

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization

International Bureau



(10) International Publication Number

WO 2015/094977 A1

(43) International Publication Date

25 June 2015 (25.06.2015)

WIPO | PCT

(51) International Patent Classification:

H04N 5/232 (2006.01) G06T 7/20 (2006.01)

(21) International Application Number:

PCT/US2014/070078

(22) International Filing Date:

12 December 2014 (12.12.2014)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/919,754 21 December 2013 (21.12.2013) US  
14/567,119 11 December 2014 (11.12.2014) US

(71) Applicant: QUALCOMM INCORPORATED [US/US];  
ATTN: International IP Administration, 5775 Morehouse  
Drive, San Diego, California 92121-1714 (US).

(72) Inventors: LI, Christopher; 210 Chiquita Avenue, Mountain View, California 94041 (US). ZHONG, Xin; 5775 Morehouse Drive, San Diego, California 92121-1714 (US). GAO, Dashan; 5775 Morehouse Drive, San Diego, California 92121-1714 (US). QI, Yingyong; 5775 Morehouse Drive, San Diego, CA 92121-1714 (US). GUO, Kai; 5775 Morehouse Drive, San Diego, California 92121-1714 (US).

(74) Agent: TOLER, JEFFREY G.; 8500 Bluffstone Cove, Suite A201, Austin, Texas 78759 (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, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, 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, 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))

(54) Title: SYSTEM AND METHOD TO STABILIZE DISPLAY OF AN OBJECT TRACKING BOX

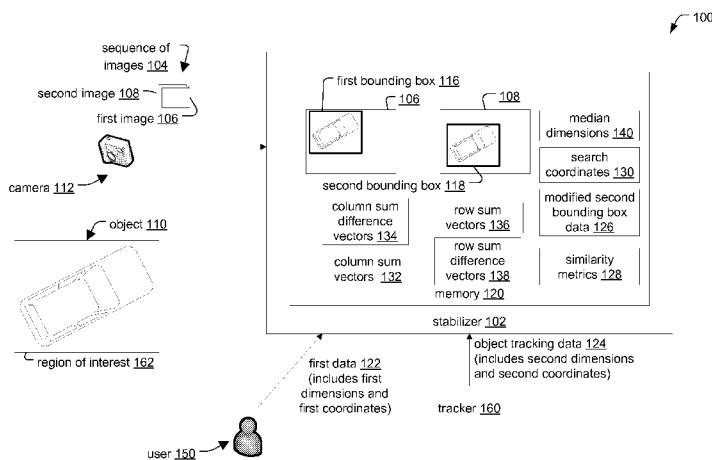


FIG. 1

(57) Abstract: A method includes receiving first data defining a first bounding box for a first image of a sequence of images. The first bounding box corresponds to a region of interest including a tracked object. The method also includes receiving object tracking data for a second image of the sequence of images, the object tracking data defining a second bounding box. The second bounding box corresponds to the region of interest including the tracked object in the second image. The method further includes determining a similarity metric for first pixels within the first bounding box and search pixels within each of multiple search bounding boxes. Search coordinates of each of the search bounding boxes correspond to second coordinates of the second bounding box shifted in one or more directions. The method also includes determining a modified second bounding box based on the similarity metric.

## SYSTEM AND METHOD TO STABILIZE DISPLAY OF AN OBJECT TRACKING BOX

### CROSS-REFERENCE TO RELATED APPLICATIONS

**[0001]** The present application claims priority from commonly owned U.S. Provisional Patent Application No. 61/919,754 filed on December 21, 2013, and U.S. Non-Provisional Patent Application No. 14/567,119 filed on December 11, 2014, the contents of which are expressly incorporated herein by reference in their entirety.

### FIELD

**[0002]** The present disclosure is generally related to stabilizing display of an object tracking box.

### DESCRIPTION OF RELATED ART

**[0003]** Advances in technology have resulted in smaller and more powerful computing devices. For example, there currently exist a variety of portable personal computing devices, including wireless computing devices, such as portable wireless telephones, personal digital assistants (PDAs), and paging devices that are small, lightweight, and easily carried by users. More specifically, portable wireless telephones, such as cellular telephones and internet protocol (IP) telephones, can communicate voice and data packets over wireless networks. Further, many such wireless telephones include other types of devices that are incorporated therein. For example, a wireless telephone can also include a digital still camera, a digital video camera, a digital recorder, and an audio file player. Also, such wireless telephones can process executable instructions, including software applications, such as a web browser application, that can be used to access the Internet. As such, these wireless telephones can include significant computing capabilities.

**[0004]** Electronic devices, such as a wireless telephone, may include a camera. The camera may capture a sequence of images that a user may view in a camera display. The user may select an arbitrary object in an image by selecting a region of the camera display. A tracking algorithm may track the object's motion over subsequent images and may display a box over the tracked object on the camera display. The displayed box may appear unstable due to rapid changes in a location and/or a size of the box between the images. For example, the user may be holding the wireless telephone in a way that

- 2 -

shakes the camera. As another example, the object may move with a high amount of displacement between images.

### **SUMMARY**

**[0005]** Systems and methods of stabilizing display of an object tracking box are disclosed. A user may select an object in an image by selecting a region (e.g., a square or a rectangle) of a camera display in which the image is displayed. The camera display may show a bounding box surrounding the selected object. The object, the camera, or both, may be moving while the sequence of images is captured. A tracker may update coordinates and/or dimensions of the bounding box such that the bounding box approximately tracks the object over subsequent images. Updating the coordinates and/or dimensions may result in the bounding box appearing to “jump” from one image to another. A stabilizer may “smooth” the display of the bounding box (e.g., reduce jitter) from a first image to a subsequent image. For example, the stabilizer may receive first coordinates of a first bounding box corresponding to the first image and may receive updated coordinates (e.g., second coordinates) of a second bounding box corresponding to the subsequent image from the tracker. The stabilizer may determine a search region around the second bounding box and may determine multiple search bounding boxes corresponding to the search region. Each of the search bounding boxes may correspond to a candidate bounding box to replace the second bounding box to reduce jitter. The stabilizer may compare search pixels of each of the search bounding boxes to first pixels of the first bounding box to select a particular search bounding box that is most similar to the first bounding box based on a similarity metric. The stabilizer may replace the second bounding box with the selected search bounding box, which reduces visual jitter associated with the display of boundary boxes in a sequence of images.

**[0006]** In a particular aspect, a method includes receiving first data defining a first bounding box for a first image of a sequence of images. The first bounding box corresponds to a region of interest including a tracked object. The method also includes receiving object tracking data for a second image of the sequence of images, the object tracking data defining a second bounding box. The second bounding box corresponds to the region of interest including the tracked object in the second image. The method further includes determining a similarity metric for first pixels within the first bounding

box and search pixels within each of multiple search bounding boxes. Search coordinates of each of the search bounding boxes correspond to second coordinates of the second bounding box shifted in one or more directions. The method also includes determining a modified second bounding box based on the similarity metric.

**[0007]** In another particular aspect, an apparatus includes a memory and a processor. The memory is configured to store instructions. The processor is configured to execute the instructions to determine a similarity metric for first pixels within a first bounding box of a first image and search pixels within each of multiple search bounding boxes. The first bounding box corresponds to a region of interest including a tracked object. Search coordinates of each of the search bounding boxes correspond to second coordinates of a second bounding box shifted in one or more directions. The first image precedes a second image in a sequence of images. The second bounding box corresponds to the region of interest including the tracked object in the second image. The processor is also configured to execute the instructions to determine a modified second bounding box based on the similarity metric.

**[0008]** In another particular aspect, a computer-readable storage device stores instructions that, when executed by a processor, cause the processor to perform operations including determining a similarity metric for first pixels within a first bounding box of a first image and search pixels within each of multiple search bounding boxes. The first bounding box corresponds to a region of interest including a tracked object. Search coordinates of each of the search bounding boxes correspond to second coordinates of a second bounding box shifted in one or more directions. The first image precedes a second image in a sequence of images. The second bounding box corresponds to the region of interest including the tracked object in the second image. The operations also include determining a modified second bounding box based on the similarity metric.

**[0009]** One particular advantage provided is that an object tracking box (e.g., the bounding box) is stabilized from one image to another in a sequence of images. For example, coordinates of a bounding box may be modified to generate a modified bounding box such that pixels within the modified bounding box are more similar to pixels within a preceding bounding box of a preceding image. As another example, a modified size (e.g., dimensions) of the bounding box may correspond to median

dimensions of a plurality of preceding images. Other aspects, advantages, and features of the present disclosure will become apparent after review of the entire application, including the following sections: Brief Description of the Drawings, Detailed Description, and the Claims.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

**[0010]** FIG. 1 is a block diagram of a particular illustrative embodiment of a system that is operable to stabilize display of an object tracking box;

**[0011]** FIG. 2 is a diagram of a particular illustrative image that includes a tracked object;

**[0012]** FIG. 3 is a diagram of another illustrative image that includes the tracked object of FIG. 2;

**[0013]** FIG. 4 is a diagram of another illustrative image that includes the tracked object of FIG. 2;

**[0014]** FIG. 5 is a flow chart of a particular illustrative embodiment of a method of stabilizing display of an object tracking box;

**[0015]** FIG. 6 is a flow chart of another illustrative embodiment of a method of stabilizing display of an object tracking box; and

**[0016]** FIG. 7 is a block diagram of a device operable to stabilize display of an object tracking box in accordance with the systems and methods of FIGS. 1-6.

### **DETAILED DESCRIPTION**

**[0017]** Referring to FIG. 1, a particular illustrative embodiment of a system operable to stabilize display of an object tracking box is disclosed and generally designated 100. The system 100 includes a stabilizer 102 coupled to a camera 112 and to a tracker 160. The stabilizer 102 may include a memory 120. In a particular embodiment, one or more components of the system 100 may be integrated into a mobile phone, a set top box, a music player, a video player, an entertainment unit, a navigation device, a communications device, a personal digital assistant (PDA), a fixed location data unit, a computer, or a combination thereof.

**[0018]** It should be noted that in the following description, various functions performed by the system 100 of FIG. 1 are described as being performed by certain

- 5 -

components or modules. However, this division of components and modules is for illustration only. In an alternate embodiment, a function performed by a particular component or module may be divided amongst multiple components or modules. Moreover, in an alternate embodiment, two or more components or modules of FIG. 1 may be integrated into a single component or module. Each component or module illustrated in FIG. 1 may be implemented using hardware (e.g., a field-programmable gate array (FPGA) device, an application-specific integrated circuit (ASIC), a digital signal processor (DSP), a controller, etc.), software (e.g., instructions executable by a processor), or any combination thereof.

**[0019]** During operation, the camera 112 may capture a sequence of images 104. In a particular embodiment, the sequence of images 104 may correspond to a video stream that a user is recording (e.g., storing in memory). In another embodiment, the sequence of images 104 may correspond to image data displayed by a camera display (e.g., corresponding to a viewfinder display) over a time period. For example, the user may view the image data and subsequently take a picture (e.g., store a particular image in memory).

**[0020]** The sequence of images 104 may include a first image 106. The first image 106 may be displayed to a user 150 via a camera display (not shown). The user 150 may select an object 110 (e.g., a car in FIG. 1) displayed in the camera display by selecting a region of interest 162 in the first image 106. The region of interest 162 may include the object 110. The stabilizer 102 may receive first data 122 from the user 150 defining a first bounding box 116 corresponding to the region of interest 162. For example, the first data 122 may include first coordinates (e.g., a horizontal axis (x-axis) coordinate and a vertical axis (y-axis) coordinate) of the first bounding box 116 in the first image 106. In a particular embodiment, the first coordinates may correspond to a top-left corner of the first bounding box 116 in the first image 106. The first data 122 may also include first dimensions of the first bounding box 116. For example, the first data may include a width (e.g., along the x-axis of the first image 106) and a height (e.g., along the y-axis of the first image 106) of the first bounding box 116. In this example, the first bounding box 116 may correspond to a square or a rectangle.

**[0021]** The sequence of images 104 may include a second image 108. The second image 108 may also include (e.g., depict) the region of interest 162 including the object

- 6 -

110. The tracker 160 may generate object tracking data 124 defining an object tracking box (e.g., a second bounding box 118). The second bounding box 118 may correspond to the region of interest 162 in the second image 108. For example, the object tracking data 124 may include second coordinates (e.g., an x-axis coordinate and a y-axis coordinate) of the second bounding box 118 in the second image 108. In a particular embodiment, the second coordinates correspond to a top-left corner of the second bounding box 118 in the second image 108. The object tracking data 124 may also include second dimensions of the second bounding box 118. For example, the object tracking data 124 may include a width (e.g., along the x-axis of the second image 108) and a height (e.g., along the y-axis of the second image 108) of the second bounding box 118.

**[0022]** The stabilizer 102 may determine a search region for the second image 108 based on the second bounding box 118. For example, the search region may include pixels within the second bounding box 118 and pixels substantially near the second bounding box 118, as described with reference to FIG. 4. The stabilizer 102 may determine a plurality of search bounding boxes within the search region, as described with reference to FIG. 4. For example, coordinates of each of the search bounding boxes may correspond to the second coordinates of the second bounding box 118 shifted in one or more directions. Each of the search bounding boxes may correspond to a candidate bounding box to replace the second bounding box 118, when displaying the second image 108, to reduce jitter. The stabilizer 102 may store search coordinates 130 of the search bounding boxes in the memory 120. Dimensions of each of the search bounding boxes may correspond to the first dimensions of the first bounding box 116.

**[0023]** The stabilizer 102 may determine a similarity metric for first pixels within the first bounding box 116 and search pixels within each of the search bounding boxes. For example, the similarity metric may include a sum of absolute differences (SAD) metric. To illustrate, the stabilizer 102 may calculate a particular similarity metric for the first pixels and candidate search pixels within a candidate search bounding box based at least in part on a SAD of first pixel characteristics (e.g., a pixel intensity, a pixel color (e.g., red, green, blue, cyan, magenta, yellow, or black) sub-component, or a combination thereof) corresponding to the first pixels and second pixel characteristics corresponding to the candidate search pixels.

**[0024]** In a particular embodiment, the stabilizer 102 may calculate a first column sum vector, a first column sum difference vector, a first row sum vector, and/or a first row sum difference vector of the first pixels of the first bounding box 116, as described with respect to FIG. 2. Similarly, the stabilizer 102 may also calculate a second column sum vector, a second column sum difference vector, a second row sum vector, and/or a second row sum difference vector of the candidate search pixels of the candidate search bounding box. The stabilizer 102 may store column sum vectors 132 (e.g., the first column sum vector and the second column sum vector), column sum difference vectors 134 (e.g., the first column sum difference vector and the second column sum difference vector), row sum vectors 136 (e.g., the first row sum vector and the second row sum vector), row sum difference vectors 138 (e.g., the first row sum difference vector and the second row sum difference vector), or a combination thereof, in the memory 120.

**[0025]** The stabilizer 102 may determine the particular similarity metric for the first pixels of the first bounding box 116 and the candidate search pixels of the candidate search bounding box by adding together a first SAD of the first column sum vector and the second column sum vector, a second SAD of the first column sum difference vector and the second column sum difference vector, a third SAD of a first row sum vector and the second row sum vector, and/or a fourth SAD of a first row sum difference vector and the second row sum difference vector. The stabilizer 102 may store similarity metrics 128 corresponding to each of the search bounding boxes in the memory 120.

**[0026]** The stabilizer 102 may select a particular search bounding box that includes search pixels that are most similar to the first pixels of the first bounding box 116. For example, the stabilizer 102 may select the particular search bounding box in response to determining that a corresponding similarity metric indicates a highest similarity with the first pixels (e.g., has a lowest value) of the similarity metrics 128.

**[0027]** The stabilizer 102 may determine a modified second bounding box based on the selected search bounding box. For example, the stabilizer 102 may generate modified second bounding box data 126. The modified second bounding box data 126 may indicate modified coordinates of a modified second bounding box. The modified coordinates may correspond to coordinates of the selected search bounding box. In a particular embodiment, the modified coordinates may be identical to the second coordinates of the second bounding box 118. Thus, the particular search bounding box

may be selected from amongst multiple search (e.g., candidate) bounding boxes because the particular search bounding box is determined to be most similar to the first bounding box 116, thereby reducing visual jitter in bounding box placement between images in the sequence of images 104.

**[0028]** The modified second bounding box data 126 may indicate dimensions of the modified second bounding box. In a particular embodiment, the dimensions of the modified second bounding box may correspond to the first dimensions of the first bounding box 116 or the second dimensions of the second bounding box 118. In another embodiment, the dimensions of the modified second bounding box may correspond to median dimensions 140 corresponding to a plurality of images preceding the second image 108. The stabilizer 102 may use the median dimensions 140 as the dimensions of the modified second bounding box in response to determining that the number of the preceding images satisfies a threshold. The threshold number of preceding images may be a default value. The stabilizer 102 may send the modified second bounding box data 126 to the camera display. For example, the camera display may display the second image 108 with the modified second bounding box.

**[0029]** In a particular embodiment, the stabilizer 102 may store pixel characteristics corresponding to the modified second bounding box in anticipation of receiving an image subsequent to the second image 108 (e.g., a third image of the sequence of images 104). The modified second bounding box may correspond to the selected search bounding box when the modified dimensions of the modified second bounding box correspond to dimensions of the selected search bounding box (i.e., the first dimensions of the first bounding box 116). The stabilizer 102 may store the pixel characteristics of the selected search bounding box in response to determining that the modified dimensions of the modified second bounding box correspond to the dimensions of the selected search bounding box (or the first dimensions of the first bounding box 116).

**[0030]** In a particular embodiment, the modified dimensions of the modified second bounding box may be distinct from the dimensions of the selected search bounding box (or the first dimensions of the first bounding box 116). For example, the modified dimensions of the modified second bounding box may correspond to the second dimensions of the second bounding box 118 or to the median dimensions 140. When the dimensions of the modified second bounding box do not correspond to the

dimensions of the selected search bounding box (or the first dimensions), the stabilizer 102 may generate and store pixel characteristics of the modified second bounding box. For example, the stabilizer 102 may generate a row sum vector, a row sum difference vector, a column sum vector, and/or a column sum difference vector, as described with reference to FIG.2, corresponding to the modified second bounding box.

**[0031]** As additional images of the sequence of images 104 are received, additional bounding boxes may be selected based on similarity to previous bounding box(es). For example, the stabilizer 102 may receive the third image and may receive object tracking data 124 defining a third bounding box corresponding to the third image. The stabilizer 102 may determine a modified third bounding box based on the modified second bounding box and the third bounding box. For example, the stabilizer 102 may use stored pixel characteristics of the modified second bounding box to generate additional similarity metrics that can be used to determine the modified third bounding box.

**[0032]** Thus, the stabilizer 102 may select a particular bounding box from amongst multiple search (e.g., candidate) bounding boxes because the particular bounding box is determined to be most similar to a bounding box of a preceding image (e.g., the first image 106), thereby reducing visual jitter in boundary box placement between images in the sequence of images 104.

**[0033]** Referring to FIG. 2, a diagram of an illustrative image that includes a tracked object is disclosed and generally designated 200. In a particular embodiment, the image 200 may correspond to the first image 106 or the second image 108 of FIG. 1. The image 200 includes a plurality of columns (e.g., column 0 – column 5) and a plurality of rows (e.g., row 0 – row 4) of pixels. The image 200 may include a first bounding box 202 including a set of columns (e.g., column 0 – column 2) and a set of rows (e.g., row 0 – row 2) of pixels, which are illustrated as boxed pixels. In a particular embodiment, the boxed pixels may correspond to pixels of the first bounding box 116, pixels of one or more search bounding boxes, or both. A particular pixel may be identified by its coordinates (e.g., x, y coordinates). The image 200 indicates a particular pixel characteristic  $I$  of each pixel (x, y). In a particular embodiment, the pixel characteristic may correspond to pixel intensity. For example,  $I(0,0)$  may indicate the pixel intensity of the pixel at column 0 and row 0, and  $I(0, 1)$  may indicate the pixel intensity of the

- 10 -

pixel at column 0 and row 1. In a particular embodiment, the image 200 may be a grayscale image.

**[0034]** The first bounding box 202 may be defined for the image 200. The pixel at (0, 0) defines a top-left corner of the first bounding box 202. Dimensions of the first bounding box 202 may include height of 3 pixels and width of 3 pixels. The stabilizer 102 may determine the bottom-right corner (e.g., (2, 2)) of the first bounding box 202 based on the coordinates of the top-left corner and the dimensions. In a particular embodiment, the first bounding box 202 may correspond to the first bounding box 116 of FIG. 1. In another embodiment, the first bounding box 202 may correspond to one or more of the search bounding boxes described with reference to FIG. 1.

**[0035]** The stabilizer 102 may determine a column sum vector  $c(x, y)$  of a bounding box. The coordinates of the top-left corner of the bounding box may correspond to (x, y). For example, the stabilizer 102 may determine a first column sum vector ( $c(0,0)$ ) 204 of the boxed pixels of the first bounding box 202. In the example shown in FIG. 2,  $I(0,0)=1$ ,  $I(1,0)=2$ ,  $I(2,0)=3$ ,  $I(3,0)=10$ ,  $I(0,1)=4$ ,  $I(1,1)=5$ ,  $I(2,1)=6$ ,  $I(3,1)=11$ ,  $I(0,2)=7$ ,  $I(1,2)=8$ ,  $I(2,2)=9$ ,  $I(3,2)=12$ ,  $I(0,3)=13$ ,  $I(1,3)=14$ ,  $I(2,3)=15$ , and  $I(3,3)=16$ . Each element  $c_j(0,0)$  of  $c(0,0)$  204 may be equal to a sum of pixel values of column  $j$  of the first bounding box 202. Thus, for the example shown in FIG. 2, the values of  $c_j(0,0)$  204 are:

$$c_0(0,0) = I(0, 0)+I(0,1)+I(0,2) = 12,$$

$$c_1(0,0) = I(1,0)+I(1,1)+I(1,2) = 15, \text{ and}$$

$$c_2(0,0) = I(2, 0)+I(2, 1)+I(2, 2) = 18.$$

**[0036]** The stabilizer 102 may determine a column sum difference vector  $c\delta(0,0)$  206 of the boxed pixels of the first bounding box 202. Each element  $c\delta_{j,0}(0,0)$  of  $c\delta(0,0)$  206 may be equal to  $c_j(0,0) - c_{j-1}(0,0)$  for  $j > 0$  and may be equal to 0 for  $j=0$ . Thus, for the example shown in FIG. 2, the values of  $c\delta(0,0)$  206 are:

$$c\delta_{0,0}(0,0) = 0,$$

$$c\delta_{1,0}(0,0) = c_1(0,0) - c_0(0,0) = 15 - 12 = 3, \text{ and}$$

$$c\delta_{2,0}(0,0) = c_2(0,0) - c_1(0,0) = 18 - 15 = 3.$$

- 11 -

**[0037]** The stabilizer 102 may determine a row sum vector  $r(0,0)$  208 of the boxed pixels of the first bounding box 202. Each element  $r_i(0,0)$  of  $r(0,0)$  208 may be equal to a sum of pixel values of row  $i$  of the first bounding box 202. Thus, for the example shown in FIG. 2, values of  $r(0,0)$  208 are:

$$\begin{aligned}r_0(0,0) &= I(0,0) + I(1,0) + I(2,0) = 6, \\r_1(0,0) &= I(0,1) + I(1,1) + I(2,1) = 15, \text{ and} \\r_2(0,0) &= I(0,2) + I(1,2) + I(2,2) = 24.\end{aligned}$$

**[0038]** The stabilizer 102 may determine a first row difference vector  $r\delta(0,0)$  210 of the boxed pixels of the first bounding box 202. Each element  $r\delta_i(0,0)$  of  $r\delta(0,0)$  210 may be equal to  $r_i(0,0) - r_{i-1}(0,0)$  for  $i > 0$  and may be equal to 0 for  $i = 0$ . Thus, for the example shown in FIG. 2, the values of  $r\delta(0,0)$  210 are:

$$\begin{aligned}r\delta_0(0,0) &= 0, \\r\delta_1(0,0) &= r_1(0,0) - r_0(0,0) = 15 - 6 = 9, \text{ and} \\r\delta_2(0,0) &= r_2(0,0) - r_1(0,0) = 24 - 15 = 9.\end{aligned}$$

**[0039]** In a particular embodiment, the stabilizer 102 may generate an integral image corresponding to the image 200. The value of each particular pixel of the integral image is equal to a sum of pixel values of the particular pixel and of pixels above and to the left of the particular pixel. The value of the pixel  $(x, y)$  in the integral image  $Int$  may be denoted by  $Int(x, y)$ .  $Int(x, y)$  may be determined by the formula  $Int(x, y) = I(x, y) + Int(x-1, y) + Int(x, y-1) - Int(x-1, y-1)$ .

**[0040]** In the example shown in FIG. 2, the values of a subset of the pixels of the integral image are:

- 12 -

$$\text{Int}(0, 0) = I(0,0) = 1$$

$$\text{Int}(0,1) = I(0,1) + I(0,0) = 5$$

$$\text{Int}(0,2) = I(0,2) + I(0,1) + I(0,0) = 12$$

$$\text{Int}(0,3) = I(0,3) + I(0,2) + I(0,1) + I(0,0) = 25$$

$$\text{Int}(1,0) = I(1,0) + I(0,0) = 3$$

$$\text{Int}(1,1) = I(1,1) + I(0,1) + I(0,0) + I(1,0) = 12$$

$$\text{Int}(1,2) = I(1,2) + I(1,1) + I(0,1) + I(0,0) + I(1,0) + I(0,2) = 27$$

$$\text{Int}(1,3) = I(1,3) + I(1,2) + I(1,1) + I(0,1) + I(0,0) + I(1,0) + I(0,2) + I(0,3) = 54$$

$$\text{Int}(2,0) = I(2,0) + I(1,0) + I(0,0) = 6$$

$$\text{Int}(2,1) = I(2,1) + I(2,0) + I(1,0) + I(0,0) + I(1,1) + I(0,1) = 21$$

$$\text{Int}(2,2) = I(2,2) + I(2,1) + I(2,0) + I(1,0) + I(0,0) + I(1,1) + I(0,1) + I(1,2) + I(0,2) = 45$$

$$\text{Int}(2,3) = I(2,3) + I(2,2) + I(2,1) + I(2,0) + I(1,0) + I(0,0) + I(1,1) + I(0,1) + I(1,2) + I(0,2) + I(0,3) + I(1,3) = 87$$

$$\text{Int}(3,0) = I(3,0) + I(2,0) + I(1,0) + I(0,0) = 16$$

$$\text{Int}(3,1) = I(3,1) + I(3,0) + I(2,1) + I(2,0) + I(1,0) + I(0,0) + I(1,1) + I(0,1) = 42$$

$$\text{Int}(3,2) = I(3,2) + I(3,1) + I(3,0) + I(2,2) + I(2,1) + I(2,0) + I(1,0) + I(0,0) + I(1,1) + I(0,1) + I(1,2) + I(0,2) = 78$$

$$\text{Int}(3,3) = I(3,3) + I(3,2) + I(3,1) + I(3,0) + I(2,3) + I(2,2) + I(2,1) + I(2,0) + I(1,0) + I(0,0) + I(1,1) + I(0,1) + I(1,2) + I(0,2) + I(0,3) + I(1,3) = 136$$

**[0041]** The stabilizer 102 may determine the values of  $c(0,0)$  204 and  $r(0,0)$  208 from the integral image. For example, a particular element of  $c(0,0)$  204 may correspond to a difference of a first element of the integral image and a second element of the integral image, where the first element and the second element correspond to adjacent columns of the integral image. To illustrate,  $c_0(0,0)$  corresponds to  $\text{Int}(0,2)$ ,  $c_1(0,0)$  corresponds to  $\text{Int}(1,2)-\text{Int}(0,2)$ , and  $c_2(0,0)$  corresponds to  $\text{Int}(2,2)-\text{Int}(1,2)$ .

**[0042]** In a particular embodiment, values of  $c_i(x,0)$  of a bounding box (e.g., the first bounding box 202) with top-left coordinates  $(x,0)$  are:

- 13 -

$c_i(x, 0) = \text{Int}(0, \text{height}-1)$ , for  $x+i=0$ , and

$c_i(x, 0) = \text{Int}(x+i, \text{height}-1) - \text{Int}(x+i-1, \text{height}-1)$ , for  $x+i>0$ ,

where height (e.g., 3) corresponds to a number of rows of the bounding box.

**[0043]** As another example, a particular element of  $r(0,0)$  208 may correspond to a difference of a first element of the integral image and a second element of the integral image, where the first element and the second element correspond to adjacent rows of the integral image. To illustrate,  $r_0(0,0)$  corresponds to  $\text{Int}(2,0)$ ,  $r_1(0,0)$  corresponds to  $\text{Int}(2,1)-\text{Int}(2,0)$ , and  $r_2(0,0)$  corresponds to  $\text{Int}(2,2)-\text{Int}(2,1)$ .

**[0044]** In a particular embodiment, values of  $r_i(0,y)$  of a bounding box (e.g., the first bounding box 202) with top-left coordinates  $(0,y)$  are:

$r_j(0,y) = \text{Int}(\text{width}-1, 0)$ , for  $y+j=0$ , and

$r_j(0,y) = \text{Int}(\text{width}-1, y+j) - \text{Int}(\text{width}-1, y+j-1)$ , for  $y+j>0$ ,

where width (e.g., 3) corresponds to a number of columns of the bounding box.

**[0045]** As another example, the stabilizer 102 may determine a row sum vector  $r(1,1)$  and a column sum vector  $c(1,1)$  for a second bounding box 204 using the integral image. For example, a particular element of  $c(1,1)$  may correspond to a difference of a first element of the integral image and a second element of the integral image, where the first element and the second element correspond to adjacent columns of the integral image. To illustrate,  $c_0(1,1)$  corresponds to  $\text{Int}(1,3) - \text{Int}(0,3) - \text{Int}(1,0) + \text{Int}(0,0)$ ,  $c_1(1,1)$  corresponds to  $\text{Int}(2,3)-\text{Int}(1,3)-\text{Int}(2,0)+\text{Int}(1,0)$ , and  $c_2(1,1)$  corresponds to  $\text{Int}(3,3)-\text{Int}(2,3)-\text{Int}(3,0)+\text{Int}(2,0)$ .

**[0046]** In a particular embodiment, values of  $c_i(x,y)$  of a bounding box (e.g., the second bounding box 204) with top-left coordinates  $(x,y)$ , where  $y>0$  are:

$c_i(x,y) = \text{Int}(0, y+\text{height}-1) - \text{Int}(0, y-1)$ , for  $x+i=0$ , and

$c_i(x,y) = \text{Int}(x+i, y+\text{height}-1) - \text{Int}(x+i-1, y+\text{height}-1) - \text{Int}(x+i, y-1) + \text{Int}(x+i-1, y-1)$ , for  $x+i>0$ ,

where height (e.g., 3) corresponds to a number of rows of the bounding box.

**[0047]** As another example, a particular element of  $r(1,1)$  may correspond to a difference of a first element of the integral image and a second element of the integral image, where the first element and the second element correspond to adjacent rows of the integral image. To illustrate,  $r_0(1,1)$  corresponds to  $\text{Int}(3,1)-\text{Int}(0,1)-\text{Int}(3,0)+\text{Int}(0,0)$ ,  $r_1(1,1)$  corresponds to  $\text{Int}(3,2)-\text{Int}(0,2)-\text{Int}(3,1)+\text{Int}(0,1)$ , and  $r_2(1,1)$  corresponds to  $\text{Int}(3,3)-\text{Int}(0,3)-\text{Int}(3,2)+\text{Int}(0,2)$ .

**[0048]** In a particular embodiment, values of  $r_j(x,y)$  of a bounding box (e.g., the second bounding box 204) with top-left coordinates  $(x,y)$ , where  $x>0$  are:

$$r_j(x,y) = \text{Int}(x+\text{width}-1, 0) - \text{Int}(x-1, 0), \text{ for } y+j=0, \text{ and}$$

$$r_j(x,y) = \text{Int}(x+\text{width}-1, y+j) - \text{Int}(x+\text{width}-1, y+j-1) - \text{Int}(x-1, y+j) + \text{Int}(x-1, y+j-1), \text{ for } y+j>0,$$

where width (e.g., 3) corresponds to a number of columns of the bounding box.

**[0049]** In a particular embodiment, the stabilizer 102 may generate an integral image corresponding to the second image 108. Calculating row sum vectors and column sum vectors corresponding to each of the search bounding boxes from the integral image may be faster (e.g., computed in constant time) and may use fewer processing resources than calculating the vectors directly from the pixel values of the second image 108.

**[0050]** Referring to FIG. 3, a diagram of an illustrative image that includes the tracked object of FIG. 2 is disclosed and generally designated 300. The image 300 may correspond to the second image 108 of FIG. 1. The top-left coordinates  $(0, 1)$  and dimensions  $4 \times 3$  (pixel width x pixel height) define the second bounding box 118 for the image 300. Each pixel value  $I_2(x, y)$  may correspond to a pixel characteristic of the corresponding pixel  $(x, y)$  of the image 300. The subscript 2 indicates that the pixel characteristic corresponds to the second image 108.

**[0051]** The stabilizer 102 may determine a search region 304 based on the second bounding box 118. For example, the search region 304 may include the second bounding box 118 and additional pixels in one or more directions relative to the second bounding box 118. For example, stabilizer 102 may add up to a first number (e.g., 1) of pixels to the right and to the left of the second bounding box 118 and up to a second

- 15 -

number (e.g., 1) of pixels to the top and to the bottom of the second bounding box 118 to generate the search region 304.

**[0052]** Referring to FIG. 4, a diagram of an illustrative image that includes the tracked object of FIG. 2 is disclosed and generally designated 400. The image 400 may correspond to the second image 108 of FIG. 1. The search region 304 may be defined for the image 400.

**[0053]** The stabilizer 102 may generate a plurality of search bounding boxes (e.g., a first search bounding box 402, a second search bounding box 404, and a third search bounding box 406) within the image 400 based on the search region 304. Dimensions of each of the search bounding boxes may be equal to the dimensions of the first bounding box 116. The stabilizer 102 may generate search bounding boxes with x-coordinates of a top-left pixel selected from a first range (e.g., 0-3) and y-coordinates of the top-left pixel selected from a second range (e.g., 0-2). The search region 304 may include pixels with x-coordinates outside the first range or y-coordinates outside the second range. However, the stabilizer 102 may refrain from generating search bounding boxes with top-left pixels having x-coordinates outside the first range, y-coordinates outside the second range, or both, because such search bounding boxes do not entirely fit within the image 400.

**[0054]** The stabilizer 102 may generate a search row sum vector, a search row sum difference vector, a search column sum vector, and/or a search column sum difference vector, as further described with reference to FIGS. 1-2, corresponding to each of the search bounding boxes 402-406. The stabilizer 102 may generate the similarity metrics 128 and may select a particular search bounding box as “most similar” to the first bounding box 116, as described with reference to FIG. 1. For example, the stabilizer 102 may select the second search bounding box 404 based on the similarity metric of the second search bounding box 404. Thus, the stabilizer 102 may generate the modified second bounding box data 126 to indicate that a modified bounding box displayed for the user 150 should have a top-left pixel at the location (1, 1).

**[0055]** In a particular embodiment, dimensions of the modified second bounding box may correspond to first dimensions of the first bounding box 116 of FIG. 1, second dimensions of the second bounding box 118 of FIG. 1, or the median dimensions 140 of FIG. 1. The modified second bounding box data 126 may indicate the coordinates (1,1)

of the top-left pixel and the dimensions of the modified second bounding box. The stabilizer 102 may provide the modified second bounding box data 126 to a camera display. The camera display may display the second image 108 with the modified second bounding box. The modified second bounding box may correspond to a more stable bounding box than the second bounding box 118. For example, pixels corresponding to the modified second bounding box may be more similar to the first pixels of the first bounding box 116. As another example, dimensions of the modified second bounding box may correspond more closely to dimensions of a plurality of preceding images.

**[0056]** Referring to FIG. 5, a flow chart of a particular illustrative embodiment of a method of stabilizing display of an object tracking box is shown and generally designated 500. The method 500 may be performed by one or more components of the system 100 of FIG. 1.

**[0057]** The method 500 includes receiving first data defining a first bounding box for a first image of a sequence of images, at 502. The first bounding box may correspond to a region of interest including a tracked object. For example, the stabilizer 102 of FIG. 1 may receive the first data 122, as described with reference to FIG. 1. The first data 122 may define the first bounding box 116 for the first image 106 of the sequences of images 104. The first bounding box 116 may correspond to the region of interest 162 including the object 110.

**[0058]** The method 500 also includes receiving object tracking data for a second image of the sequence of images, at 504. The object tracking data may define a second bounding box. The second bounding box may correspond to the region of interest including the tracked object in the second image. For example, the stabilizer 102 of FIG. 1 may receive the object tracking data 124 for the second image 108, as described with reference to FIG. 1. The object tracking data 124 may define the second bounding box 118. The second bounding box 118 may correspond to the region of interest 162 including the object 110 in the second image 108.

**[0059]** The method 500 further includes determining a similarity metric for first pixels within the first bounding box and search pixels within each of multiple search bounding boxes, at 506. Search coordinates of each of the search bounding boxes may correspond to second coordinates of the second bounding box shifted in one or more directions. For

example, the stabilizer 102 of FIG. 1 may determine a similarity metric for the first pixels within the first bounding box 116 and search pixels within each of multiple search bounding boxes, as described with reference to FIGS. 1-2 and 4.

**[0060]** The method 500 also includes determining a modified second bounding box based on the similarity metric, at 508. For example, the stabilizer 102 of FIG. 1 may determine a modified second bounding box based on the similarity metric, as described with reference to FIGS. 1 and 4.

**[0061]** The method 500 further includes determining median dimensions corresponding to a plurality of images, at 510. The plurality of images may precede the second image in the sequence of images. Second dimensions of the modified second bounding box may correspond to the median dimensions. For example, the stabilizer 102 of FIG. 1 may determine the median dimensions 140 corresponding to a plurality of images preceding the second image 108, as described with reference to FIG. 1. Second dimensions of the modified second bounding box may correspond to the median dimensions 140, as described with reference to FIGS. 1 and 4.

**[0062]** Thus, the method 500 includes selection of a particular bounding box from amongst multiple search (e.g., candidate) bounding boxes because the particular bounding box is determined to be most similar to a bounding box of a preceding image (e.g., the first image 106), thereby reducing visual jitter in boundary box placement between images in the sequence of images 104.

**[0063]** The method 500 of FIG. 5 may be implemented by a field-programmable gate array (FPGA) device, an application-specific integrated circuit (ASIC), a processing unit such as a central processing unit (CPU), a digital signal processor (DSP), a controller, another hardware device, firmware device, or any combination thereof. As an example, the method 500 of FIG. 5 can be performed by a processor that executes instructions, as described with respect to FIG. 7.

**[0064]** Referring to FIG. 6, a flow chart of a particular illustrative embodiment of a method of stabilizing display of an object tracking box is shown and generally designated 600. The method 600 may be performed by one or more components of the system 100 of FIG. 1. In a particular embodiment, the method 600 may correspond to the operation illustrated at step 506 of FIG. 5.

- 18 -

**[0065]** The method 600 includes determining a first column sum vector of the first pixels, at 602. For example, the stabilizer 102 of FIG. 1 may determine a first column sum vector of the first pixels of the first bounding box 116, as described with reference to FIGS. 1-2. A particular element of the first column sum vector may correspond to a sum of pixel values corresponding to a particular column of the first bounding box 116.

**[0066]** The method 600 also includes determining a first column sum difference vector of the first pixels, at 604. For example, the stabilizer 102 of FIG. 1 may determine a first column difference vector of the first pixels of the first bounding box 116, as described with reference to FIGS. 1-2. A particular element of the first column sum difference vector may correspond to a difference between a first element of the first column sum vector and a second element of the first column sum vector. The first element and the second element may correspond to adjacent columns of the first bounding box 116.

**[0067]** The method 600 further includes determining a first row sum vector of the first pixels, at 606. For example, the stabilizer 102 of FIG. 1 may determine a first row sum vector of the first pixels of the first bounding box 116, as described with reference to FIG. 1. A particular element of the first row sum vector may correspond to a sum of pixel values corresponding to a particular row of the first bounding box 116.

**[0068]** The method 600 also includes determining a first row sum difference vector of the first pixels, at 608. For example, the stabilizer 102 of FIG. 1 may determine a first row sum difference vector of the first pixels of the first bounding box 116, as described with reference to FIGS. 1-2. A particular element of the first row sum difference vector may correspond to a difference between a first element of the first row sum vector and a second element of the row sum vector. The first element and the second element may correspond to adjacent rows of the first bounding box 116.

**[0069]** The method 600 further includes calculating a particular similarity metric for the first pixels and particular search pixels within a particular search bounding box by adding: a first sum of absolute differences (SAD) of the first column sum vector of the first pixels and a second column sum vector of the particular search pixels, a second SAD of the first column sum difference vector of the first pixels and a second column sum difference vector of the particular search pixels, a third SAD of the first row sum vector of the first pixels and a second row sum vector of the particular search pixels,

and a fourth SAD of the first row sum difference vector of the first pixels and a second row sum difference vector of the particular search pixels, at 610. For example, the stabilizer 102 of FIG. 1 may calculate a particular similarity metric for the first pixels of the first bounding box 116 and particular search pixels within a particular search bounding box, as described with reference to FIGS. 1-2 and 4.

**[0070]** It should be noted that although various embodiments are described as utilizing column sum vectors, column sum difference vectors, row sum vectors, and row sum difference vectors, this is for example only and not to be considered limiting. In alternate embodiments, similarity may be determined based on fewer, more, or different computations and data structures.

**[0071]** The method 600 of FIG. 6 may be implemented by a field-programmable gate array (FPGA) device, an application-specific integrated circuit (ASIC), a processing unit such as a central processing unit (CPU), a digital signal processor (DSP), a controller, another hardware device, firmware device, or any combination thereof. As an example, the method 600 of FIG. 6 can be performed by a processor that executes instructions, as described with respect to FIG. 7.

**[0072]** Referring to FIG. 7, a block diagram of a particular illustrative embodiment of a device (e.g., a wireless communication device) is depicted and generally designated 700. The device 700 includes a processor 710, such as a digital signal processor (DSP) or a central processing unit (CPU), coupled to a memory 732. The processor 710 may include and/or execute the stabilizer 102 of FIG. 1 and the tracker 160 of FIG. 1, or both.

**[0073]** One or more components of the device 700 may be implemented via dedicated hardware (e.g., circuitry), by a processor executing instructions to perform one or more tasks, or a combination thereof. As an example, the memory 732 or one or more components of the stabilizer 102 and/or the tracker 160 may be a memory device, such as a random access memory (RAM), magnetoresistive random access memory (MRAM), spin-torque transfer MRAM (STT-MRAM), flash memory, read-only memory (ROM), programmable read-only memory (PROM), erasable programmable read-only memory (EPROM), electrically erasable programmable read-only memory (EEPROM), registers, hard disk, a removable disk, or a compact disc read-only memory (CD-ROM). The memory device may include instructions that, when executed by a computer (e.g., the processor 710), may cause the computer to perform at least a portion

of the method 500 of FIG. 5, the method 600 of FIG. 6, or a combination thereof. As an example, the memory 732 or the one or more components of the stabilizer 102 may be a non-transitory computer-readable medium that includes instructions that, when executed by a computer (e.g., the processor 710), cause the computer perform at least a portion of the method 500 of FIG. 5, the method 600 of FIG. 6, or a combination thereof.

**[0074]** FIG. 7 also shows a display controller 726 that is coupled to the processor 710 and to a display 728. A camera, such as, the camera 112 of FIG. 1, may be coupled to the processor 710. A coder/decoder (CODEC) 734 can also be coupled to the processor 710. A speaker 736 and a microphone 738 can be coupled to the CODEC 734.

**[0075]** FIG. 7 also indicates that a wireless controller 740 can be coupled to the processor 710 and to a wireless antenna 742. In a particular embodiment, the processor 710, the display controller 726, the memory 732, the CODEC 734, and the wireless controller 740 are included in a system-in-package or system-on-chip device 722. In a particular embodiment, the camera 112, an input device 730 and a power supply 744 are coupled to the system-on-chip device 722. Moreover, in a particular embodiment, as illustrated in FIG. 7, the display 728, the camera 112, the stabilizer 102, the tracker 160, the input device 730, the speaker 736, the microphone 738, the wireless antenna 742, and the power supply 744 are external to the system-on-chip device 722. However, each of the display 728, the camera 112, the stabilizer 102, the tracker 160, the input device 730, the speaker 736, the microphone 738, the wireless antenna 742, and the power supply 744 can be coupled to a component of the system-on-chip device 722, such as an interface or a controller.

**[0076]** In conjunction with the described embodiments, a system is disclosed that includes means for receiving first data. The first data may define a first bounding box for a first image of a sequence of images. The first bounding box may correspond to a region of interest including a tracked object. The means for receiving may include the input device 730 of FIG. 7, one or more other devices or circuits configured to receive the data defining a bounding box (e.g., a touchscreen of a mobile phone), or any combination thereof.

**[0077]** The system may also include means for generating object tracking data. The object tracking data may correspond to a second image of the sequence of images. The object tracking data may define a second bounding box. The second bounding box may

correspond to the region of interest including the tracked object in the second image. The means for generating may include the tracker 160 of FIG. 1 and FIG. 7, one or more other devices or circuits configured to generate object tracking data, or any combination thereof.

**[0078]** The system may further include means for determining a similarity metric and a modified second bounding box based on the similarity metric. The similarity metric may be determined for first pixels within the first bounding box and search pixels within each of multiple search bounding boxes. Search coordinates of each of the search bounding boxes may correspond to second coordinates of the second bounding box shifted in one or more directions. The means for determining may include the processor 710 of FIG. 7, the stabilizer 102 of FIG. 1 and 7, one or more other devices or circuits configured to determine a similarity metric and a modified bounding box, or any combination thereof.

**[0079]** Those of skill would further appreciate that the various illustrative logical blocks, configurations, modules, circuits, and algorithm steps described in connection with the embodiments disclosed herein may be implemented as electronic hardware, computer software executed by a processor, or combinations of both. Various illustrative components, blocks, configurations, modules, circuits, and steps have been described above generally in terms of their functionality. Whether such functionality is implemented as hardware or processor executable instructions depends upon the particular application and design constraints imposed on the overall system. Skilled artisans may implement the described functionality in varying ways for each particular application, but such implementation decisions should not be interpreted as causing a departure from the scope of the present disclosure.

**[0080]** The steps of a method or algorithm described in connection with the embodiments disclosed herein may be embodied directly in hardware, in a software module executed by a processor, or in a combination of the two. A software module may reside in random access memory (RAM), flash memory, read-only memory (ROM), programmable read-only memory (PROM), erasable programmable read-only memory (EPROM), electrically erasable programmable read-only memory (EEPROM), registers, hard disk, a removable disk, a compact disc read-only memory (CD-ROM), or any other form of non-transient storage medium known in the art. An exemplary

storage medium is coupled to the processor such that the processor can read information from, and write information to, the storage medium. In the alternative, the storage medium may be integral to the processor. The processor and the storage medium may reside in an application-specific integrated circuit (ASIC). The ASIC may reside in a computing device or a user terminal. In the alternative, the processor and the storage medium may reside as discrete components in a computing device or user terminal.

**[0081]** The previous description of the disclosed embodiments is provided to enable a person skilled in the art to make or use the disclosed embodiments. Various modifications to these embodiments will be readily apparent to those skilled in the art, and the principles defined herein may be applied to other embodiments without departing from the scope of the disclosure. Thus, the present disclosure is not intended to be limited to the embodiments shown herein but is to be accorded the widest scope possible consistent with the principles and novel features as defined by the following claims.

**CLAIMS:**

1. A method comprising:

receiving first data defining a first bounding box for a first image of a sequence of images, wherein the first bounding box corresponds to a region of interest including a tracked object;

receiving object tracking data for a second image of the sequence of images, the object tracking data defining a second bounding box, wherein the second bounding box corresponds to the region of interest including the tracked object in the second image;

determining a similarity metric for first pixels within the first bounding box and search pixels within each of multiple search bounding boxes, wherein search coordinates of each of the search bounding boxes correspond to second coordinates of the second bounding box shifted in one or more directions; and

determining a modified second bounding box based on the similarity metric.

2. The method of claim 1, wherein the first data includes first coordinates and first dimensions of the first bounding box and wherein the object tracking data includes the second coordinates and second dimensions of the second bounding box.

3. The method of claim 2, wherein search dimensions of each of the search bounding boxes correspond to the first dimensions of the first bounding box.

4. The method of claim 1, wherein the similarity metric includes a sum of absolute differences (SAD) metric.

5. The method of claim 1, further comprising:  
selecting a particular search bounding box of the search bounding boxes based on a particular similarity metric for the first pixels and particular search pixels within the particular search bounding box,  
wherein modified second coordinates of the modified second bounding box correspond to particular search coordinates of the particular search bounding box.
6. The method of claim 5, further comprising:  
calculating the particular similarity metric based at least in part on a sum of absolute differences (SAD) of first pixel characteristics corresponding to the first pixels and second pixel characteristics corresponding to the particular search pixels.
7. The method of claim 5, further comprising:  
calculating the particular similarity metric by adding:
  - a first sum of absolute differences (SAD) of a first column sum vector of the first pixels and a second column sum vector of the particular search pixels,
  - a second SAD of a first column sum difference vector of the first pixels and a second column sum difference vector of the particular search pixels,
  - a third SAD of a first row sum vector of the first pixels and a second row sum vector of the particular search pixels, and
  - a fourth SAD of a first row sum difference vector of the first pixels and a second row sum difference vector of the particular search pixels.
8. The method of claim 1, further comprising determining a column sum vector of the first pixels, wherein a particular element of the column sum vector corresponds to a sum of pixel values corresponding to a particular column of the first bounding box.
9. The method of claim 8, wherein each of the pixel values identifies a pixel intensity of a corresponding pixel.

- 25 -

10. The method of claim 1, further comprising determining a column sum difference vector of the first pixels, wherein a particular element of the column sum difference vector corresponds to a difference between a first element of the column sum vector and a second element of the column sum vector, and wherein the first element and the second element correspond to adjacent columns of the first bounding box.
11. The method of claim 1, further comprising determining a row sum vector of the first pixels, wherein a particular element of the row sum vector corresponds to a sum of pixel values corresponding to a particular row of the first bounding box.
12. The method of claim 1, further comprising determining a row sum difference vector of the first pixels, wherein a particular element of the row sum difference vector corresponds to a difference between a first element of the row sum vector and a second element of the row sum vector, and wherein the first element and the second element correspond to adjacent rows of the first bounding box.
13. The method of claim 1, further comprising determining median dimensions corresponding to a plurality of images, wherein the plurality of images precede the second image in the sequence of images and wherein second dimensions of the modified second bounding box correspond to the median dimensions.

14. An apparatus comprising:

a memory configured to store instructions; and  
a processor configured to execute the instructions to:

determine a similarity metric for first pixels within a first bounding box of a first image and search pixels within each of multiple search bounding boxes, wherein the first bounding box corresponds to a region of interest including a tracked object, wherein search coordinates of each of the search bounding boxes correspond to second coordinates of a second bounding box shifted in one or more directions, wherein the first image precedes a second image in a sequence of images, and wherein the second bounding box corresponds to the region of interest including the tracked object in the second image; and

determine a modified second bounding box based on the similarity metric.

15. The apparatus of claim 14, wherein the processor is further configured to:

select a particular search bounding box of the search bounding boxes based on a particular similarity metric for the first pixels and particular search pixels within the particular search bounding box,  
wherein modified second coordinates of the modified second bounding box correspond to particular search coordinates of the particular search bounding box.

16. The apparatus of claim 15, wherein the processor is further configured to calculate the particular similarity metric based at least in part on a sum of absolute differences (SAD) of first pixel characteristics of the first pixels and second pixel characteristics of the particular search pixels.

17. The apparatus of claim 15, wherein the processor is further configured to calculate the particular similarity metric based at least in part on first pixel intensities of the first pixels and second pixel intensities of the particular search pixels.

- 27 -

18. The apparatus of claim 14, wherein the sequence of images corresponds to a video stream.

19. The apparatus of claim 14,  
wherein the processor is further configured to obtain first data for the first image  
and to obtain object tracking data for the second image,  
wherein the first data defines the first bounding box, and  
wherein the object tracking data defines the second bounding box.

20. The apparatus of claim 19, wherein the first data includes first coordinates and first dimensions of the first bounding box, and wherein the object tracking data includes the second coordinates and second dimensions of the second bounding box.

21. The apparatus of claim 14, wherein search dimensions of each of the search bounding boxes correspond to first dimensions of the first bounding box.

22. The apparatus of claim 14, wherein the processor is further configured to determine median dimensions corresponding to a plurality of images, wherein the plurality of images precedes the second image in the sequence of images, and wherein second dimensions of the modified second bounding box correspond to the median dimensions.

23. A computer-readable storage device storing instructions that, when executed by a processor, cause the processor to perform operations comprising:

determining a similarity metric for first pixels within a first bounding box of a first image and search pixels within each of multiple search bounding boxes, wherein the first bounding box corresponds to a region of interest including a tracked object, wherein search coordinates of each of the search bounding boxes correspond to second coordinates of a second bounding box shifted in one or more directions, wherein the first image precedes a second image in a sequence of images, and wherein the second bounding box corresponds to the region of interest including the tracked object in the second image; and

determining a modified second bounding box based on the similarity metric.

24. The computer-readable storage device of claim 23, wherein the operations further comprise determining median dimensions corresponding to a plurality of images, wherein the plurality of images precede the second image in the sequence of images, and wherein second dimensions of the modified second bounding box correspond to the median dimensions.

25. The computer-readable storage device of claim 23, wherein the operations further comprise:

receiving first data for the first image, wherein the first data defines the first bounding box; and

receiving object tracking data for the second image, wherein the object tracking data defines the second bounding box.

26. The computer-readable storage device of claim 25, wherein the first data includes first coordinates and first dimensions of the first bounding box, and wherein the object tracking data includes the second coordinates and second dimensions of the second bounding box.

- 29 -

27. The computer-readable storage device of claim 23, wherein search dimensions of each of the search bounding boxes correspond to first dimensions of the first bounding box.

28. The computer-readable storage device of claim 23, wherein the operations further comprise:

selecting a particular search bounding box of the search bounding boxes based on a particular similarity metric for the first pixels and particular search pixels within the particular search bounding box,

wherein the similarity metric is determined based at least in part on first pixel intensities of the first pixels and second pixel intensities of the particular search pixels, and

wherein modified second coordinates of the modified second bounding box correspond to particular search coordinates of the particular search bounding box.

- 30 -

29. An apparatus including:

means for receiving first data defining a first bounding box for a first image of a sequence of images, wherein the first bounding box corresponds to a region of interest including a tracked object;

means for generating object tracking data for a second image of the sequence of images, the object tracking data defining a second bounding box, wherein the second bounding box corresponds to the region of interest including the tracked object in the second image; and

means for determining:

a similarity metric for first pixels within the first bounding box and search pixels within each of multiple search bounding boxes, wherein search coordinates of each of the search bounding boxes correspond to second coordinates of the second bounding box shifted in one or more directions; and

a modified second bounding box based on the similarity metric.

30. The apparatus of claim 29, wherein the means for receiving, the means for generating, and the means for determining are integrated into at least one of a mobile phone, a set top box, a music player, a video player, an entertainment unit, a navigation device, a communications device, a personal digital assistant (PDA), a fixed location data unit, or a computer.

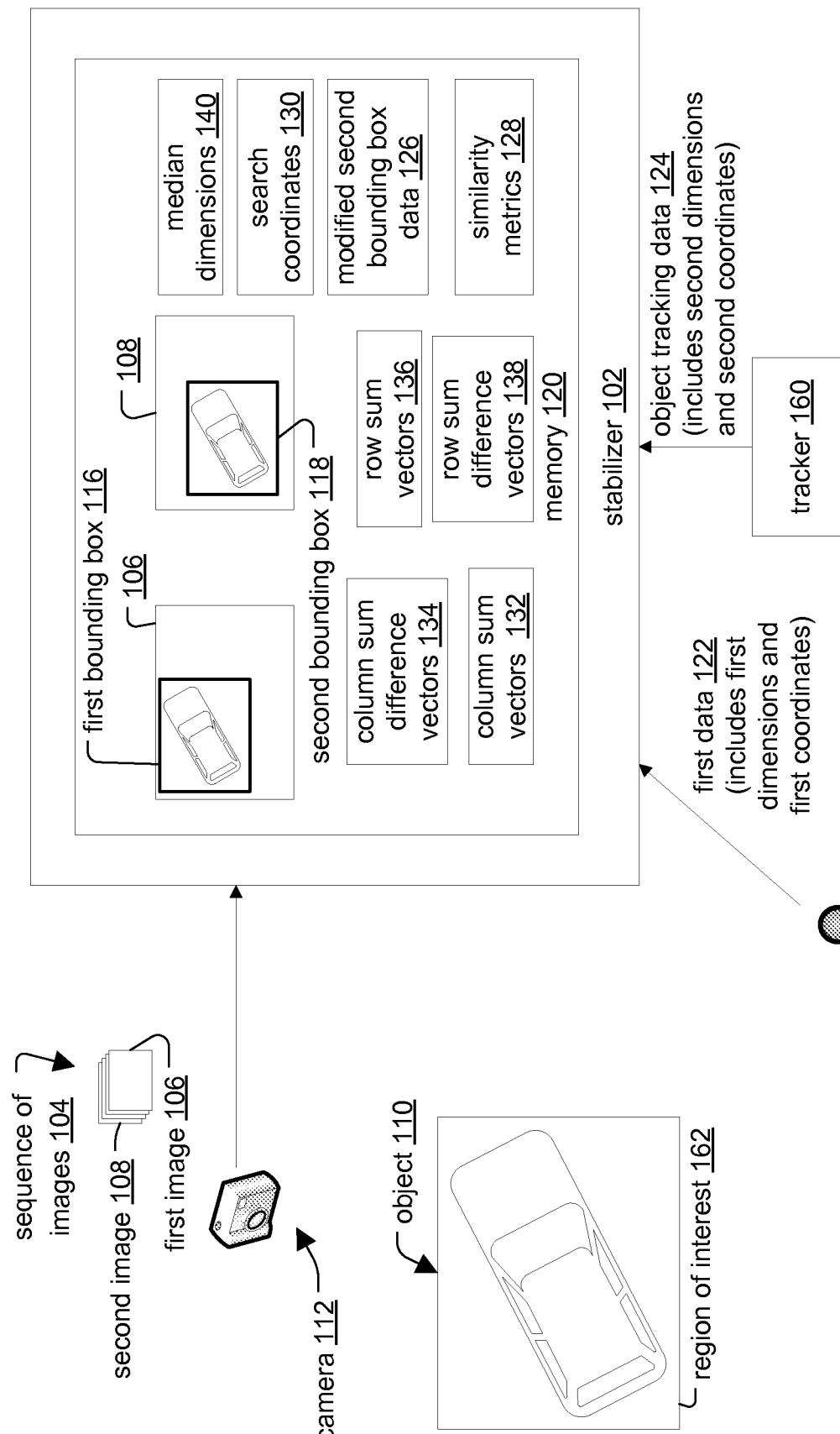


FIG. 1

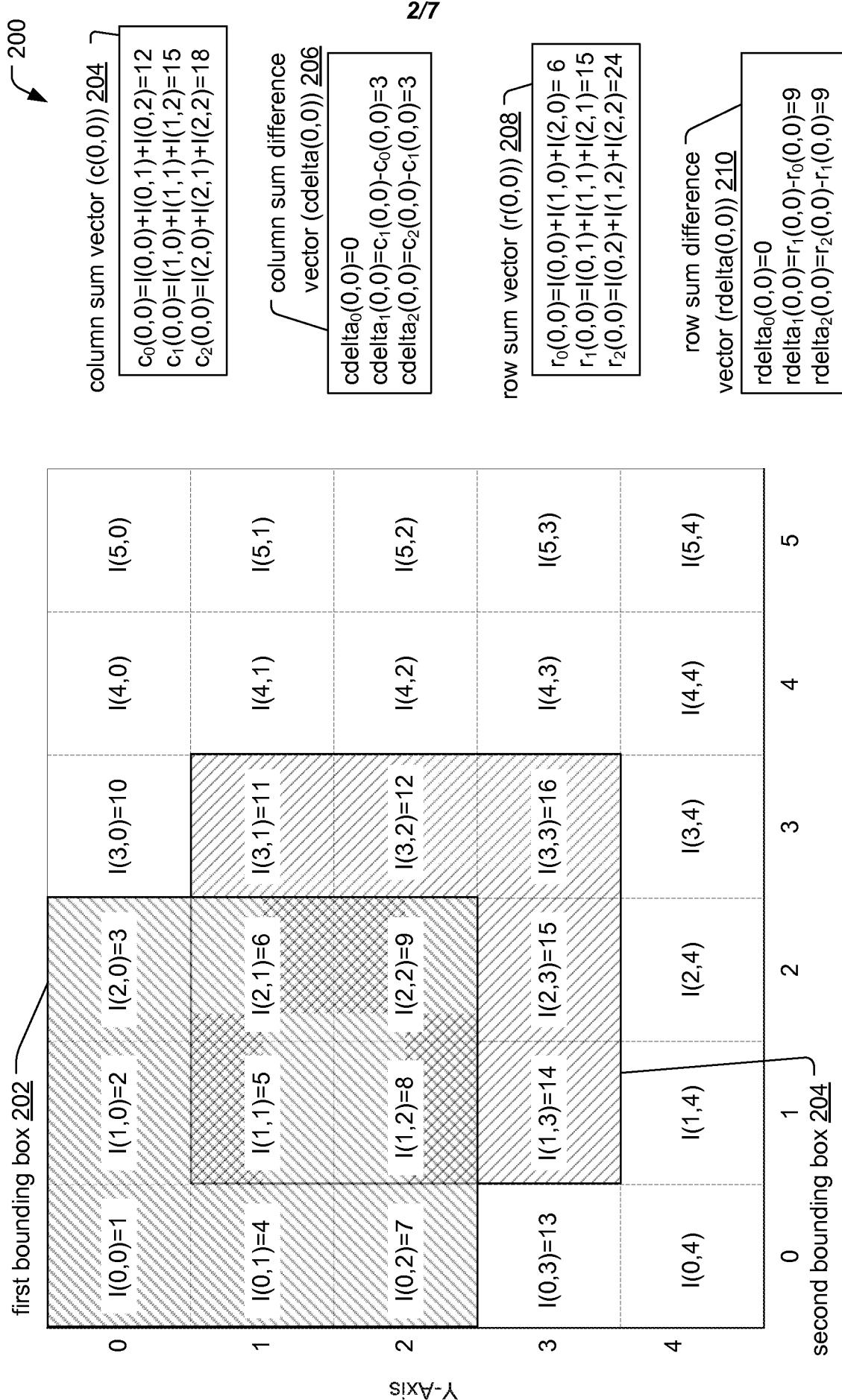


FIG. 2

3/7

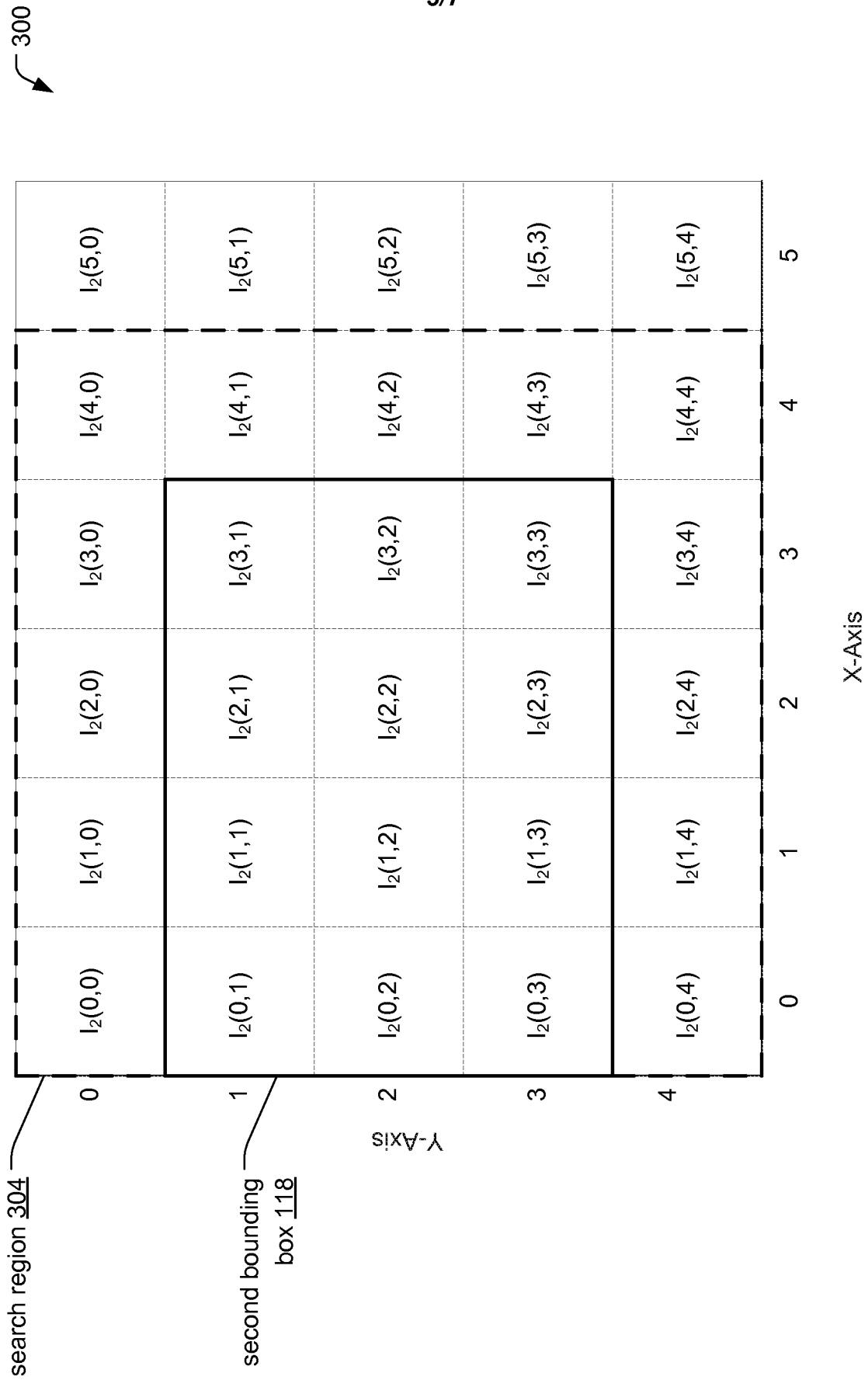
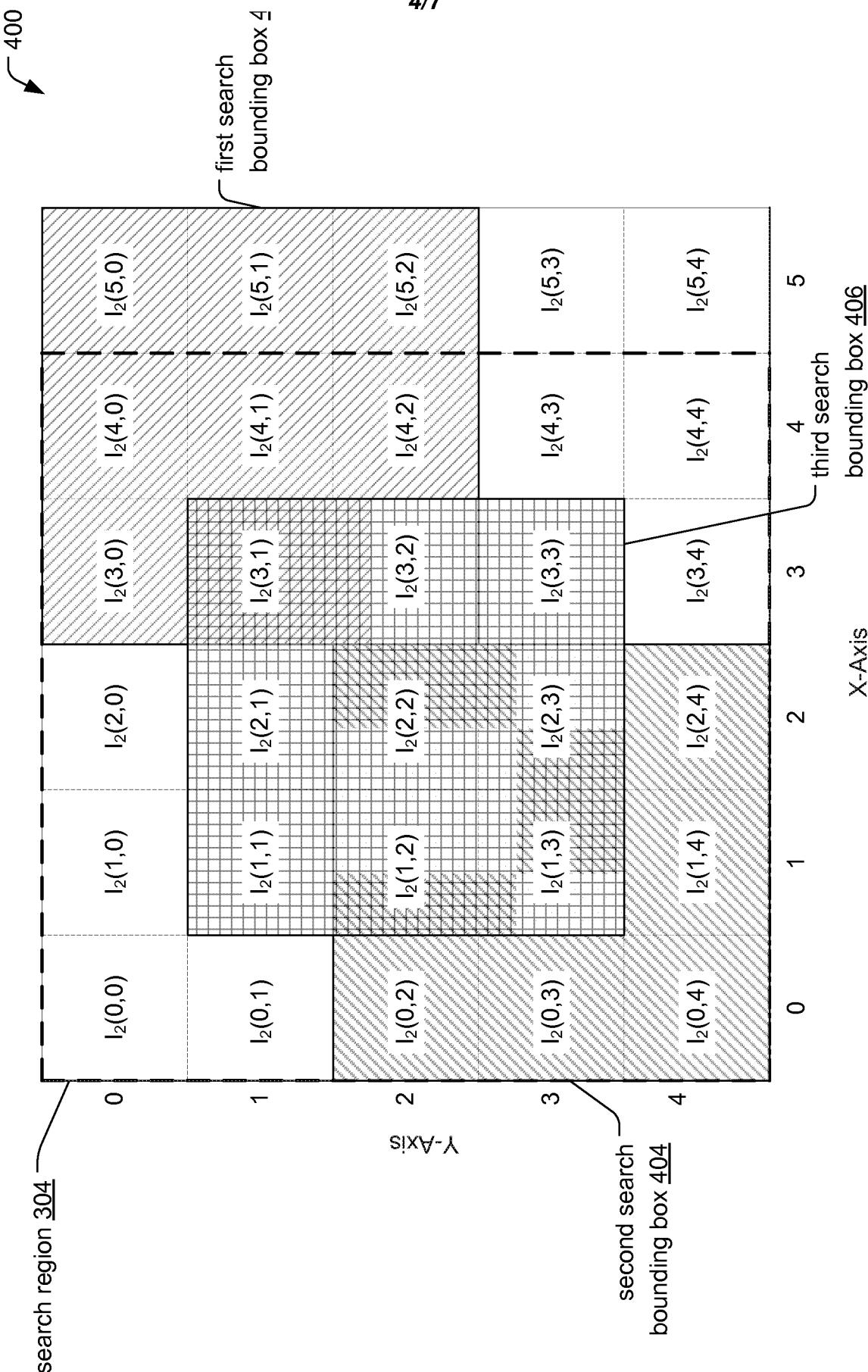


FIG. 3

4/7

**FIG. 4**

Receive first data defining a first bounding box for a first image of a sequence of images, where the first bounding box corresponds to a region of interest including a tracked object

Receive object tracking data for a second image of the sequence of images, the object tracking data defining a second bounding box, where the second bounding box corresponds to the region of interest including the tracked object in the second image

Determine a similarity metric for first pixels within the first bounding box and search pixels within each of multiple search bounding boxes, where search coordinates of each of the search bounding boxes correspond to second coordinates of the second bounding box shifted in one or more directions

Determine a modified second bounding box based on the similarity metric

Determine median dimensions corresponding to a plurality of images, where the plurality of images precede the second image in the sequence of images and where second dimensions of the modified second bounding box correspond to the median dimensions

6/7

600

506

602

Determine a first column sum vector of the first pixels, where a particular element of the first column sum vector corresponds to a sum of pixel values corresponding to a particular column of the first bounding box

604

Determine a first column sum difference vector of the first pixels, where a particular element of the first column sum difference vector corresponds to a difference between a first element of the first column sum vector and a second element of the first column sum vector, and where the first element and the second element correspond to adjacent columns of the first bounding box

606

Determine a first row sum vector of the first pixels, where a particular element of the first row sum vector corresponds to a sum of pixel values corresponding to a particular row of the first bounding box

608

Determine a first row sum difference vector of the first pixels, where a particular element of the first row sum difference vector corresponds to a difference between a first element of the first row sum vector and a second element of the row sum vector, and where the first element and the second element correspond to adjacent rows of the first bounding box

610

Calculate a particular similarity metric for the first pixels and particular search pixels within a particular search bounding box by adding: a first sum of absolute differences (SAD) of the first column sum vector of the first pixels and a second column sum vector of the particular search pixels, a second SAD of the first column sum difference vector of the first pixels and a second column sum difference vector of the particular search pixels, a third SAD of the first row sum vector of the first pixels and a second row sum vector of the particular search pixels, and a fourth SAD of the first row sum difference vector of the first pixels and a second row sum difference vector of the particular search pixels

**FIG. 6**

7/7

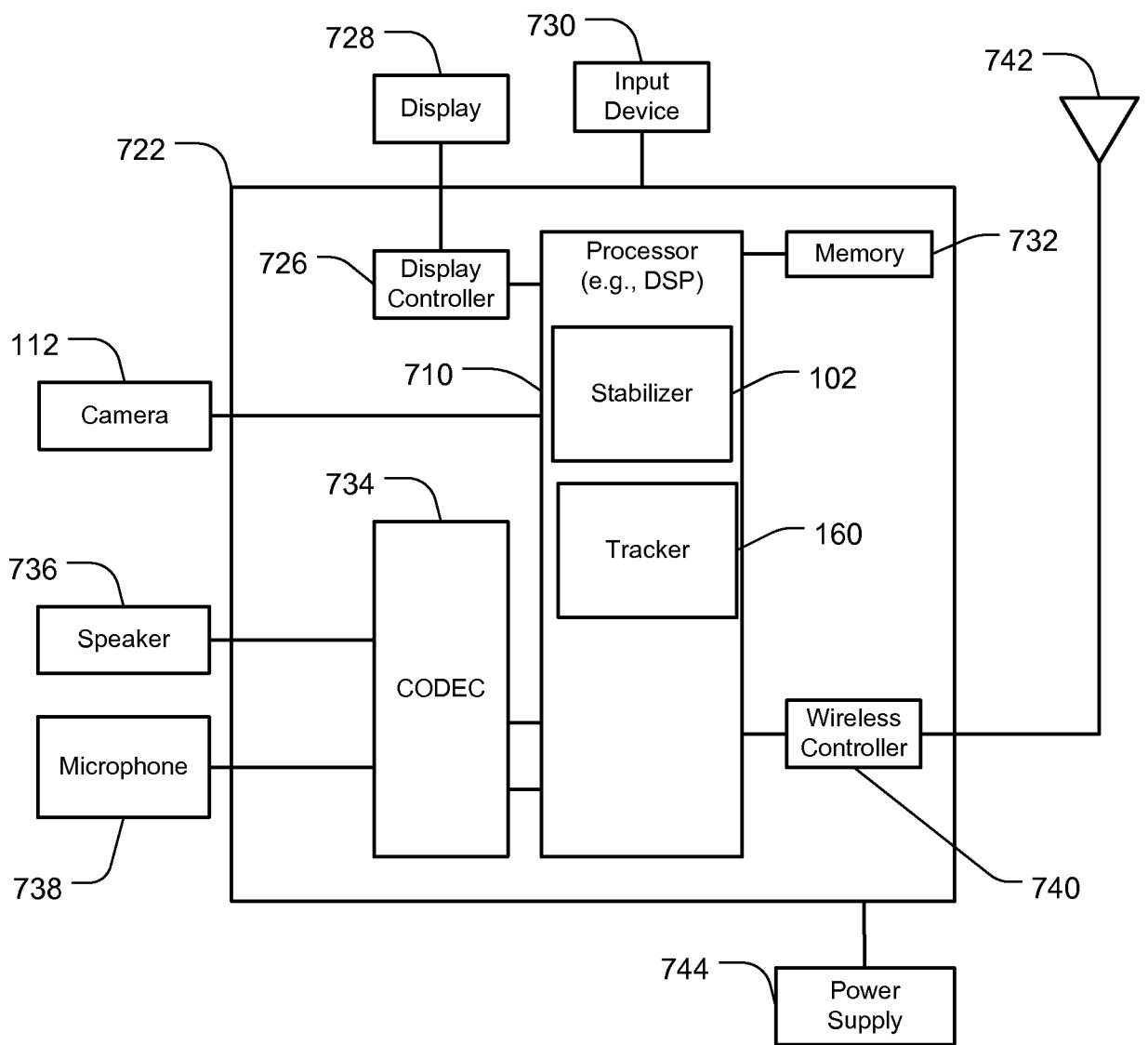


FIG. 7

# INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2014/070078

**A. CLASSIFICATION OF SUBJECT MATTER**  
INV. H04N5/232 G06T7/20  
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)  
H04N G06T

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 2010/073502 A1 (AN MYUNG-SEOK [KR] ET AL) 25 March 2010 (2010-03-25) equations (1) and (2), (12) and (15); paragraphs [0014], [0043], [0077]; claims 7, 8, 12; figures 2-5, 9 -----	1-30
X	US 2011/188758 A1 (KINOSHITA MASAYA [JP] ET AL) 4 August 2011 (2011-08-04) paragraphs [0020], [0165] - [0170], [0229] - [0230], [0264] - [0268]; figures 9, 22, 23 -----	1-30
X	WO 2007/077283 A1 (NOKIA CORP [FI]; CHEBIL FEHMI [US]; MILED MOHAMED KHAMES BEN HADJ [US]) 12 July 2007 (2007-07-12) figures 5-9 -----	1-30



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

Date of mailing of the international search report

27 February 2015

09/03/2015

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

Trimeche, Mejdi

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No

PCT/US2014/070078

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
US 2010073502	A1 25-03-2010	CN 101686322 A		31-03-2010
		KR 20100034631 A		01-04-2010
		US 2010073502 A1		25-03-2010
<hr/>				
US 2011188758	A1 04-08-2011	CN 102196170 A		21-09-2011
		EP 2372655 A1		05-10-2011
		JP 2011160379 A		18-08-2011
		KR 20110090787 A		10-08-2011
		US 2011188758 A1		04-08-2011
<hr/>				
WO 2007077283	A1 12-07-2007	EP 1966648 A1		10-09-2008
		JP 2009522591 A		11-06-2009
		US 2010045800 A1		25-02-2010
		WO 2007077283 A1		12-07-2007
<hr/>				