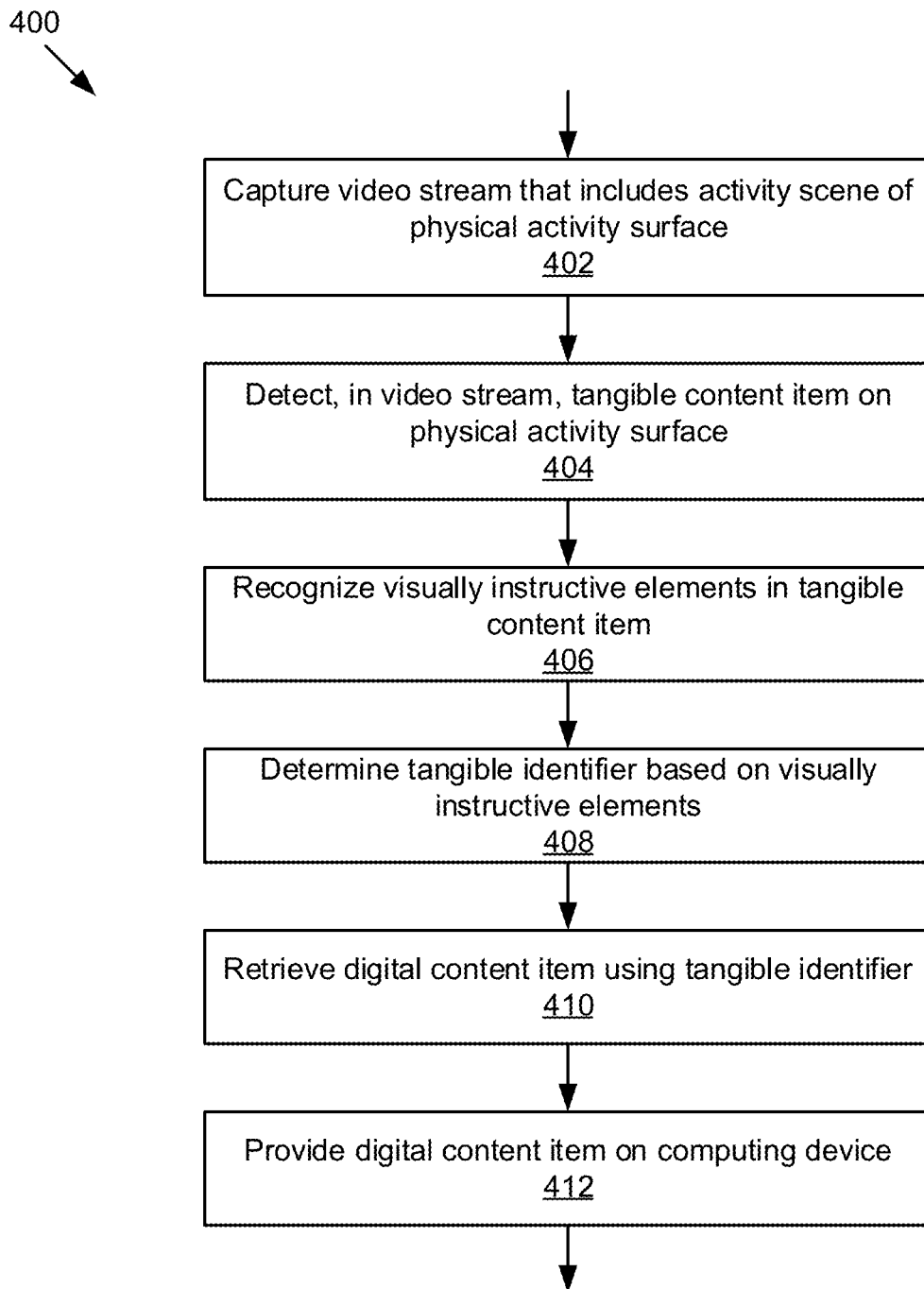


Figure 4

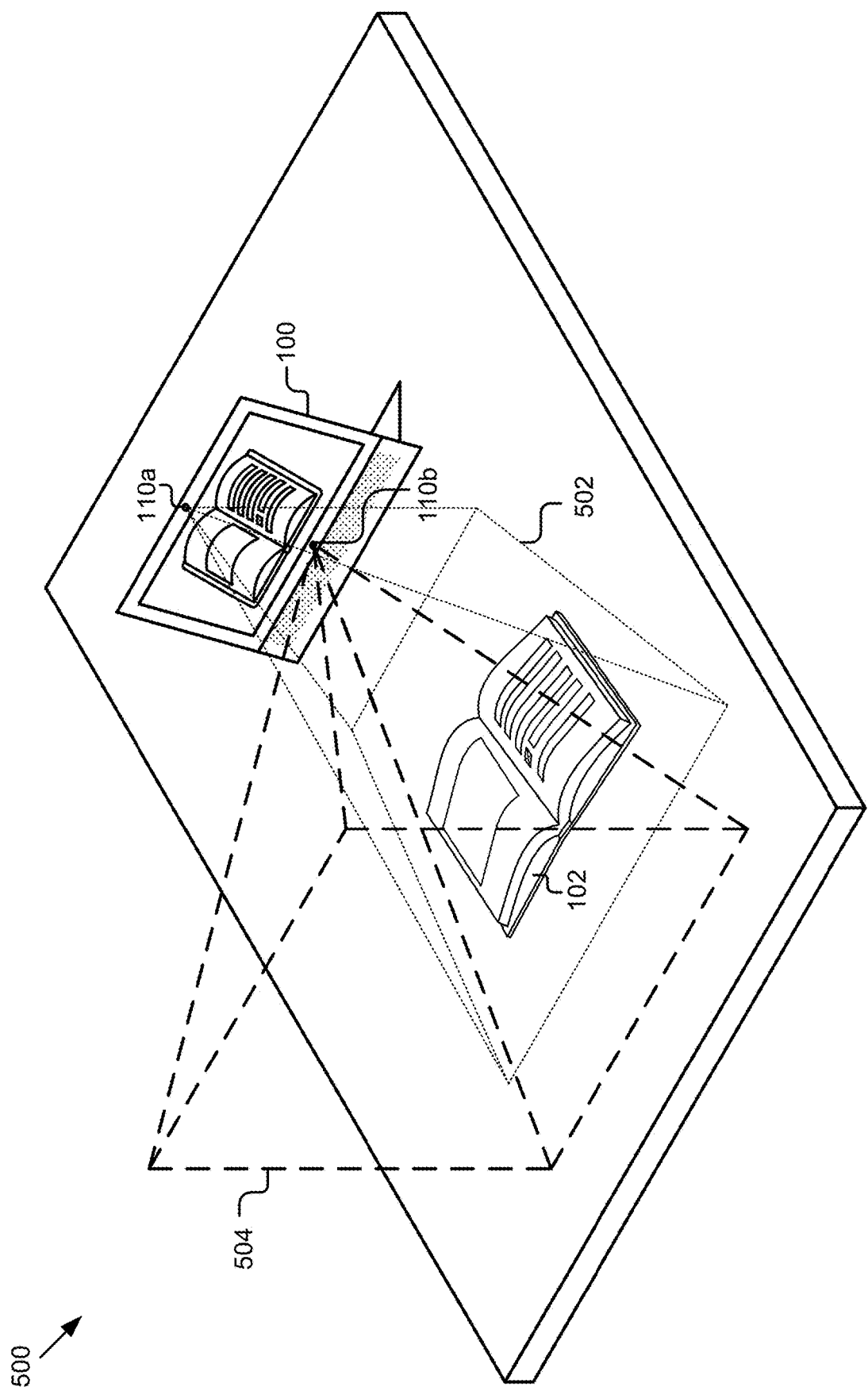


Figure 5

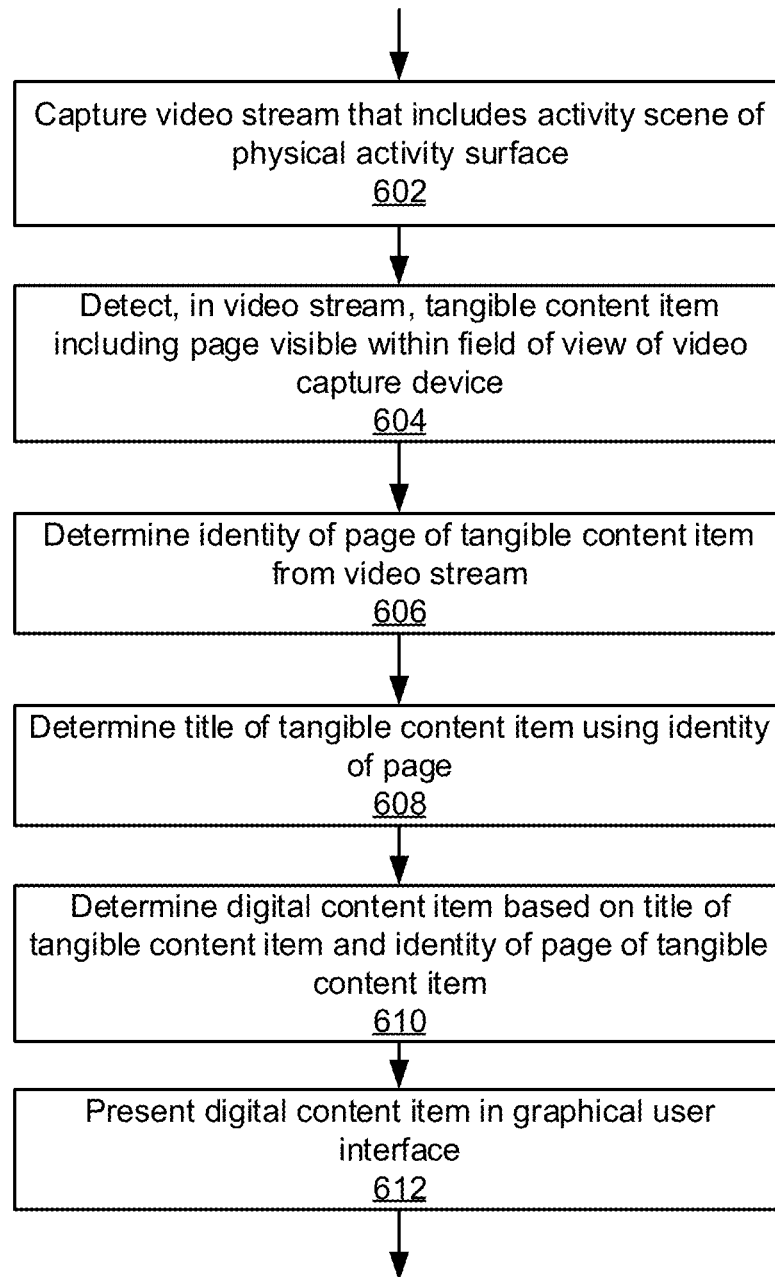
600
↓

Figure 6

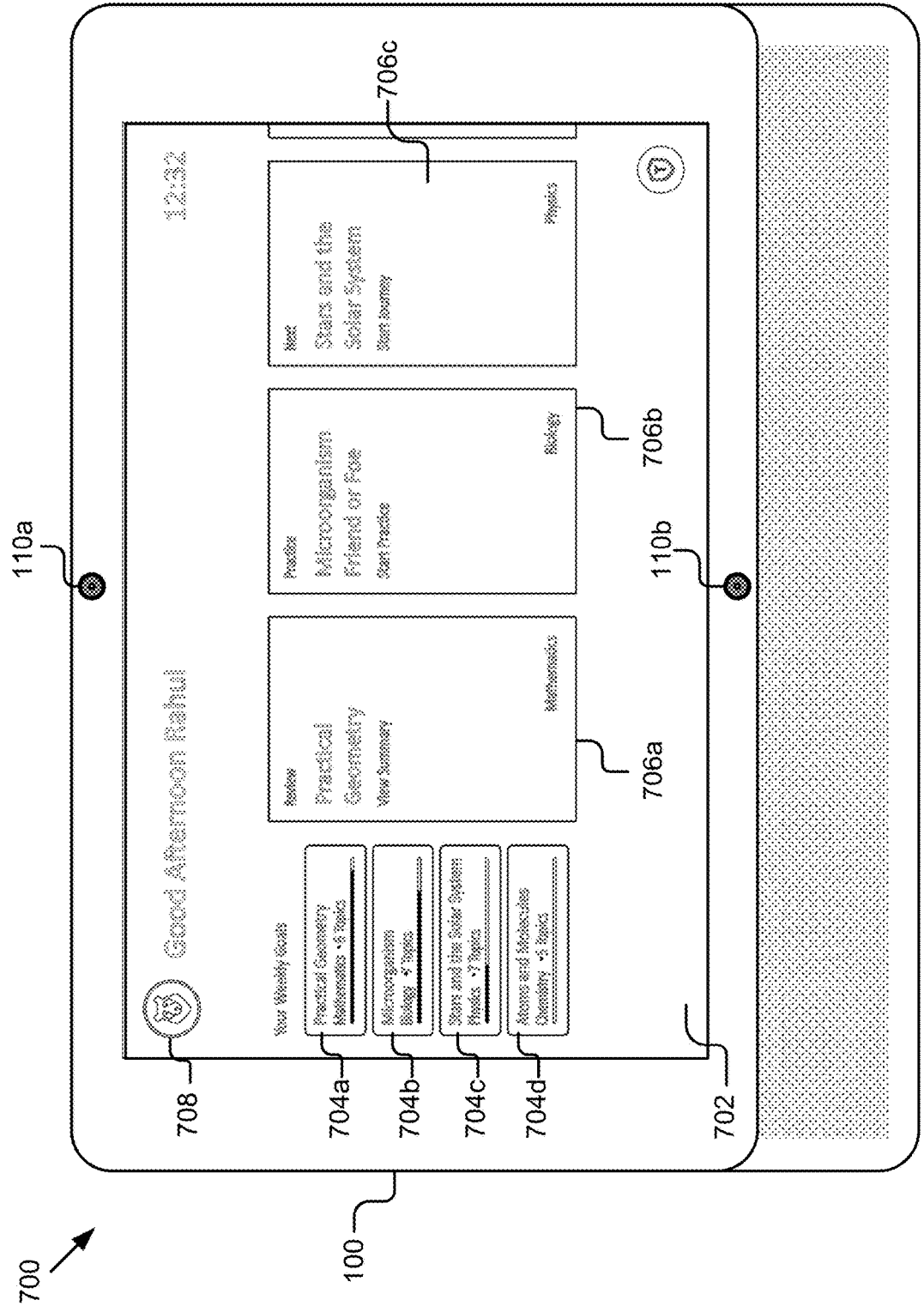


Figure 7A

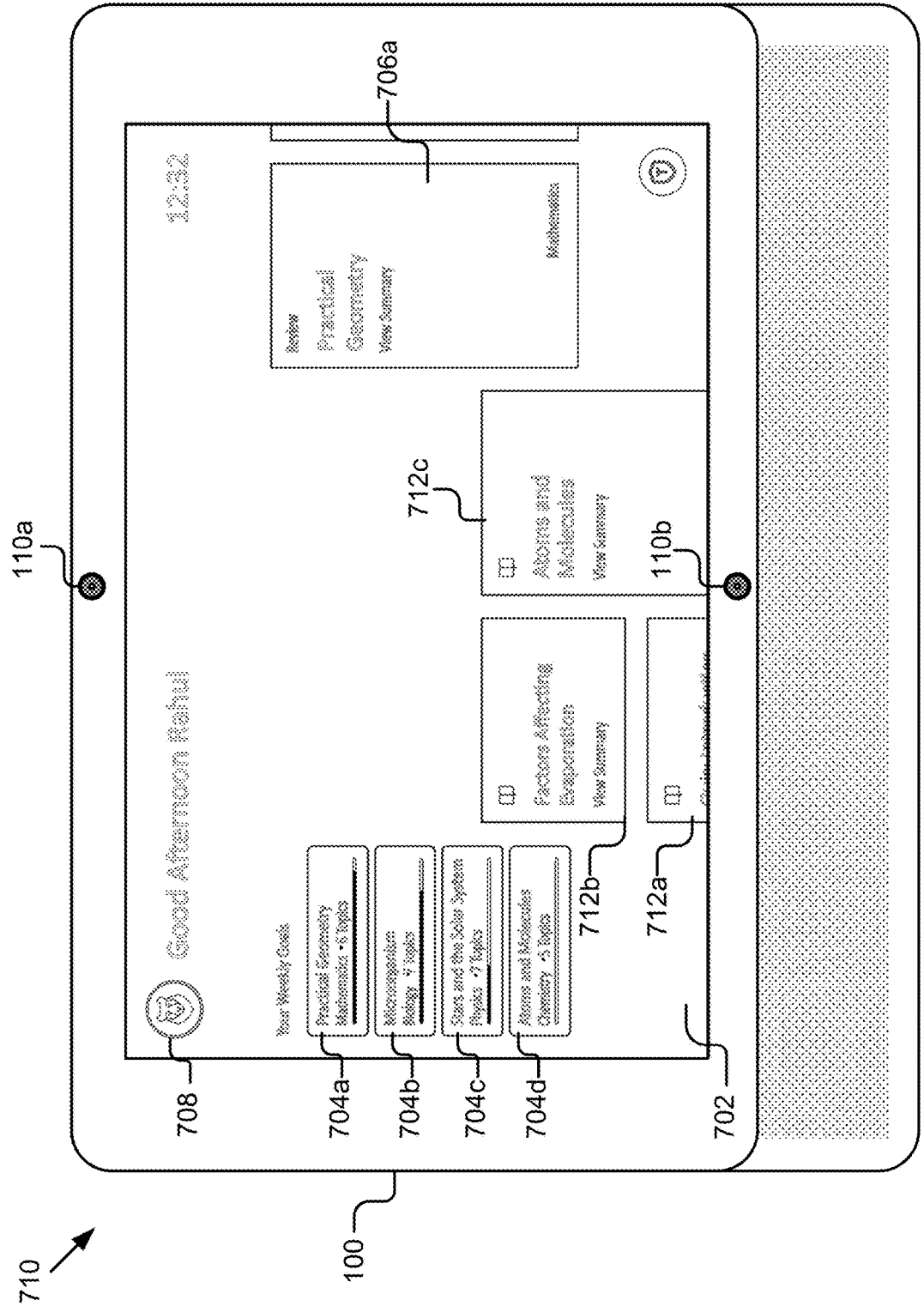


Figure 7B

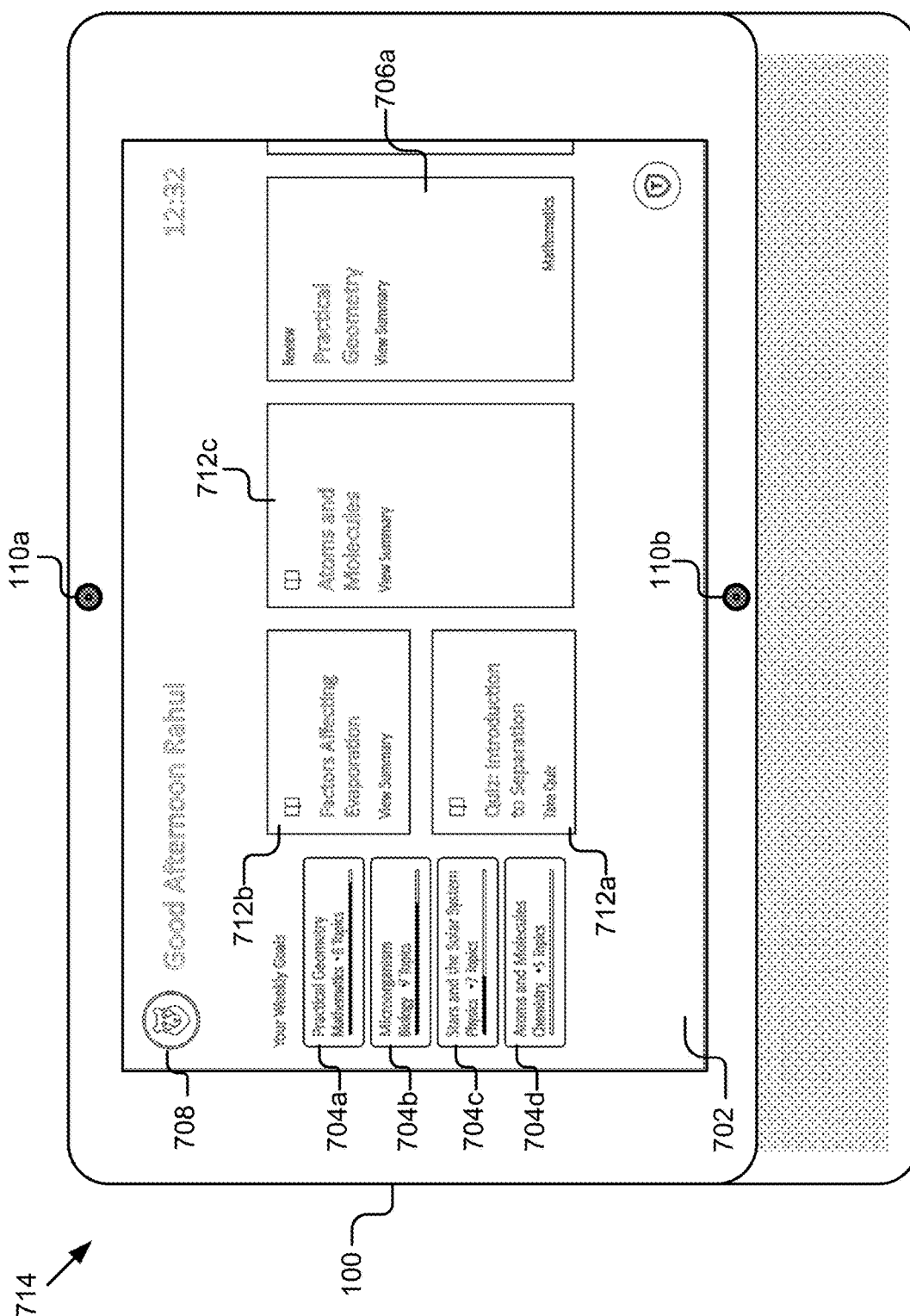


Figure 7C

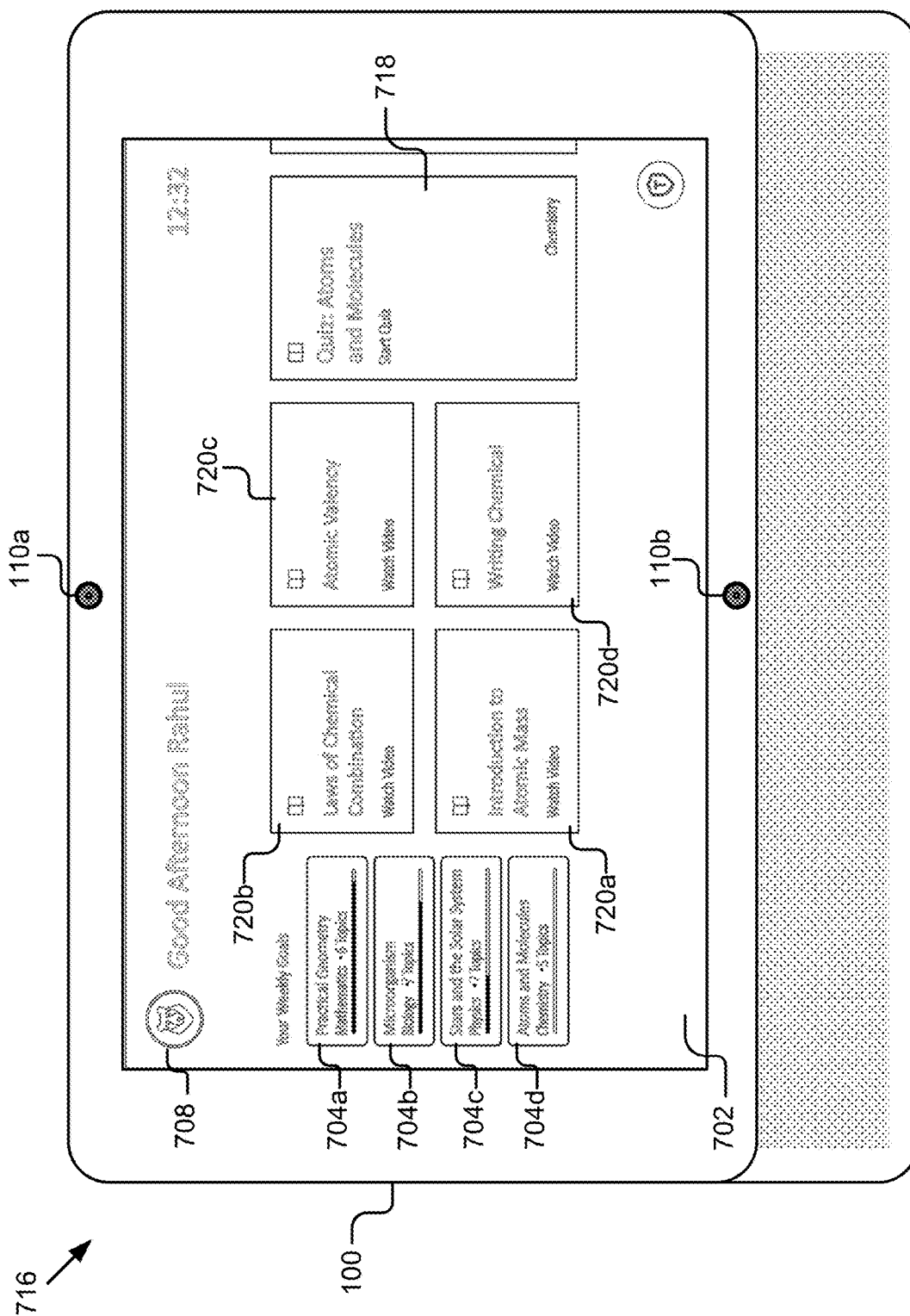


Figure 7D

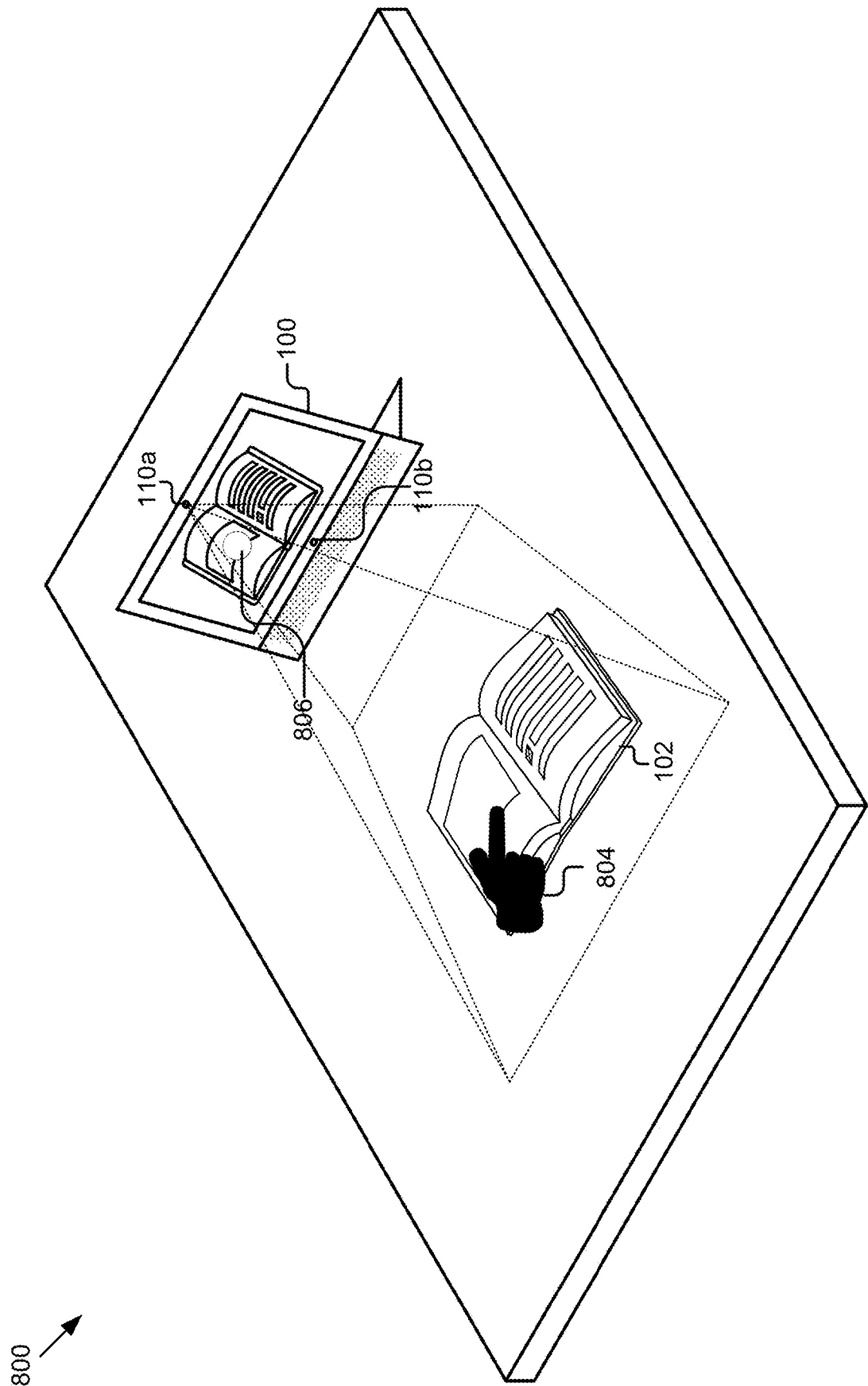


Figure 8

ENHANCING TANGIBLE CONTENT ON PHYSICAL ACTIVITY SURFACE

BACKGROUND

[0001] The present disclosure relates to a computing device.

[0002] Computing devices are often used to access and view digital information. This digital information can be present within an application, viewed in a graphical user interface on an internet webpage, or even created by the user, such as a document or image. This digital information is accessible to the user via the computing device to allow the user to learn from the digital information.

[0003] Physical information has historically been used to gain access to information. Physical information can be included in books, such as textbooks, encyclopedias, and other publications. A user may access this physical information by physical interacting with the book and viewing the information on individual pages of the book.

[0004] If a user wants to access both physical information and digital information that is related to each other, the user has to manually access the physical information, such as by opening and viewing the content of a book and also has to manually retrieve the digital information, such as by accessing a webpage to view the webpages contents. This results in an inconvenient and time intensive process to supplement physical information with digital information. Existing solutions create digital applications that present digital information to the user and may link or reference to physical content, such as a specific student workbook designed to be used with the digital applications, but these existing solutions are not able to intuitively combine the digital information experience and the physical information experience together beyond this limited use of specially designed software applications and specific student workbooks or other specially designed specific physical content.

SUMMARY

[0005] According to one innovative aspect of the subject matter in this disclosure, a method for enhancing tangible content on a physical activity surface is described. In an example implementation, a method includes capturing, using a video capture device of a computing device, a video stream that includes an activity scene of a physical activity surface; detecting in the video stream, using a detector executable on the computing device, a tangible content item on the physical activity surface; recognizing, from the video stream, one or more visually instructive elements in the tangible content item; determining a tangible identifier based on the one or more recognized visually instructive elements in the tangible content item; automatically retrieving a digital content item using the tangible identifier; and providing the digital content item for display in a user interface of the computing device.

[0006] Implementations may include one or more of the following features. The method may include: detecting, in the video stream, a distinct visual element on the tangible content item; and where determining the tangible identifier includes determining the tangible identifier based on the distinct visual element on the tangible content item. The distinct visual element includes one or more of an image, a drawing, a diagram, and a vision marker visible on the tangible content item. Determining the tangible identifier

includes determining the tangible identifier based on one or more of the document title, the page number, the section title, and the section number associated with the tangible content item. Determining the tangible identifier includes determining the tangible identifier based on the one or more characters associated with the user marking. The user marking includes one or more of a highlighting effect and a user note created by a user on the tangible content item. The tangible content item is a first tangible content item: the first tangible content item is included in a tangible object that also includes a second tangible content item; and the method includes: determining that a user interacted with the second tangible content item of the tangible object; recognizing one or more visually instructive elements in the second tangible content item; and determining the tangible identifier based on the one or more visually instructive elements in the first tangible content item and the one or more visually instructive elements in the second tangible content item of the tangible object. Determining the tangible identifier includes determining a first tangible identifier and a second tangible identifier based on the one or more visually instructive elements in the tangible content item; and retrieving the digital content item using the tangible identifier includes: determining a contextual relationship between the first tangible identifier and the second tangible identifier; and retrieving the digital content item based on the contextual relationship between the first tangible identifier and the second tangible identifier. The computing device is placed on a stand situated on the physical activity surface; and a field of view of the video capture device of the computing device is redirected by a camera adapter situated over the video capture device of the computing device. The tangible content item on the physical activity surface includes a first tangible content section and a second tangible content section, the method may include: detecting, using a detector executable on the computing device, an arrangement of the first tangible content section relative to the second tangible content section on the physical activity surface; determining a supplemental digital content item based on the arrangement of the first tangible content section relative to the second tangible content section; and displaying in the user interface on the computing device, the supplemental digital content along with the digital content item. The first tangible content section is a first programming tile and the second tangible content section is a second programming tile, the first programming tile being coupled with the second programming tile to form the tangible content item. The supplemental digital content is an instruction set for how to use the first programming tile and the second programming tile. Displaying the supplemental digital content further may include: automatically displaying a first digital representation of the first programming tile in the user interface; automatically displaying a second digital representation of the second programming tile in the user interface; and automatically displaying an animation of the first digital representation interacting with the second digital representation to provide instructions on how to arrange the first digital programming tile and the second digital programming tile on the activity surface. Providing the digital content item for display in the user interface on the computing device further may include: generating a visualization of the digital content item for display in the user interface; and displaying an animation of the digital content item appearing in the user interface. The animation of the digital

content item depicts a first portion of the visualization of the digital content item extending out from a bottom side of the display screen and a second portion of the visualization of the digital content item appearing as the animation causes the visualization of the digital content item to move into the user interface. The animation of the digital content item corresponds to an appearance of the tangible content item being detected on the physical activity surface. The visualization of the digital content item is selectable by a user and responsive to the digital content item being selected, the digital content item may be displayed.

[0007] The method also includes capturing, using a video capture device of a computing device, a video stream that includes an activity scene of a physical activity surface; detecting in the video stream, using a detector executable on the computing device, a tangible content item on the physical activity surface, the tangible content item including a page visible within a field of view of the video capture device; determining, using a processor of the computing device, an identity of the page of the tangible content item from the video stream; determining, using the processor of the computing device, a concept of the tangible content item using the identity of the page; determining, using the processor of the computing device, a digital content item based on the concept of the tangible content item and the identity of the page of the tangible content item; and presenting, in a display of the computing device, the digital content item in a graphical user interface.

[0008] Implementations may include one or more of the following features. The method where determining the identity of the page of the tangible content item further may include: detecting a page identity anchor on the page; and matching the page identity anchor to a database of page identities. The tangible content item is a book that includes a plurality of pages and each of the plurality of pages has a unique page identity anchor. Determining the digital content item based on the concept of the tangible content item and the identity of the page of the tangible content item further may include: identifying visually instructive elements that are present on the page of the tangible content; determining a subject context based on the visually instructive elements; and identifying the digital content item based on the subject context. Identifying the digital content item based on the subject context further may include: determining a category level of a user based on the concept of the tangible content item and the identity of the page of the tangible content item; accessing a node of a knowledge graph that is related to the subject context, the node of the knowledge graph including links to a plurality of digital content items related to the subject context; and determining a relevant digital content item from the plurality of digital content items based on the category level of the user. Detecting a page identity anchor on the page further may include: performing a matching search to identify the page identity anchor, the matching search including ignoring all text that does not include the page identity anchor. The digital content item is a video related to the identified page of the tangible content item. Presenting the digital content item further may include: embedding an external link to the video related to a topic present on the identified page of the tangible content item; and causing the external link to be executed responsive to an input received from a user. The digital content item is one or more of a video, a question, a quiz, and a worksheet. Determining an identity of the page of the tangible content

item from the video stream further may include: detecting, using a processor of a computing device, a finger of a user present in the video stream of the physical activity surface; determining an area of the tangible content item that is being pointed to by a point of the finger; determining a tangible identifier included within the area of the tangible content item; and determining an identity of the page of the tangible content item based on the tangible identifier.

[0009] The visualization system also includes a computing device positioned on a physical activity surface; a video capture device coupled to the computing device, the video capture device including a field of view directed towards the physical activity surface and capable of capturing a video stream of the physical activity surface; a detector of the computing device configured to detect a tangible content item within the video stream and recognize one or more visually instructive elements in the tangible content item; a processor of the computing device configured to determine a tangible identifier based on the recognized visually instructive elements in the tangible content item and automatically retrieve a digital content item using the tangible identifier; and a display screen coupled to the computing device, the display screen being configured to display the digital content item in a graphical user interface.

[0010] Other implementations of one or more of these aspects and other aspects described in this document include corresponding systems, apparatus, and computer programs, configured to perform the actions of the methods, encoded on computer storage devices. The above and other implementations are advantageous in a number of respects as articulated through this document. Moreover, it should be understood that the language used in the present disclosure has been principally selected for readability and instructional purposes, and not to limit the scope of the subject matter disclosed herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] The disclosure is illustrated by way of example, and not by way of limitation in the figures of the accompanying drawings in which like reference numerals are used to refer to similar elements.

[0012] FIG. 1A illustrates an example tangible content item and an example computing device that captures the tangible content item.

[0013] FIG. 1B illustrates an example tangible content item and an example computing device that provides digital content items related to the tangible content item.

[0014] FIG. 1C is an example configuration for enhancing tangible content on a physical activity surface.

[0015] FIG. 1D is an example configuration for enhancing tangible content on a physical activity surface.

[0016] FIG. 2 is a block diagram illustrating an example computer system for enhancing tangible content on a physical activity surface.

[0017] FIG. 3 is a block diagram illustrating an example computing device.

[0018] FIG. 4 is a flowchart of an example method for enhancing tangible content on a physical activity surface.

[0019] FIG. 5 is an example configuration for enhancing tangible content on a physical activity surface.

[0020] FIG. 6 is a flowchart of an example method for enhancing tangible content on a physical activity surface.

[0021] FIGS. 7A-7D are a graphical representation of an example user interface for enhancing tangible content.

[0022] FIG. 8 is an example configuration for finger detection to enhance tangible content.

DETAILED DESCRIPTION

[0023] The technology described herein is capable of providing data enhancement for a tangible content item that is present on a physical activity surface. As described in detail below, a computing device may capture a video stream of the physical activity surface, and detect the tangible content item in the video stream. The computing device may apply various recognition techniques on the tangible content item to recognize visually instructive elements (e.g., characters, letters, numbers, symbols, etc.), distinct visual elements (e.g., images, drawings, diagrams, etc.), user markings (e.g., notes, highlights), etc., in the tangible content item. The computing device may determine one or more tangible identifiers based on the visually instructive elements, the distinct visual elements, the user markings, etc., that are recognized in the tangible content item, and retrieve digital content items using the one or more visually instructive elements, the distinct visual elements, the user markings, etc., and provide the digital content items to a user via the computing device. Thus, the present technology can provide on the computing device the digital content items that are related to the tangible content item physically present on the physical activity surface, thereby effectively enhancing the content data provided to a user and improving user experience. In some implementations, this process may happen automatically when a tangible content item is detected and without any prompting or request by a user.

[0024] The technology described herein is advantageous, because it can facilitate the user in obtaining additional data relevant to the content of interest, especially in the context of education. As an example, the tangible content item may be a text book that is placed on the physical activity surface and opened to a page presenting content about airplanes. In this example, the computing device may provide digital content items (e.g., images, video, online articles, research papers, etc.) related to the airplanes on the computing device. Thus, the user may be provided with additional information about various aspects of the airplanes that is potentially useful or interesting but the user may not be aware or may not think of. Furthermore, this presentation of the digital content related to airplanes may happen automatically and allow for the user to supplement their learning without having to do anything other than open the page of the book.

[0025] FIGS. 1A-1D illustrate example configurations for enhancing a tangible content item on a physical activity surface. As depicted in FIG. 1A, the configuration may include a tangible content item 102 situated on a physical activity surface 150 and a computing device 100 including a video capture device 110. The video capture device 110 may be adapted to capture an image or a video stream that includes an activity scene 132 of the physical activity surface 150.

[0026] In some embodiments, the tangible content item 102 may include one or more portions that are exposed and visible to the video capture device 110 of the computing device 100, and thus the video capture device 110 can capture the visible portions of the tangible content item 102 in an image or in a video frame of a video stream. As an example, the tangible content item 102 may be a book opened at one or more pages, and the portions of the tangible

content item 102 may include the one or more pages of the book that are visually exposed to the video capture device 110 of the computing device 100. In another example, the tangible content item 102 may be a collection of paper sheets, and the visible portions of the tangible content item 102 may be one or more paper sheets in the collection that are visually exposed to the video capture device 110 of the computing device 100. In another example, the tangible content item 102 may be an electronic device (e.g., tablet, mobile phone, etc.) displaying images and/or text in a graphical user interface of a display screen resting on the physical activity surface 150, and the visible portions of the tangible content item 102 may be the displayed content items that are visually exposed to the video capture device 110 of the computing device 100. In another example, the tangible content item 102 may be a content item that is projected by a projector onto the physical activity surface 150, and the visible portions of the tangible content item 102 may be the projected content item that is visually exposed to the video capture device 110 of the computing device 100. Thus, the tangible content item 102 may be the content item that is displayed or otherwise physically present on the physical activity surface 150 and can be captured by the video capture device 110 of the computing device 100.

[0027] FIG. 1B illustrates an example configuration in which the computing device 100 may provide a digital content item based on the content detected in the tangible content item 102. As shown in the example, the tangible content item 102 may be a book that includes information about airplanes. The book may include a first portion 104 that include text related to the content of the tangible content item 102, a second portion 106 that includes a graphic related to the content of the tangible content item 102, and a third portion 108 that includes another graphic related to the content of the tangible content item 102. The book may be open to a page that is visible within a field of view of a video capture device 110 of the computing device and the video capture device may capture an image and/or video stream of the open page including the first portion 104, second portion 106, and/or third portion 108. A detector 304 of the computing device 100 may perform image recognition processes to detect one or more visually instructive elements, such as the text, the graphic, or the additional graphic included in the portions of the open page of the tangible content item 102.

[0028] In further implementations, the detector 304 may detect an identity of the open page, such as by detecting a page number and/or a page identity anchor visible within the field of view of the camera 110. The detector 304 may then use the identity of the open page to identify concept of the tangible content item 102. The concept of the tangible content item 102 may be a topic that the tangible content item 102 covers, such as a title of the book, a title of the chapter the open page is included within, a course subject of the book, a specific topic being presented on the open page, etc. The concept of the tangible content item 102 may be based on the context of the various information being presented, such as a relationship between the content in the first portion 104, the second portion 106, and/or the third portion 108 of the tangible content item 102.

[0029] The activity application 214 of the computing device 100 may then identify one or more digital content items based on the concept of the tangible content item 102, the identity of the open page, and/or the topic being pre-

sented on the open page. The digital content item may be supplemental content related to the content present on the open page and may enhance the experience of the user as the interact with the tangible content item 102. In the example shown in FIG. 1B, the digital content item 138 may include additional text related to “aircraft structure,” “principles of aerodynamics,” and/or “what happens if all engines fail on a plane” as shown. These additional texts may be selectable links embedded into the GUI of the display screen and when selected by the user may direct the display screen to the link and display the contents of the resulting selection. In further implementations, the GUI of the display screen may automatically display the content in detail without requiring the user to do any addition interactions with the digital content item 138. In further implementations, the digital content item 138 may also include videos and/or images of airplanes that may be automatically displayed or may be presented as selectable links for the user to interact with. By identifying the concept being presented in the tangible content item 102 and then presenting related digital content items 138 to the user, the user gains an enhanced experience of being presented both physical and digital information that is related without having to navigate to the information themselves. This enhances the educational opportunities by reducing distractions caused by having to find the additional information on a student own. It further enhances educational opportunities by being able to specifically provide curated content directly related to a specific page a student is working on that may be causing confusion or needing additional guidance.

[0030] FIG. 1C illustrates a configuration in which the computing device 100 may be part of a visualization system to provide data enhancement for the tangible content item 102 present on the physical activity surface. As depicted, the computing device 100 may be placed on the stand 140 situated on the physical activity surface 150. The computing device 100 may include the video capture device 110 with the camera adapter 130 situated over the video capture device 110. In some embodiments, the camera adapter 130 may redirect the field of view of the video capture device 110, thereby enabling the video capture device 110 to capture the video stream including the activity scene 132 of the physical activity surface 150 located in front of the display of the computing device as depicted in FIG. 1C. As a result of this configuration, the video capture device 110 may capture the video stream of the activity scene 132 that includes the tangible object 102 on the physical activity surface 150 without the user manually capturing the video stream using the computing device 100. In some embodiments, as the user continues to interact with various tangible content items 102, the activity application 214 implemented on the computing device 100 may perform the method 400 in FIG. 4 for each tangible content item 104 with which the user is currently interacting. As a result, the computing device 100 may automatically present the digital content items that are relevant to the tangible content item 102 in real-time. Thus, when the user interacts with a different tangible content item 102 in the physical activity surface 150, the digital content items being displayed to the user on the computing device 100 may be automatically updated accordingly.

[0031] In some implementations, the visualization system depicted in FIG. 1C may be an integrated system where the stand 140, computing device 100, and video capture device

110 are all integrated into housing of the computing device 100. In some implementations, the video capture device 110 may be positioned so that the field of view encompass the physical activity surface without redirecting the field of view using the camera adapter 130. In further implementations, multiple video capture devices 110 may be included within the computing device 100 and each of the different video capture devices 110 may be directed to capture a different field of view. For example, a first video capture device 110a may be directed downward towards the physical activity surface 150 and a second video capture device 110b may be directed outward towards a user facing the computing device 100 in order to capture a video stream that includes the user, such as the user’s face and hand motions.

[0032] FIG. 1D illustrates a configuration in which the computing device 100 may be part of a visualization system to provide data enhancement for the tangible content item 102 present on the physical activity surface 150. In the example, the tangible content item 102 may be one more tangible interface objects that are arranged on the physical activity surface 150. The user may create a specific arrangement of the tangible interface objects and the detector 304 may be able to detect the arrangement over time, including when specific tangible interface objects were present in a first image from a video stream compared to a later image from a video stream. By detecting the arrangement and manipulation of the tangible interface objects, the detector 304 may be able to provide to the activity application 214 various events, such as when an arrangement is started being built, completed, as well as if the arrangement matches an expected arrangement of the activity application 214 or if the arrangement is different than expected (e.g., the arrangement may be incorrect). The activity application 214 may then be able to retrieve and display one or more digital content items 146 based on the arrangement and/or manipulation of the tangible content items on the physical activity surface.

[0033] In one example, the tangible interface objects may be programming tiles that include one or more visually instructive elements that can be identified by the detector 304. Based on the visually instructive elements, the detector 304 may be able to determine a tangible identifier of each of the programming tiles arranged on the physical activity surface 150. As a user arranges the programming tiles to depict a series of commands to be executed step by step based on the way the programming tiles are arranged, the activity application 214 may search for related digital content items 146 that are conceptually related to the arrangement of the programming tiles. The activity application 214 may then present the digital content items 146 to the user to enhance how the user is arranging the programming tiles.

[0034] In the example in FIG. 1D, the activity application 214 may identify digital content items 146 that are an instruction set for how to use the programming tiles together. The user may be able to view an animation of what happens when two specific programming tiles are arranged in a specific way and enhance their learning of the programming tiles and commands. By providing this instruction automatically, without requiring the user to find and display the digital content items 146, the user may be able quickly resolve questions or confusion around how to use various objects or items, as well as receiving supplemental material on a concept related to what they are interacting with in the physical activity scene. In some implementations, these

digital content items may be instruction sets, such as animations on how to arrange the different programming tiles, tasks that can be presented for the user to complete, such as specific arrangements based on the detected tangible interface objects, videos of other users interacting with the tangible interface objects, a product description on a web page listing what items are included in a set of tangible interface objects the user is interacting with, etc.

[0035] FIG. 2 is a block diagram illustrating an example computer system 200 that is capable of providing data enhancement for tangible content present on a physical activity surface. As depicted, the system 200 may include computing devices 100a . . . 100n and servers 202a . . . 202n communicatively coupled via a network 206. In FIG. 2 and the remaining figures, a letter after a reference number, e.g., “100a”, represents a reference to the element having that particular reference number. A reference number in the text without a following letter, e.g., “100”, represents a general reference to instances of the element bearing that reference number. It should be understood that the system 200 depicted in FIG. 2 is provided by way of example and that the system 200 and/or further systems contemplated by this present disclosure may include additional and/or fewer components, may combine components and/or divide one or more of the components into additional components, etc. For example, the system 100 may include any number of servers 202, computing devices 100, or networks 206. As depicted in FIG. 2, the computing device 100 may be coupled to the network 206 via the signal line 208 and the server 202 may be coupled to the network 206 via the signal line 204. The computing device 100 may be accessed by user 190.

[0036] The network 206 may include any number of networks and/or network types. For example, the network 206 may include, but is not limited to, one or more local area networks (LANs), wide area networks (WANs) (e.g., the Internet), virtual private networks (VPNs), mobile (cellular) networks, wireless wide area network (WWANs), WiMAX® networks, Bluetooth® communication networks, peer-to-peer networks, other interconnected data paths across which multiple devices may communicate, various combinations thereof, etc.

[0037] The computing device 100 may be a computing device that has data processing and communication capabilities. In some embodiments, the computing device 100 may include a processor (e.g., virtual, physical, etc.), a memory, a power source, a network interface, and/or other software and/or hardware components, such as front and/or rear facing cameras, display screen, graphics processor, wireless transceivers, keyboard, firmware, operating systems, drivers, various physical connection interfaces (e.g., USB, HDMI, etc.). In some embodiments, the computing device 100 may be coupled to and communicate with one another and with other entities of the system 200 via the network 206 using a wireless and/or wired connection. As discussed elsewhere herein, the system 200 may include any number of computing devices 100 and the computing devices 100 may be the same or different types of devices (e.g., tablets, mobile phones, desktop computers, laptop computers, etc.).

[0038] As depicted in FIG. 2, the computing device 100 may include the video capture device 110, a detection engine 212, and one or more activity applications 214. In some implementations, the computing device 100 and/or the video capture device 110 may be equipped with the camera adapter

130 as discussed elsewhere herein, while in other embodiments, the video capture device 110 may be directed to the appropriate field of view without the camera adapter 130. In some embodiments, the camera adapter 130 may include one or more optical elements (e.g., mirrors and/or lenses) to adapt the standard field of view of the video capture device 110 of the computing device 100 to capture substantially and only the activity scene 132 of the physical activity surface 150 in front of the display of the computing device 100, although other implementations are also possible and contemplated. To adapt the field of view of the video capture device 110, the mirrors and/or lenses of the camera adapter 130 may be positioned at an angle to redirect and/or modify the light being reflected from the physical activity surface 150 into the video capture device 110. In some embodiments, the camera adapter 130 may include a slot adapted to receive an edge of the computing device 100 and retain (e.g., secure, grip, etc.) the camera adapter 130 on the edge of the computing device 100. In some embodiments, the camera adapter 130 may be positioned over the video capture device 110 to direct the field of view of the video capture device 110 toward the physical activity surface 150.

[0039] In some embodiments, the detection engine 212 may detect and/or recognize tangible objects located in the activity scene 132 of the physical activity surface 150, and cooperate with the activity application(s) 214 to provide the user with a virtual experience that incorporates in real-time the tangible objects and the user manipulation of the tangible objects in the physical environment. As an example, the detection engine 212 may detect a tangible object located in the activity scene 132 of the physical activity surface 150 and/or recognize visually instructive elements, distinct visual elements, user markings in the tangible content item of the tangible object, and cooperate with the activity application(s) 214 to provide the user with digital content items that are relevant to the tangible content item on the physical activity surface 150. In another example, the detection engine 212 may process the video stream captured by the video capture device 110 to detect and recognize a tangible object created by the user on the activity scene 132. The activity application 214 may generate a visualization of the tangible object created by the user, and display to the user a virtual scene in which an animated character may interact with the visualization of the tangible object. The components and operations of the detection engine 212 and the activity application 214 are described in details with reference to at least FIGS. 3 and 4.

[0040] The server 202 may include one or more computing devices that have data processing, storing, and communication capabilities. In some embodiments, the server 202 may include one or more hardware servers, server arrays, storage devices and/or storage systems, etc. In some embodiments, the server 202 may be a centralized, distributed and/or a cloud-based server. In some embodiments, the server 202 may include one or more virtual servers that operate in a host server environment and access the physical hardware of the host server (e.g., processor, memory, storage, network interfaces, etc.) via an abstraction layer (e.g., a virtual machine manager).

[0041] The server 202 may include software applications operable by one or more processors of the server 202 to provide various computing functionalities, services, and/or resources, and to send and receive data to and from the computing devices 100. For example, the software applica-

tions may provide the functionalities of internet searching, social networking, web-based email, blogging, micro-blogging, photo management, video/music/multimedia hosting/sharing/distribution, business services, news and media distribution, user account management, or any combination thereof. It should be understood that the server 202 may also provide other network-accessible services.

[0042] In some embodiments, the server 202 may include a search engine capable of retrieving results that match one or more search criteria from a data store. As an example, the search criteria may include an image and the search engine may compare the image to product images in its data store (not shown) to identify a product that matches the image. In another example, the detection engine 212 and/or the storage 310 (e.g., see FIG. 3) may request the search engine to provide information that matches a physical drawing, an image, and/or a tangible object extracted from a video stream.

[0043] It should be understood that the system 200 illustrated in FIG. 2 is provided by way of example, and that a variety of different system environments and configurations are contemplated and are within the scope of the present disclosure. For example, various functionalities may be moved from a server to a client, or vice versa and some implementations may include additional or fewer computing devices, services, and/or networks, and may implement various client or server-side functionalities. In addition, various entities of the system 200 may be integrated into a single computing device or system or divided into additional computing devices or systems, etc.

[0044] FIG. 3 is a block diagram of an example computing device 100. As depicted, the computing device 100 may include a processor 312, a memory 314, a communication unit 316, an input device 318, a display 320, and the video capture device 110 communicatively coupled by a communications bus 308. It should be understood that the computing device 100 is not limited to such and may include other components, including, for example, those discussed with reference to the computing devices 100 in FIGS. 1A-1C, 2, 4, and 5.

[0045] The processor 312 may execute software instructions by performing various input/output, logical, and/or mathematical operations. The processor 312 may have various computing architectures to process data signals including, for example, a complex instruction set computer (CISC) architecture, a reduced instruction set computer (RISC) architecture, and/or an architecture implementing a combination of instruction sets. The processor 312 may be physical and/or virtual, and may include a single core or plurality of processing units and/or cores.

[0046] The memory 314 may be a non-transitory computer-readable medium that is configured to store and provide access to data to other components of the computing device 100. In some embodiments, the memory 314 may store instructions and/or data that are executable by the processor 312. For example, the memory 314 may store the detection engine 212, the activity applications 214, and the camera driver 306. The memory 314 may also store other instructions and data, including, for example, an operating system, hardware drivers, other software applications, data, etc. The memory 314 may be coupled to the bus 308 for communication with the processor 312 and other components of the computing device 100.

[0047] The communication unit 316 may include one or more interface devices (I/F) for wired and/or wireless connectivity with the network 206 and/or other devices. In some embodiments, the communication unit 316 may include transceivers for sending and receiving wireless signals. For example, the communication unit 316 may include radio transceivers for communication with the network 206 and for communication with nearby devices using close-proximity connectivity (e.g., Bluetooth®, NFC, etc.). In some embodiments, the communication unit 316 may include ports for wired connectivity with other devices. For example, the communication unit 316 may include a CAT-5 interface, Thunderbolt™ interface, FireWire™ interface, USB interface, etc.

[0048] The display 320 may display electronic images and data output by the computing device 100 for presentation to the user 190. The display 320 may include any display device, monitor or screen, including, for example, an organic light-emitting diode (OLED) display, a liquid crystal display (LCD), etc. In some embodiments, the display 320 may be a touch-screen display capable of receiving input from one or more fingers of the user 190. For example, the display 320 may be a capacitive touch-screen display capable of detecting and interpreting multiple points of contact with the display surface. In some embodiments, the computing device 100 may include a graphic adapter (not shown) for rendering and outputting the images and data for presentation on display 320. The graphic adapter may be a separate processing device including a separate processor and memory (not shown) or may be integrated with the processor 312 and memory 314.

[0049] The input device 318 may include any device for inputting information into the computing device 100. In some embodiments, the input device 318 may include one or more peripheral devices. For example, the input device 318 may include a keyboard (e.g., a QWERTY keyboard), a pointing device (e.g., a mouse or touchpad), a microphone, a camera, etc. In some implementations, the input device 318 may include a touch-screen display capable of receiving input from the one or more fingers of the user 190. In some embodiments, the functionality of the input device 318 and the display 320 may be integrated, and the user 190 may interact with the computing device 100 by contacting a surface of the display 320 using one or more fingers. For example, the user 190 may interact with an emulated keyboard (e.g., soft keyboard or virtual keyboard) displayed on the touch-screen display 320 by contacting the display 320 in the keyboard regions using his or her fingers.

[0050] The detection engine 212 may include a calibrator 302 and a detector 304. The components 212, 302, and 304 may be communicatively coupled to one another and/or to other components 214, 306, 310, 312, 314, 316, 318, 320, and/or 110 of the computing device 100 by the bus 308 and/or the processor 312. In some embodiments, the components 212, 302, and 304 may be sets of instructions executable by the processor 312 to provide their functionality. In some embodiments, the components 212, 302, and 304 may be stored in the memory 314 of the computing device 100 and may be accessible and executable by the processor 312 to provide their functionality. In any of the foregoing implementations, these components 212, 302, and 304 may be adapted for cooperation and communication with the processor 312 and other components of the computing device 100.

[0051] The calibrator 302 includes software and/or logic for performing image calibration on the video stream captured by the video capture device 110. In some embodiments, to perform the image calibration, the calibrator 302 may calibrate the images in the video stream to adapt to the capture position of the video capture device 110, which may be dependent on the configuration of the stand 140 on which the computing device 100 is situated. When the computing device 100 is placed into the stand 140, the stand 140 may position the video capture device 110 of the computing device 100 at a camera height relative to the physical activity surface and a tilt angle relative to a horizontal line. Capturing the video stream from this camera position may cause distortion effects on the video stream. Therefore, the calibrator 302 may adjust one or more operation parameters of the video capture device 110 to compensate for these distortion effects. Examples of the operation parameters being adjusted include, but are not limited to, focus, exposure, white balance, aperture, f-stop, image compression, ISO, depth of field, noise reduction, focal length, etc. Performing image calibration on the video stream is advantageous, because it can optimize the images of the video stream to accurately detect the objects depicted therein, and thus the operations of the activity applications 214 based on the objects detected in the video stream can be significantly improved.

[0052] In some embodiments, the calibrator 302 may also calibrate the images to compensate for the characteristics of the activity surface (e.g., size, angle, topography, etc.). For example, the calibrator 302 may perform the image calibration to account for the discontinuities and/or the non-uniformities of the activity surface, thereby enabling accurate detection of objects on the activity surface when the stand 140 and the computing device 100 are set up on various activity surfaces (e.g., bumpy surface, beds, tables, whiteboards, etc.). In some embodiments, the calibrator 302 may calibrate the images to compensate for optical effect caused by the camera adapter 130 and/or the optical elements of the video capture device 110. In some embodiments, the calibrator 302 may also calibrate the video capture device 110 to split its field of view into multiple portions with the user being included in one portion of the field of view and the activity surface being included in another portion of the field of view of the video capture device 110.

[0053] In some embodiments, different types of computing device 100 may use different types of video capture devices 110 that have different camera specifications. For example, the tablets made by Apple may use a different type of video capture device 110 from the tablets made by Amazon. In some embodiments, the calibrator 302 may use the camera information specific to the video capture device 110 of the computing device 100 to calibrate the video stream captured by the video capture device 110 (e.g., focal length, distance between the video capture device 110 to the bottom edge of the computing device 100, etc.). The calibrator 302 may also use the camera position at which the video capture device 110 is located to perform the image calibration.

[0054] The detector 304 includes software and/or logic for processing the video stream captured by the video capture device 110 to detect the tangible objects present in the activity surface in the video stream. In some embodiments, to detect an object in the video stream, the detector 304 may

analyze the images of the video stream to determine line segments, and determine the object that has the contour matching the line segments using the object data in the storage 310. In some embodiments, the detector 304 may provide the tangible objects detected in the video stream to the activity applications 214. In some embodiments, the detector 304 may store the tangible objects detected in the video stream in the storage 310 for retrieval by other components. In some embodiments, the detector 304 may determine whether the line segments and/or the object associated with the line segments can be identified in the video stream, and instruct the calibrator 302 to calibrate the images of the video stream accordingly. In some embodiments, the detector 304 may perform character recognition (such as OCR as described elsewhere herein) to identify visually instructive elements, such as word or reference marks that may be used to determine tangible identifiers of the tangible content item 102.

[0055] The activity application 214 includes software and/or logic executable on the computing device 100. In some embodiments, the activity application 214 may receive the characters, distinct visual elements, user markings in the tangible objects detected in the video stream of the activity surface from the detector 304. The activity application 214 may determine one or more visually instructive elements (such as keywords or text, etc.) based on these factors, retrieve one or more digital content items using the visually instructive elements, and display the digital content items on the computing device 100. Non-limiting examples of the activity application 214 include video games, learning applications, assistive applications, storyboard applications, collaborative applications, productivity applications, etc. Other types of activity application are also possible and contemplated.

[0056] The camera driver 306 includes software storable in the memory 314 and operable by the processor 312 to control/operate the video capture device 110. For example, the camera driver 306 may be a software driver executable by the processor 312 for instructing the video capture device 110 to capture and provide a video stream and/or a still image, etc. In some embodiments, the camera driver 306 may be capable of controlling various features of the video capture device 110 (e.g., flash, aperture, exposure, focal length, etc.). In some embodiments, the camera driver 306 may be communicatively coupled to the video capture device 110 and other components of the computing device 100 via the bus 308, and these components may interface with the camera driver 306 to capture video and/or still images using the video capture device 110.

[0057] As discussed elsewhere herein, the video capture device 110 (also referred to herein as a camera) is a video capture device adapted to capture video streams and/or images of the physical activity surface. In some embodiments, the video capture device 110 may be coupled to the bus 308 for communication and interaction with the other components of the computing device 100. In some embodiments, the video capture device 110 may include a lens for gathering and focusing light, a photo sensor including pixel regions for capturing the focused light, and a processor for generating image data based on signals provided by the pixel regions. The photo sensor may be any type of photo sensor (e.g., a charge-coupled device (CCD), a complementary metal-oxide-semiconductor (CMOS) sensor, a hybrid CCD/CMOS device, etc.). In some embodiments, the video cap-

ture device **110** may include a microphone for capturing sound. Alternatively, the video capture device **110** may be coupled to a microphone coupled to the bus **308** or included in another component of the computing device **100**. In some embodiments, the video capture device **110** may also include a flash, a zoom lens, and/or other features. In some embodiments, the processor of the video capture device **110** may store video and/or still image data in the memory **314** and/or provide the video and/or still image data to other components of the computing device **100**, such as the detection engine **212** and/or the activity applications **214**.

[0058] In some embodiments, multiple video capture devices (such as **110a** and **110b**) may be included in the computing device **100**. These multiple video capture devices **110** may include separate fields of view directed to different portions of the area around the computing device **100**. In some embodiments, the fields of view of the multiple video capture devices may overlap and allow for stereo images or video streams to be created by an activity application(s) **214**, where the different two-dimensional video streams may be combined to form a three-dimensional representation of an tangible content item **102** in a graphical user interface of the display of the computing device **100**.

[0059] The storage **310** is a non-transitory storage medium that stores and provides access to various types of data. Non-limiting examples of the data stored in the storage **310** include video stream and/or still images captured by the video capture device **110**, various calibration profiles associated with each camera position of the video capture device **110**, object data describing various tangible objects, image recognition models and their model parameters, textual recognition models and their model parameters, visually instructive elements corresponding to each tangible content item, etc. In some embodiments, the storage **310** may be included in the memory **314** or another storage device coupled to the bus **308**. In some embodiments, the storage **310** may be or included in a distributed data store, such as a cloud-based computing and/or data storage system. In some embodiments, the storage **310** may include a database management system (DBMS). The DBMS may be a structured query language (SQL) DBMS. For example, the storage **310** may store data in an object-based data store or multi-dimensional tables including rows and columns, and may manipulate (i.e., insert, query, update, and/or delete) data entries stored in the storage **310** using programmatic operations (e.g., SQL queries and statements or a similar database manipulation library). Other implementations of the storage **310** with additional characteristics, structures, acts, and functionalities are also possible and contemplated. In some implementations, the storage **310** may include a database of previously scanned tangible content, such as books, that may be scanned in and then the detector **304** may OCR the tangible content for later identification and retrieval as described elsewhere herein.

[0060] An example method **400** for enhancing tangible content on a physical activity surface is illustrated in FIG. 4. The method **400** may be performed by a detector **304** and an activity application **214** implemented on the computing device **100** (see FIGS. 2 and 3). In block **402**, the video capture device **110** of the computing device **100** may capture the video stream that includes the activity scene **132** of the physical activity surface **150**. As discussed above, the physical activity surface **150** may include a tangible content item **102** situated on the physical activity surface **150**. The

tangible content item **102** may be a tangible object that a user can place and manipulate in the physical activity scene **132**. The tangible content item **102** may include one or more portions that are visually exposed to the video capture device **110** of the computing device **100**, and thus the video capture device **110** may capture the tangible content item **102** in the video stream as depicted in FIG. 1A.

[0061] In block **404**, the detector **304** may detect in the video stream the tangible content item **102** situated on the physical activity surface **150**. In some embodiments, the detector **304** may apply an object detection algorithm to the image of the video stream to detect the tangible content item **102** depicted in the image. For example, as depicted in FIG. 1A, the detector **304** may detect the tangible content item **102** that includes textual content and/or graphical content in the image **112** of the video stream. In some implementations, this image **112** may be displayed as a live feed on a display of the computing device **100**. In further implementations, the image **112** may be processed by the detector **304** from the video stream and may not be displayed to the user.

[0062] In block **406**, the detector **304** may recognize from the video stream one or more visually instructive elements in the tangible content item **102**. In some embodiments, the detector **304** may apply a textual recognition technique (e.g., Optical Character Recognition—OCR) to the image of the video stream to detect the visually instructive elements in the tangible content item **102**. The visually instructive elements may include various letters, characters, numbers, symbols, etc., that form the textual content of the tangible content item **102**. In some embodiments, in addition to the visually instructive elements, the detector **304** may detect in the video stream one or more distinct visual elements in the tangible content item **102**. Non-limiting examples of the distinct visual element include images, drawings, diagrams, vision markers specifically designed to indicate certain information, etc. In some embodiments, the detector **304** may apply an image recognition technique to the image of the video stream to determine the objects depicted in the distinct visual element included in the tangible content item **102**. For example, as depicted in FIGS. 1A and 1B, the detector **304** may apply the image recognition technique to the image **112** of the video stream. The detector **304** may determine that the tangible content item **104** includes a diagram **106** and a drawing **108**, and determine that the diagram **106** depicts an airplane and the drawing **108** depicts an airplane turbine.

[0063] In some embodiments, the detector **304** may recognize from the video stream the metadata of the tangible content item **102**. The metadata of the tangible content item **102** may provide general information about the tangible content item **102** and may be positioned on the top portion and/or the bottom portion of the tangible content item **104**. In some embodiments, the detector **304** may apply the textual recognition technique (e.g., OCR) on the top portion and/or the bottom portion of the tangible content item **102** to detect the metadata of the tangible content item **102**. Non-limiting examples of the metadata of the tangible content item **102** include, but are not limited to, the document title of the tangible object **102** including the tangible content item **102** (e.g., “Airplane Structure for Engineering Students”), the page number associated with the tangible content item **102** (e.g., pages 25 and 26), the section title associated with the tangible content item **102** (e.g., “General Components of

Airplane”), the section number associated with the tangible content item 102 (e.g., Chapter 1), etc.

[0064] In some embodiments, the detector 304 may detect in the video stream one or more user markings on the tangible content item 104. The user markings may be an emphasize effect, such as highlighting created by a user to emphasize a portion of the tangible content item 102 (e.g., highlighting, underlining, etc.) or a user note added to a portion of the tangible content item 102 by the user (e.g., hand-writing text, sketch, diagram, etc.). In some embodiments, the detector 304 may recognize from the video stream one or more characters associated with the user marking. For the user marking that is an emphasize effect, the detector 304 may determine an emphasized portion of the tangible content item 102 indicated by the user marking, and apply the textual recognition technique (e.g., OCR) on the emphasized portion of the tangible content item 104 to determine the visually instructive elements associated with the user marking. For the user marking that is a user note added to the tangible content item 102 by the user, the detector 304 may apply a machine learning model for handwriting recognition on the user note to recognize the handwriting characters in the user note. The detector 304 may also apply the image recognition technique on the user note to detect the distinct visual element in the user note (e.g., sketch, diagram, etc.) and determine the objects depicted by the distinct visual element in the user note. For example, the detector 304 may detect a sketch of airflows created by the user around the body of the airplane in the tangible content item 102.

[0065] In some embodiments, the detector 304 may generate a detection result including one or more of the visually instructive elements recognized in the tangible content item 102, the distinct visual element included in the tangible content item 102 and/or the objects detected in the distinct visual element, the metadata of the tangible content item 102 that is recognized in the tangible content item 102 (e.g., document title, page number, section title, section number, etc.), the characters associated with the user markings in the tangible content item 102, the distinct visual element associated with the user markings and/or the objects detected in the distinct visual element, etc. In some embodiments, the detector 304 may transmit the detection result to the activity application 214. The activity application 214 may be capable of providing additional content related to the tangible content item 104 based on the detection result.

[0066] In block 408, the activity application 214 may determine one or more tangible identifiers based on the characters in the tangible content item 104. To determine the tangible identifiers, the activity application 214 may analyze the detection result generated by the detector 304 to obtain the visually instructive elements recognized in the tangible content item 102. The activity application 214 may aggregate these recognized visually instructive elements into words or other categories, each word may include the characters located between two sequential space characters in the tangible content item 102. In some embodiments, the activity application 214 may select one or more tangible identifiers from the words in the tangible content item 102, such as a keyword. For example, the activity application 214 may determine a heading line in the tangible content item 102, and select one or more descriptive words in the heading line to be the tangible identifier. In another example, the activity application 214 may determine one or more empha-

sized words that are presented in different format from other words in the tangible content item 102 (e.g., different font, size, color, etc.), and select the emphasized words to be the tangible identifiers. In some embodiments, the activity application 214 may determine the descriptive words in the tangible content item 102, compute a repeat frequency indicating a number of times each descriptive word is repeated in the tangible content item 102, and select the descriptive words that have the repeat frequency satisfying a repeat frequency threshold (e.g., more than seven times) to be the tangible identifier.

[0067] In some embodiments, instead of using the words included in the tangible content item 102, the activity application 214 may determine the content category of the tangible content item 102 based on the descriptive words and/or the distinct visual element in the tangible content item 102, and select the terminologies associated with the content category of the tangible content item 102 to be the visually instructive elements. As an example, the activity application 214 may categorize the tangible content item 102 into the content category “airplane,” and select the terminologies associated with the content category “airplane” to be the keywords (e.g., “aircraft,” “aerodynamics,” “airplane engine,” etc.).

[0068] In some embodiments, the activity application 214 may determine the tangible identifier based on the distinct visual element included in the tangible content item 102 and/or the objects detected in the distinct visual element. As an example, the activity application 214 may determine that the tangible content item 102 includes an image of an airplane engine. The activity application 214 may also determine the image description of the image from the recognized characters of the tangible content item 102 (e.g., “Gas-turbine engine of jet aircraft”). In this example, the activity application 214 may determine the tangible identifier based on the airplane engine depicted in the image and the image description of the image (e.g., “airplane engine,” “aircraft turbine,” “gas-turbine,” “jet aircraft,” etc.).

[0069] In some embodiments, the activity application 214 may determine the tangible identifier based on the metadata of the tangible content item 102. The activity application 214 may analyze the detection result generated by the detector 304 to obtain the metadata including the document title, the page number, the section title, and/or the section number associated with the tangible content item 102. If the section title associated with the tangible content item 102 is not included in the detection result, the activity application 214 may retrieve the table of content of the tangible object 102 that contains the tangible content item 102 from the data storage 310 (see FIG. 3), and determine the section title from the table of content of the tangible content item 102 using the page number and/or the section number associated with the tangible content item 102. In some embodiments, the document title and/or the section title associated with the tangible content item 102 may provide a general summary of the content presented in the tangible content item 102. Therefore, the activity application 214 may determine the descriptive words in the document title and/or the section title associated with the tangible content item 102, and select the descriptive words to be the tangible identifier.

[0070] In some embodiments, the activity application 214 may determine the tangible identifier based on the characters associated with the user marking. The user marking may indicate the information that the user considered as impor-

tant and should pay attention to in the tangible content item **102**. In some embodiments, the activity application **214** may analyze the detection result generated by the detector **304** to obtain the characters associated with the user marking. As discussed elsewhere herein, the characters associated with the user marking may be the characters in the portion of the tangible content item **102** that is subjected to the emphasize effect created by the user and/or the characters in the user note that is added to the tangible content item **102** by the user. In some embodiments, the activity application **214** may aggregate the characters associated with the user marking into words, each word may include the characters located between two sequential space characters. The activity application **214** may determine one or more descriptive words among the words formed by the characters associated with the user marking, and select the descriptive words to be the tangible identifier. As an example, the user may highlight a word “turbulence modeling” and add a handwriting note “use k-omega model with wall function” in the tangible content item **102**. In this example, the activity application **214** may determine the descriptive words in the user markings, and determine the keywords to be “turbulence modeling,” “k-omega model,” and “wall function.”

[0071] In some embodiments, the activity application **214** may analyze the detection result generated by the detector **304** to also obtain the distinct visual element associated with the user markings and/or the objects detected in the distinct visual element. In some embodiments, the activity application **214** may determine the keywords based on the distinct visual element associated with the user markings and/or the objects detected therein. As an example, the user may create a sketch of an airplane with arrows illustrating airflows around the body of the airplane in the tangible content item **102**. In this example, the activity application **214** may determine that the user note includes the airplane and the airflows, and determine the keywords to be “airplane,” “airflows,” and “aerodynamic.”

[0072] In some embodiments, the tangible object may include a plurality of tangible content items **102** and the user may interact with different tangible content items **102** of the tangible object. For example, the tangible content item **102** may be a book, and the user may read a first portion of the first tangible content item **102** associated with pages 11 and 12 of the book, and then turn to a second portion of the second tangible content item **102** associated with pages 3 and 4 of the book. In some embodiments, the activity application **214** may monitor the tangible content items **102** with which the user interacted, and determine the tangible identifier and/or visually instructive elements based on these tangible content items **102**. To determine the tangible identifier, the activity application **214** may determine a first tangible content item **102** with which the user interacted at a first timestamp, and recognize one or more visually instructive elements in the first tangible content item **102**. The activity application **214** may determine a second tangible content item **102** with which the user interacted at a second timestamp, and recognize one or more visually instructive elements in the second tangible content item **102**. The first timestamp at which the user interacted with the first tangible content item **102** may be prior to or subsequent to the second timestamp at which the user interacted with the second tangible content item **102**. In some embodiments, the

time distance between the first timestamp and the second timestamp may satisfy a time distance threshold (e.g., less than 30 s).

[0073] In some embodiments, the activity application **214** may determine the tangible identifier based on the characters in the first tangible content item **102** and the characters in the second tangible content item **102**. In addition to the characters, the activity application **214** may also determine the tangible identifier based on the distinct visual elements, the user markings, etc., in the first tangible content item **102** and the second tangible content item **102** as discussed above. In some embodiments, the activity application **214** may select the descriptive words included in both the first tangible content item **102** and the second tangible content item **102** to be the tangible identifiers. Alternatively, the activity application **214** may select the tangible identifiers that are relevant to the content of the first tangible content item **102** and the content of the second tangible content item **102**. For example, the first tangible content item **102** may present information about airflows around airplane body, and the second tangible content item **102** may present information about shape and size of commercial airplanes. In this example, the activity application **214** may determine the tangible identifier to be “airflows around commercial airplanes.” Determining the tangible identifier based on multiple tangible content items **102** with which the user interacted is advantageous, because this implementation can increase the likelihood of the tangible identifier being relevant to the content that the user is interested in.

[0074] In block **410**, the activity application **214** may retrieve one or more digital content items using one or more tangible identifiers. In some embodiments, the activity application **214** may query the data storage **310** and/or other content databases using the tangible identifier to retrieve the digital content items. The digital content items returned as the query result may include the tangible identifiers, and thus may be relevant or related to the tangible content item **102** that is present on the physical activity surface **150**. Non-limiting examples of the digital content item includes images, videos, audios, webpages, electronic documents, instructions, etc. Other types of digital content item are also possible and contemplated.

[0075] In some embodiments, the activity application **214** may retrieve the digital content items using one tangible identifier in the query. Alternatively, the activity application **214** may retrieve the digital content items using multiple tangible identifiers. In some embodiments, the activity application **214** may determine a first tangible identifier and a second tangible identifier based on the recognized visually instructive elements, the distinct visual elements, the user markings, etc., in the tangible content item **102** as discussed above. In some embodiments, the activity application **214** may determine a contextual relationship between the first tangible identifier and the second tangible identifier in describing subject matters similar to the tangible content item **102**. For example, the contextual relationship between the first tangible identifier and the second tangible identifier may indicate the likelihood of the first tangible identifier and the second tangible identifier being in the same content item that belongs to the content category of the tangible content item **102**, the average distance between the first tangible identifier and the second tangible identifier in the content item, etc. In some embodiments, the contextual relationship

between the first tangible identifier and the second tangible identifier may be indicated by a contextual relationship metric.

[0076] In some embodiments, the activity application **214** may retrieve the digital content items based on the contextual relationship between the first tangible identifier and the second tangible identifier. In some embodiments, the activity application **214** may determine whether the contextual relationship metric between the first tangible identifier and the second tangible identifier satisfies a contextual relationship metric threshold (e.g., more than 0.75). If the contextual relationship metric between the first tangible identifier and the second tangible identifier satisfies the contextual relationship metric threshold, the activity application **214** may retrieve the digital content items using both the first tangible identifier and the second tangible identifier. For example, the activity application **214** may query the data storage **310** and/or other content databases using the first tangible identifier and the second tangible identifier in the same query to retrieve the digital content items. The digital content items returned as the query result may include both the first tangible identifier and the second tangible identifier. In some embodiments, the relative distance between the first tangible identifier and the second tangible identifier in the digital content items may be proportional to the contextual relationship metric between the first tangible identifier and the second tangible identifier.

[0077] In some embodiments, the activity application **214** may retrieve the digital content item from a knowledge graph stored in the storage **310**. The knowledge graph may include one or more nodes related to different topics and each node may be linked to other nodes related to different topics within the knowledge graph. Each node may include digital content items related to the topic of that node in the knowledge graph. In some embodiments, the knowledge graph may link sub-nodes with sub-topics to the main node with a main-topic in the knowledge graph and as the activity application **214** provides the one or more tangible identifiers to the knowledge graph, the nodes related to those tangible identifiers may be identified and digital content items within that node may be retrieved. In some embodiments, the nodes may also include category levels for different users, where the different category levels relate to depth of the knowledge on the topic. For example, if the node of the knowledge graph is related to aerodynamics, a user with a lower category level (such as in a second year of school) will be different than a user with a higher category level (such as in a college level year of school). Based on the category level of a user, different digital content items related to the topic of the node may be provided that are relevant to that user based on the category level.

[0078] In block **412**, once the digital content items are retrieved, the activity application **214** may provide the digital content items on the computing device **100**. The digital content items may include videos, images, webpages, electronic documents, etc., displayed on the computing device **100**, and may provide content data relevant to the tangible content item **104** of the tangible object **102** that is present on the physical activity surface **150**. An example of the digital content items being provided on the computing device **100** is illustrated in FIG. 1B. In this example, the tangible content item **104** on the physical activity surface **150** may include the characters, the distinct visual elements, and/or the user markings related to airplane, and the digital

content items displayed on the computing device **100** may present content data relevant to airplane (e.g., airplane parts, aircraft structure, principle of aerodynamics, airplane engine, etc.) as depicted in FIG. 1B.

[0079] FIG. 5 is an example configuration **500** for enhancing tangible content on a physical activity surface. As shown in the example, the computing device **100** may include a first camera **110a** that includes a first field of view **502** directed downwards towards a surface of in front of the computing device **100**. The computing device **100** may also include a second camera **110b** that includes a second field of view **504**. The second field of view **504** may be directed outward from the computing device **100** and include the area in front of the computing device **100**. In some implementations, the second field of view **504** may be able to capture an image of at least a portion user interacting with the physical activity surface in front of the computing device **100**. For example, the second field of view **504** may be able to capture an upper torso of a first user including their head and arms. This may allow the activity application **214** to view facial expressions and/or hand gestures performed by the user in the second field of view **504**. In some implementations, the second field of view **504** may be displayed on the computing device **100** or another computing device (not shown). In this example, the display screen may display multiple items on the display screen, one of more items may be based on the physical content items **102** included in the first field of view **502** and another item may be a video stream of the user from the second field of view **504**.

[0080] In some implementations, the activity application **214** may be able to use the second field of view **504** to identify a specific user present in the video stream. For example, the activity application **214** may identify a first user as being in a first grader and a second user as being a parent of the first user and may present appropriate content based on the identity of the user. If the first user is detected in the second field of view **504**, then the activity application may retrieve the current homework tasks for the first grader and present them on the display screen for the first grader to begin working on. If the second user is detected in the second field of view **504**, then the activity application may display a status of the various tasks that have been assigned to the first grader and easily facilitate helping the parent know how the first grader is doing on various tasks, without having to navigate to specific content manually.

[0081] In some implementations, the activity application **214** may use one or more of the fields of view of the camera **110** to execute various applications. For example, when no user is detected in a field of view of the camera **110**, then the activity application **214** may cause the computing device **100** to go into a passive or sleep mode. In some implementations, in the passive mode, the activity application **214** and/or the detector **304** may routinely check the fields of view of the cameras **110** and if something is detected in one or more of the fields of view, then the activity application **214** may execute an appropriate program. For example, the detector **304** may detect a user positioning themselves (such as by sitting down in a chair in front of the computing device **100**) and may cause the activity application **214** to determine an identity of the user and/or one or more relevant applications to being displaying on the display screen.

[0082] For example, when a first user is detected, then activity application **214** may retrieve a profile of the first user and review commonly executed applications. The activ-

ity application **214** may then launch one or more of these applications before receiving a selection by the user. These applications may be launched in the background without the user's express knowledge and if the user requests to run one of the pre-launched applications, then the activity application **214** may display the pre-launched application and/or close down the other applications. This allows the activity application **214** to quickly execute applications and the users doesn't have to wait for common applications to load up based on the profile of the identified user.

[0083] In some implementations, the detector **304** may detect an object, such as tangible content item **102** being positioned in one of the fields of view of the camera **100** and may cause the activity application **214** to retrieve one or more digital content items based on a tangible identifier of the tangible content item **102**. The activity application **214** may further display a graphic and/or animation on the display screen to guide the user on how to begin using the tangible content item **102**, such as displaying a current page number for the user to open the tangible content item **102**, etc. In some implementations, the detector **304** may sense the presence of a user and/or tangible content item **102** and display a query, prompt, or detection awareness graphic to signal to the user that they and/or the object have been detected and can begin interacting with the computing device **100**.

[0084] FIG. 6 is a flowchart **600** of an example method for enhancing tangible content on a physical activity surface. At **602**, the video capture device may capture a video stream that includes an activity scene **132** of a physical activity surface **150**. In some implementations, as described elsewhere herein, the computing device **100** may be positioned on the physical activity surface **100**, such as in a stand or an integrated device. In further implementations, the computing device **100** may be portable and positioned by a user **100** to capture a video stream of the physical activity surface **150**.

[0085] At **604**, the detector **304** may detect in the video stream one or more tangible content items **102** included in the video stream. In some implementations, the tangible content item **102** may include one or more pages that have been opened to be visible within the field of view of the video capture device. In further implementations, the tangible content item **102** may be a single page, such as a worksheet, that is visible within the field of view of the video capture device.

[0086] At **606**, the activity application **214** may determine an identity of the page of the tangible content item **102** from the video stream. The identity of the page may be determined based on one or more visually instructive elements included on the page, such as a page number, a title, a graphic, etc. In further implementations, the identity of the page may be determined based on a page identity anchor on the page that may be a unique graphic or textual icon visible to the detector **304**. The activity application **214** may determine the identity of the page based on the page identity anchor or other visually instructive elements included within the page. For example, each page of a tangible content item **102** may have a unique page identity anchor and each page identity anchor may match a database of page identity anchors that are associated with different pages. In some implementations, pages of the tangible content item **102** may be captured as images and image recognition may be performed on the pages to determine content of the page and

one or more concepts the pages are related to. The pages may then be catalogued in a database for future retrieval when their unique page identity anchor has been identified. In some implementations, the detector **304** may be configured to only search for page identity anchors in the images captured by the video stream to increase the speed at which the activity application **304** can identify the page. By only looking for the page identity anchor, the detector **304** can quickly ignore the other text and/or graphics present on a tangible content item **102** until the page identity anchor has been identified and the profile of the page has been retrieved.

[0087] At **608**, the activity application **214** may determine a title of the tangible content item using the identity of the page. For example, the activity application **214** may know that the page is page 26 of a specific book and the tangible content item **102** may be the specific book based on this identity of the page 26. The title of the book may then signal the subject and/or context of the tangible content item **102**.

[0088] At **610**, the activity application **214** may determine digital content items based on the title of the tangible content item and/or the identity of the page of the tangible content item **102**. For example, where the tangible content item **102** is a textbook on math title "3rd Grade Math" and the specific page identity is page 26 related to line segments. The activity application **214** may use the contextual information related to the title "3rd Grade Math" and the "line segments" to identify a digital content item representing a video of how to solve a problem with line segments by a different third grade teacher. In some implementations, the digital content item can be one or more of a video, a question, a quiz, a worksheet, etc. At **612**, the digital content item is then presented in the graphical user interface of the display screen. This supplemental digital content item can be presented on the display of the computing device **100** for the user to watch and assist in the learning they are accomplishing.

[0089] FIGS. 7A-7D are graphical representations of an example user interface for enhancing tangible content. As shown in FIG. 7A, the graphical representation **700** depicts a user interface **702** that includes a personalization portion **708**. The personalization portion may be based on a detected identity of a user in a field of view of the camera **110**. The personalization portion may include various settings that the user may personalize to create their own experience.

[0090] The user interface **702** may include progression icons **704** displaying a progress of the specific user through various topics. For example, progression icon **704a** depicts a progress in "Practical Geometry Mathematics" and includes a progress bar showing the majority of the topics related to that progression icon **704a** have been completed. As different tasks or topics are assigned to a user, such as by a teacher as homework or supplemental learning, the progression icons **704** may update to include an additional topic based on the new tasks. The progression icons may be selectable by a user and may include nested icons representing different subtopics within the progression icons. The progressions icons may be displayed in an easily accessible area for the user to begin selecting topics quickly. In some implementations, the activity application **214** may cause one of the topics of the progression icons **704** to be displayed before a user begins interacting with the display screen.

[0091] The user interface **702** may include topic icons **706** representing different topics and/or tasks for the user to interact with. For example, the user may be able to select

between “practical geometry” in topic icon **706a** or “stars and the solar system” in topic icon **706c**. The topic icons **706** may extend beyond an edge of the screen and the user may be able to scroll through the various topic icons **706** to select various icons.

[0092] As shown in FIG. 7B, the graphical representation **710** may update the user interface **702** when the activity application detects a tangible content item **102** (not shown) has been positioned within a field of view of the camera **100**. For example, a user may place a text book related to “Atoms and Molecules” on the physical activity surface and the activity application **214** may cause a detected textbook icon **712** to be displayed on the user interface **702**. In some implementations, the detected textbook icon **712** may scroll into the user interface **702** so that it appears it is being pushed into the field of view as the user moves the physical textbook related to “Atoms and Molecules” into the field of view of the camera **110**. By imitating the physical placement of the tangible content item **102**, the detected textbook icon **712** may create a more immersive experience for the user. In further implementations, as different textbooks are moved on positioned on the physical activity surface **150**, the user interface **702** may cause the virtual representations of the topic icons **706** or detected textbook icons **712** to mirror those movements and move in and out of the user interface **702** similar to how the physical items are moving in and out of the field of view of the camera **102** in order to create a virtual tabletop that mimics what is present in the physical activity surface **150**.

[0093] In some implementations, the activity application **214** may generate the detected textbook icon **712** as a visualization based on a digital content item that may be appearing in the graphical user interface **702**. The detected textbook icon **702** may be displayed as an animation of the digital content item appearing in the user interface. In some implementations, this animation may mimic the appearance of a corresponding tangible content item **102** that is detected by the detector **304** as it appears in the video stream of the physical activity surface. In some implementations, the animation may start by displaying only a first portion of the visualization of the detected textbook icon **712**, such as shown in FIG. 7B, where the first portion of the detected textbook icon **712c** is displayed and a second portion is not shown in the user interface. The animation may then appear to scroll the rest of the visualization of the detected textbook icon **712** to appear on the screen, such as by presenting the second portion of the detected textbook icon **712c** as shown in FIG. 7C.

[0094] As shown in FIG. 7C, the graphical representation **714** may display the user interface **702** where the detected textbook icon **712** is fully displayed in the user interface **712**. In some implementations, the topic icons **706a-706c** may be shifted to allow for the detected textbook icon **712** to be displayed. In some implementations, the user interface **702** may also cause one or more subtopics, such as the portion of the detected textbook icons **712a-712b** to be displayed along with an image of the full topic of the detected textbook icon **712c**.

[0095] As shown in FIG. 7D, the graphical representation **716** may display the user interface **702** where a user is navigating into one or more of the detected textbook icons **712** as shown to display task icons **720** related to the selected topic **718**. Using this nested topic option, a user can quickly navigate through various tasks and assignments based on

topic and each of the topics can be triggered for display based on the detection of a tangible content item **102** in the field of view of the camera **110**. This automatic appearance of digital content based on the interaction with a tangible content item **102** allows users to quickly access various relevant information. For example, a student can quickly view various tasks and assignments for the day, just by viewing the progression icons **704** or the user can place one of their textbooks or activity books into the field of view of the camera **110** and it will automatically display the current assignments related to that book for the student to begin working on.

[0096] FIG. 8 is an example configuration **800** for finger detection to enhance tangible content. As shown in the example, a computing device **100** may be positioned on an activity surface with one or more cameras **110**. The detector **304** may detect an appearance of a hand **804** and/or finger of a user in an image of the physical activity surface. The detector **304** may identify the hand **804** and/or finger based on expected contours or shapes described elsewhere herein related to finger and hand shapes. In some implementations, the detector **304** may be configured to identify a point of a finger of the hand **804** within the image and may determine an area of the tangible content item **102** that the point of the finger of the hand **804** is pointing to. As shown in FIG. 8, the area **806** of the tangible content item **102** may be a portion of the tangible content item **102** proximate to where the detector **304** has detected the point of the finger of the hand **804** to be pointing. In some implementations, the activity application **214** may map the position of the tangible content item **102** into a coordinate plane and divide the mapped tangible content item **102** into various areas within each section of the coordinate plane. The activity application **214** may map the position of the point of the finger of the hand **804** into the same coordinate plane relative to the position of the tangible content item **102** and identify which areas of the coordinate plane include both the point of the finger and the corresponding area of the tangible content item **102**.

[0097] In some implementations, the area **806** that the finger is pointing to of the tangible content item **102** may be analyzed by the detector **304** to identify one or more tangible identifiers within the area. For example, a tangible content item **102** may be a textbook and the user may be pointing to an area **806** with their finger to a specific word or image on the page. Using the finger detection, the detector **304** may focus on only the content within the area **806** to identify a tangible identifier, such as only the specific word or image being pointed to by the point of the finger. This allows the interaction between the tangible content item **102** and the digital content items to be enhanced where a user can simply point to specific areas that they need assistance with or supplemental content provided for. For example, a user can be reading a book and not know the definition for a specific word. The user can point at the word and the activity application **214** may cause the definition of the word to appear and/or be audibly output using the computing device **100**. In another example, when a user is stuck on a math problem, they can point to a specific step in the math problem and the user can quickly provide suggestions for that specific step rather than the context of the entire math problem.

[0098] It should be understood that the above-described example activities are provided by way of illustration and not limitation and that numerous additional use cases are

contemplated and encompassed by the present disclosure. In the above description, for purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the present disclosure. However, it should be understood that the technology described herein may be practiced without these specific details. Further, various systems, devices, and structures are shown in block diagram form in order to avoid obscuring the description. For instance, various implementations are described as having particular hardware, software, and user interfaces. However, the present disclosure applies to any type of computing device that can receive data and commands, and to any peripheral devices providing services.

[0099] In some instances, various implementations may be presented herein in terms of algorithms and symbolic representations of operations on data bits within a computer memory. An algorithm is here, and generally, conceived to be a self-consistent set of operations leading to a desired result. The operations are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

[0100] It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the following discussion, it is appreciated that throughout this disclosure, discussions utilizing terms including “processing,” “computing,” “calculating,” “determining,” “displaying,” or the like, refer to the action and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (electronic) quantities within the computer system’s registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

[0101] Various implementations described herein may relate to an apparatus for performing the operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general-purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, including, but is not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, and magnetic disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, flash memories including USB keys with non-volatile memory or any type of media suitable for storing electronic instructions, each coupled to a computer system bus.

[0102] The technology described herein can take the form of a hardware implementation, a software implementation, or implementations containing both hardware and software elements. For instance, the technology may be implemented in software, which includes but is not limited to firmware, resident software, microcode, etc. Furthermore, the technology can take the form of a computer program product accessible from a computer-usable or computer-readable medium providing program code for use by or in connection

with a computer or any instruction execution system. For the purposes of this description, a computer-usable or computer readable medium can be any non-transitory storage apparatus that can contain, store, communicate, propagate, or transport the program for use by or in connection with the instruction execution system, apparatus, or device.

[0103] A data processing system suitable for storing and/or executing program code may include at least one processor coupled directly or indirectly to memory elements through a system bus. The memory elements can include local memory employed during actual execution of the program code, bulk storage, and cache memories that provide temporary storage of at least some program code in order to reduce the number of times code must be retrieved from bulk storage during execution. Input/output or I/O devices (including but not limited to keyboards, displays, pointing devices, etc.) can be coupled to the system either directly or through intervening I/O controllers.

[0104] Network adapters may also be coupled to the system to enable the data processing system to become coupled to other data processing systems, storage devices, remote printers, etc., through intervening private and/or public networks. Wireless (e.g., Wi-Fi™) transceivers, Ethernet adapters, and modems, are just a few examples of network adapters. The private and public networks may have any number of configurations and/or topologies. Data may be transmitted between these devices via the networks using a variety of different communication protocols including, for example, various Internet layer, transport layer, or application layer protocols. For example, data may be transmitted via the networks using transmission control protocol/Internet protocol (TCP/IP), user datagram protocol (UDP), transmission control protocol (TCP), hypertext transfer protocol (HTTP), secure hypertext transfer protocol (HTTPS), dynamic adaptive streaming over HTTP (DASH), real-time streaming protocol (RTSP), real-time transport protocol (RTP) and the real-time transport control protocol (RTCP), voice over Internet protocol (VOIP), file transfer protocol (FTP), Web Socket (WS), wireless access protocol (WAP), various messaging protocols (SMS, MMS, XMS, IMAP, SMTP, POP, WebDAV, etc.), or other known protocols.

[0105] Finally, the structure, algorithms, and/or interfaces presented herein are not inherently related to any particular computer or other apparatus. Various general-purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct more specialized apparatus to perform the required method blocks. The required structure for a variety of these systems will appear from the description above. In addition, the specification is not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the specification as described herein.

[0106] The foregoing description has been presented for the purposes of illustration and description. It is not intended to be exhaustive or to limit the specification to the precise form disclosed. Many modifications and variations are possible in light of the above teaching. It is intended that the scope of the disclosure be limited not by this detailed description, but rather by the claims of this application. As will be understood by those familiar with the art, the specification may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. Likewise, the particular naming and division of the

modules, routines, features, attributes, methodologies and other aspects are not mandatory or significant, and the mechanisms that implement the specification or its features may have different names, divisions and/or formats.

[0107] Furthermore, the modules, routines, features, attributes, methodologies and other aspects of the disclosure can be implemented as software, hardware, firmware, or any combination of the foregoing. Also, wherever an element, an example of which is a module, of the specification is implemented as software, the element can be implemented as a standalone program, as part of a larger program, as a plurality of separate programs, as a statically or dynamically linked library, as a kernel loadable module, as a device driver, and/or in every and any other way known now or in the future. Additionally, the disclosure is in no way limited to implementation in any specific programming language, or for any specific operating system or environment. Accordingly, the disclosure is intended to be illustrative, but not limiting, of the scope of the subject matter set forth in the following claims.

What is claimed is:

1. A method comprising:

capturing, using a video capture device of a computing device, a video stream that includes an activity scene of a physical activity surface located in front of a display of the computing device;

detecting in the video stream, using a detector executable on the computing device, a tangible content item on the physical activity surface;

recognizing, from the video stream, one or more visually instructive elements in the tangible content item;

determining a tangible identifier based on the one or more recognized visually instructive elements in the tangible content item;

automatically retrieving a digital content item using the tangible identifier; and

providing the digital content item for display in a user interface of the computing device.

2. The method of claim 1, further comprising:

detecting, in the video stream, a distinct visual element on the tangible content item; and wherein

determining the tangible identifier includes determining the tangible identifier based on the distinct visual element on the tangible content item.

3. The method of claim 2, wherein:

the distinct visual element includes one or more of an image, a drawing, a diagram, and a vision marker visible on the tangible content item.

4. The method of claim 1, further comprising:

recognizing, from the video stream, one or more of a document title, a page number, a section title, and a section number associated with the tangible content item; and wherein

determining the tangible identifier includes determining the tangible identifier based on one or more of the document title, the page number, the section title, and the section number associated with the tangible content item.

5. The method of claim 1, further comprising:

detecting, in the video stream, a user marking visible on the tangible content item;

recognizing, from the video stream, one or more characters associated with the user marking; and wherein

determining the tangible identifier includes determining the tangible identifier based on the one or more characters associated with the user marking.

6. The method of claim 5, wherein:

the user marking includes one or more of a highlighting effect and a user note created by a user on the tangible content item.

7. The method of claim 1, wherein the tangible content item is a first tangible content item:

the first tangible content item is included in a tangible object that also includes a second tangible content item; and

the method includes:

determining that a user interacted with the second tangible content item of the tangible object;

recognizing one or more visually instructive elements in the second tangible content item; and

determining the tangible identifier based on the one or more visually instructive elements in the first tangible content item and the one or more visually instructive elements in the second tangible content item of the tangible object.

8. The method of claim 1, wherein:

determining the tangible identifier includes determining a first tangible identifier and a second tangible identifier based on the one or more visually instructive elements in the tangible content item; and

retrieving the digital content item using the tangible identifier includes:

determining a contextual relationship between the first tangible identifier and the second tangible identifier; and

retrieving the digital content item based on the contextual relationship between the first tangible identifier and the second tangible identifier.

9. The method of claim 1, wherein:

the computing device is placed on a stand situated on the physical activity surface; and

a field of view of the video capture device of the computing device is redirected by a camera adapter situated over the video capture device of the computing device.

10. The method of claim 1, wherein the tangible content item on the physical activity surface includes a first tangible content section and a second tangible content section, the method further comprising:

detecting, using a detector executable on the computing device, an arrangement of the first tangible content section relative to the second tangible content section on the physical activity surface;

determining a supplemental digital content item based on the arrangement of the first tangible content section relative to the second tangible content section; and

displaying in the user interface on the computing device, the supplemental digital content along with the digital content item.

11. The method of claim 10, wherein the first tangible content section is a first programming tile and the second tangible content section is a second programming tile, the first programming tile being coupled with the second programming tile to form the tangible content item.

12. The method of claim 11, wherein the supplemental digital content is an instruction set for how to use the first programming tile and the second programming tile.

13. The method of claim **12**, wherein displaying the supplemental digital content further comprises:

- automatically displaying a first digital representation of the first programming tile in the user interface;
- automatically displaying a second digital representation of the second programming tile in the user interface; and
- automatically displaying an animation of the first digital representation interacting with the second digital representation to provide instructions on how to arrange the first digital programming tile and the second digital programming tile on the activity surface.

14. The method of claim **1**, wherein providing the digital content item for display in the user interface on the computing device further comprises:

- generating a visualization of the digital content item for display in the user interface; and
- displaying an animation of the digital content item appearing in the user interface.

15. The method of claim **14**, wherein the animation of the digital content item depicts a first portion of the visualization of the digital content item extending out from a bottom side of the display screen and a second portion of the visualization of the digital content item appearing as the animation causes the visualization of the digital content item to move into the user interface.

16. The method of claim **15**, wherein the animation of the digital content item corresponds to an appearance of the tangible content item being detected on the physical activity surface.

17. The method of claim **14**, wherein the visualization of the digital content item is selectable by a user and responsive to the digital content item being selected, the digital content item may be displayed.

18. A method comprising:

- capturing, using a video capture device of a computing device, a video stream that includes an activity scene of a physical activity surface;
- detecting in the video stream, using a detector executable on the computing device, a tangible content item on the physical activity surface, the tangible content item including a page visible within a field of view of the video capture device;
- determining, using a processor of the computing device, an identity of the page of the tangible content item from the video stream;
- determining, using the processor of the computing device, a concept of the tangible content item using the identity of the page;
- determining, using the processor of the computing device, a digital content item based on the concept of the tangible content item and the identity of the page of the tangible content item; and
- presenting, in a display of the computing device, the digital content item in a graphical user interface.

19. The method of claim **18**, wherein determining the identity of the page of the tangible content item further comprises:

- detecting a page identity anchor on the page; and
- matching the page identity anchor to a database of page identities.

20. The method of claim **19**, wherein the tangible content item is a book that includes a plurality of pages and each of the plurality of pages has a unique page identity anchor.

21. The method of claim **18**, wherein the digital content item is a video related to the identified page of the tangible content item.

22. The method of claim **21**, wherein presenting the digital content item further comprises:

- embedding an external link to the video related to a topic present on the identified page of the tangible content item; and
- causing the external link to be executed responsive to an input received from a user.

23. The method of claim **19**, wherein determining the digital content item based on the concept of the tangible content item and the identity of the page of the tangible content item further comprises:

- identifying visually instructive elements that are present on the page of the tangible content;
- determining a subject context based on the visually instructive elements; and
- identifying the digital content item based on the subject context.

24. The method of claim **23**, wherein identifying the digital content item based on the subject context further comprises:

- determining a category level of a user based on the concept of the tangible content item and the identity of the page of the tangible content item;
- accessing a node of a knowledge graph that is related to the subject context, the node of the knowledge graph including links to a plurality of digital content items related to the subject context; and
- determining a relevant digital content item from the plurality of digital content items based on the category level of the user.

25. The method of claim **19**, wherein detecting a page identity anchor on the page further comprises:

- performing a matching search to identify the page identity anchor, the matching search including ignoring all text that does not include the page identity anchor.

26. The method of claim **18** wherein the digital content item is one or more of a video, a question, a quiz, and a worksheet.

27. The method of claim **18**, wherein determining an identity of the page of the tangible content item from the video stream further comprises:

- detecting, using a processor of a computing device, a finger of a user present in the video stream of the physical activity surface;
- determining an area of the tangible content item that is being pointed to by a point of the finger;
- determining a tangible identifier included within the area of the tangible content item; and
- determining an identity of the page of the tangible content item based on the tangible identifier.

28. A visualization system comprising:

- a computing device positioned on a physical activity surface;
- a video capture device coupled to the computing device, the video capture device including a field of view directed towards the physical activity surface and capable of capturing a video stream of the physical activity surface;

- a detector of the computing device configured to detect a tangible content item within the video stream and recognize one or more visually instructive elements in the tangible content item;
- a processor of the computing device configured to determine a tangible identifier based on the recognized visually instructive elements in the tangible content item and automatically retrieve a digital content item using the tangible identifier; and
- a display screen coupled to the computing device, the display screen being configured to display the digital content item in a graphical user interface.

* * * * *