METHOD AND APPARATUS FOR APPLYING A MASK ON AN OBJECT IN AN IMAGE

FIELD OF INVENTION

[0001] The present invention relates broadly, but not exclusively, to a method and an apparatus for applying a mask on an object in an image.

BACKGROUND

[0002] Video surveillance system is widely used around the world by law and security enforcement for monitoring and protection against crimes. Public privacy concerns on video surveillance have been growing with the popularity of video surveillance systems. There is a risk of revealing identities of individuals in public areas by, for example, security officers while monitoring surveillance cameras.

[0003] To address the above privacy concerns, privacy masking is applied to the surveillance camera footages to protect sensitive and private information, such as a person’s identity, before displaying on users' devices. Object detection is often used to detect such sensitive information before selecting the detected object regions for privacy masking. This is due to fast detection speed and ability to detect both a front-facing person and a back-facing person. To further optimize the privacy masking process, some solutions may choose to skip privacy masking for smaller objects which may not contain enough data to reveal the sensitive information. However, the methods to detect small objects might accidentally skip privacy masking for objects which might still contain sensitive information.

[0004] Herein disclosed are embodiments of a method and apparatus for applying a mask on an object in an image that addresses one or more of the above problems.

[0005] Furthermore, other desirable features and characteristics will become apparent from the subsequent detailed description and the appended claims, taken in conjunction with the accompanying drawings and this background of the disclosure.

SUMMARY

[0006] In a first aspect, the present disclosure provides a method for applying a mask on an object in an image, comprising: determining, by a processor, whether to apply a mask on an object in an image based on a comparison with another object in the image; and applying, by the processor, the mask on the object in the image based on the determination.

[0007] In a second aspect, the present disclosure provides an apparatus for applying a mask on an object in an image, comprising: at least one processor; and at least one memory including computer program code; the at least one memory and the computer program code configured to, with at least one processor, cause the apparatus at least to: determine whether to apply a mask on an object in an image based on a comparison with another object in the image; and apply the mask on the object in the image based on the determination.

[0008] In a third aspect, the present disclosure provides a system for applying a mask on an object in an image, comprising the apparatus of the second aspect and one or more image capturing apparatuses configured to capture one or more images or video frames, wherein the one or more images or video frames comprises an image of a person, the person being the object.

[0009] Additional benefits and advantages of the disclosed embodiments will become apparent from the specification and drawings. The benefits and/or advantages may be individually obtained by the various embodiments and features of the specification and drawings, which need not all be provided in order to obtain one or more of such benefits and/or advantages.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] The accompanying Figures, where like reference numerals refer to identical or functionally similar elements throughout the separate views and which together with the detailed description below are incorporated in and form part of the specification, serve to illustrate various embodiments and to explain various principles and advantages in accordance with a present embodiment, by way of non-limiting example only.

[0011] Embodiments of the invention will be better understood and readily apparent to one of ordinary skill in the art from the following written description, by way of example only, and in conjunction with the drawings, in which:

[0012] Figure 1 shows an illustration of an object detection and privacy masking process on a sample video according to an example.

[0013] Figure 2 shows an illustration of object detection and privacy masking on an image according to an example.

[0014] Figure 3 shows an exemplary illustration of setting body proportion for object detection and privacy masking according to various embodiments of the present disclosure.

[0015] Figure 4 shows an exemplary illustration for setting a two-dimension (2D) plane for object detection and privacy masking according to various embodiment of the present disclosure.

[0016] Figure 5 shows a flow chart illustrating a method for applying a mask on an object in an image according to various embodiments of the present disclosure.

[0017] Figure 6 shows an exemplary illustration for object detection and privacy masking on an image according to an embodiment of the present disclosure.

[0018] Figure 7 shows an exemplary illustration for object detection and privacy masking on an image according to another embodiment of the present disclosure.

[0019] Figure 8 shows an exemplary illustration for object detection and privacy masking on an image according to another embodiment of the present disclosure.

[0020] Figure 9 shows an exemplary illustration for object detection and privacy masking on an image according to another embodiment of the present disclosure.

[0021] Figure 10 shows a flow chart for object detection and privacy masking according to various embodiments of the present disclosure.

[0022] Figure 11 shows an exemplary illustration of an image in which objects are detected and privacy masks are applied according to an embodiment of the present disclosure.

[0023] Figure 12 shows a schematic diagram of an exemplary computing device suitable for use to execute the method in Figure 5 and 10.

DETAILED DESCRIPTION

[0024] Embodiments of the present invention will be described, by way of example only, with reference to the drawings. Like reference numerals and characters in the drawings refer to like elements or equivalents.

[0025] Some portions of the description which follows are explicitly or implicitly presented in terms of algorithms and functional or symbolic representations of

operations on data within a computer memory. These algorithmic descriptions and functional or symbolic representations are the means used by those skilled in the data processing arts to convey most effectively the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities, such as electrical, magnetic or optical signals capable of being stored, transferred, combined, compared, and otherwise manipulated.

[0026] Unless specifically stated otherwise, and as apparent from the following, it will be appreciated that throughout the present specification, discussions utilizing terms such as “detecting”, “estimating”, “comparing”, “receiving”, “calculating”, “determining”, “updating”, “generating”, “initializing”, “outputting”, “receiving”, “retrieving”, “identifying”, “dispersing”, “authenticating” or the like, refer to the action and processes of a computer system, or similar electronic device, that manipulates and transforms data represented as physical quantities within the computer system into other data similarly represented as physical quantities within the computer system or other information storage, transmission or display devices.

[0027] The present specification also discloses apparatus for performing the operations of the methods. Such apparatus may be specially constructed for the required purposes, or may comprise a computer or other device selectively activated or reconfigured by a computer program stored in the computer. The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various machines may be used with programs in accordance with the teachings herein. Alternatively, the construction of more specialized apparatus to perform the required method steps may be appropriate. The structure of a computer will appear from the description below.

[0028] In addition, the present specification also implicitly discloses a computer program, in that it would be apparent to the person skilled in the art that the individual steps of the method described herein may be put into effect by computer code. The computer program is not intended to be limited to any

particular programming language and implementation thereof. It will be appreciated that a variety of programming languages and coding thereof may be used to implement the teachings of the disclosure contained herein. Moreover, the computer program is not intended to be limited to any particular control flow. There are many other variants of the computer program, which can use different control flows without departing from the spirit or scope of the invention.

[0029] Furthermore, one or more of the steps of the computer program may be performed in parallel rather than sequentially. Such a computer program may be stored on any computer readable medium. The computer readable medium may include storage devices such as magnetic or optical disks, memory chips, or other storage devices suitable for interfacing with a computer. The computer readable medium may also include a hard-wired medium such as exemplified in the Internet system, or wireless medium such as exemplified in the GSM mobile telephone system. The computer program when loaded and executed on such a computer effectively results in an apparatus that implements the steps of the preferred method.

Exemplary embodiments

[0030] Various embodiments of the present disclosure relate to a method and an apparatus for applying a mask on an object in an image.

[0031] A privacy mask (also referred to herein as a mask) is one that is applied on a specific area in an image, a video, or other similar media to blur or pixelate the specific area, so that the image, video or similar media can be viewed without compromising privacy of an entity associated with the specific area. Using object detection to detect people or objects such as car plates in, for example, dashcam or security camera video footage for privacy masking can be resource intensive and time consuming. Each individual frame of the video footage would need object detection engine to detect the object-of interest to be privacy masked. A privacy masking algorithm will then be applied to each detected object-of-interest before merging them back to the final frame to be re-encoded back to a privacy-masked dashcam video footage for sharing. It will be appreciated that an object

may refer to a person, a car plate, an address plate, or other similar objects that may require a privacy mask to be applied in order to protect against privacy invasion.

[0032] Referring to illustration 100 of Fig. 1, performing a privacy-masking of one detected object in an input image 102 typically takes about 50 ms (miliseconds). The input image 102 may be a video frame from a FHD (full high definition) video (recorded via, for example, a security camera) that typically runs at 20 fps (frames per second). Thus, the input image 102 may be one of 20 video frames in one second of the FHD video. Accordingly, 1 second of the FHD video with 20 fps requires (total frames) * (processing time of privacy masking for 1 frame) = (1 second * 20 fps) * (50 ms) = 20 * 50ms = 1000 ms in order to perform privacy masking per person per second of the FHD video. One minute (e.g., 60 seconds) of the FHD video running at 20 fps and with 5 persons in the video would require 60 seconds * (5 person * 1000 ms) = 60 * (5000 ms) = 60 * 5 seconds = 300 seconds = 5 minutes to complete object detection for 5 persons (e.g., as shown in image 104) and privacy masking of all 5 persons (e.g., as shown in image 106), which is not feasible for a video running in real time.

[0033] One of the ways to optimize the privacy-masking process is to skip any unnecessary privacy masking based on a person's detected height (e.g., if a detected height of an object in an image is shorter than a certain threshold, it may be determined that the object is too small in the image for a privacy mask to be necessary), for example utilizing object and pose detection. As pose detection is a resource intensive process, 2 types of a pre-trained machine learning model may be used. A first type is based on single pose detection in which a high accuracy of detection based on more key-points (e.g., a point on an image of a person corresponding to a body part position of the person) may be achieved for single person pose detection. A second type is based on multiple pose detection that is limited to detection of 4 to 6 persons with an averaged number of key-points to strike a balance between accuracy and speed of detection. A pre-trained pose estimation machine learning (ML) model provides a plurality of key-points between eyes and toes of a person, which may be used to identify a person and estimate a height of the person. For example referring to image 200 of Fig. 2, it

may be configured such that privacy masking is skipped for a detected person whose height in the image 200 is lower than 25 px (pixels). As 3 persons (e.g., persons 202, 204 and 206) are detected in the image 200 whose height is below 25 px (pixels), privacy masking is skipped for these 3 persons. Although pose estimation allows better estimation of a person's height for, for example, person 206 who is bending down and whose actual height may actually not be below 25 px, it requires higher computing power and has the constraint of being limited to detection of only 4-6 poses within an image.

[0034] In the present disclosure, a method is proposed to improve object detection accuracy for skipping privacy masking for selected persons detected in an image based on a pre-configured minimum detection height (e.g., a threshold height), and further based on detection of a tallest person within a 2D (2-dimension) plane in the image (e.g., a person with a height that is highest among one or more persons within the 2D plane in the image). For example referring to illustration 302 of Fig. 3, an average body proportion may be determined as 1:7 for a head to body ratio and 1:4 for a head to arm ratio. Thus, based on a stand height of 7 units and a width of 5 units (e.g., based on head to arm ratio of 1 unit : 4 units, a width comprising a head and an arm is set as 1 + 4 = 5 units) as shown in illustration 304, a minimum stand height to width ratio may be determined as 7/5 = 1.4.

[0035] Further assumptions are as follows. A minimum height of a detected person in an image for applying a privacy mask may be set as 25 px (25 pixels). An upper 2D plane distance may be set as 3 px. A lower 2D plane distance may be set as 2 px. A search in a Y-axis of an image may be configured to start from a height of a detected person minus the upper 2D plane distance, and configured to end at the height of the detected person plus the lower 2D plane distance. For example, referring to illustration 400 of Fig. 4, a person 402 may be detected from pixel 1200 of an X-axis in an image and from pixel 510 of a Y-axis in the image. The person 402 may be below the minimum height of 25 px (e.g., having a detected height of 24 px) and also below a minimum height to width ratio (e.g., a threshold ratio) of 1.4 for applying a privacy mask. This triggers a search for a person with a highest height within a 2D plane distance of the person 402 (e.g.,

a 2D plane distance 406 of between 512 to 507 pixels in the Y-axis of the image). A person 404 is detected from pixel 1550 of the X-axis and pixel 511 of the Y-axis in the image as the highest person within the 2D plane distance with the person 402, having a height of 35 px which is above the minimum height of 25 px. Thus, a privacy mask flag of the person 402 is set to “true” e.g., a privacy mask is applied on the person 402. In an implementation, one or more algorithms may be utilized to determine a privacy mask flag of every detected person in an image. The privacy mask flag may carry an initial value of “?”, and then set to either “TRUE” or “FALSE” based on the detected height and ratio, as well as comparison with other detected person(s) that are located within a 2D plane as explained above. Advantageously, it is possible to skip privacy masking on a detected person without resource intensive process such as pose detection and without the limitations of keypoints during pose detection. It will be appreciated that the threshold height and ratio may be adjusted based on application and factors such as image or video resolution, angle of view of the camera from which the video or image is taken, the scene being captured in the image or video, and other similar factors.

[0036] Figure 5 shows a flow chart 500 illustrating a method for applying a mask on an object in an image according to various embodiments of the present disclosure. In step 502, it is determined, by a processor, whether to apply a mask on an object in an image based on a comparison with another object in the image. In step 504, the mask is applied by the processor on the object in the image based on the determination.

[0037] In an implementation, determining whether to apply the mask may comprise: comparing a height of the object in the image with a threshold height; if the height of the object is lower than the threshold height, further comparing a ratio of the height to a width of the object in the image with a threshold ratio; and determining whether to apply the mask based on the further comparison. The method may further comprise: if the ratio is lower than the threshold ratio, identifying the another object from one or more other objects in the image, the another object having a height that is highest among the object and the one or more other objects, the one or more other objects being located within a two-

dimensional plane distance from the object; and determining whether a mask is to be applied on the another object. The method may further comprise applying the mask on the object if it is determined that a mask is to be applied on the another object.

[0038] In a implementation, the method may further comprise applying the mask on the object if the height of the object is higher than or equal to the threshold height. The method may further comprise determining not to apply the mask on the object if the ratio is higher than the threshold ratio.

[0039] In an implementation, the method may further comprise applying the mask on the object in a plurality of frames of a video, the image being a frame of the plurality of frames of the video.

[0040] In an implementation, an apparatus may comprise at least one processor and at least one memory including computer program code, the at least one memory and the computer program code configured to, with at least one processor, cause the apparatus at least to perform the method 500. In an implementation, a system for applying a mask on an object in an image may comprise the apparatus and one or more image and video capturing apparatuses configured to capture one or more images or video frames, wherein the one or more images or video frames comprises an image of a person, the person being the object.

[0041] Figure 6 shows an exemplary illustration for object detection and privacy masking on an image 600 according to an embodiment of the present disclosure. Two persons 602 and 604 are detected in the image 600. A table 608 provides, in a first row and second row, the various detected dimensions of persons 602 and 604 respectively. For example referring to table 608, person 602 is detected from pixel 1200 of the X-axis and pixel 485 of the Y-axis to pixel 1223 of the X-axis and pixel 509 of the Y-axis in the image 600, has a height of 24 px, a width of 23 px, a height to width ratio of 1.0434, a corresponding 2D plane 606 within pixels 506 to 511 of the Y-axis of the image 600, and a privacy mask flag of “?”. Further, person 604 is detected from pixel 1550 of the

X-axis and pixel 472 of the Y-axis to pixel 1570 of the X-axis and pixel 507 of the Y-axis in the image 600, has a height of 35 px, a width of 20 px, a height to width ratio of 1.75, a corresponding 2D plane within pixels 504 to 509 of the Y-axis of the image 600, and a privacy mask flag of “?”. It will be appreciated that there may be one or more other persons that are detected in image 600 but not shown for simplicity.

[0042] The detected persons 602 and 604 may be processed one after another in a sequential sequence to determine the privacy mask flag of each person. For example, it is checked if a privacy mask flag of person 602 is ”?”. If it is not (e.g., the privacy mask flag is “TRUE” or “FALSE”, meaning that the detected person associated with the privacy mask flag has already been processed and determined to apply a privacy mask or to skip the masking), person 602 will not be processed (since the privacy mask flag which is not ”?” indicates that the person has already been processed). Otherwise, person 602’s height is compared against a threshold height (e.g., 25 px). As shown in table 608 in Fig. 6, the height of person 602 is 24 px which is smaller than the threshold height. Thus, the person 602’s height to width ratio is also compared against a threshold ratio (e.g., 1.4). Since person 602’s ratio of 1.0434 is also smaller than the threshold ratio, a search for one or more persons within the same 2D plane as person 602 (e.g., a corresponding 2D plane 606 within pixels 506 to 511 of the Y-axis of the image 600) is made. The one or more persons that are found in the search will be further compared to determine one person with the highest height among the one or more persons, and the determined person with the highest height will be used as a basis for further comparison with the person 602 (e.g., to determine if the privacy mask flag of person 602 should be updated to “TRUE” or “FALSE”). Since person 604 is the only person within the same 2D plane 606 as person 602 in image 600, the search returns person 604 as a search result. It will be appreciated that, if there are one or more other persons besides person 604 within the same 2D plane 606, the search will return person 604 as the search result if it is determined that person 604 has the highest height among the one or more other persons.

[0043] Since the privacy mask flag of person 604 is “?” (e.g., it is not yet determined if a privacy mask should be applied to person 604), the person 604’s height is checked. Based on table 608, since person 604’s height is 35 px which is higher than the threshold height of 25 px, the privacy mask flag of person 604 is updated to “TRUE” (e.g., a privacy mask is to be applied on person 604). Further, the privacy mask flag of person 602 is also updated to be the same as that of person 604 (person 604 being the only detected person within the same 2D plane 606 as person 602) to “TRUE” (e.g., a privacy mask is to be applied on person 602).

[0044] Figure 7 shows an exemplary illustration for object detection and privacy masking on an image 700 according to an embodiment of the present disclosure. Two persons 702 and 704 are detected in the image 700. A table 708 provides, in a first row and second row, the various detected dimensions of persons 702 and 704 respectively. For example referring to table 708, person 702 is detected from pixel 1550 of the X-axis and pixel 472 of the Y-axis to pixel 1570 of the X-axis and pixel 507 of the Y-axis in the image 700, has a height of 35 px, a width of 20 px, a height to width ratio of 1.75, a corresponding 2D plane within pixels 504 to 509 of the Y-axis of the image 700, and a privacy mask flag of “TRUE” (e.g., person 702 has already been processed and determined that a privacy mask is to be applied). Further, person 604 is detected from pixel 1200 of the X-axis and pixel 485 of the Y-axis to pixel 1223 of the X-axis and pixel 509 of the Y-axis in the image 700, has a height of 24 px, a width of 23 px, a height to width ratio of 1.0434, a corresponding 2D plane within pixels 506 to 511 of the Y-axis of the image 700, and a privacy mask flag of “?”. It will be appreciated that there may be one or more other persons (not shown in image 700 for simplicity) that are detected in image 700 but are not within the 2D plane 706 with the person 702.

[0045] The detected persons 702 and 704 may be processed one after another in a sequential sequence to determine the privacy mask flag of each person. For example, it is checked if a privacy mask flag of person 702 is ”?”. If it is not (e.g., the privacy mask flag is “TRUE” or “FALSE”, meaning that the detected person associated with the privacy mask flag has already been

processed and determined to apply a privacy mask or to skip the masking), person 702 will not be processed (since the privacy mask flag which is not "?" indicates that the person has already been processed). Otherwise, person 702's height is compared against a threshold height (e.g., 25 px). Since person 702's privacy mask flag already indicates "TRUE", processing for person 702 is skipped, and person 704 is processed next. As shown in table 708, the height of person 704 is 24 px which is smaller than the threshold height. Thus, the person 704's height to width ratio is also compared against a threshold ratio (e.g., 1.4). Since person 704's ratio of 1.0434 is also smaller than the threshold ratio, a search for one or more persons within the same 2D plane as person 704 (e.g., a corresponding 2D plane within pixels 506 to 511 of the Y-axis of the image 700) is made. The one or more persons that are found in the search will be further compared to determine one person with the highest height among the one or more persons, and the determined person with the highest height will be used as a basis for further comparison with the person 704 (e.g., to determine if the privacy mask flag of person 704 should be updated to "TRUE" or "FALSE"). Since person 702 is the only person within the same 2D plane 706 as person 704 in image 700, the search returns person 702 as a search result.

[0046] Since the privacy mask flag of person 702 is "TRUE" (e.g., it is already determined that a privacy mask should be applied to person 702), the privacy mask flag of person 704 is also updated to be the same as that of person 702 (person 702 being the only detected person within the same 2D plane 706 as person 704) to "TRUE" (e.g., a privacy mask is to be applied on person 704).

[0047] Figure 8 shows an exemplary illustration for object detection and privacy masking on an image 800 according to an embodiment of the present disclosure. Two persons 802 and 804 are detected in the image 800. A table 808 provides, in a first row and second row, the various detected dimensions of persons 802 and 804 respectively. For example referring to table 808, person 802 is detected from pixel 300 of the X-axis and pixel 250 of the Y-axis to pixel 324 of the X-axis and pixel 274 of the Y-axis in the image 800, has a height of 24 px, a width of 24 px, a height to width ratio of 1.0, a corresponding 2D plane 806 within pixels 271 to 276 of the Y-axis of the image 800, and a

privacy mask flag of “?”. Further, person 804 is detected from pixel 450 of the X-axis and pixel 248 of the Y-axis to pixel 473 of the X-axis and pixel 271 of the Y-axis in the image 800, has a height of 23 px, a width of 23 px, a height to width ratio of 1.0, a corresponding 2D plane within pixels 268 to 273 of the Y-axis of the image 800, and a privacy mask flag of “?”.It will be appreciated that there may be one or more other persons (not shown in image 800 for simplicity) that are detected in image 800 but are not within the 2D plane 806 with the person 802.

[0048] The detected persons 802 and 804 may be processed one after another in a sequential sequence to determine the privacy mask flag of each person. For example, it is checked if a privacy mask flag of person 802 is ”?”. If it is not (e.g., the privacy mask flag is “TRUE” or “FALSE”, meaning that the detected person associated with the privacy mask flag has already been processed and determined to apply a privacy mask or to skip the masking), person 802 will not be processed (since the privacy mask flag which is not ”?” indicates that the person has already been processed). Otherwise, person 802’s height is compared against a threshold height (e.g., 25 px). As shown in table 808, the height of person 802 is 24 px which is smaller than the threshold height. Thus, the person 802’s height to width ratio is also compared against a threshold ratio (e.g., 1.4). Since person 802’s ratio of 1.0 is also smaller than the threshold ratio, a search for one or more persons within the same 2D plane as person 802 (e.g., a corresponding 2D plane within pixels 271 to 276 of the Y-axis of the image 800) is made. The one or more persons that are found in the search will be further compared to determine one person with the highest height among the one or more persons, and the determined person with the highest height will be used as a basis for further comparison with the person 802 (e.g., to determine if the privacy mask flag of person 802 should be updated to “TRUE” or “FALSE”). Since person 804 is the only person within the same 2D plane 806 as person 802 in image 800, the search returns person 804 as a search result.

[0049] Since the privacy mask flag of person 804 is “?” (e.g., it is not yet determined if a privacy mask should be applied to person 804), the person 804’s height is checked. Based on table 808, the person 804’s height is 23 px which is smaller than the threshold height of 25 px. Thus, a search for one or more persons

within the same 2D plane as person 804 (e.g., a corresponding 2D plane within pixels 268 to 273 of the Y-axis of the image 800) is made. The one or more persons that are found in the search will be further compared to determine one person with the highest height among the one or more persons, and the determined person with the highest height will be used as a basis for further comparison with the person 804 (e.g., to determine if the privacy mask flag of person 804 should be updated to “TRUE” or “FALSE”). Since person 802 is the only person within the same 2D plane as person 804 in image 800, the search returns person 802 as a search result. Since no other person is available for comparison, the privacy mask flag of person 804 is updated to “TRUE” (e.g., a privacy mask is to be applied on person 804) as a default value. Further, the privacy mask flag of person 802 is also updated to be the same as that of person 804 (person 804 being the only detected person within the same 2D plane 806 as person 802) to “TRUE” (e.g., a privacy mask is to be applied on person 802). It will be appreciated that the default value may be set to “FALSE” (e.g., a privacy mask is not required) depending on application.

[0050] Figure 9 shows an exemplary illustration for object detection and privacy masking on an image 900 according to an embodiment of the present disclosure. A person 902 is detected in the image 900. A table 906 provides the various detected dimensions of person 902. For example referring to table 906, person 902 is detected from pixel 1200 of the X-axis and pixel 485 of the Y-axis to pixel 1223 of the X-axis and pixel 509 of the Y-axis in the image 900, has a height of 24 px, a width of 23 px, a height to width ratio of 1.0434, a corresponding 2D plane 904 within pixels 506 to 511 of the Y-axis of the image 900, and a privacy mask flag of “?”. It will be appreciated that there may be one or more other persons (not shown in image 900 for simplicity) that are detected in image 900 but are not within the 2D plane 904 with the person 902.

[0051] The detected person 902 may be processed one after another (e.g., along with one or more other persons that are in image 900 but not shown for simplicity purposes) in a sequential sequence to determine the privacy mask flag of each person. For example, it is checked if a privacy mask flag of person 902 is ”?”. If it is not (e.g., the privacy mask flag is “TRUE” or “FALSE”, meaning

that the detected person associated with the privacy mask flag has already been processed and determined to apply a privacy mask or to skip the masking), person 902 will not be processed (since the privacy mask flag which is not "?" indicates that the person has already been processed). Otherwise, person 902's height is compared against a threshold height (e.g., 25 px). As shown in table 906, the height of person 902 is 24 px which is smaller than the threshold height. Thus, the person 902's height to width ratio is also compared against a threshold ratio (e.g., 1.4). Since person 902's ratio of 1.0434 is also smaller than the threshold ratio, a search for one or more persons within the same 2D plane as person 902 (e.g., a corresponding 2D plane 904 within pixels 506 to 511 of the Y-axis of the image 900) is made. The one or more persons that are found in the search will be further compared to determine one person with the highest height among the one or more persons, and the determined person with the highest height will be used as a basis for further comparison with the person 902 (e.g., to determine if the privacy mask flag of person 902 should be updated to "TRUE" or "FALSE"). Since there is no other person within the same 2D plane 904 in image 900, the privacy mask flag of person 902 is also updated to "TRUE" (e.g., a privacy mask is to be applied on person 902) as a default value.

[0052] Figure 10 shows a flow chart 1000 for object detection and privacy masking according to various embodiments of the present disclosure. In a step 1002, a video file may be uploaded into, for example, an apparatus or system for applying a mask on an object in an image according to various embodiments of the present disclosure. In step 1004, the video file may be decoded into a plurality of video frames, each video frame being an image on which one or more objects may be detected and a mask may or may not be applied on a detected object. In step 1006, it is determined whether the plurality of video frames comprises more video frames e.g., new video frames that are different from video frames that are previously processed. If it is determined that the plurality of video frames does not comprise more video frames, the process proceeds to step 1034 in which one or more privacy masks are applied to the plurality of video frames according to results obtained from previously processed video frames. For example, results of one or more detected persons and their corresponding privacy mask flag may be obtained from

previously processed video frames and stored in, for example, a database, and these results may be reused when same video frames are encountered. This advantageously saves processing time and resources that would otherwise be required for the plurality of video frames. Thereafter, in step 1036, the plurality of processed video frames (e.g., having privacy masks that may be applied in step 1034) are encoded into the video file and generated as output, and the process ends.

[0053] If it is determined that there are new video frames in step 1006, the process proceeds instead to step 1008 in which object detection is performed on each video frame of the plurality of video frames. In step 1010, it is determined whether there is any detected person in a video frame having a privacy mask flag with an initial value of “?” (e.g., indicating that the detected person has not been processed to determine if a privacy mask should be applied). If it is determined that there is no detected person with a privacy mask flag with an initial value of “?”, the process returns to step 1006. Otherwise, the process proceeds to step 1012 in which it is determined whether a height of a detected person is less than a threshold height. If the height is greater than or equal to the threshold height, the process proceeds to step 1030 in which the privacy mask flag of the detected person is set to “TRUE”. In step 1032, an identifier (ID) of the video frame being processed and results of the processing and detection (e.g., results of all detected persons in the video frame and corresponding privacy mask flag) are stored in, for example, a database, which may be retrieved for use during processing of other video frames in step 1010. The process then proceeds to steps 1032 and 1034 and ends accordingly.

[0054] If the height is determined to be less than the threshold height in step 1012, the process proceeds to step 1014 instead, in which it is determined whether a height to width ratio of the detected person is less than a threshold ratio. If it is determined that the height to width ratio of the detected person is greater than or equal to the threshold ratio, the process proceeds to step 1026 in which the privacy mask flag of the detected person is set to “FALSE” (e.g., no privacy mask shall be applied to the detected person). The process then

proceeds to steps 1032, 1034 and 1036 as previously explained, and then the process ends. If it is determined at step 1014 that the height to width ratio of the detected person is smaller than the threshold ratio, the process proceeds to step 1016 in which a Y-coordinate search range is calculated to determine a 2D plane corresponding to the detected person.

[0055] In step 1018, it is determined whether a person with a highest height within the Y-coordinate search range (e.g., highest height among one or more other persons within the same 2D plane as the detected person) is found. If such a person is not found (e.g., there is no other person within the same 2D plane as the detected person), the process proceeds to step 1028 in which the privacy mask flag of the detected person is set to “TRUE” as a default value, and then proceeds to steps 1032, 1034 and 1036 as previously explained, and then the process stops. Otherwise, the process proceeds from step 1018 to step 1020 in which the privacy mask flag of the person with the highest height is checked. If it is checked that the privacy mask flag is “?”, the process proceeds to step 1022 to determine whether the privacy mask flag of the person with the highest height should be “TRUE” or “FALSE” (e.g., determine whether a privacy mask should be applied on the person based on comparison of the person’s height against the threshold height), and the process then proceeds to step 1024 in which the privacy mask flag is updated according to the determination in step 1022. If, on the other hand, it is checked that the privacy mask flag is not “?” in step 1020, step 1024 is skipped. After step 1024, the process proceeds to steps 1030, 1032, 1034 and 1036 as previously explained, and then the process ends.

[0056] Figure 11 shows an exemplary illustration of a resultant image 1100 in which objects are detected and privacy masks are applied. For example, person 1102 may have a height that is less than a threshold height and a height to width ratio that is less than a threshold ratio. Thus, a Y-coordinate search (e.g., search for a person with a highest height among one or more other persons within a 2D plane 1104 as person 1102) is performed and a person 1110 is found from the search. As the detected person 1110 has a height that is greater than or equal to the threshold height, person 1110’s corresponding privacy mask flag is set to

"TRUE" (e.g., a privacy mask flag is applied on person 1110 to protect his/her identity). Further, the privacy mask flag of person 1102 is also set to "TRUE" (e.g., privacy masking is necessary for person 1102) since person 1110's corresponding privacy mask flag is set to "TRUE". Further, each detected person 1106, 1108 and 1112 has a height that is greater than or equal to the threshold height, and thus their corresponding privacy mask flag is set to "TRUE" (e.g., a privacy mask flag is applied on each of them to protect their identities). Furthermore, detected persons 1114 and 1116 do not have privacy mask applied because their corresponding height is less than the threshold height.

[0057] Figure 12 depicts an exemplary computing device 1200, hereinafter interchangeably referred to as a computer system 1200, where one or more such computing devices 1200 may be used to execute the method of Figure 5 and Figure 10. The exemplary computing device 1200 can be used to implement an apparatus for applying a mask on an object in an image. In an implementation, a system for applying a mask on an object in an image may comprise the apparatus and one or more image and video capturing apparatuses configured to capture one or more images or video frames, wherein the one or more images or video frames comprises an image of a person. The following description of the computing device 1200 is provided by way of example only and is not intended to be limiting.

[0058] As shown in Figure 12, the example computing device 1200 includes a processor 1204 for executing software routines. Although a single processor is shown for the sake of clarity, the computing device 1200 may also include a multi-processor system. The processor 1204 is connected to a communication infrastructure 1206 for communication with other components of the computing device 1200. The communication infrastructure 1206 may include, for example, a communications bus, cross-bar, or network.

[0059] The computing device 1200 further includes a main memory 1208, such as a random access memory (RAM), and a secondary memory 1210. The secondary memory 1210 may include, for example, a storage drive 1212, which may be a hard disk drive, a solid state drive or a hybrid drive and/or a removable

storage drive 1214, which may include a magnetic tape drive, an optical disk drive, a solid state storage drive (such as a USB flash drive, a flash memory device, a solid state drive or a memory card), or the like. The removable storage drive 1214 reads from and/or writes to a removable storage medium 1218 in a well-known manner. The removable storage medium 1218 may include magnetic tape, optical disk, non-volatile memory storage medium, or the like, which is read by and written to by removable storage drive 1214. As will be appreciated by persons skilled in the relevant art(s), the removable storage medium 1618 includes a computer readable storage medium having stored therein computer executable program code instructions and/or data.

[0060] In an alternative implementation, the secondary memory 1210 may additionally or alternatively include other similar means for allowing computer programs or other instructions to be loaded into the computing device 1200. Such means can include, for example, a removable storage unit 1222 and an interface 1220. Examples of a removable storage unit 1222 and interface 1220 include a program cartridge and cartridge interface (such as that found in video game console devices), a removable memory chip (such as an EPROM or PROM) and associated socket, a removable solid state storage drive (such as a USB flash drive, a flash memory device, a solid state drive or a memory card), and other removable storage units 1222 and interfaces 1220 which allow software and data to be transferred from the removable storage unit 1222 to the computer system 1200.

[0061] The computing device 1200 also includes at least one communication interface 1224. The communication interface 1224 allows software and data to be transferred between computing device 1200 and external devices via a communication path 1226. In various embodiments of the inventions, the communication interface 1224 permits data to be transferred between the computing device 1200 and a data communication network, such as a public data or private data communication network. The communication interface 1224 may be used to exchange data between different computing devices 1200 which such computing devices 1200 form part an interconnected computer network. Examples of a communication interface 1224 can include a modem, a network

interface (such as an Ethernet card), a communication port (such as a serial, parallel, printer, GPIB, IEEE 1394, RJ45, USB), an antenna with associated circuitry and the like. The communication interface 1224 may be wired or may be wireless. Software and data transferred via the communication interface 1224 are in the form of signals which can be electronic, electromagnetic, optical or other signals capable of being received by communication interface 1224. These signals are provided to the communication interface via the communication path 1226.

[0062] As shown in Figure 12, the computing device 1200 further includes a display interface 1202 which performs operations for rendering images or videos to an associated display 1230 and an audio interface 1232 for performing operations for playing audio content via associated speaker(s) 1234.

[0063] As used herein, the term "computer program product" may refer, in part, to removable storage medium 1218, removable storage unit 1222, a hard disk installed in storage drive 1212, or a carrier wave carrying software over communication path 1226 (wireless link or cable) to communication interface 1224. Computer readable storage media refers to any non-transitory, non-volatile tangible storage medium that provides recorded instructions and/or data to the computing device 1200 for execution and/or processing. Examples of such storage media include magnetic tape, CD-ROM, DVD, Blu-ray Disc, a hard disk drive, a ROM or integrated circuit, a solid state storage drive (such as a USB flash drive, a flash memory device, a solid state drive or a memory card), a hybrid drive, a magneto-optical disk, or a computer readable card such as a PCMCIA card and the like, whether or not such devices are internal or external of the computing device 1200. Examples of transitory or non-tangible computer readable transmission media that may also participate in the provision of software, application programs, instructions and/or data to the computing device 1200 include radio or infra-red transmission channels as well as a network connection to another computer or networked device, and the Internet or Intranets including e-mail transmissions and information recorded on Websites and the like.

[0064] The computer programs (also called computer program code) are stored in main memory 1208 and/or secondary memory 1210. Computer programs can also be received via the communication interface 1224. Such computer programs, when executed, enable the computing device 1200 to perform one or more features of embodiments discussed herein. In various embodiments, the computer programs, when executed, enable the processor 1204 to perform features of the above-described embodiments. Accordingly, such computer programs represent controllers of the computer system 1200.

[0065] Software may be stored in a computer program product and loaded into the computing device 1200 using the removable storage drive 1214, the storage drive 1212, or the interface 1220. The computer program product may be a non-transitory computer readable medium. Alternatively, the computer program product may be downloaded to the computer system 1200 over the communications path 1226. The software, when executed by the processor 1204, causes the computing device 1200 to perform the necessary operations to execute the method as shown in Figures 5 and 10.

[0066] It is to be understood that the embodiment of Figure 12 is presented merely by way of example to explain the operation and structure of an apparatus for applying a mask on an object in an image. Therefore, in some embodiments one or more features of the computing device 1200 may be omitted. Also, in some embodiments, one or more features of the computing device 1200 may be combined together. Additionally, in some embodiments, one or more features of the computing device 1200 may be split into one or more component parts.

[0067] It will be appreciated by a person skilled in the art that numerous variations and/or modifications may be made to the present invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects to be illustrative and not restrictive.

CLAIMS

1. A method for applying a mask on an object in an image comprising:

determining, by a processor, whether to apply a mask on an object in an image based on a comparison with another object in the image; and

applying, by the processor, the mask on the object in the image based on the determination.

2. The method of claim 1, wherein determining whether to apply the mask comprises:

comparing a height of the object in the image with a threshold height;

if the height of the object is lower than the threshold height, further comparing a ratio of the height to a width of the object in the image with a threshold ratio; and

determining whether to apply the mask based on the further comparison.

3. The method of claim 2, further comprising:

applying the mask on the object if the height of the object is higher than or equal to the threshold height.

4. The method of claim 2, wherein the comparison with the another object comprises:

if the ratio is lower than the threshold ratio, identifying the another object from one or more other objects in the image, with the another object having a height that is highest among the object and the one or more other objects, the one or more other objects being located within a two-dimensional plane distance from the object; and

determining whether a mask is to be applied on the another object.

5. The method of claim 4, further comprising applying the mask on the object if it is determined that a mask is to be applied on the another object.

6. The method of claim 2, further comprising: determining not to apply the mask on the object if the ratio is higher than the threshold ratio.

7. The method of any one of claims 1-6, further comprising applying the mask on the object in a plurality of frames of a video, the image being a frame of the plurality of frames of the video.

8. An apparatus for applying a mask on an object in an image, comprising:

at least one processor; and

at least one memory including computer program code; the at least one memory and the computer program code configured to, with at least one processor, cause the apparatus at least to:

determine whether to apply a mask on an object in an image based on a comparison with another object in the image; and

apply the mask on the object in the image based on the determination.

9. The apparatus of claim 8, wherein determining whether to apply the mask comprises:

comparing a height of the object in the image with a threshold height;

if the height of the object is lower than the threshold height, further comparing a ratio of the height to a width of the object in the image with a threshold ratio; and

determining whether to apply the mask based on the further comparison.

10. The apparatus of claim 9, further configured to:

apply the mask on the object if the height of the object is higher than or equal to the threshold height.

11. The apparatus of claim 9, wherein the comparison with the another object comprises:

if the ratio is lower than the threshold ratio, identifying the another object from one or more other objects in the image, the another object having

a height that is highest among the object and the one or more other objects, the one or more other objects being located within a two-dimensional plane distance from the object; and

determining whether a mask is to be applied on the another object.

12. The apparatus of claim 11, further configured to apply the mask on the object if it is determined that a mask is to be applied on the another object.

13. The apparatus of claim 9, further configured to: determine not to apply the mask on the object if the ratio is higher than the threshold ratio.

14. The apparatus of any one of claims 8-13, further configured to apply the mask on the object in a plurality of frames of a video, the image being a frame of the plurality of frames of the video.

15. A system for applying a mask on an object in an image, comprising the apparatus of any one of claims 8-14 and one or more image and video capturing apparatuses configured to capture one or more images or video frames, wherein the one or more images or video frames comprises an image of a person, the person being the object.