



(51) International Patent Classification:

G06T 11/00 (2006.01) G06Q 30/0601 (2023.01)
G06T 7/11 (2017.01)

(21) International Application Number:

PCT/US2022/081875

(22) International Filing Date:

16 December 2022 (16.12.2022)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/291,276 17 December 2021 (17.12.2021) US
17/666,997 08 February 2022 (08.02.2022) US

(71) Applicant: PINTEREST, INC. [US/US]; 651 Brannan Street, San Francisco, California 94107 (US).

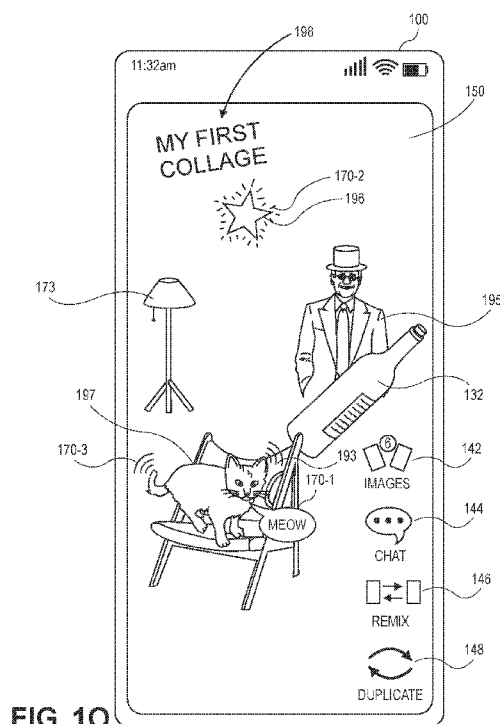
(72) Inventors: TEMPLE, David; 651 Brannan Street, San Francisco, California 94107 (US). PROBASCO, Ryan, Wilson; 651 Brannan Street, San Francisco, California 94107 (US). FARRE, Albert, Pereta; 651 Brannan Street,

San Francisco, California 94107 (US). RAMKUMAR, Steven; 651 Brannan Street, San Francisco, California 94107 (US). MAHADEV, Rohan; 651 Brannan Street, San Francisco, California 94107 (US). ZHAO, Luke; 651 Brannan Street, San Francisco, California 94107 (US). KIS-LYUK, Dmitry, Olegovich; 651 Brannan Street, San Francisco, California 94107 (US). XUE, David, Ding-Jia; 651 Brannan Street, San Francisco, California 94107 (US).

(74) Agent: HARRIS, Larry, T.; Athorus, PLLC, P.O. Box 579, Mount Vernon, Washington 98273 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH,

(54) Title: EXTRACTED IMAGE SEGMENTS COLLAGE



(57) Abstract: Described are systems and methods to extract image segments from an image and include those extracted image segments in a collage. The origin information, such as the source image, source image location, etc., from which the extracted image segment is generated is maintained as metadata so that interaction with the extracted image segment on the collage can be used to determine and/or return to the origin of the extracted image segment. Collages may be updated, shared, adjusted, etc., by the creator of the collage or other users.

TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
ZA, ZM, ZW.

- (84) Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

EXTRACTED IMAGE SEGMENTS COLLAGE

PRIORITY CLAIM

[0001] This application claims priority to U.S. Patent Application No. 17/666,997, filed February 08, 2022, and titled “Extracted Image Segments Collage,” which claims priority to U.S. Provisional Patent Application No. 63/291,276, filed December 17, 2021, and titled “Extracted Image Segments Collage,” the contents of each of which are herein incorporated by reference in their entirety.

BACKGROUND

[0002] With the ever expanding amount of accessible digital content available to users and customers, it continues to become more and more difficult for users to organize and maintain information relating to digital content of interest and/or discovered by the user. For example, some systems allow users to maintain links or bookmarks to websites or specific webpages discovered by a user. Other systems also allow users to store images of items discovered by users.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] FIGS. 1A through 1O are representations of a graphical user interface, the creation of a collage of extracted image segments, and a remix of the created collage, in accordance with disclosed implementations.

[0004] FIG. 2 is an example collage process, in accordance with disclosed implementations.

[0005] FIG. 3 is an example image processing subprocess, in accordance with disclosed implementations.

[0006] FIG. 4 is an example image segment modification subprocess, in accordance with disclosed implementations.

[0007] FIG. 5 is an example collage transformation process, in accordance with disclosed implementations.

[0008] FIG. 6 is an example buyable objects collage process, in accordance with disclosed implementations.

[0009] FIG. 7A is an example buyable objects collage, in accordance with disclosed implementations.

[0010] FIG. 7B is an example buyable object detail page, in accordance with disclosed implementations.

[0011] FIG. 8 is an example image segmentation deep neural network update process, in accordance with disclosed implementations.

[0012] FIG. 9 illustrates an example computing device, in accordance with disclosed implementations.

[0013] FIG. 10 illustrates an example configuration of components of a computing device, such as that illustrated in FIG. 9.

[0014] FIG. 11 is a pictorial diagram of an illustrative implementation of a server system that may be used for various implementations.

DETAILED DESCRIPTION

[0015] Described are systems and methods to extract image segments, referred to herein as extracted image segments, from an image and include those image segments in a collage. The origin information, such as the source image, source image location, etc., from which the extracted image segment is generated, is maintained as metadata so that interaction with the extracted image segment on the collage can be used to determine and/or return to the origin of the extracted image segment. For example, if an extracted image segment on a collage originated from an e-commerce website, the address to the e-commerce website may be maintained in metadata of the extracted image segment when generated and added to the collage.

[0016] Extracted image segments may be positioned anywhere on a collage that is presented on a user device. For example, extracted image segments may be visually stacked with respect to other extracted image segments of the collage, extracted image segments may be rotated, extracted image segments may be adjusted in size, etc. Likewise, in some implementations, extracted image segments may be animated or otherwise distinguished when presented as part of a collage when presented.

[0017] In some implementations, an object represented in an extracted image segment may be buyable. For example, a seller of an item represented in an extracted image segment of a collage may be determined and associated with and/or identified in the metadata of the extracted image segment. Likewise, an indicator may be presented with the extracted image segment to indicate that the object represented in the extracted image segment may be purchased from the seller. A user, when viewing the collage, may interact with the extracted image segment and, for example, be redirected to the e-commerce website of the seller of the object and complete a purchase of the object. In other implementations, the user may interact with the extracted image segment and directly purchase the object represented in the extracted image segment.

[0018] As discussed further below, an image segment and/or extracted image segment may be any portion of an image and may correspond to an object represented in the image segment/extracted image segment. In some implementations, an image may be processed by a deep neural network (“DNN”) that is trained to detect object(s) in an image and segment the image such that each object represented in the image corresponds to an image segment of the image. When viewing the image, the image segments determined for an image may be presented such that they are visually distinguished from the image. A user may select an image segment and the pixels of the image corresponding to the selected image segment are extracted to generate an extracted image segment. Likewise, metadata, such as an indication of the image, the location of the image, a link to a website from which the object represented by the extracted image segment can be purchased or obtained, additional information about the object, reviews of the object, a link to a second collage from which the image or the extracted image segment were obtained, a popularity of the extracted image segment, an indication of a user that created the extracted image segment, etc., may be included in the extracted image segment.

[0019] FIGs. 1A through 1O are representations of a graphical user interface, the creation of a collage of extracted image segments, and a remix of the created collage, in accordance with disclosed implementations. As illustrated, the disclosed implementations may be performed in whole or in part on a user device 100 such as a cell phone, smart phone, tablet, wearable, laptop, desktop, etc. In other implementations, portions of the disclosed implementations, such as image processing, image segmentation, and/or extraction of image segments, may be performed on one or more remote computing resources and images, image segments, extracted image segments, collage generation, etc., performed on the user device. As will be appreciated, any variation of processing and/or other operations of the disclosed implementations may be performed on one or many different devices. Likewise, the disclosed implementations, may, for example, be provided as part of a social networking environment, e-commerce environment, or any other form of interactive computing.

[0020] Turning first to FIG. 1A, a user interface 111 is presented on a display of a user device 100. In the illustrated example, the user interface 111 includes a plurality of images that may be viewed and optionally selected by a user through interaction with the user device 100. In the example illustrated with respect to FIG. 1A, the user interface includes three columns of images. A user may view any number of images through the user interface 111 and select one or more images.

[0021] Images may be provided from a remote data store that is accessible to the user device 100, such as a social networking service, the Internet, etc., may be provided from a memory on the user device, may be generated from a camera or other imaging element of the user device, etc. In general, an image may be obtained from any source and utilized with the disclosed implementations.

[0022] In the illustrated example, the user selects image 112, for example through physical interaction with a touch-based display of the user device. In response to selection of the image 112, and turning to FIG. 1B, an image segment of the image 112 may be determined and presented with the image such that the image segment 112-2 is distinguished from the remainder of the image 112-1. In the illustrated example, the image segment 112-2 of the image 112 (FIG. 1A) includes a wine bottle and when presented by the user device, the image segment 112-2 is presented such that it is

visually distinguished from the remainder of the image 112-1. While the example discussed with respect to FIG. 1A and 1B indicates the wine bottle as the image segment 112-2, in other implementations, other image segments, or all image segments, such as an image segment of the wine glass 112-3 and an image segment for the table 112-4, may be determined and visually presented such that the image segments are visually distinguished.

[0023] In some implementations, additional images 124, image segments, and/or extracted image segments, such as images/extracted image segments that are visually similar to the image segment 112-2 may also be presented on the user interface of the device 100 in response to a user selection of an image 112. For example, in some implementations, the popularity or frequency of extracted image segments used on other collages by the same or other users may be monitored and popular or trending extracted image segments presented to the user as additional images 124. Alternatively, or in addition thereto, and as another example, existing extracted images that are similar to other extracted images included on a collage by the user and/or that are determined to be of potential interest to the user may be presented to the user as additional images 124. Other additional images 124 that may be presented include, but are not limited to extracted image segments that enable purchase of an object represented in the extracted image segments, extracted image segments that are related to an extracted image segment of the collage and/or the image segment, extracted image segments generated by the user that selected the image segment, etc.

[0024] In this example, the user interacting with the device selects the image segment 112-2. Upon selection of the image segment 112-2, pixels of the image 112 corresponding to the selected image segment 112-2 are extracted from the image 112 and an extracted image segment that includes the pixels is generated. In addition, as discussed further below, metadata, including but not limited to an indication of the image, the location of the image, a link to a website from which the object represented by the extracted image segment can be purchased or obtained, additional information about the object, reviews of the object, a link to a second collage from which the image or the extracted image segment were obtained, a popularity of the extracted image segment, an indication of a user that created the extracted image segment, etc., may be included in the extracted image segment.

[0025] Referring now to FIG. 1C, upon selection of an image segment and generation of the extracted image segment, the extracted image segment 132 is presented on the display of the device 100 as part of a collage 150. A user may interact with the extracted image segment 132 included on the collage 150. For example, the user may crop the extracted image segment, rotate the extracted image segment, increase/decrease the size of the extracted image segment, etc. In some implementations, the object represented in the extracted image segment 132, may be further processed to determine a three-dimensional (“3D”) mesh of the object such that a user can rotate the extracted image segment in three-dimensions.

[0026] In addition to interacting with the extracted image segment 132, in some implementations, the user may select to lock the extracted image segment so that it cannot be further interacted with, cannot be transformed, the position/size of the extracted image segment cannot be changed, etc., through selection of the lock control 133-1. Alternatively, or in addition thereto, the user may select to generate a duplicate of the extracted image segment 132 through selection of the duplication control 133-2. Finally, if the user decides they do not want to include the extracted image segment 132 in the collage, the user may remove or delete the extracted image segment through selection of the delete control 133-3.

[0027] In the illustrated example and referring to FIG. 1D, the user has adjusted the first extracted image segment 132 by decreasing the size of the extracted image segment 132, rotating the extracted image segment 132 and moving the extracted image segment 132 up and to the left portion of the collage 150.

[0028] In addition to viewing extracted image segments presented on a collage, additional information indicators 142, 144, 146, 148 may also be presented. The additional information indicator 142 may provide information indicating the number of extracted image segments included on the collage, in this example, one. The additional information indicator 144 may provide the opportunity for the user that created the collage to invite a second user to view the collage, for example making the collage a collaborative collage (as discussed below) and to chat with the second user. The additional information indicator 146 may be a re-mix indicator that, when selected by the user, or another user, remixes the presentation of the extracted image segments of the collage. Remixing may include adjusting the position, size, orientation, stack

position, etc. of one or more extracted image segment of a collage. The additional information indicator 148 may be a duplication indicator that, when selected by the user, or another user, causes a duplicate (also referred to as a child copy) of the collage to be generated. Similar to generating a duplicate of a collage in response to a transformation request by another user, as discussed below, a duplicate collage generated in response to selection of the indicator 148 may visually appear the same but the metadata for the collage and image segments may be updated to link back to or otherwise reference the collage from which it was generated.

[0029] Continuing with the current example and referring now to FIG. 1E, after adding a first extracted image segment 132 to the collage, the user may again view any number of images on a user interface of the device 100. As before, the user may scroll through any number and/or source of images and select an image, such as image 152. In this example, upon selection of the image 152 and referring now to FIG. 1F, an image segment 152-2 corresponding to an object represented in the image, in this example a floor-standing lamp, may be determined and presented to the user via the display of the device 100 such that the image segment 152-2 is visually distinguished from the image 152-1. In the illustrated example, as part of processing of the image to determine objects and generate image segments, processing may likewise be performed to determine if a portion of the object of an image segment is occluded by another object in the image 152. In this example, one of the legs 152-3 of the lamp represented in the image segment 152-2 is determined to be occluded. In such an example, one or more image processing algorithms, such as an inpainting algorithm, may be used to determine the look, shape, and position of the occluded part of the object and present the occluded part of the object as part of the image segment 152-2. Likewise, in some implementations, the user interface may also include a modification control 162 that may be selected by the user to modify the image segment by adjusting the pixels of the image that are included in the image segment. Likewise, for occluded portions of an object corresponding to the image segment, the modification control 162 may be utilized by the user to in-paint or correct a portion of the image segment determined through processing of the image segment.

[0030] Referring now to FIG. 1G, upon selection of an image segment and generation of the extracted image segment, the extracted image segment 173 is

presented on the display of the device 100 as part of a collage 150 along with the other extracted image segments of the collage, in this example, extracted image segment 132. A user may interact with the extracted image segment included on the collage 150. For example, the user may crop the extracted image segment 173, rotate the extracted image segment, increase/decrease the size of the extracted image segment, etc. In some implementations, the object represented in the extracted image segment may be further processed to determine a 3D mesh of the object such that a user can rotate the extracted image segment in three-dimensions. Likewise, when the collage 150 has more than one extracted image segment, a stack control 172 may be included that allows the user to move extracted image segments up or down in the stack with respect to other image segments of the collage. For example, in the collage 150 presented in FIG. 1G, the second extracted image segment 173 is visually presented on the top of a stack such that it is presented on top of the first extracted image segment 132. The user may adjust the position of the extracted image segments through selection of an extracted image segment and interaction with the stack control 172, such as selection of the move up stack control 172-2 or the move down stack control 172-1. For example, comparing FIG. 1G to 1H, in FIG. 1G, the second extracted image segment 173 is presented at the top of the stack and in front of the first extracted image segment 132. In response to the user selecting the second extracted image segment 173 and the move down control 172-1, referring now to FIG. 1H, the second extracted image segment is moved down in the stack and presented behind the first extracted image segment 132.

[0031] In the illustrated example and referring to FIG. 1I, the user has adjusted the second extracted image segment 173 by decreasing the size of the extracted image segment 173, moving it to the bottom of the stack such that it is presented behind or beneath the first extracted image segment 132, and moved to be centered behind the first extracted image segment 132 in the upper left corner of the collage 150.

[0032] Continuing with the above example and referring now to FIG. 1J, the user has generated an image 182 using an imaging element of the device 100. In this example, the image 182 includes a representation of a chair and other objects. Upon processing the image, as discussed below, an image segment 182-2 corresponding to the chair is determined and presented such that the image segment is visually distinguished from the remainder of the image 182-1, as illustrated in FIG. 1K.

[0033] Similar to the above, the user interface may include an alteration control 192 that may be selected by the user to alter pixels of the image 182 that are to be included or excluded from the image segment when extracted. For example, in the example illustrated with respect to FIG. 1K, processing of the image 182 to determine the image segment 182-2 incorrectly excluded pixels representative of the rear leg 182-3 of the chair from the image segment 182-2. Through interaction with the alteration control 192, the user may choose to add or remove pixels from the image segment. For example, and referring to FIG. 1L, the user may interact with the alteration control 192 and select to either add pixels to the image segment, by selecting the add alteration control 192-1 or select to remove pixels from the image segment, by selecting the remove alteration control 192-2. In this example, the user has selected the add alteration control 192-1 and through interaction with the image, such as through a touch-based display of the device 100, indicated which portion of the image and thus the pixels to include or add into the image segment 182-2. Specifically, in this example, the user has selected to include pixels corresponding to the leg 182-3 of the chair. Likewise, the user has also selected the remove alteration control 192-2 and selected to remove the pixels corresponding to the space 182-4 between the seat of the chair and the back of the chair that were originally included in the image segment as a result of a processing of the image segment. Upon altering the image segment to include or exclude pixels, in some implementations, image processing of the image, a portion of the image, such as the altered portion and/or the portion corresponding to the object of interest, may be re-run, as discussed further below.

[0034] In some implementations, rather than adjusting an image to include/exclude pixels of an object of interest that is then extracted as an extracted image segment, as discussed herein, a user may select to remove an object from all or a portion of the image. In such an example, the indicated object may be removed from the image or portion of the image and an in-fill or in-painting process, as is known in the art, utilized to assign pixel values to the pixels that previously represented the removed object. As a result, the image may be adjusted to appear as if the object was not included in the image. For example, and referring to FIG. 1K, if the user selected to remove the chair (object) from the image, an in-fill or in-painting process may be utilized to assign pixel values such that the floor and wall of the room represented in the object are represented by the pixels that previously represented the chair.

[0035] After altering the image segment, the user may select the image segment and an extracted image segment 193 may be generated that includes the pixels of the image corresponding to the image segment 182-2 and metadata for the image segment. Likewise, the extracted image segment 193 is presented on the collage 150 with the other extracted image segments 132, 173, and the user may adjust the extracted image segment, as discussed. Referring now to FIG. 1M, in this example, the user has adjusted the third extracted image segment 193 to be positioned near the first extracted image segment 132 and the second extracted image segment 173, as well as in-front of the second extracted image segment 173. In addition, the image's additional information 142 now includes metadata for each of the three extracted image segments 132, 173, 193 that may be selected and viewed by the user, or other viewers of the collage. Likewise, the image's additional information 142 now indicates that there are three extracted images included in the collage 150.

[0036] A user may go through the process of extracting image segments and including extracted images segments onto the collage 150 for any number of extracted image segments, each of which may be placed anywhere on the collage. Likewise, in some implementations, the user may draw or write on the collage and/or choose to animate one or more of the extracted image segments. Referring to FIG. 1N, illustrated is a collage that includes six extracted image segments 132, 173, 193, 195, 196, 197, and a typed note 198 ("MY FIRST COLLAGE") added by the user. Likewise, the user has selected animation 170-2 for the star extracted image segment 196 so that the star extracted image segment has a flashing animation effect. Additionally, the user has added animation to the cat extracted image segment 197 that includes both movement 170-3 of the cat extracted image segment 197 and audio (e.g., "MEOW") animation 170-1. Any of a variety of forms of animation may be added to an extracted image segment included on a collage. For example, animation may include, but is not limited to, color changes, movement, flashing, sound, haptics, etc.

[0037] FIG. 1O is an example illustration of the collage 150 illustrated in FIG. 1N after selection of the remix control 146, in accordance with disclosed implementations.

[0038] As illustrated, in response to selection of the remix control 146, the position of the extracted image segments 132, 173, 193, 195, 196, 197 on the collage and with respect to each other having been re-arranged or remixed. In some implementations,

the rearrangement or remixing of the extracted image segments may be random. In other implementations, rearrangement may be based on, for example, a popularity of the extracted images segments, user preference, cross-pattern configuration, layout, etc.

[0039] FIG. 2 is an example collage process 200, in accordance with disclosed implementations.

[0040] The example process 200 begins upon receipt of an image, as in 202. As discussed above, the image can be from any source such as a camera or other imaging element, from a website, from photos stored in a memory of a user device or stored in memory that is accessible by the user device (local or remote), a video frame from a video, etc. Likewise, in some implementations, the image received by the example process may already be an extracted image segment. For example, in some implementations, the popularity or frequency of extracted image segments used on other collages by the same or other users may be monitored and popular or trending extracted image segments presented to a user for selection and inclusion in the collage. Alternatively, or in addition thereto, and as another example, existing extracted image segments that are similar to other extracted image segments included in a collage by a user and/or that are determined to be of potential interest to the user may be presented and/or selected by the user as the image.

[0041] A determination may then be made as to whether a region of interest is indicated by the user, as in 204. For example, in addition to receiving an image, a user may indicate, for example through interaction with the image, a region or portion of the image that is of interest to the user. If it is determined that a region of interest is indicated, the portion of the image included in the indicated region of interest is provided as the image, as in 206. If it is determined that a region of interest is not provided, or after providing the portion of the image included in an indicated region of interest as the image, the example image processing subprocess 300 is performed on the image, as in 300. The example image processing subprocess 300 is discussed in more detail below with respect to FIG. 3. As discussed below, the image processing subprocess 300 generates and returns one or more image segments of an image, each image segment corresponding to an object represented in the image. In some implementations, the image processing subprocess 300 may be performed independent of the example process 200 such that image segments are already determined for and

associated with an image. In such an example, upon receipt of the image, the image segments already determined for the image may be utilized without again processing the image with the image processing subprocess.

[0042] One or more of the image segments returned by the image processing subprocess may then be presented to a user such that the image segment(s) are distinguished from the rest of the image, as in 208. An example of a presentation of an image segment such that the image segment is distinguished from other portions of the image is illustrated in FIGs. 1B, 1F, 1K, and 1L, and discussed above. In some implementations, all image segments determined for an image may be presented to the user as distinguished from other portions of the image. In other examples, only a subset (e.g., 3-5) of the image segments determined for an image may be presented to the user as distinguished from other portions of the image. In still other examples, only a single image segment, such as a primary or central image segment of the image, may be presented to the user as distinguished from other portions of the image.

[0043] After presenting the image segment(s), a determination is made as to whether a modification to a presented image segment has been received, as in 210. As discussed above with respect to FIGs. 1F, 1K, and 1L, a user may interact with an image segment to modify the image segment to include/exclude pixels from the image segment. For example, if initial processing of the image inaccurately included/excluded a portion of an object represented in an image segment, a user may interact with the image segment to adjust the pixels included/excluded from the image segment such that the image segment corresponds to the object of the image segment.

[0044] If it is determined that a modification to the image segment is received, the example image segment modification subprocess may be performed, as in 400 (FIG. 4). The example image segment modification subprocess 400 is discussed further below with respect to FIG. 4. Upon completion of the image segment modification subprocess 400, the example process 200 returns to block 208 and continues.

[0045] If it is determined that a modification to an image segment is not received, a determination is made as to whether a selection of an image segment of the image has been received, as in 212. If a selection of an image segment has not been received, the example process 200 returns to block 210 and continues. If a selection of an image

segment is received, pixel data of the selected image segment and corresponding metadata are extracted and used to create an extracted image segment for the selected image segment, as in 216. As discussed above, the metadata may include, but is not limited to, an indication of the image from which the image segment was extracted, the location of the image from which the image segment was extracted, a link to a website from which the object represented by the extracted image segment can be purchased or obtained, additional information about the object, reviews of the object, a link to a second collage from which the image or the extracted image segment was obtained, a popularity of the extracted image segment, an indication of a user that created the extracted image segment, etc. The metadata included in the extracted image segment may be used for attribution information with respect to the extracted image segment, to enable purchase of the object represented in the extracted image segment, etc.

[0046] The extracted image segment may also be presented on a collage, as in 218. If this is the first extracted image segment of the collage, the extracted image segment may be presented on a blank collage. If other extracted image segments are already included on the collage, the extracted image segment may be initially presented in the center of the collage such that the user can adjust the size, orientation, position, etc., of the image in the collage.

[0047] After presenting the extracted image segment on the collage, a determination is made as to whether any adjustments to the extracted image segment have been received, as in 220. Adjustments may include, for example, adjustments to the size, position, orientation, and/or rotation of the extracted image segment, and/or animation of the extracted image segment.

[0048] If it is determined that an adjustment to the extracted image segment has been received, the extracted image segment is adjusted in accordance with the received adjustment, as in 222. After adjusting the extracted image segment, the example process 200 returns to decision block 220 and continues. If it is determined that an adjustment to the extracted image segment has not been received, the collage of extracted image segments is presented, as in 224, and a determination is made as to whether another extracted image segment is to be added to the collage, as in 226. As discussed, any number of extracted image segments may be added to a collage. If it is determined that another extracted image segment is to be added to the collage, the

example process 200 returns to block 202 and continues with receipt of another image. If it is determined that another extracted image segment is not to be added to the collage, the example process 200 completes, as in 228.

[0049] FIG. 3 is an example image processing subprocess 300, in accordance with disclosed implementations.

[0050] The example subprocess 300 begins by segmenting an image, in 302. Any variety of segmentation techniques, such as circle packing algorithm, super-pixels, etc., may be used. The segments may then be processed to remove background portions of the image from consideration, in 304. Determining background segments may be done, for example, using a combination of attentive constraints (e.g., salient objects are likely to be at the center of the image) and unique constraints (e.g., salient objects are likely to be different from the background). In one implementation, for each segment (S_i), a unique constraint may be computed using a combination of color, texture, shape and/or other feature detection. The pairwise Euclidian distances for all pairs of segments: $L2(S_i, S_j)$ may also be computed for $\forall S_i \in S, \forall S_j \in S$. The unique constraint U for segment S_i , or U_i , may be computed as $U_i = \sum_j L2(S_i, S_j)$. The attentive constraint for each Segment S_i may be computed as $A = [X(s) - X']^2 + [Y(s) - Y']^2$, where X' and Y' are the center coordinates of the image.

[0051] One or more of the segments S' , a subset of S , may then be selected such that $U(s) - A(s) > t$, where t is a threshold set manually or learned from the data. The threshold t may be any defined number or amount utilized to distinguish segments as background information or potential objects. Alternatively, $Similarity(s'_i \in S', r_i \in R-)$ and $Similarity(s'_i \in S', r_i \in R+)$, where s'_i is an element of S' and r_i is an element $R-$, and $R-$ is a set of image non-salient regions (background), may be computed and used as the similarity between each segment to a labelled database of labelled salient segments and non-salient segments.

[0052] Returning to FIG. 3, upon removing the background segments, the objects remaining in the image are determined, in 306. Objects remaining in the image may be determined, for example, by using a sliding window approach to compute the score of each possible hypothesis of the location of the object. Using approaches such as boosted selection of Haar-like wavelets, or multiple-parts based models, each segment

may be processed to determine potentially matching objects. For example, an image vector may be determined for a segment and compared to information stored for objects. Based on the image vector and the stored information, a determination may be made as to how similar the image vector is to the stored information for particular objects.

[0053] The DNN may perform a sliding window approach N times, each with a different trained object classifier (e.g., person, bag, shoes, face, arms, etc.). After determining a hypothesis for each object classifier, the output is a set of best hypotheses for each object class. Because objects do not generally appear randomly in images (e.g., eyes and noses typically appear together), position-sensitive constraints may also be considered. For example, positions of the root object (e.g., person) may be defined as $W(\text{root})$ and each geometric constraint for each object k may be denoted with respect to each other as $\lambda_k^{\rightarrow}(O)$, a 6-element vector. The geometric “fit” of each landmark W_{oi} with respect to the root object W_{root} may be defined by $L(\vec{\lambda}(i)(W_o) * \theta)$, $\theta_i = [dy, dx^2, dy^2, dxdy, z]$, where dx, dy are the average geometric distance between each pixel in the object box W_{oi} and each pixel in the root object box. The problem of finding optimal $\lambda_k^{\rightarrow}$ can be formulated as, $\arg \min_{\lambda_i} L(\vec{\lambda}(i) * D_{\text{train}}(\theta_i))$ where $D_{\text{train}}(\theta_i)$ is the observed value of θ_i in training images.

[0054] To optimize this function, the location of the objects in the image may be determined, in 308. For example, the center of a root object (e.g., person) in the image is marked as (0, 0), and the location of other objects in the processed images is shifted with respect to the root object. A linear-Support Vector Machine (SVM) is then applied with θ_i as parameters. The input to the SVM is $D_{\text{train}}(\theta_i)$. Other optimizing approaches, such as linear programming, dynamic programming, convex optimizations, and the like, may also be used alone or in combination with the optimization discussed herein. The training data $D_{\text{train}}(\theta_k)$, can be collected by having users place a bounding box on top of both the entire object and the landmarks. Alternatively, semi-automated approaches, such as facial detection algorithms, edge detection algorithms, etc., may be utilized to identify objects. In some implementations, other shapes, such as ovals, ellipses, and/or irregular shapes may be used to represent objects.

[0055] Finally, image segments for each detected object are maintained, as in 310. As will be appreciated, the example subprocess 300 of processing images may be performed by a trained DNN that processes an image to generate image segments corresponding to objects represented in the image. For example, a DNN such as a convolution neural network may be trained, for example using labeled and/or unlabeled data, to process an input image and output one or more image segments of the image corresponding to objects detected in the image. Likewise, as discussed further below, as image segments are adjusted by users, those adjusted image segments and corresponding images may be utilized as additional labeled training data to continue training the DNN, thereby further improving the accuracy of the DNN based on user provided inputs.

[0056] FIG. 4 is an example image segment modification subprocess 400, in accordance with disclosed implementations.

[0057] The example process 400 begins by adjusting the image segment based on user input, such as through a touch-based display, to include and/or exclude pixels from the image, thereby generating an adjusted image segment, as in 402. For example, as discussed above with respect to FIGs. 1K and 1L, a user may interact with the disclosed implementations to add pixels to an image segment, remove pixels from the image segment originally determined and presented to the user, and/or remove an object from the image by replacing pixel values using an in-fill or in-painting process. For example, as illustrated in FIGs. 1K and 1L, the originally determined image segment 182-2 for the chair excluded pixels of the image that represent the rear leg 183-3 of the chair. Likewise, the original image segment included pixels of the image corresponding to the space 182-4 between the seat of the chair and the back of the chair. A user may interact with the presentation of the image to include pixels corresponding to the rear leg 182-3 of the chair in the image segment 182-2 and exclude pixels corresponding to the space 182-4 between the back of the chair and the seat of the chair.

[0058] A determination may then be made as to whether the adjusted segment is to be again processed to identify object(s) included in the adjusted image segment, as in 404. If it is determined that the adjusted image segment is to be processed to determine the object included in the image segment, the example image processing

subprocess 300 discussed above with respect to FIG. 3 is performed with the adjusted image segment.

[0059] After processing the adjusted image segment or if it is determined that the adjusted image segment is not to be again processed, metadata for the adjusted image segment is updated to include/exclude an indication of the pixels to/from the metadata, as in 406. Likewise, if the image is processed again, information resultant from the example process 300 may be updated in the metadata for the image segment. Finally, the adjusted image segment, or data corresponding to the adjusted image segment is returned, as in 408.

[0060] FIG. 5 is an example collage transformation process 500, in accordance with disclosed implementations.

[0061] The example collage transformation process 500 begins by presenting a collage that includes one or more extracted image segments, as in 502. For example, a collage, such as the collage 150 illustrated and discussed above with respect to FIG. 1N, which includes six image segments and a text input, may be presented to a user. In some implementations, a collage may be designed as private such that only the user that created the collage may view and/or transfer the collage. In other implementations, the user may designate the collage as collaborative and invite other users to view and/or collaborate on the collage. Collaboration between users and a collage may be in real-time or near real-time such that each of the users collaborating on the collage can see changes to the collage and/or chat about the collage/changes to the collage. In other examples, collaboration may be incremental such that a second user may transform the collage and the first user may later view the collage and see the changes made by the second user.

[0062] In still other examples, the user may make the collage public such that any user may view the collage. A collaborative collage is a collage in which an invited user, or if allowed by the collage creator, any other users other than the creator of the collage, may modify the collage.

[0063] After presenting the collage, a transformation request to transform one or more aspects of the collage may be received, as in 504. A transformation request may be any input to transform one or more aspects of the collage, such as an extracted image

segment of the collage. For example, a transformation request may include, but is not limited to, a request to remix the visual placement and presentation of the extracted image segments of the collage, a request to add an extracted image segment to the collage, a request to remove an extracted image segment from the collage, a request to adjust a size, shape, and/or position of an extracted image segment of the collage, a request to add, remove, or change an animation of an extracted image segment of the collage, etc.

[0064] In response to receiving the transformation request, a determination is made as to whether the transformation request is from the creator of the collage (a first user), as in 506. For example, a user identifier or user identifier that is associated with an application executing on a user device that is used to create the collage may be indicated as the creator of the collage. If the user is utilizing the same user device, another user device associated with the user or the user account, or otherwise accessing the user account, it may be determined that the transformation request was from the creator of the collage.

[0065] If it is determined that the request is from the creator of the collage, the collage is transformed in accordance with the transformation request, as in 508. If it is determined that the transformation request is not from the creator of the collage, a determination is made as to whether the collage is a collaborative collage, as in 509. As noted above, the creator of a collage may indicate a collage as collaborative such that other users may transform the collage. In such an example, the collage may be transformed by the user and/or other users and those transformations to the collage may be presented to the user and/or the other users. If it is determined that the collage is a collaborative collage, the collage is transformed in accordance with the transformation request, as in 508. In some implementations, the user may specify which other users may transform the collage, such that the collage is only considered a collaborative collage for those specific users. For any other user that submits a transformation request to the collage, a duplicate collage may be generated, as discussed below, for which the transformation request may be applied such that the transformation does not impact the collage generated by the user.

[0066] If it is determined that the collage is not a collaborative collage or not a collaborative collage for the user that submitted the transformation request, a duplicate

collage is generated for the other user, referred to herein as a second user, as in 510. A duplicate collage may include the same extracted image segments in the same position, orientation, size, etc., as the collage such that the user transforming the collage cannot determine the difference between the duplicate collage and the collage. However, the metadata of the collage and each extracted image segment may be updated to indicate that the collage is a duplicate collage and include information, a link, and/or other reference to the collage from which the duplicate was generated, as in 512. Likewise, the metadata of each extracted image segment may be updated to indicate the original collage as a source of the extracted image segment. Such information may be in addition to any source information already included in the metadata for the original collage and/or the extracted image segments.

[0067] Finally, the duplicate collage may be transformed in accordance with the received transformation request, as in 514. The duplicate collage becomes another collage maintained by the system, the second user is identified as the creator of the duplicate collage, and there is a link or other reference maintained between the duplicate collage, the original collage, as well as any other source information for extracted image segments included in the collage and/or the duplicate collage. Likewise, the second user may transform the duplicate collage without transforming the original collage. In addition, the second user may also specify the duplicate collage as a private collage, a duplicate collage, etc., just as if the second user had been the original creator of the duplicate collage.

[0068] FIG. 6 is an example buyable objects collage process 600, in accordance with disclosed implementations.

[0069] The example process 600 begins by determining an object represented by an extracted image segment that is included in a collage, as in 602. For example, any of a plurality of image processing algorithms or DNNs may be utilized to process an image and detect an object, or an object type represented in the image. Alternatively, or in addition thereto, metadata about the extracted image segment may be utilized to determine an object represented in the extracted image segment. For example, if the extracted image segment is originally obtained from a website, the metadata of that extracted image segment may include an indication of the object represented in the extracted image segment.

[0070] In addition to determining the object represented in the extracted image segment, one or more sellers of the object may be determined, as in 604. For example, if the extracted image segment was originally obtained from a website, such as an e-commerce website, metadata of the extracted image segment may indicate the seller of the object. In other examples, sellers of objects may provide information, such as catalogs indicating objects offered for sale by that seller. In still other examples, websites of sellers may be processed to determine objects offered for sale by those sellers and that information used to determine one or more sellers of the object represented in the extracted image segment. In still another example, a seller or other user may provide an indication of the seller of the object represented in the image segment.

[0071] Each determined seller may then be associated with the extracted image segment, as in 606. For example, if the seller corresponds to an e-commerce website, a detail page for the object may be associated with the extracted image segment, thereby indicating the seller of the object.

[0072] In response to determining one or more sellers of the object represented in the image segment, a buyable indication may be presented with the extracted image segment as part of the collage, as in 608. For example, FIG. 7A is an illustration of a collage 740 presented on a user device 700 that includes buyable indicators, in accordance with the disclosed implementations. In this example, the collage 740 includes five extracted image segments 743-1, 743-2, 743-4, 743-5, 743-6, and a typed text input 743-3 of “MY CHRISTMAS LIST.” The extracted image segments of the collage 740 may be processed by the example process 600 and a determination made that extracted image segments 743-1 (bicycle), 743-2 (cowboy hat), and 743-5 (book) correspond to buyable objects. As such, a buyable indication 745-1, 745-2, and 745-3 are presented next to the respective extracted image segment. In this example, the object of a sweater that is represented by the extracted image segment 743-4 may have been previously indicated as buyable and now indicated as purchased, through presentation of the purchased indicator 747. For example, if the user that created the collage purchases an object represented by an extracted image segment, the buyable indicator for that extracted image segment may be replaced with a purchased indicator 747, indicating that the item has been purchased.

[0073] In other examples, a collage may be created by a first user and shared with other users to indicate items the first user would like to receive, such as Christmas gifts, birthday gifts, wedding gifts, etc., in accordance with the disclosed implementations. In such an example, the collage may be shared with one or more other users. The one or more other users may interact with the collage 740 and optionally purchase items corresponding to extracted image segments included in the collage. In such an example, as items are purchased or otherwise obtained, the buyable indicator may change to a purchased indicator, thereby indicating to other users that the item has already been purchased for the first user.

[0074] Returning to FIG. 6, after presenting a buyable indicator in the collage with the corresponding extracted image segment, a determination may be made as to whether another extracted image segment of the collage remains that is to be processed by the example process 600, as in 610. If it is determined that additional extracted image segments of the collage remain, a next extracted image segment is selected, as in 612, the example process 600 returns to block 602, and continues. If it is determined that no additional extracted image segments of the collage remain, the example process 600 completes, as in 614.

[0075] In some implementations, if a user selects one of the extracted image segments that are indicated as buyable, such as the extracted image segment 743-1, a buyable object detail page corresponding to the object represented by the extracted image segment may be presented.

[0076] For example, FIG. 7B is an example buyable object detail page, in accordance with disclosed implementations. In particular, in response to a user selecting the extracted image segment 743-1 of the bicycle, the buyable object detail page 755 is presented that includes additional information about the object represented by the selected extracted image segment, in this example the extracted image segment 743-1. For example, the buyable object detail page 755 may include an indication of the price 755-1 of the object, a delivery timeframe 755-2 of the object when purchased, a seller 755-3 of the object, etc. Additionally, the buyable object detail page 755 may include a purchase control, such as a “Buy” button 757 that, when selected, enables a purchase of the object represented by the extracted image segment 743-1 from the seller 755-3, in this example, Company A.

[0077] FIG. 8 is an example image segmentation deep neural network update process 800, in accordance with disclosed implementations. As discussed above, in some implementations, the DNN used to determine image segments of an image may be continually or periodically updated as image segments are adjusted by users to include/exclude pixels of those image segments. In some implementations, if an image segment is adjusted on a user device, the adjusted image segment may be sent to a remote computing resource and compiled with other user feedback (other image segment adjustments) and the adjusted image segments used for ongoing training to update the DNN. As a DNN is updated, the updated DNN may be sent to user devices for operation on those user devices. Alternatively, the updated DNN may reside on one or more remote computing resources and operate on those remote computing resources.

[0078] The example process 800 begins by initially training a DNN to generate one or more image segments for an input image, as in 802. In some implementations, the DNN may be trained to perform the image processing subprocess 300 discussed above with respect to FIG. 3. In other implementations, other training techniques or processes may be used to train a DNN to receive an input image and determine one or more image segments corresponding to objects represented in the input image.

[0079] At some point after the DNN is initially trained, one or more adjusted image segments may be obtained based on user input that caused the adjustment to image segments originally determined by the DNN, as in 804. With a significantly large set of users, a large set of adjusted image segments may be received as different users interact with images and image segments determined and presented in accordance with the disclosed implementations.

[0080] The adjusted image segments and the corresponding image may be utilized as labeled training data for the DNN. Accordingly, the adjusted image segments may be used to update the DNN, as in 806.

[0081] FIG. 9 illustrates an example user device 900 that can be used in accordance with various implementations described herein. In this example, the user device 900 includes a display 902 and optionally at least one input component 904, such as a camera, on a same side and/or opposite side of the device as the display 902. The user device 900 may also include an audio transducer, such as a speaker 906, and optionally

a microphone 908. Generally, the user device 900 may have any form of input/output components that allow a user to interact with the user device 900. For example, the various input components for enabling user interaction with the device may include a touch-based display 902 (e.g., resistive, capacitive, Interpolating Force-Sensitive Resistance (IFSR)), camera (for gesture tracking, etc.), microphone, global positioning system (GPS), compass or any combination thereof. One or more of these input components may be included on a user device or otherwise in communication with the user device. Various other input components and combinations of input components can be used as well within the scope of the various implementations as should be apparent in light of the teachings and suggestions contained herein.

[0082] In order to provide the various functionality described herein, FIG. 10 illustrates an example set of basic components 1000 of a user device, such as the user device 900 described with respect to FIG. 9 and discussed herein. In this example, the device includes one or more processors 1002 for executing instructions that can be stored in at least one memory device or element 1004. As would be apparent to one of ordinary skill in the art, the device can include many types of memory, data storage or computer-readable storage media, such as a first data storage for program instruction for execution by the one or more processors 1002. Removable storage memory can be available for sharing information with other devices, etc. The device typically will include some type of display 1006, such as a touch-based display, electronic ink (e-ink), organic light emitting diode (OLED), liquid crystal display (LCD), etc.

[0083] As discussed, the device in many implementations will include at least one image capture element 1008, such as one or more cameras that are able to image objects in the vicinity of the device. An image capture element can include, or be based at least in part upon, any appropriate technology, such as a CCD or CMOS image capture element having a determined resolution, focal range, viewable area, and capture rate. The device can include at least one application component 1010 for performing the implementations discussed herein, such as the generation of collages. The user device may be in constant or intermittent communication with one or more remote computing resources and may exchange information, such as collages, extracted image segments, transformed image segments, metadata, updated DNNs, etc., with the remote computing system(s) as part of the disclosed implementations.

[0084] The device also can include at least one location component, such as GPS, NFC location tracking, Wi-Fi location monitoring, etc. Location information obtained by the location component may be used with the various implementations discussed herein as a factor in, for example, determining a seller of an object represented in an extracted image segment. For example, if the user is located in a Store A department store and generates an extracted image segment from an image generated by the image capture element 1008 of the user device while located in the Store A department store, the location information may be used as a factor in determining a seller of an object represented in the extracted image segment.

[0085] The user device may also include a DNN 1012, as discussed herein, that is operable to receive an image as an input and determine one or more image segments corresponding to objects represented in the input image. Likewise, the user device may also include a collage management component 1014 that maintains, for example, collages created and/or viewed by the user of the user device, extracted image segments, etc., and/or performs some or all of the implementations discussed herein.

[0086] The example user device may also include at least one additional input device able to receive conventional input from a user. This conventional input can include, for example, a push button, touch pad, touch-based display, wheel, joystick, keyboard, mouse, trackball, keypad or any other such device or element whereby a user can submit an input to the device. These I/O devices could be connected by a wireless, infrared, Bluetooth, or other link as well in some implementations. In some implementations, however, such a device might not include any buttons at all and might be controlled only through touch inputs (e.g., touch-based display), audio inputs (e.g., spoken), or a combination thereof.

[0087] FIG. 11 is a pictorial diagram of an illustrative implementation of a server system 1100, such as a remote computing resource, that may be used with one or more of the implementations described herein. The server system 1100 may include one or more processors 1101, such as one or more redundant processors, a video display adapter 1102, a disk drive 1104, an input/output interface 1106, a network interface 1108, and a memory 1112. The processor(s) 1101, the video display adapter 1102, the disk drive 1104, the input/output interface 1106, the network

interface 1108, and the memory 1112 may be communicatively coupled to each other by a communication bus 1110.

[0088] The video display adapter 1102 provides display signals to a local display permitting an operator of the server system 1100 to monitor and configure operation of the server system 1100. The input/output interface 1106 likewise communicates with external input/output devices not shown in FIG. 11, such as a mouse, keyboard, scanner, or other input and output devices that can be operated by an operator of the server system 1100. The network interface 1108 includes hardware, software, or any combination thereof, to communicate with other computing devices. For example, the network interface 1108 may be configured to provide communications between the server system 1100 and other computing devices, such as the user device 900.

[0089] The memory 1112 generally comprises random access memory (RAM), read-only memory (ROM), flash memory, and/or other volatile or permanent memory. The memory 1112 is shown storing an operating system 1114 for controlling the operation of the server system 1100. The server system 1100 may also include a trained DNN 1116, as discussed herein. In some implementations, the DNN may determine object segments on the server. In other implementations, the DNN 1012 (FIG. 10) may determine image segments on a user device. In still other examples, a DNN may exist on both the server and each user device.

[0090] The memory 1112 additionally stores program code and data for providing network services that allow user devices 900 and external sources to exchange information and data files with the server system 1100. The memory 1112 may also include a collage management application 1118 that maintains collage and/or collage information for different users that utilize the disclosed implementations. The collage management application 1118 may communicate with a data store manager application 1120 to facilitate data exchange and mapping between the data store 1103, user devices, such as the user device 900, external sources, etc.

[0091] As used herein, the term “data store” refers to any device or combination of devices capable of storing, accessing and retrieving data, which may include any combination and number of data servers, databases, data storage devices and data storage media, in any standard, distributed or clustered environment. The server

system 1100 can include any appropriate hardware and software for integrating with the data store 1103 as needed to execute aspects of one or more applications for the user device 900, the external sources, etc.

[0092] The data store 1103 can include several separate data tables, databases or other data storage mechanisms and media for storing data relating to a particular aspect. For example, the data store 1103 may include digital items (e.g., images) and corresponding metadata (e.g., image segments, popularity, source) about those items. Collage data and/or user information and/or other information may likewise be stored in the data store.

[0093] It should be understood that there can be many other aspects that may be stored in the data store 1103, which can be stored in any of the above listed mechanisms as appropriate or in additional mechanisms of any of the data store. The data store 1103 may be operable, through logic associated therewith, to receive instructions from the server system 1100 and obtain, update or otherwise process data in response thereto.

[0094] The server system 1100, in one implementation, is a distributed environment utilizing several computer systems and components that are interconnected via communication links, using one or more computer networks or direct connections. However, it will be appreciated by those of ordinary skill in the art that such a system could operate equally well in a system having fewer or a greater number of components than are illustrated in FIG. 11. Thus, the depiction in FIG. 11 should be taken as being illustrative in nature and not limiting to the scope of the disclosure.

[0095] Implementations disclosed herein may include a computer-implemented method that includes one or more of processing a first image to determine a first image segment that corresponds to less than all of the first image and corresponds to an object represented in the first image, presenting, on a user device, the first image and the first image segment such that the first image segment is visually distinguished from the first image, and in response to a selection of the first image segment, forming a first extracted image segment that includes at least pixel data corresponding to pixels of the first image segment, and metadata indicating at least one of the first image or a source location of the first image. The computer implemented may also include one or more

of presenting the first extracted image segment as part of a collage that includes at least a second extracted image segment representative of a second object extracted from a second image, and wherein the first extracted image segment may be transformed with respect to the second extracted image segment, the first extracted image segment may be transformed with respect to a size of the first extracted image segment when presented on the collage, and the first extracted image segment may be transformed with respect to a position of the first extracted image segment when presented on the collage.

[0096] Optionally, the computer-implemented method may further include receiving a modification of the first image segment to modify which pixels of the first image are included in the first image segment. Optionally, the metadata further includes one or more of: a link to a website from which the object can be purchased, additional information about the object, reviews of the object, a link to a second collage that includes the first extracted image segment, a popularity of the first extracted image segment, a comment corresponding to the first extracted image segment, a rating of the first extracted image segment, an indication of other collages that include the first extracted image segment, or an indication of a user that created the first extracted image segment. Optionally, processing the first image may include processing the first image to determine a plurality of image segments, wherein the first image segment is included in the plurality of image segments, and/or presenting includes presenting, on the user device, each of the plurality of image segments such that each of the plurality of image segments are visually distinguished from the first image. Optionally, the computer-implemented method may further include receiving an indication of a region of interest within the first image, and wherein processing includes processing a portion of the first image within the region of interest to determine the first image segment. Optionally, processing the first image may be performed on the user device.

[0097] Implementations disclosed herein may include a computing system that includes one or more processors and a memory storing program instructions. The program instructions, when executed by the one or more processors, may cause the one or more processors to at least, in response to receipt of a first selection of a first image segment of a first image and generate a first extracted image segment that includes at least pixel data corresponding to pixels of the first image segment and metadata

indicating at least one of the first image or a source location of the first image. The program instructions may further cause the one or more processors to at least present the first extracted image segment as part of a collage, receive a second selection of a second extracted image segment, wherein the second extracted image segment corresponds to a second image, and present the second extracted image segment with the first extracted image segment as part of the collage, wherein the first extracted image segment and the second extracted image segment are independently transformable when presented as part of the collage.

[0098] Optionally, the program instructions, that when executed by the one or more processors, may further cause the one or more processors to at least present a plurality of extracted image segments, wherein the second extracted image segment is included in the plurality of extracted image segments and selected from the plurality of extracted image segments. Optionally, at least one of the plurality of extracted image segments may be at least one of a popular extracted image segment, a trending extracted image segment, an extracted image segment that enables a purchase of an object represented in the extracted image segment, visually similar to the first extracted image segment, related to the first extracted image segment, or an extracted image segment generated by a user that selected the first extracted image segment. Optionally, the program instructions, that when executed by the one or more processors, may further cause the one or more processors to at least determine a seller of an object represented in the first extracted image segment, associate the seller with the first extracted image segment, and present, with the first extracted image segment and concurrent with the presentation of the first extracted image segment and the second extracted image segment as part of the collage, a buyable indication indicating that the object represented in the first extracted image segment is buyable through an interaction with the first extracted image segment. Optionally, the program instructions, that when executed by the one or more processors, may further cause the one or more processors to at least receive a request to transform the collage, determine that the request is from a second user that is different than a first user that created the collage, in response to determination that the request is from the second user: generate a duplicate collage from the collage, maintain, as part of a metadata for the duplicate collage, a link or other identifier indicating that the duplicate collage was generated based on the collage, and transform the duplicate collage in response to the request. Optionally, the request may be at least one of a

request to add a third extracted image segment to the collage, remove the second extracted image segment from the collage, re-mix a presentation of the first extracted image segment of the collage and the second extracted image segment of the collage, transform a position of the first extracted image segment, or transform a size of the first extracted image segment. Optionally, the program instructions, that when executed by the one or more processors, may further cause the one or more processors to at least receive a request to transform the collage, determine that the request is from a user that created the collage or that the collage is a collaborative collage, and in response to determination that the request is from the user or that the collage is a collaborative collage, transform the collage in response to the request. Optionally, the program instructions, that when executed by the one or more processors, may further cause the one or more processors to at least receive a modification to the first image segment to modify which pixels of the first image are included in the first image segment, in response to the modification, generate, based at least in part on the modification, a modified image segment, and update, based at least in part with the modified image segment, a training of a deep neural network (“DNN”), wherein the DNN is trained to generate one or more image segments from an input image. Optionally, the program instructions, that when executed by the one or more processors, may further cause the one or more processors to at least, in response to the modification, reprocess the first image segment to determine an object represented in the first image segment.

[0099] Implementations disclosed herein may include a method that includes one or more of processing, with a DNN, a first image to determine a first image segment corresponding to an object represented in the first image, wherein the first image segment is less than all of the first image, presenting, on a display, the first image segment and the first image such that the first image segment is visually distinguished from the first image, in response to receiving a modification of the first image segment to modify which pixels of the first image are included in the first image segment, generating an updated first image segment, in response to a selection of the updated first image segment, forming a first extracted image segment that includes pixel data corresponding to pixels of the updated first image segment, and presenting the first extracted image segment as part of a collage that includes at least a second extracted image segment representative of a second object extracted from a second image, wherein the first extracted image segment is adjustable with respect to the second

extracted image segment, the first extracted image segment is adjustable with respect to a size of the first extracted image segment when presented on the collage, and the first extracted image segment is adjustable with respect to a position of the first extracted image segment when presented on the collage. Optionally, the method may further include updating a training of the DNN based at least in part on the first image and the updated first image segment. Optionally, the display may be included on a user device, and/or the DNN may operate on the user device. Optionally, the first extracted image segment may include metadata, wherein the metadata may include one or more of an indication of the first image, an indication of a source location of the first image, a link to a website from which the object can be purchased, additional information about the object, reviews of the object, a link to a second collage that includes the first extracted image segment, a popularity of the first extracted image segment, or an indication of a user that created the first extracted image segment. Optionally, the method may further include presenting, as part of the collage, an indication of a total number of extracted image segments included in the collage.

[0100] The above aspects of the present disclosure are meant to be illustrative. They were chosen to explain the principles and application of the disclosure and are not intended to be exhaustive or to limit the disclosure. Many modifications and variations of the disclosed aspects may be apparent to those of skill in the art. Persons having ordinary skill in the field of computers, communications, media files, and machine learning should recognize that components and process steps described herein may be interchangeable with other components or steps, or combinations of components or steps, and still achieve the benefits and advantages of the present disclosure. Moreover, it should be apparent to one skilled in the art that the disclosure may be practiced without some, or all of the specific details and steps disclosed herein.

[0101] Moreover, with respect to the one or more methods or processes of the present disclosure shown or described herein, including but not limited to the flow charts shown in FIGs. 2 through 6 and 8, orders in which such methods or processes are presented are not intended to be construed as any limitation on the claims, and any number of the method or process steps or boxes described herein can be combined in any order and/or in parallel to implement the methods or processes described herein. In

addition, some process steps or boxes may be optional. Also, the drawings herein are not drawn to scale.

[0102] Aspects of the disclosed system may be implemented as a computer method or as an article of manufacture such as a memory device or non-transitory computer readable storage medium. The computer readable storage medium may be readable by a computer and may comprise instructions for causing a computer or other device to perform processes described in the present disclosure. The computer readable storage media may be implemented by a volatile computer memory, non-volatile computer memory, hard drive, solid-state memory, flash drive, removable disk, and/or other media. In addition, components of one or more of the modules and engines may be implemented in firmware or hardware.

[0103] Disjunctive language such as the phrase “at least one of X, Y, or Z,” or “at least one of X, Y and Z,” unless specifically stated otherwise, is otherwise understood with the context as used in general to present that an item, term, etc., may be any of X, Y, or Z, or any combination thereof (e.g., X, Y, and/or Z). Thus, such disjunctive language is not generally intended to, and should not, imply that certain implementations require at least one of X, at least one of Y, or at least one of Z to each be present.

[0104] Unless otherwise explicitly stated, articles such as “a” or “an” should generally be interpreted to include one or more described items. Accordingly, phrases such as “a device configured to” or “a device operable to” are intended to include one or more recited devices. Such one or more recited devices can also be collectively configured to carry out the stated recitations. For example, “a processor configured to carry out recitations A, B and C” can include a first processor configured to carry out recitation A working in conjunction with a second processor configured to carry out recitations B and C.

[0105] Language of degree used herein, such as the terms “about,” “approximately,” “generally,” “nearly” or “substantially” as used herein, represent a value, amount, or characteristic close to the stated value, amount, or characteristic that still performs a desired function or achieves a desired result. For example, the terms “about,” “approximately,” “generally,” “nearly” or “substantially” may refer to an

amount that is within less than 10% of, within less than 5% of, within less than 1% of, within less than 0.1% of, and within less than 0.01% of the stated amount.

[0106] Conditional language, such as, among others, “can,” “could,” “might,” or “may,” unless specifically stated otherwise, or otherwise understood within the context as used, is generally intended to convey in a permissive manner that certain implementations could include, or have the potential to include, but do not mandate or require, certain features, elements and/or steps. In a similar manner, terms such as “include,” “including” and “includes” are generally intended to mean “including, but not limited to.” Thus, such conditional language is not generally intended to imply that features, elements and/or steps are in any way required for one or more implementations or that one or more implementations necessarily include logic for deciding, with or without user input or prompting, whether these features, elements and/or steps are included or are to be performed in any particular implementation.

[0107] Although the invention has been described and illustrated with respect to illustrative implementations thereof, the foregoing and various other additions and omissions may be made therein and thereto without departing from the spirit and scope of the present disclosure.

CLAIMS

WHAT IS CLAIMED IS:

1. A computer-implemented method, comprising:
 - processing a first image to determine a first image segment that corresponds to less than all of the first image and corresponds to an object represented in the first image;
 - presenting, on a user device, the first image and the first image segment such that the first image segment is visually distinguished from the first image;
 - in response to a selection of the first image segment, forming a first extracted image segment that includes at least:
 - pixel data corresponding to pixels of the first image segment; and
 - metadata indicating at least one of the first image or a source location of the first image; and
 - presenting the first extracted image segment as part of a collage that includes at least a second extracted image segment representative of a second object extracted from a second image; and
 - wherein:
 - the first extracted image segment may be transformed with respect to the second extracted image segment;
 - the first extracted image segment may be transformed with respect to a size of the first extracted image segment when presented on the collage, and
 - the first extracted image segment may be transformed with respect to a position of the first extracted image segment when presented on the collage.
2. The computer-implemented method of claim 1, further comprising:
 - receiving a modification of the first image segment to modify which pixels of the first image are included in the first image segment.

3. The computer-implemented method of any one of claims 1 or 2, further comprising:
 - receiving an indication of a region of interest within the first image; and
 - wherein processing includes processing a portion of the first image within the region of interest to determine the first image segment.
4. The computer-implemented method of any one of claims 1, 2, or 3, wherein processing the first image is performed on the user device.
5. A computing system, comprising:
 - one or more processors;
 - a memory storing program instructions that, when executed by the one or more processors, cause the one or more processors to at least:
 - in response to receipt of a first selection of a first image segment of a first image, generate a first extracted image segment that includes at least:
 - pixel data corresponding to pixels of the first image segment; and
 - metadata indicating at least one of the first image or a source location of the first image;
 - present the first extracted image segment as part of a collage;
 - receive a second selection of a second extracted image segment, wherein the second extracted image segment corresponds to a second image; and
 - present the second extracted image segment with the first extracted image segment as part of the collage; and
 - wherein the first extracted image segment and the second extracted image segment are independently transformable when presented as part of the collage.
6. The computing system of claim 5, wherein the program instructions, that when executed by the one or more processors, further cause the one or more processors to at least:
 - present a plurality of extracted image segments, wherein the second extracted image segment is included in the plurality of extracted image segments and selected from the plurality of extracted image segments.

7. The computing system of claim 6, wherein at least one of the plurality of extracted image segments is at least one of a popular extracted image segment, a trending extracted image segment, an extracted image segment that enables a purchase of an object represented in the extracted image segment, visually similar to the first extracted image segment, related to the first extracted image segment, or an extracted image segment generated by a user that selected the first extracted image segment.

8. The computing system of any one of claims 5, 6, or 7, wherein the program instructions, that when executed by the one or more processors, further cause the one or more processors to at least:

- determine a seller of an object represented in the first extracted image segment;
- associate the seller with the first extracted image segment; and
- present, with the first extracted image segment and concurrent with the presentation of the first extracted image segment and the second extracted image segment as part of the collage, a buyable indication indicating that the object represented in the first extracted image segment is buyable through an interaction with the first extracted image segment.

9. The computing system of any one of claims 5, 6, 7, or 8, wherein the program instructions, that when executed by the one or more processors, further cause the one or more processors to at least:

- receive a request to transform the collage;
- determine that the request is from a second user that is different than a first user that created the collage;
- in response to determination that the request is from the second user:
 - generate a duplicate collage from the collage;
 - maintain, as part of a metadata for the duplicate collage, a link or other identifier indicating that the duplicate collage was generated based on the collage; and
 - transform the duplicate collage in response to the request.

10. The computing system of claim 9, wherein the request is at least one of a request to:
- add a third extracted image segment to the collage;
 - remove the second extracted image segment from the collage;
 - re-mix a presentation of the first extracted image segment of the collage and the second extracted image segment of the collage;
 - transform a position of the first extracted image segment; or
 - transform a size of the first extracted image segment.
11. The computing system of any one of claims 5, 6, 7, 8, 9, or 10, wherein the program instructions, that when executed by the one or more processors, further cause the one or more processors to at least:
- receive a request to transform the collage;
 - determine that the request is from a user that created the collage or that the collage is a collaborative collage; and
 - in response to determination that the request is from the user or that the collage is a collaborative collage:
 - transform the collage in response to the request.
12. The computing system of any one of claims 5, 6, 7, 8, 9, 10, or 11 wherein the program instructions, that when executed by the one or more processors, further cause the one or more processors to at least:
- receive a modification to the first image segment to modify which pixels of the first image are included in the first image segment;
 - in response to the modification:
 - generate, based at least in part on the modification, a modified image segment; and
 - update, based at least in part with the modified image segment, a training of a deep neural network (“DNN”), wherein the DNN is trained to generate one or more image segments from an input image.

13. A method, comprising:

processing, with a deep neural network (“DNN”), a first image to determine a first image segment corresponding to an object represented in the first image, wherein the first image segment is less than all of the first image;

presenting, on a display, the first image segment and the first image such that the first image segment is visually distinguished from the first image;

in response to receiving a modification of the first image segment to modify which pixels of the first image are included in the first image segment, generating an updated first image segment;

in response to a selection of the updated first image segment, forming a first extracted image segment that includes pixel data corresponding to pixels of the updated first image segment; and

presenting the first extracted image segment as part of a collage that includes at least a second extracted image segment representative of a second object extracted from a second image, wherein:

the first extracted image segment is adjustable with respect to the second extracted image segment;

the first extracted image segment is adjustable with respect to a size of the first extracted image segment when presented on the collage; and

the first extracted image segment is adjustable with respect to a position of the first extracted image segment when presented on the collage.

14. The method of claim 13, further comprising:
updating a training of the DNN based at least in part on the first image and the updated first image segment.
15. The method of any one of claims 13 or 14, further comprising:
presenting, as part of the collage, an indication of a total number of extracted image segments included in the collage.

1/26



FIG. 1A

2/26

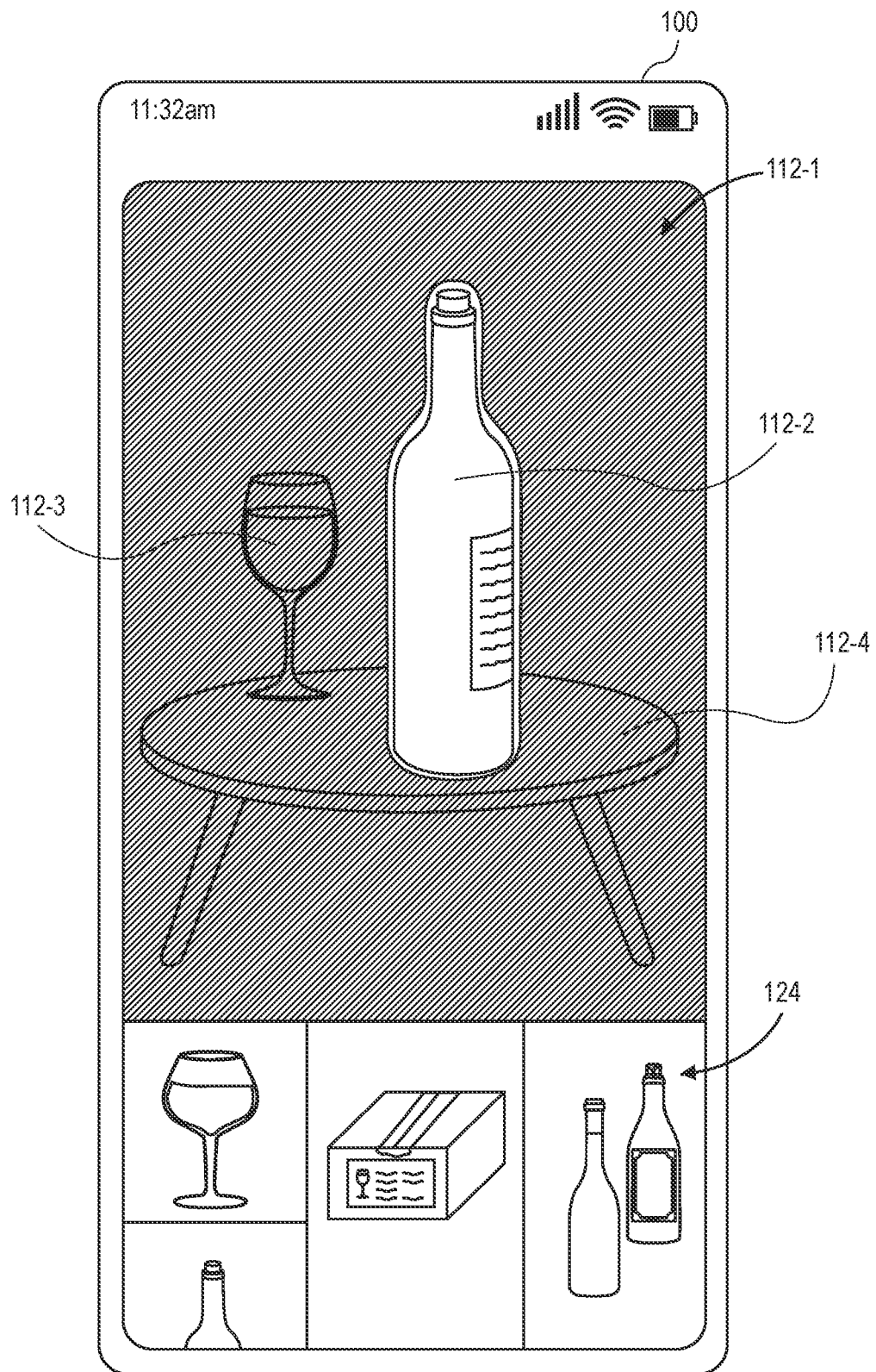
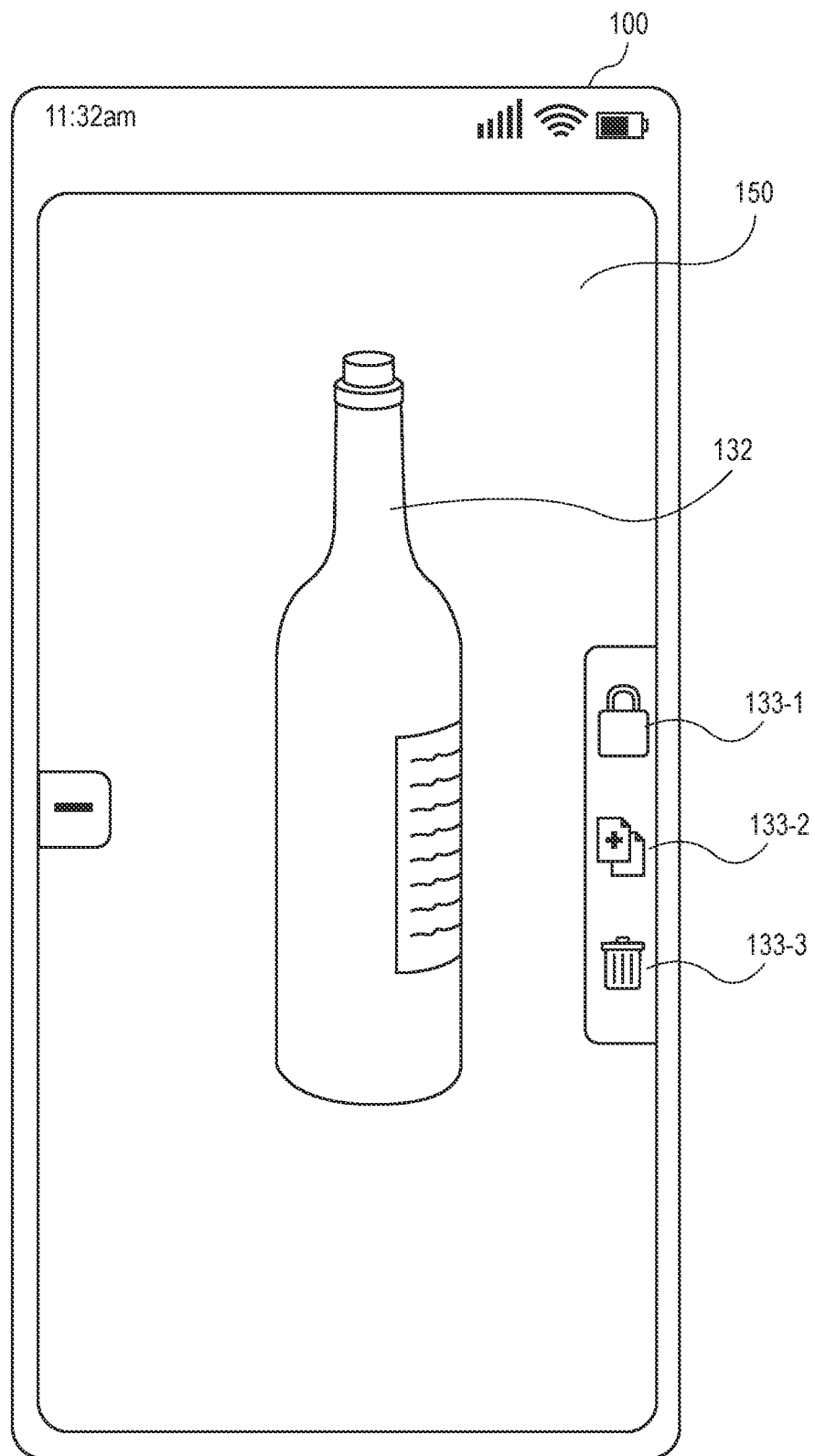


FIG. 1B

3/26

**FIG. 1C**

4/26

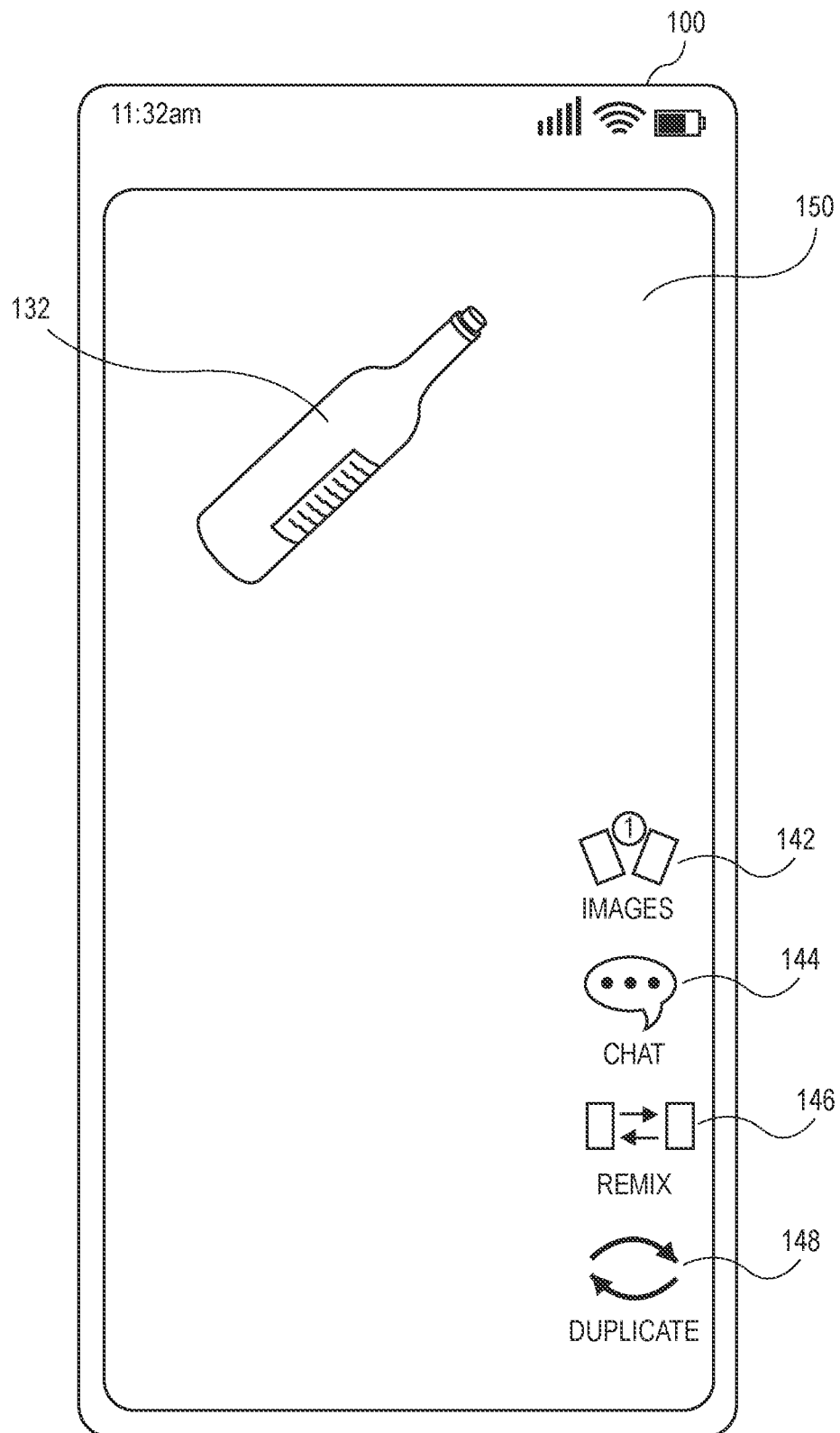


FIG. 1D

5/26

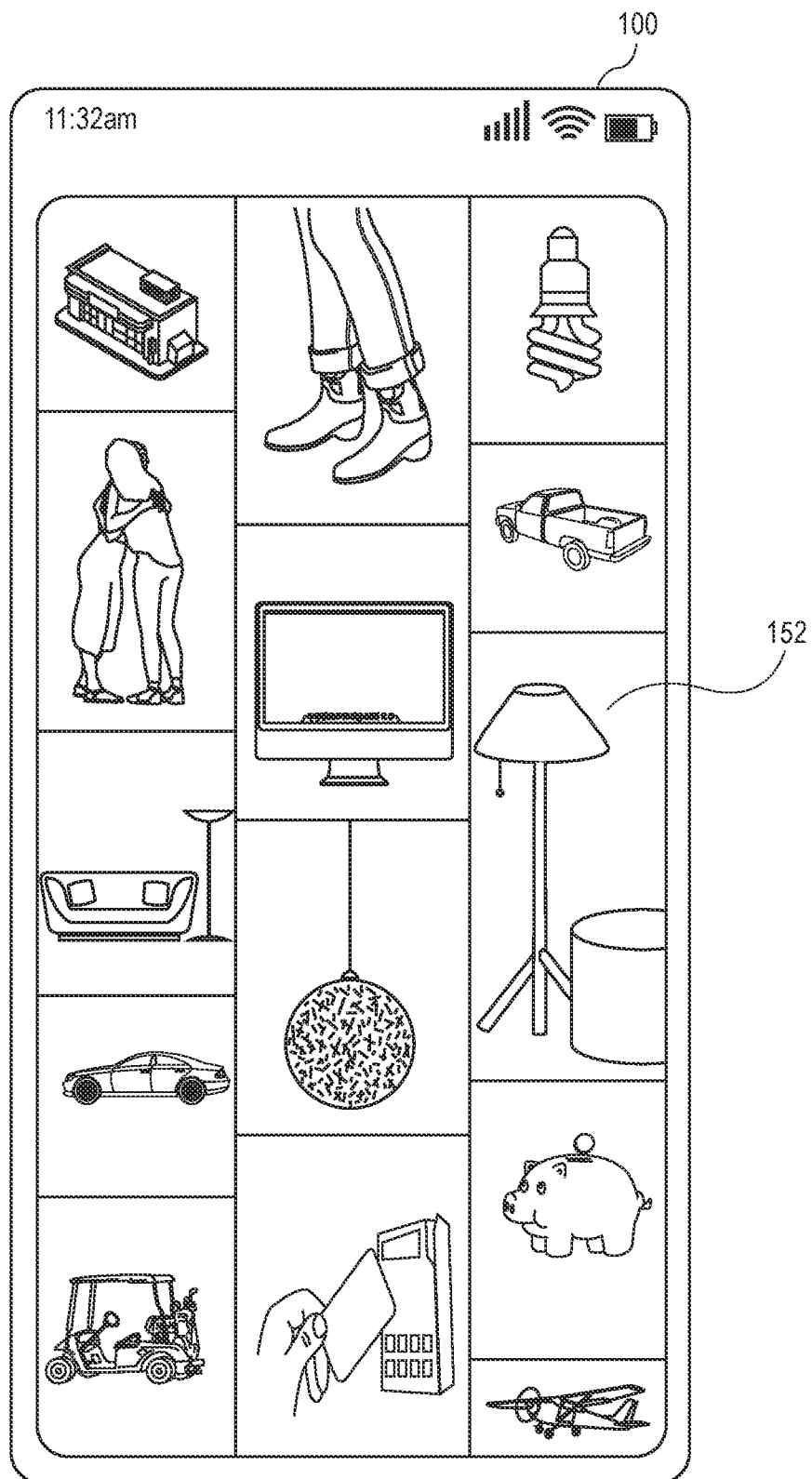


FIG. 1E

6/26

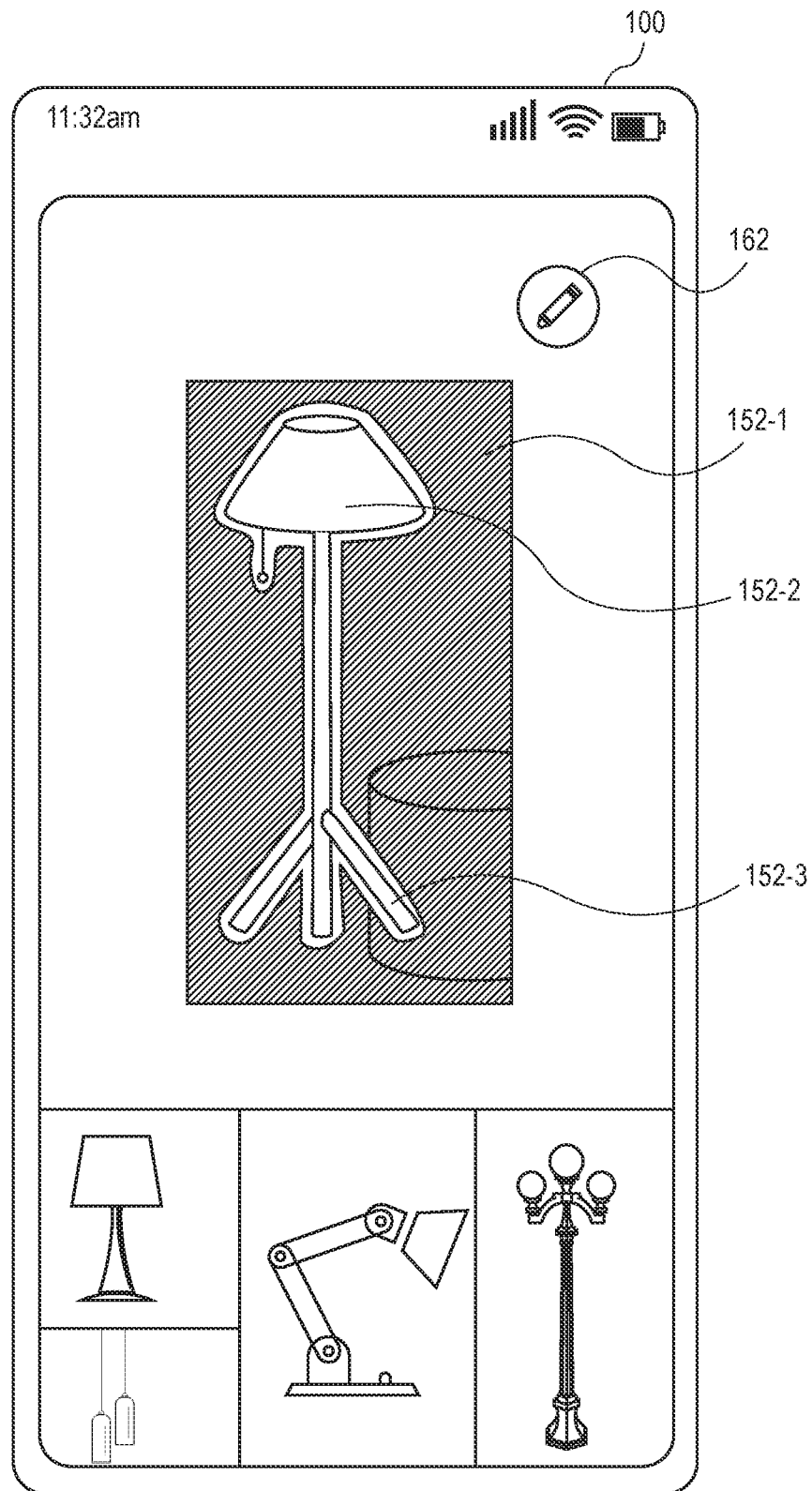


FIG. 1F

7/26

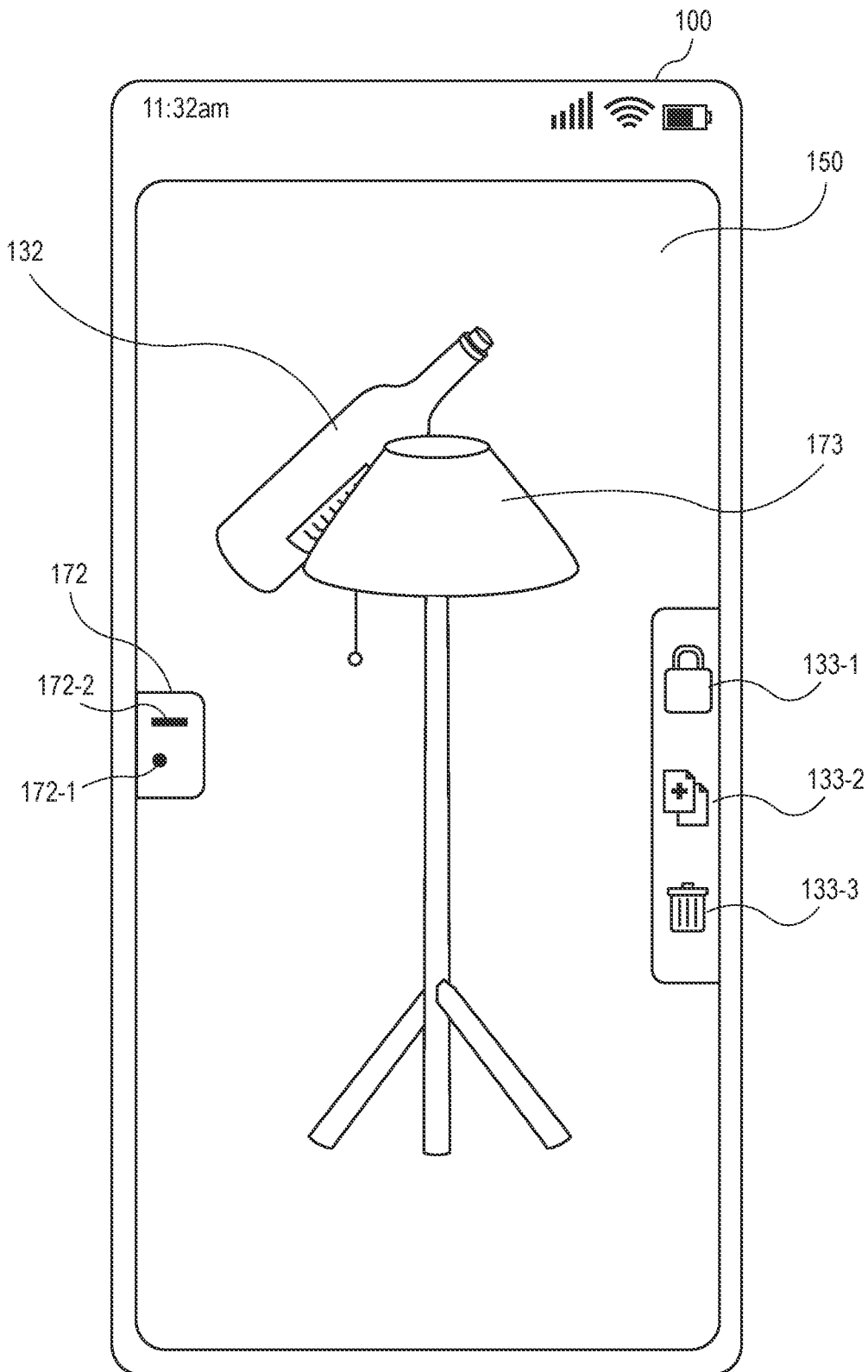
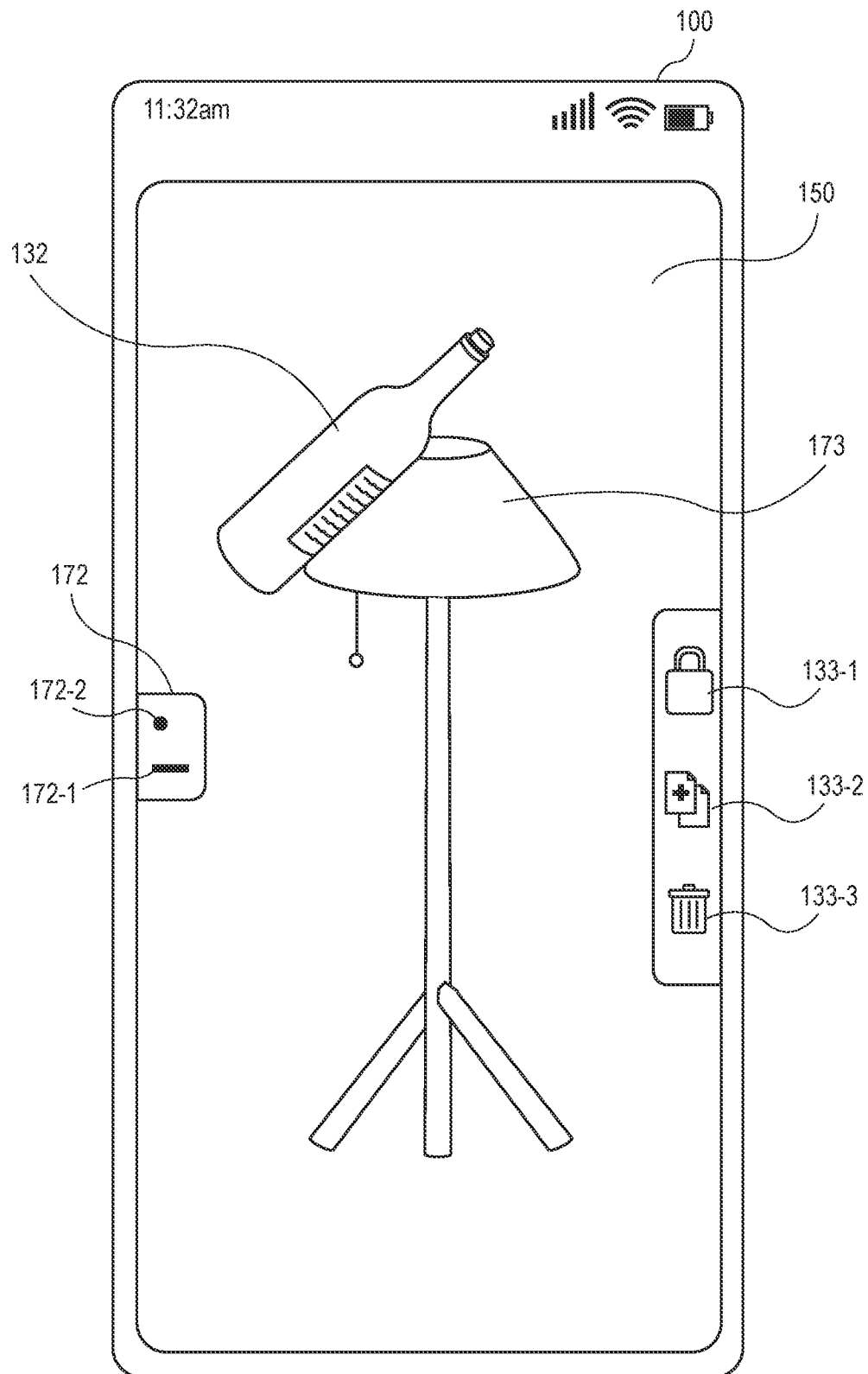


FIG. 1G

8/26

**FIG. 1H**

9/26

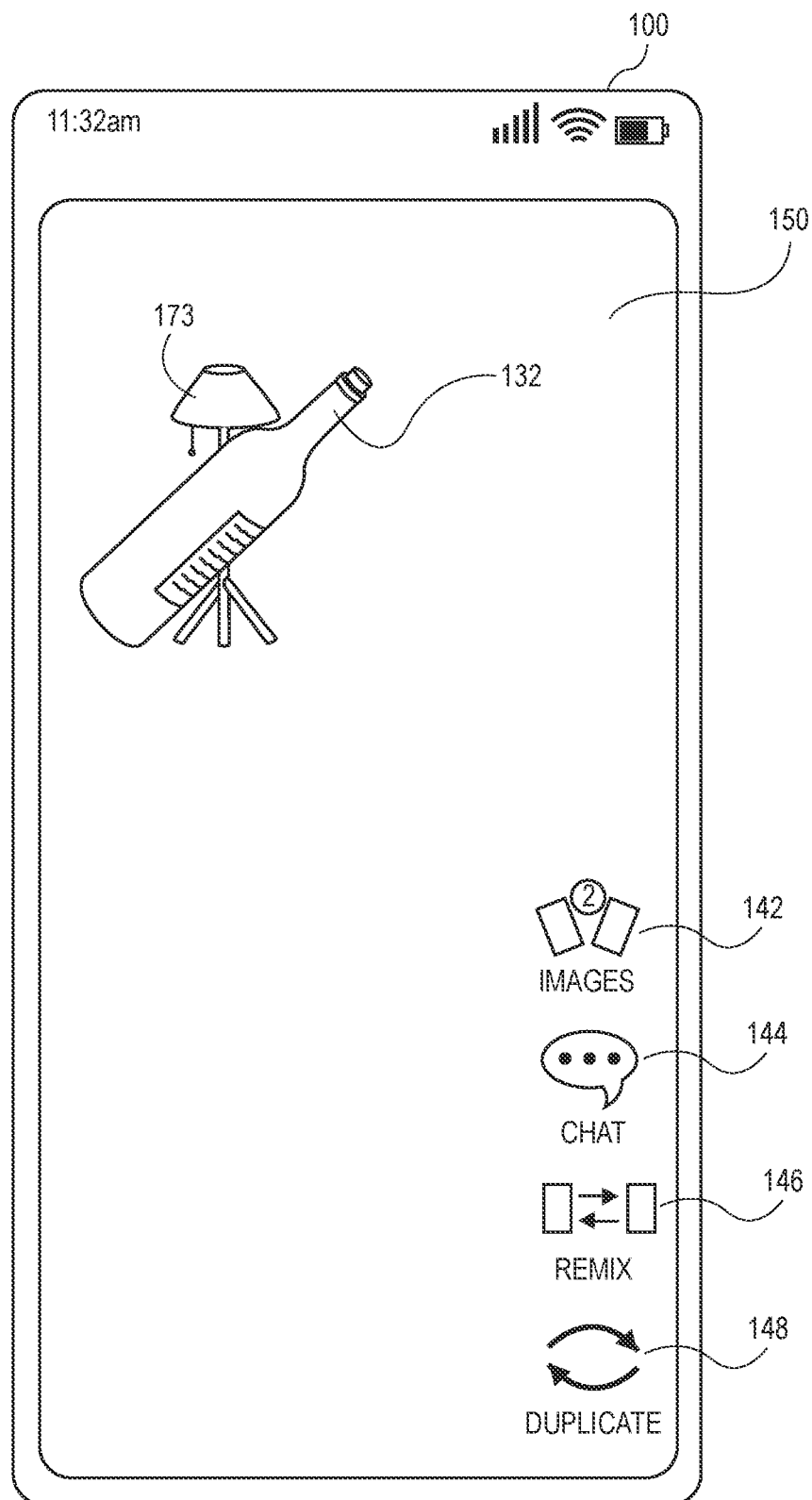


FIG. 11

10/26

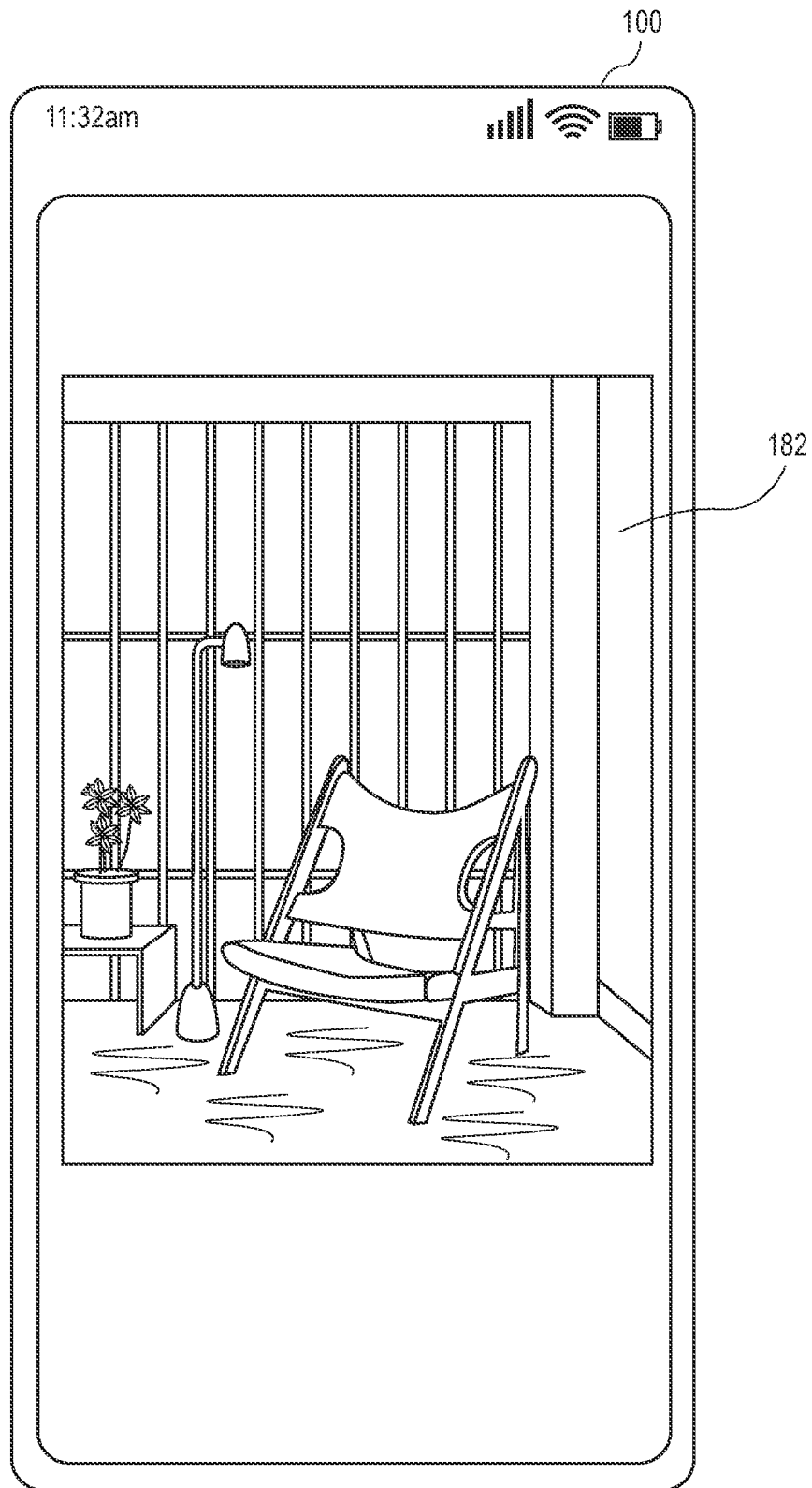


FIG. 1J

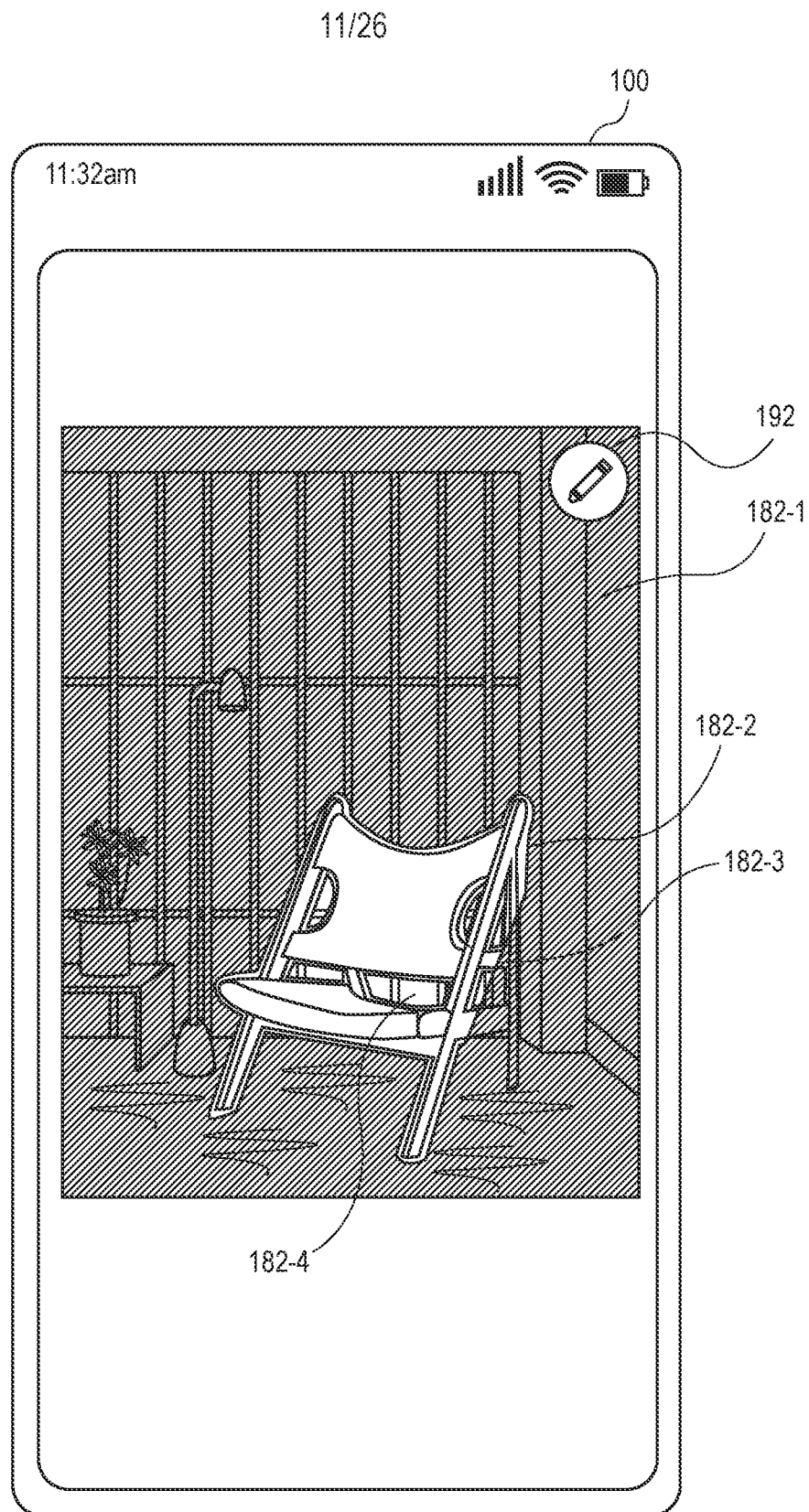


FIG. 1K

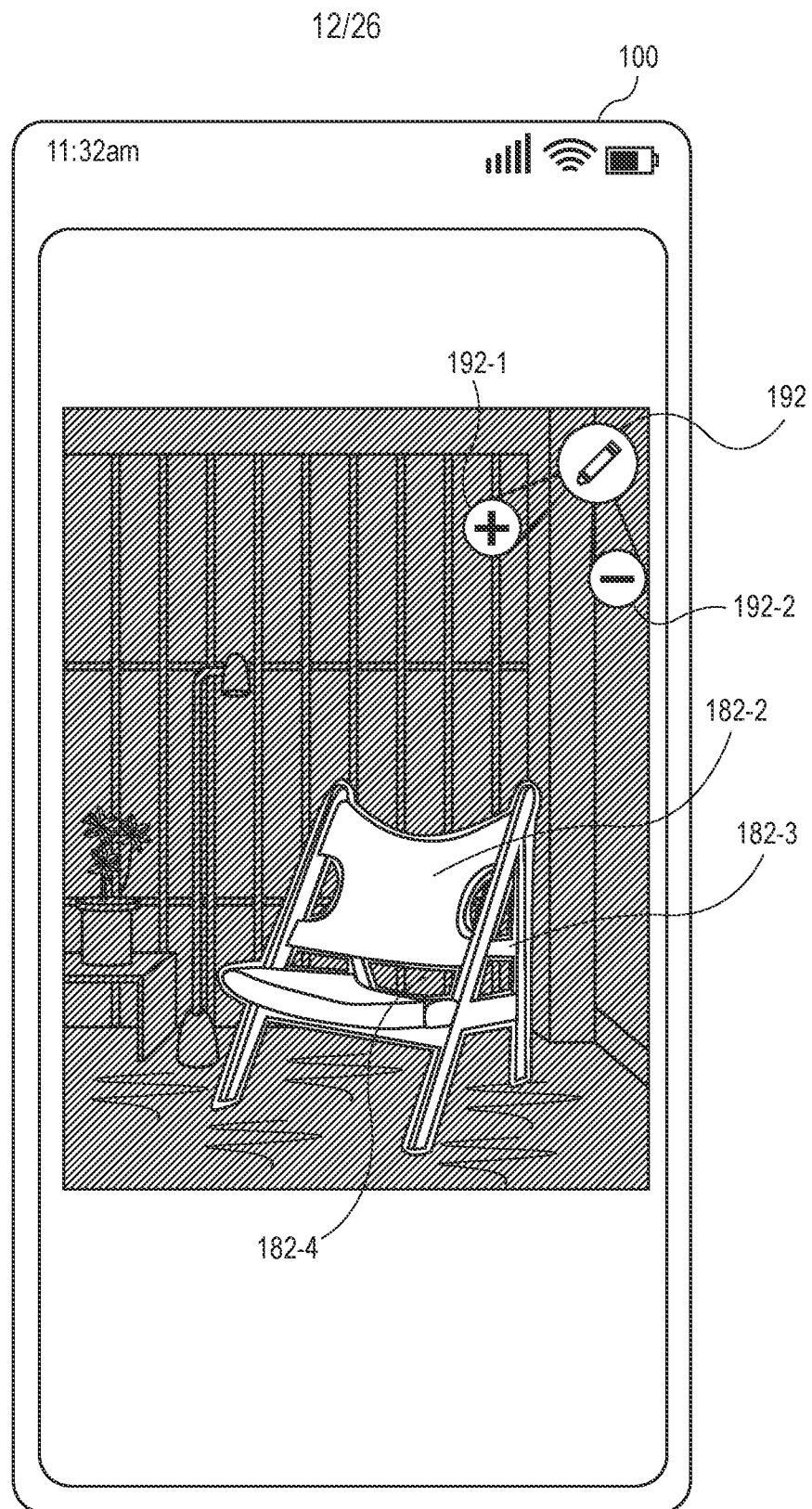


FIG. 1L

13/26

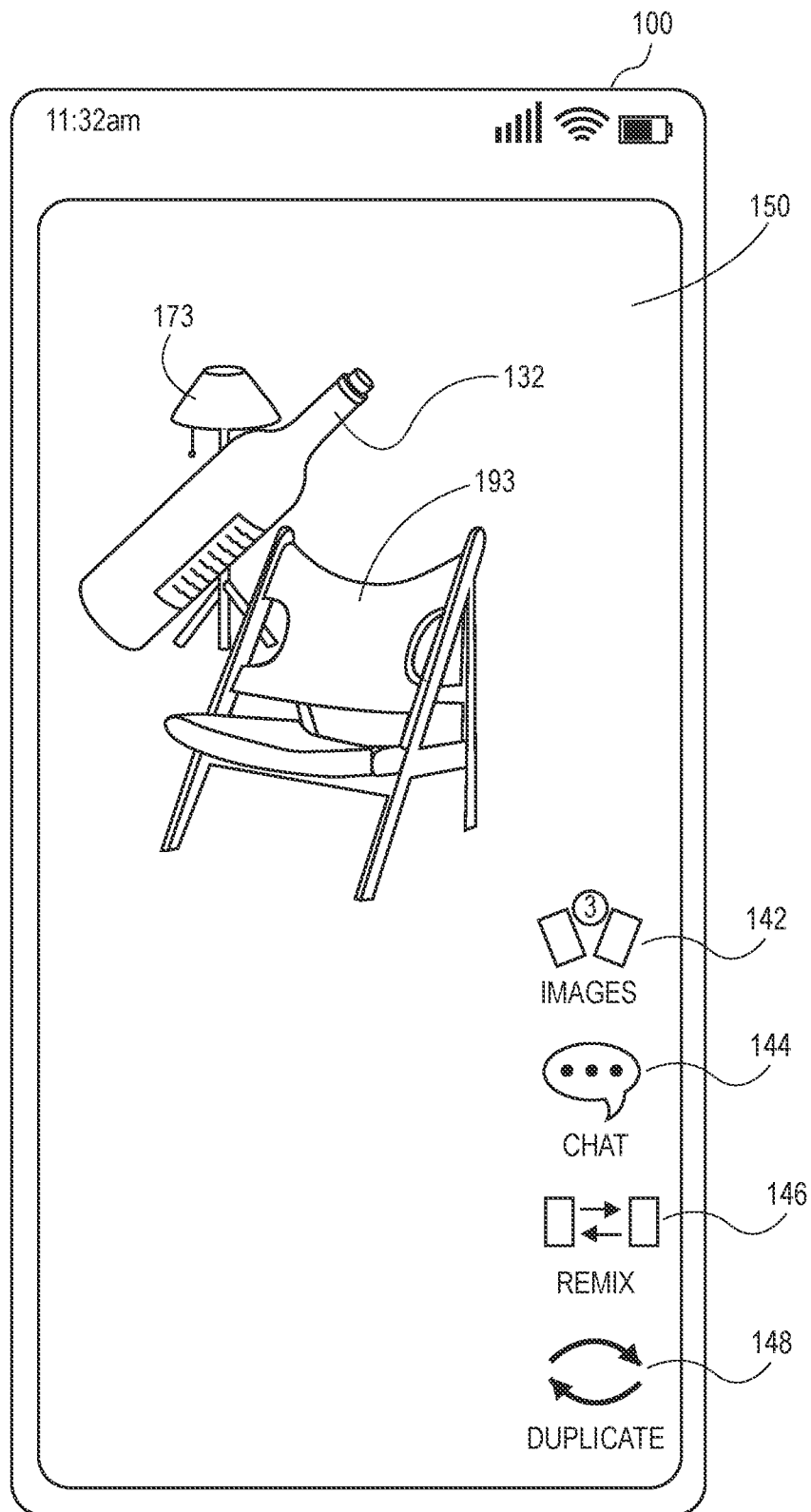


FIG. 1M

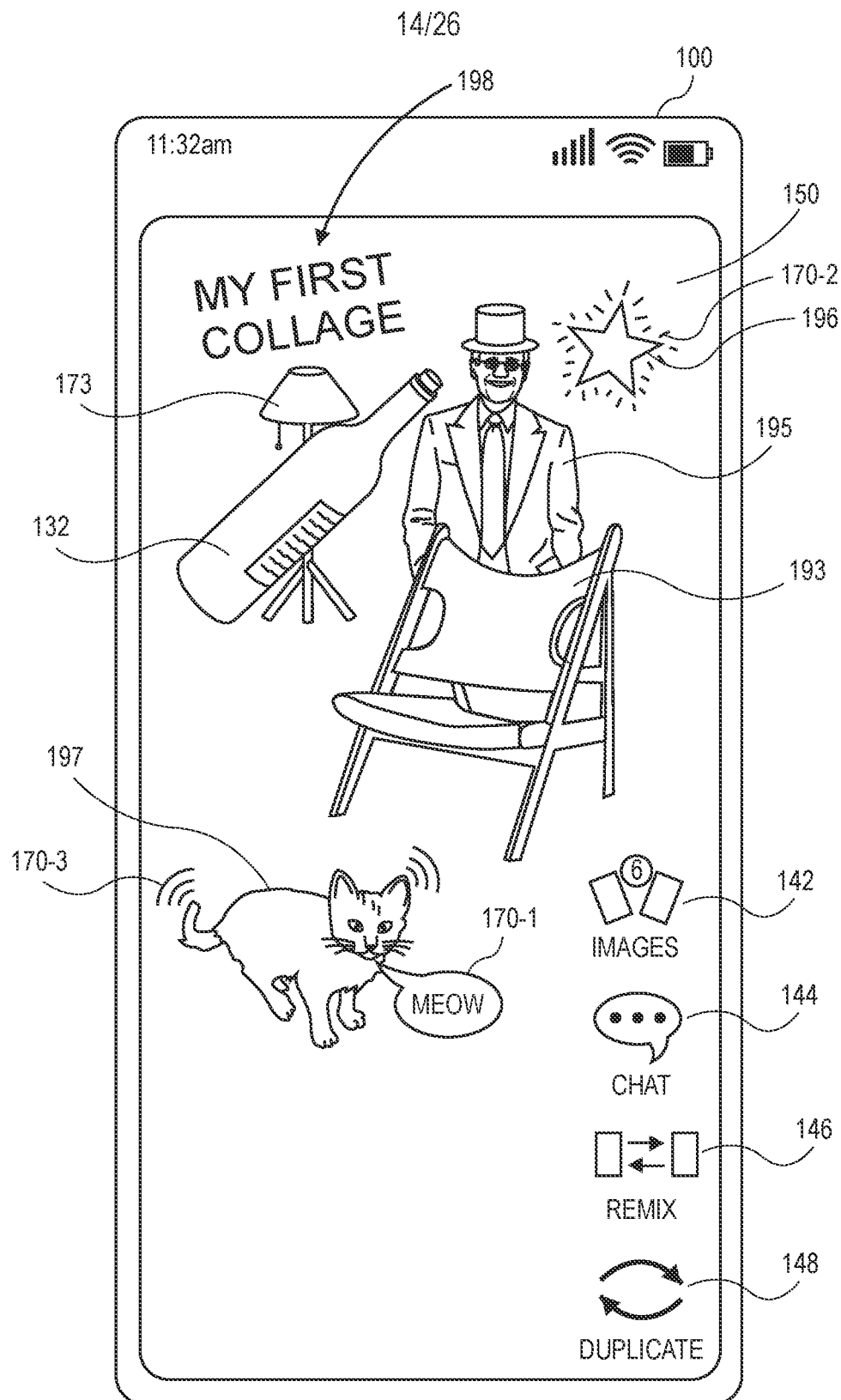


FIG. 1N

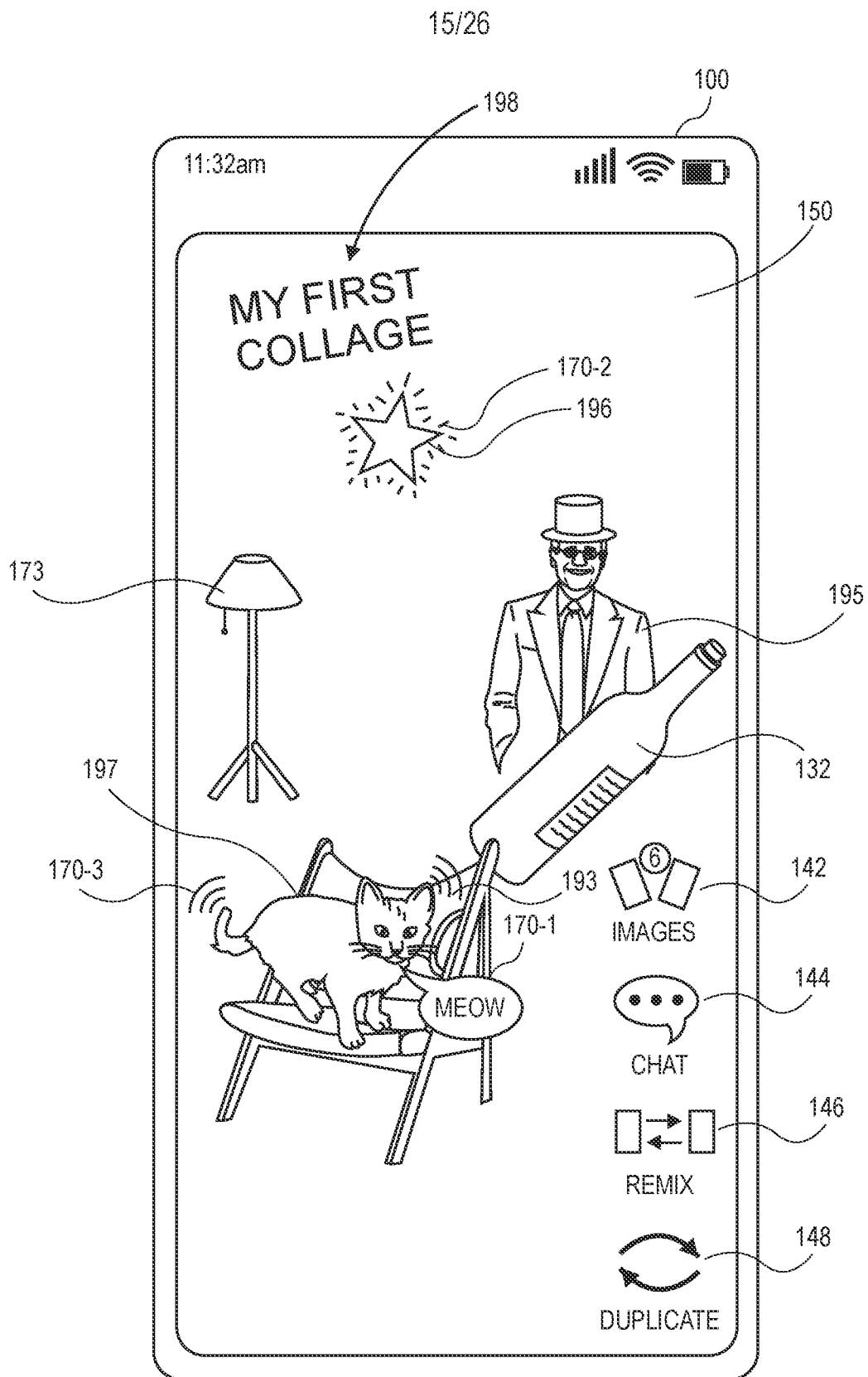


FIG. 10

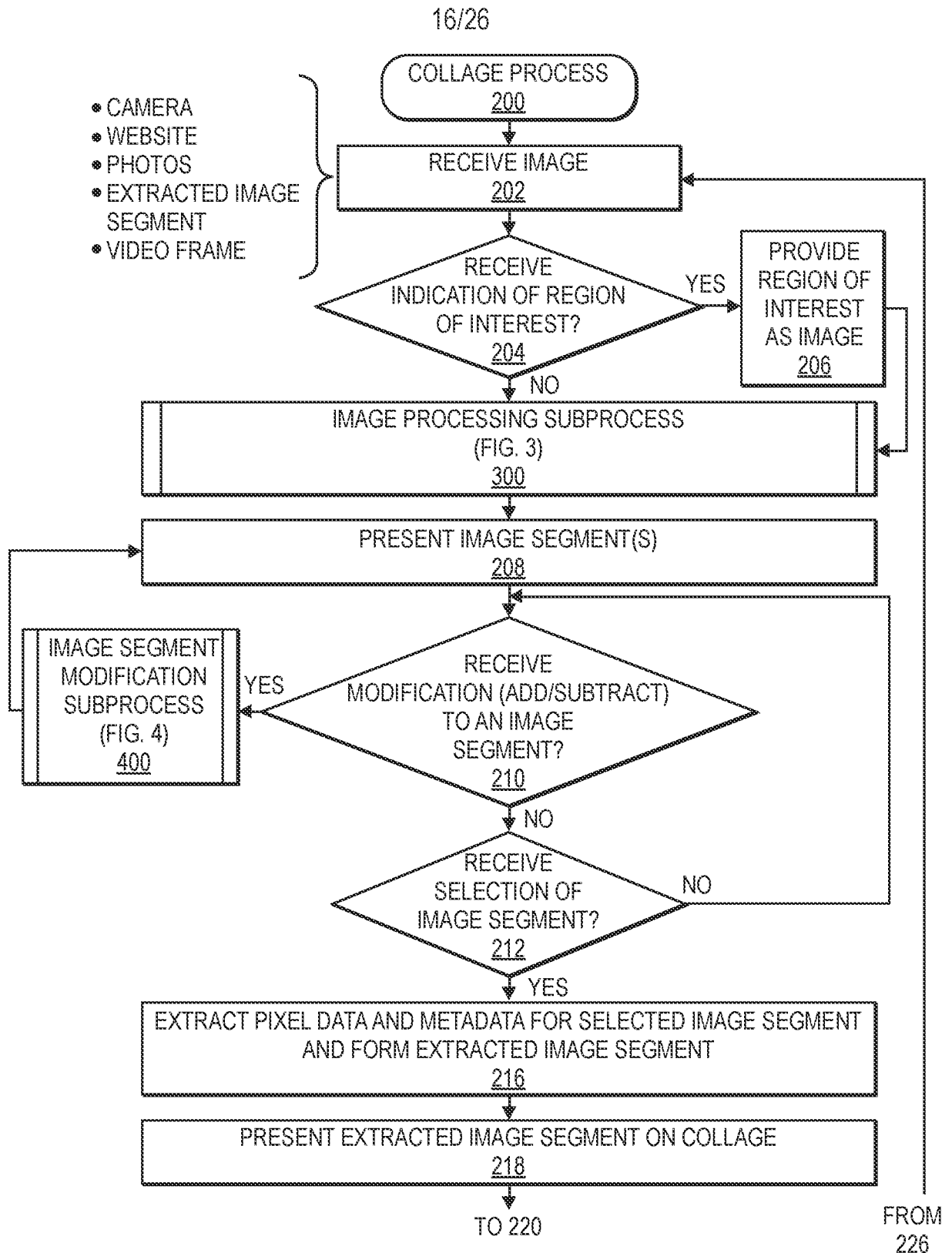
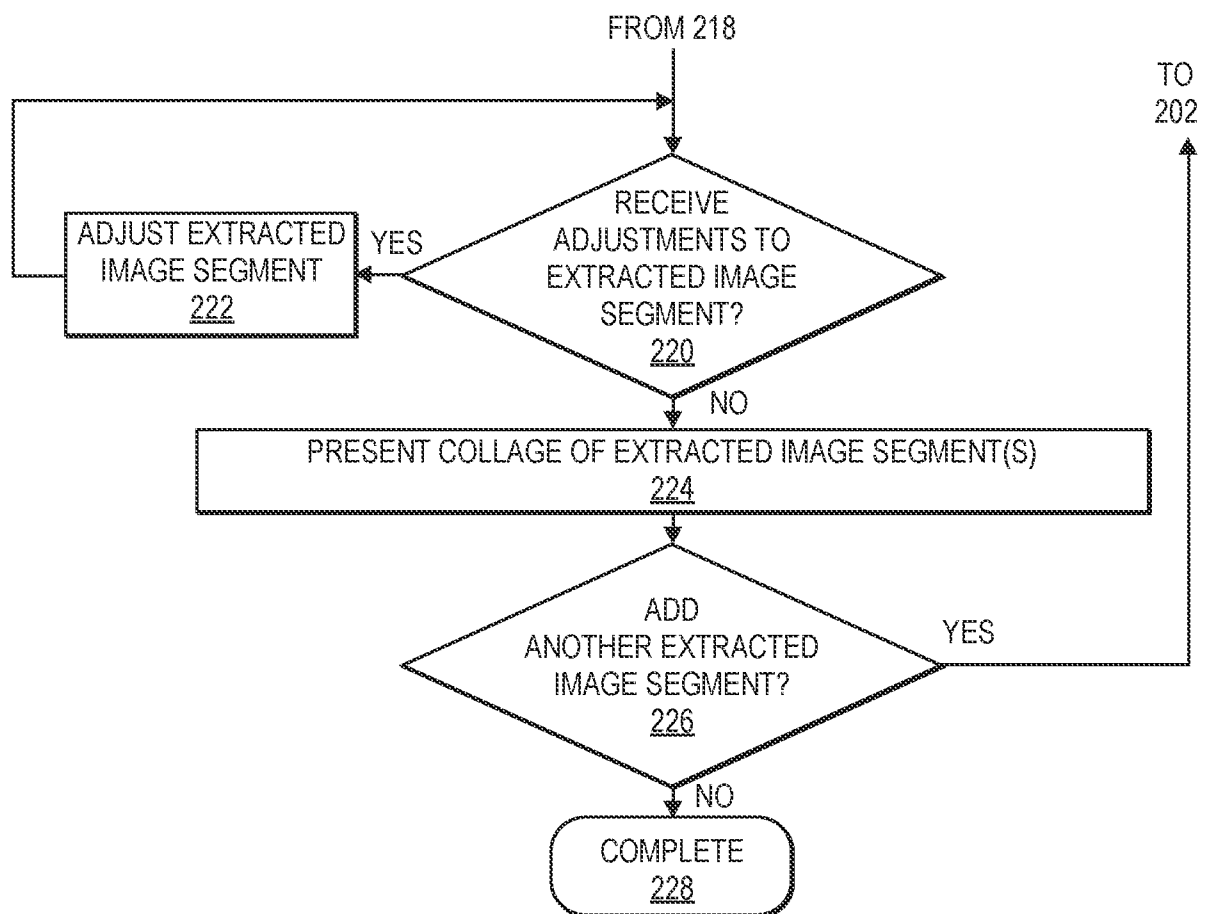
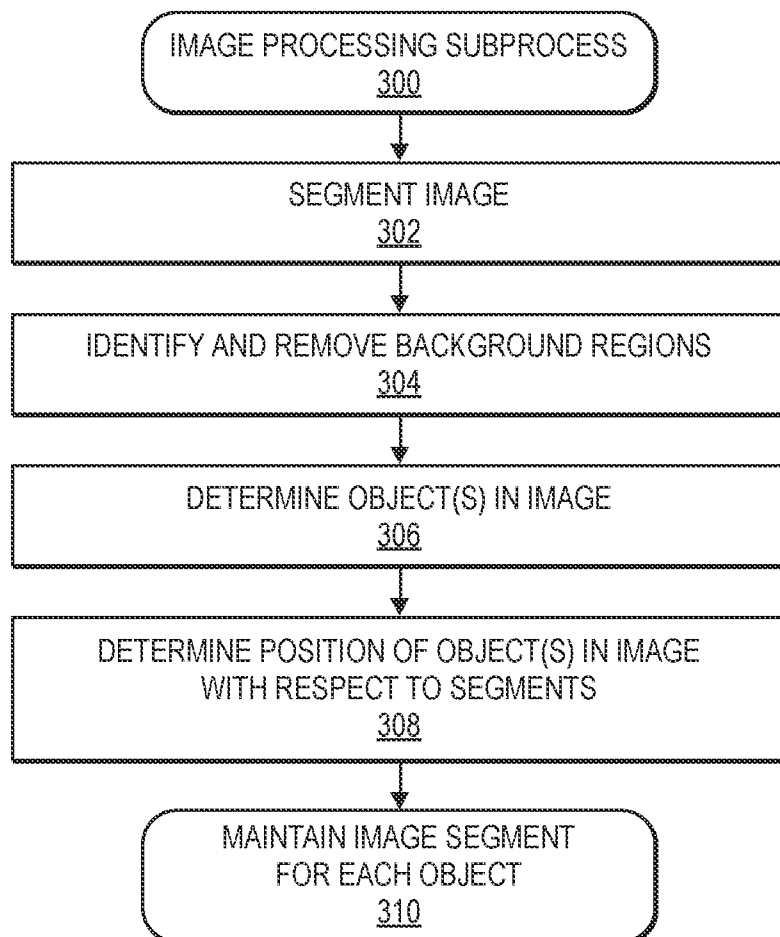


FIG. 2

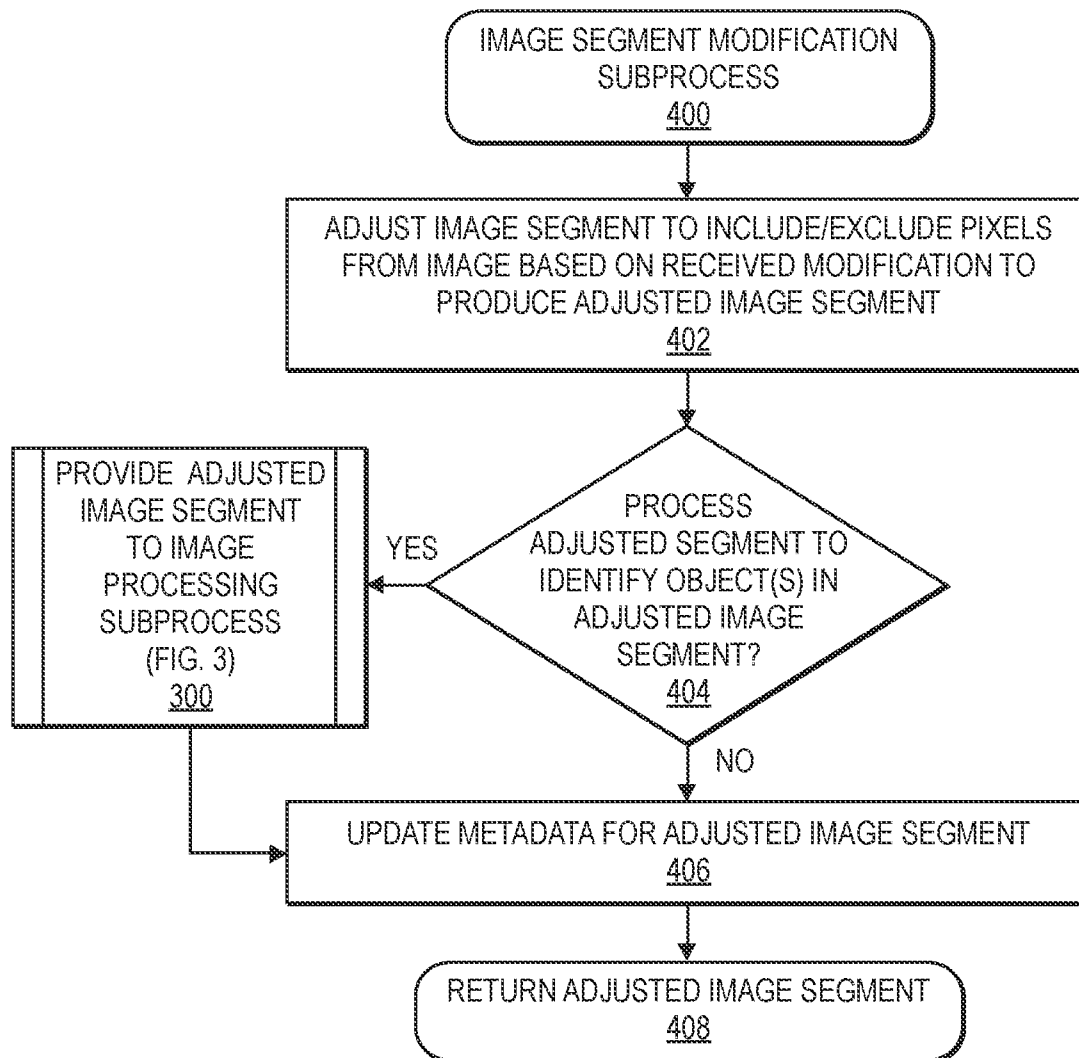
17/26

**FIG. 2 (cont.)**

18/26

**FIG. 3**

19/26

**FIG. 4**

20/26

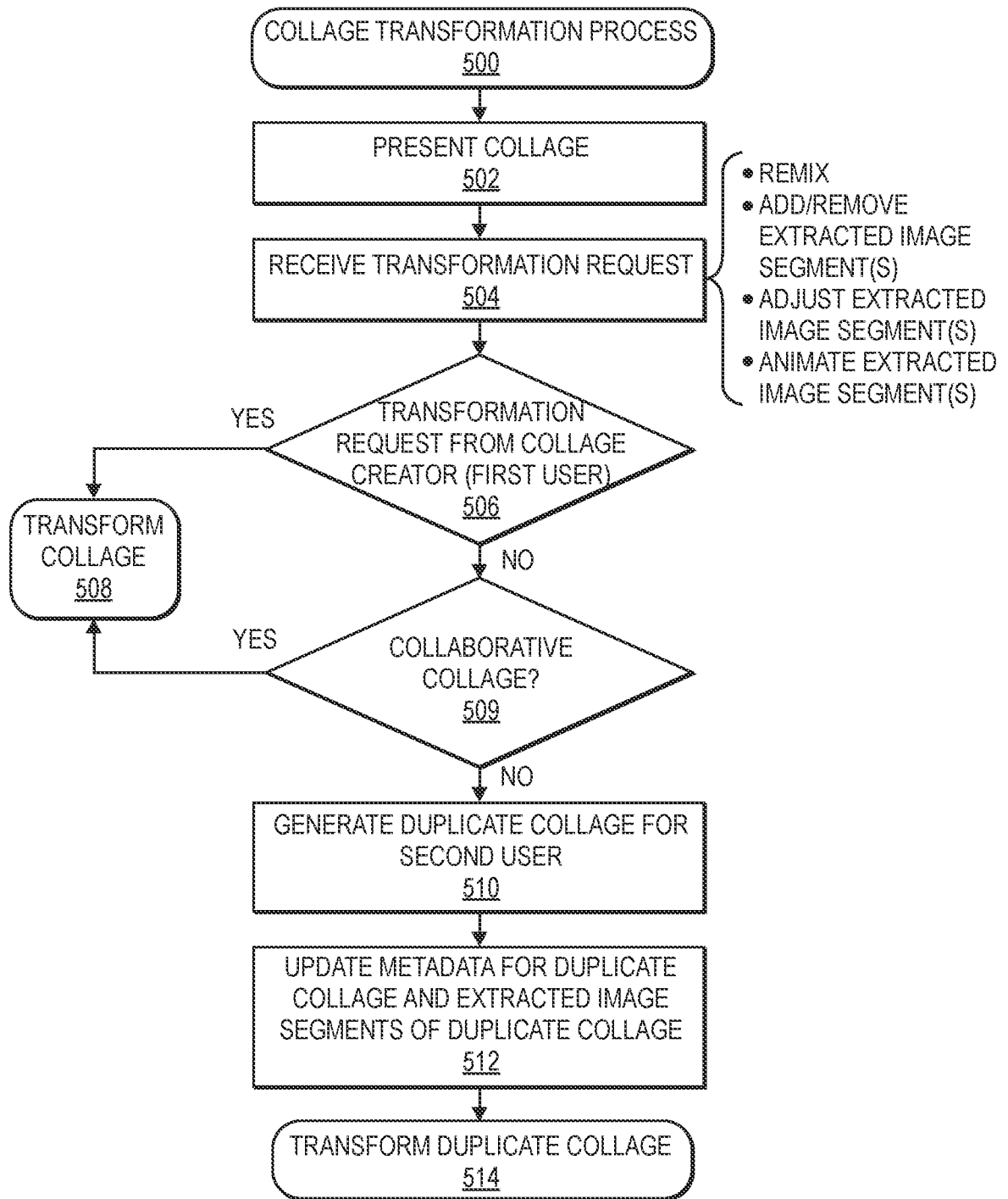
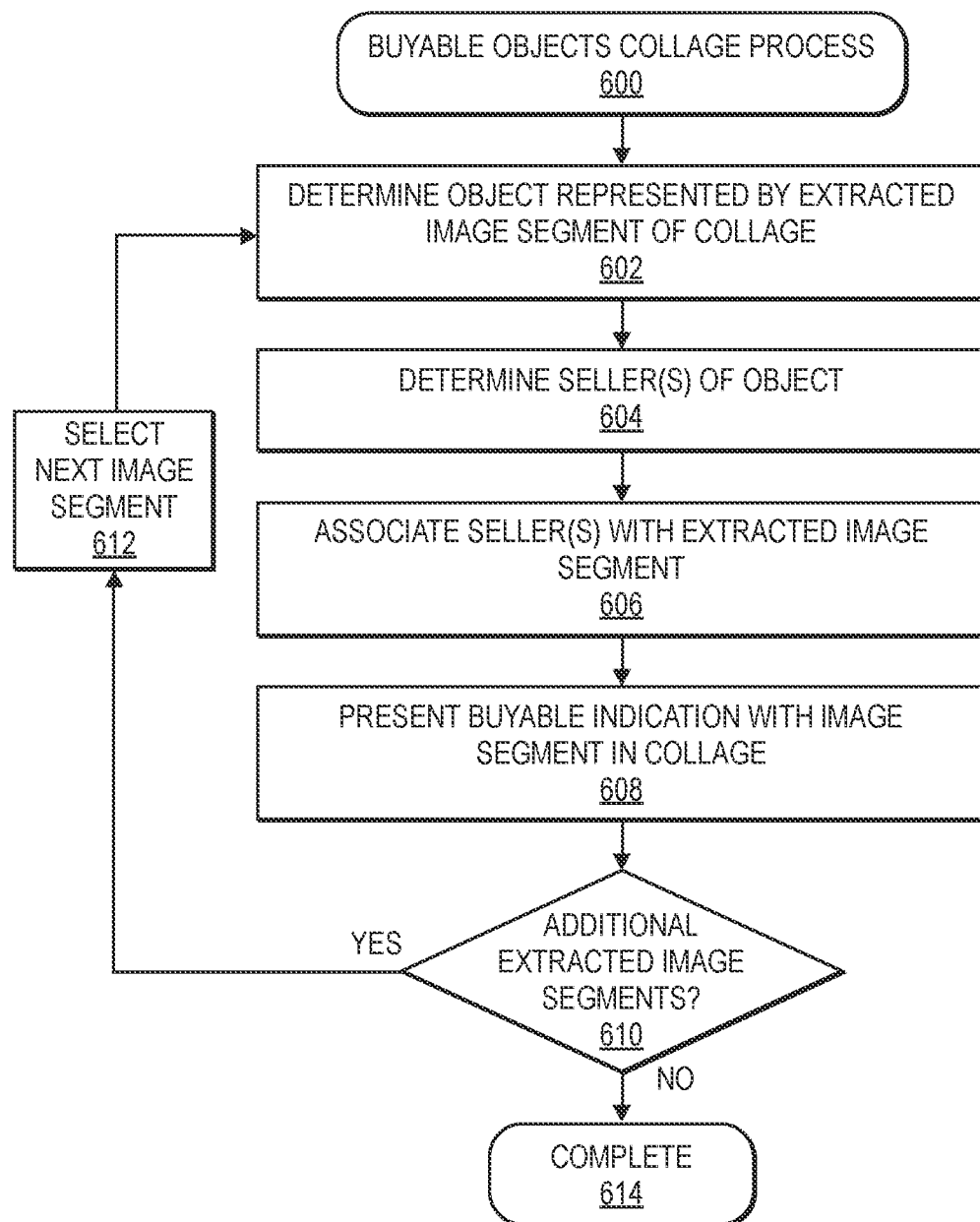


FIG. 5

21/26

**FIG. 6**

22/26

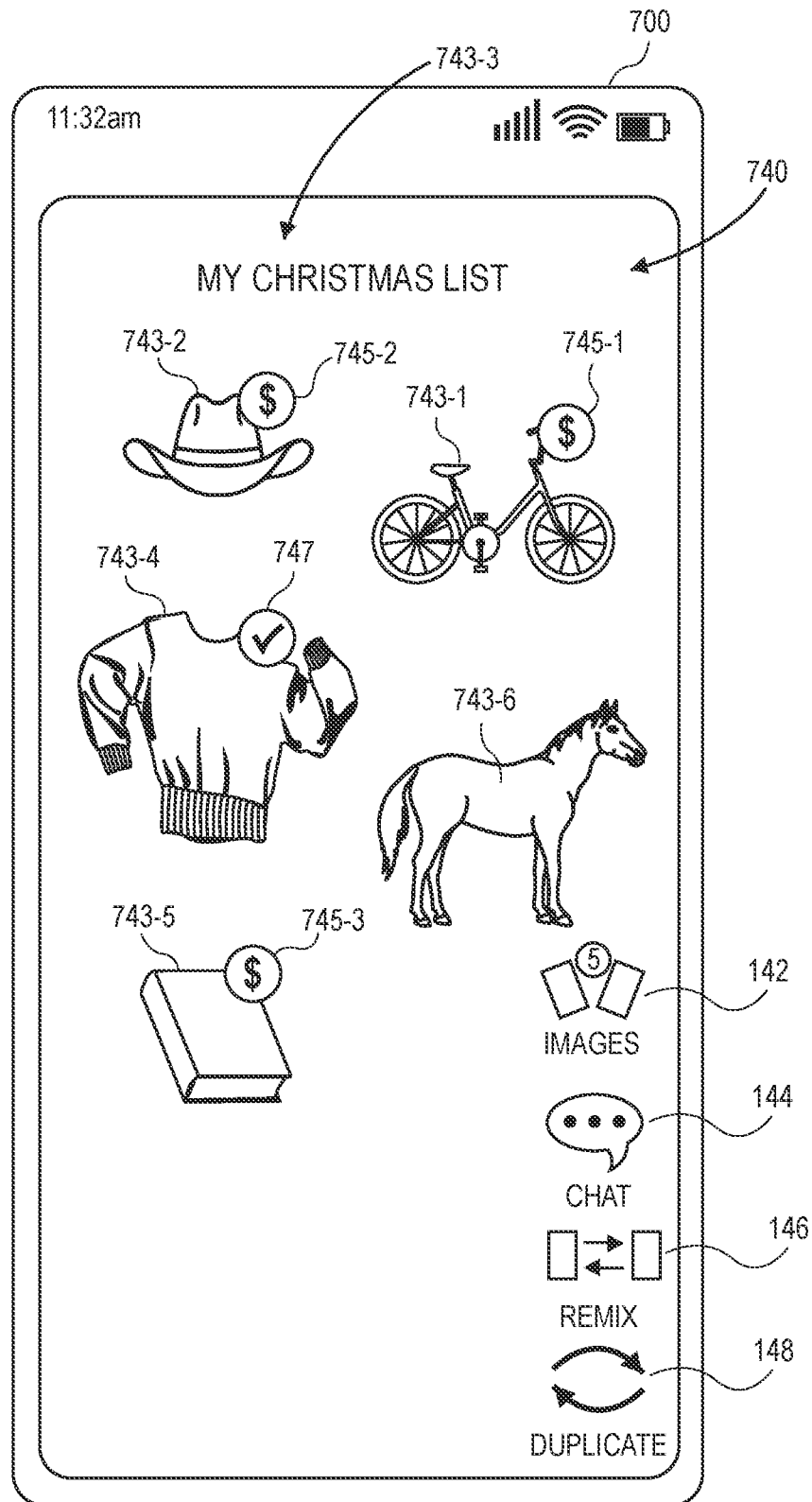


FIG. 7A

23/26

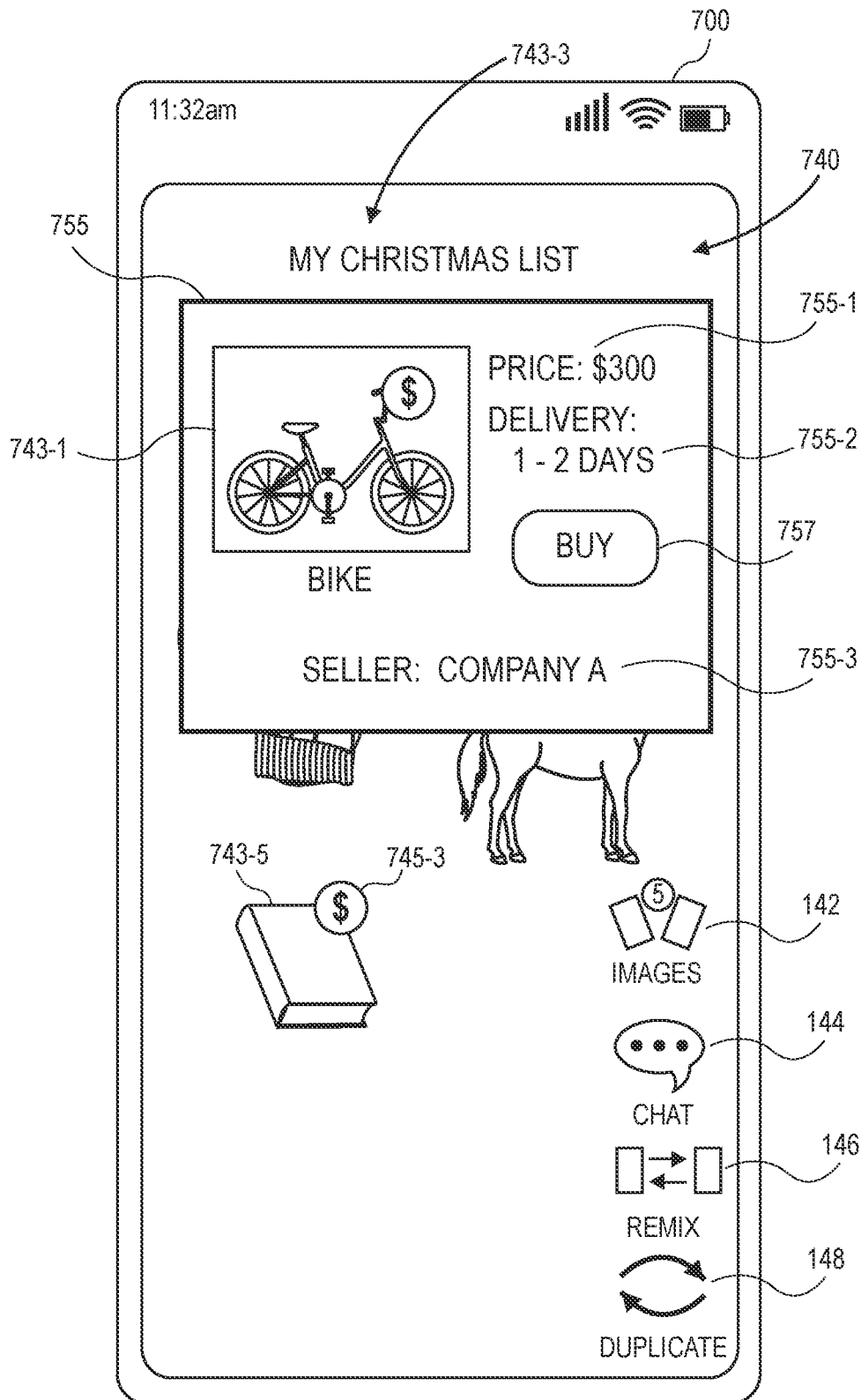
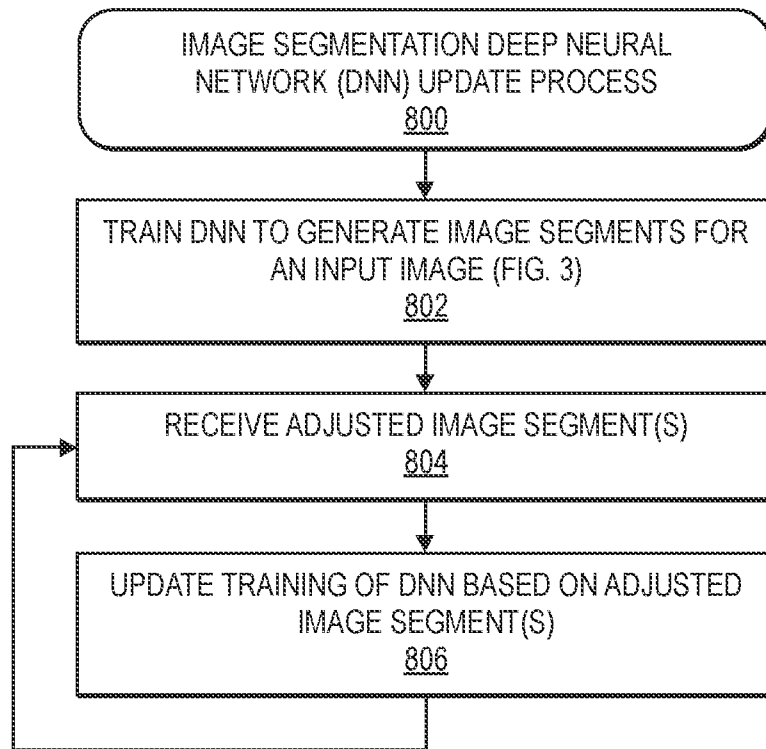


FIG. 7B

24/26

**FIG. 8**

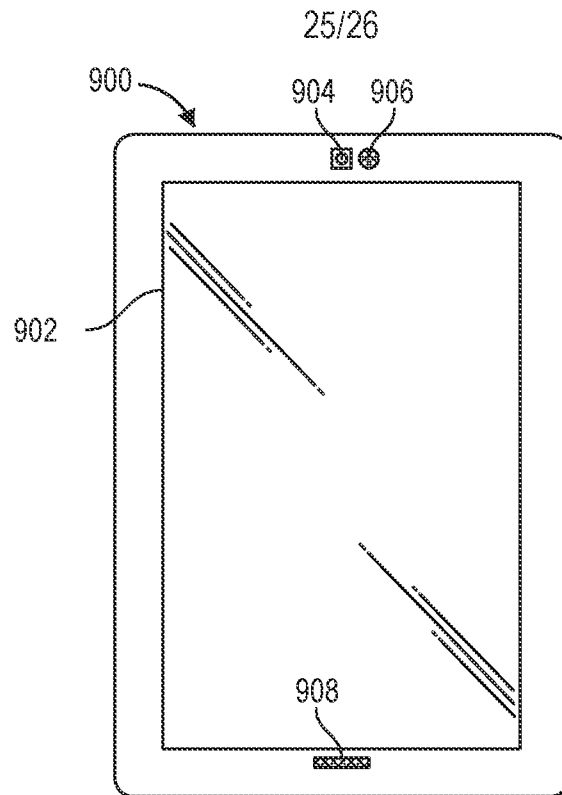


FIG. 9

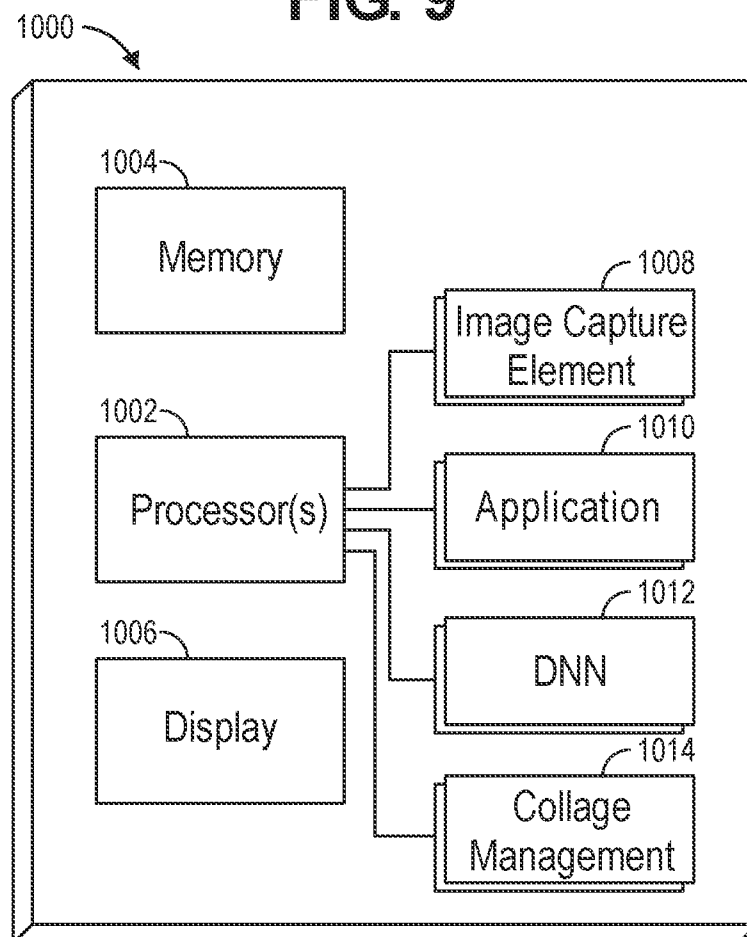
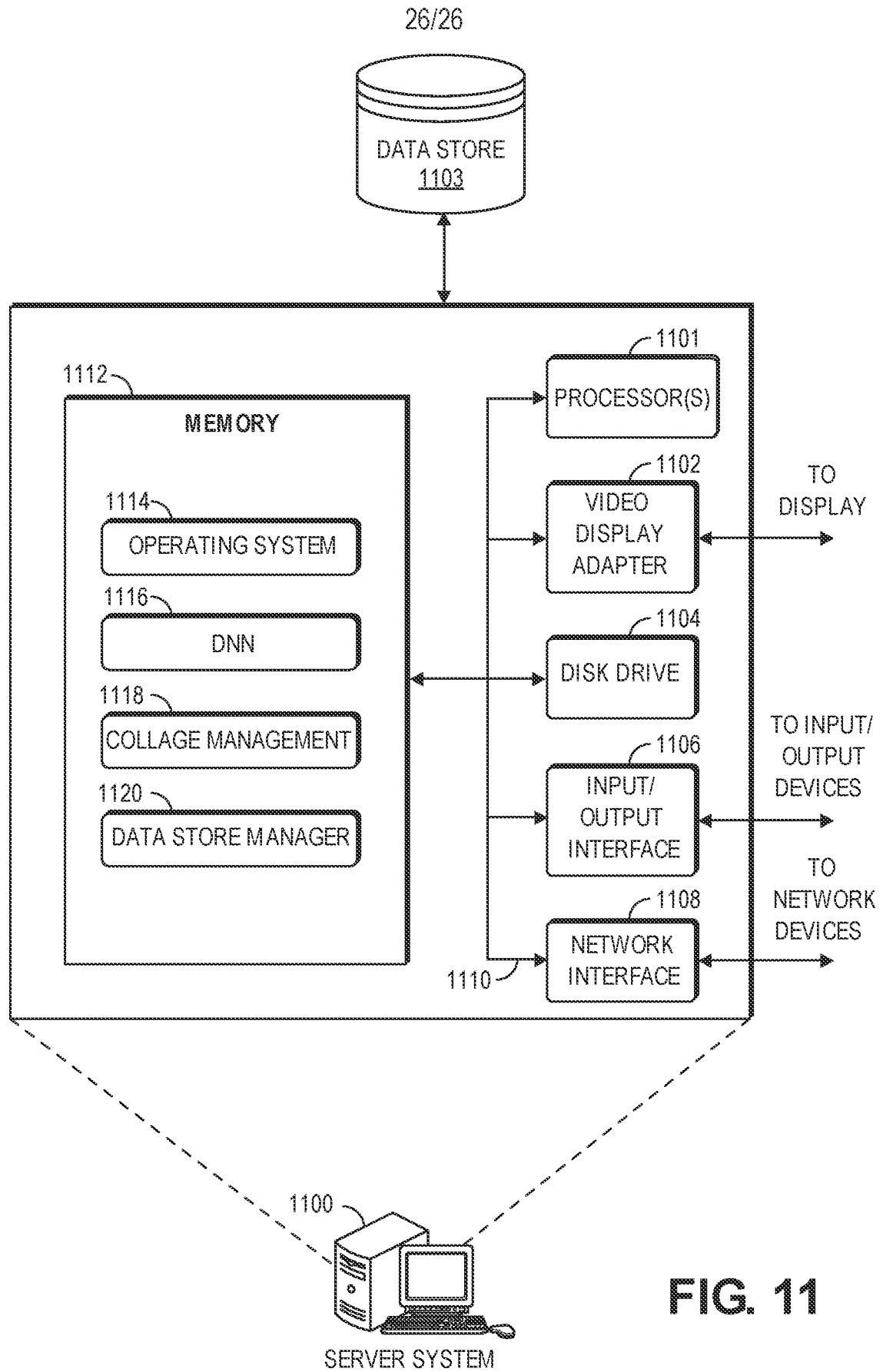


FIG. 10



INTERNATIONAL SEARCH REPORT

International application No
PCT/US2022/081875

A. CLASSIFICATION OF SUBJECT MATTER

INV. **G06T11/00** **G06T7/11** **G06Q30/0601**
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06T G06Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2019/164322 A1 (KONG SARAH AYE [US] ET AL) 30 May 2019 (2019-05-30)	1, 3-7, 12
Y	abstract paragraph [0001] - paragraph [0005] paragraph [0022] - paragraph [0060] figures 1-4	2, 8-11, 13-15
Y	US 2015/178786 A1 (CLAESSENS CATHARINA A J [NL]) 25 June 2015 (2015-06-25) abstract	2, 8-11, 13-15

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

29 March 2023

Date of mailing of the international search report

11/04/2023

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Katartzis, Antonios

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2022/081875

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2019164322 A1	30-05-2019	US 2019164322 A1	30-05-2019
		US 2021217216 A1	15-07-2021

US 2015178786 A1	25-06-2015	NONE	
