



- (51) International Patent Classification:
G06T 15/20 (2006.01) *G06K 9/00* (2006.01)
- (21) International Application Number:
PCT/US2014/058142
- (22) International Filing Date:
29 September 2014 (29.09.2014)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
61/883,927 27 September 2013 (27.09.2013) US
61/949,999 7 March 2014 (07.03.2014) US
- (71) Applicant: **PELICAN IMAGING CORPORATION**
[US/US]; 451 El Camino Real, Suite 100, Santa Clara, CA
95050 (US).
- (72) Inventors: **YANG, Samuel**; 451 El Camino Real, Suite
100, Santa Clara, CA 95050 (US). **SRIKANTH, Mano-
har**; 451 El Camino Real, Suite 100, Santa Clara, CA
95050 (US). **LELESCU, Dan**; 18325 Serra Avenida, Mor-

gan Hill, CA 95037 (US). **VENKATARAMAN, Kartik**;
1736 Indigo Oak Lane, San Jose, CA 95121 (US).

- (74) Agent: **BAILEY, David, J.**; KPPB LLP, 2400 E. Katella,
Suite 1050, Anaheim, CA 92806 (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, MZ, NA, RW, SD, SL, ST, SZ,
TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU,
TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE,
DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,

[Continued on next page]

- (54) Title: SYSTEMS AND METHODS FOR DEPTH-ASSISTED PERSPECTIVE DISTORTION CORRECTION

FIG. 1A



FIG. 1B



FIG. 1C



FIG. 1D



FIG. 1E



FIG. 1F



FIG. 1G



FIG. 1H



(57) Abstract: Systems and methods for automatically cor-
recting apparent distortions in close range photographs that
are captured using an imaging system capable of capturing
images and depth maps are disclosed. In many embodiments,
faces are automatically detected and segmented from images
using a depth-assisted alpha matting. The detected faces can
then be re-rendered from a more distant viewpoint and com-
posited with the background to create a new image in which
apparent perspective distortion is reduced.



LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, **Published:**

SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,

GW, KM, ML, MR, NE, SN, TD, TG).

— *without international search report and to be republished
upon receipt of that report (Rule 48.2(g))*

SYSTEMS AND METHODS FOR DEPTH-ASSISTED PERSPECTIVE DISTORTION CORRECTION

FIELD OF THE INVENTION

[0001] The present invention relates generally to correcting of perspective distortion and more specifically to automatic depth-assisted face perspective distortion correction.

BACKGROUND OF THE INVENTION

[0002] Close range portraiture photographs, such as self-portraits, are often perceived as having apparent perspective distortions at typical image viewing distances, even if the optics produce a geometrically accurate perspective projection of the scene. A mismatch in the field of view of the camera and that of the viewing display configuration can result in the perceived distortion, which in portraiture photographs, tends to magnify the size of the nose and chin, among other features. These distortions are especially common with photographs taken with mobile device cameras due to the wide angular field of view typical of such cameras and close range nature of many self-portraits and candid portraits. Examples of perceived distortion in portraiture photos can be seen in FIGS. 1A – 1H. FIGS. 1A, 1C, 1E, and 1G are portraiture images captured at a distance of 0.3 m. FIGS. 1B, 1D, 1F, and 1H are portraiture images captured at a distance of 0.6 m.

SUMMARY OF THE INVENTION

[0003] Systems and methods in accordance with embodiments of the invention automatically correct apparent distortions in close range photographs that are captured using an imaging system capable of capturing images and depth maps. In many embodiments, faces are automatically detected and segmented from images using a depth-assisted alpha matting. The detected faces can then be re-rendered from a more distant viewpoint and composited with the background to create a new image in which apparent perspective distortion is reduced.

[0004] One embodiment of the invention includes a processor; and memory containing an image processing application, and image data for an image captured from an initial viewpoint and depth map data indicating distances to objects within the scene from the initial viewpoint. In addition, the image processing application stored in the memory directs the processor to: detect an object within the image data and a distance from the initial viewpoint to the object from the depth map data; segment object image data from background image data using the depth map data; rerender the object from a synthetic viewpoint by warping the segmented object image data based upon the depth map data to generate warped object image data, where the synthetic viewpoint is a greater distance from the object relative to the distance from the initial viewpoint to the object; and combine the warped object image data with the background image data to create perspective distortion corrected image data.

[0005] As can be readily appreciated, the embodiment of the invention can include a number of optional additional aspects of the invention depending upon the requirements of specific applications. Accordingly, one or more of any of the following optional aspects of the invention may be present in an embodiment of the invention and the invention should be understood as involving disclosure of separate embodiments in which each of the possible combinations of the following optional aspects of the invention are present.

[0006] In a further optional aspect of the embodiment, the object detected within the image data is a face.

[0007] In another optional aspect of the embodiment, the image processing application stored in the memory directs the processor to segment object image data from background image data using the depth map data by identifying a trimap for the image data using the image data and the depth map data.

[0008] In a still further optional aspect of the embodiment, the trimap includes pixels within the image data identified as: foreground pixels; background pixels; or unknown pixels.

[0009] In still another optional aspect of the embodiment, the image processing application stored in the memory directs the processor to segment object image data from background image data using the depth map data by segmenting the object image data by generating an alpha matte using the trimap.

[0010] In a yet further optional aspect of the embodiment, the image processing application stored in the memory directs the processor to warp the segmented object image data to create warped object image data by: selecting a desired viewpoint distance; projecting the segmented object image data to 3D locations based upon distances to pixels within the segmented object image data contained within the depth map data; re-projecting the 3D locations to new 2D pixel locations based upon the desired viewpoint distance to create warped object image data; and filling holes in the warped object image data.

[0011] In yet another optional aspect of the embodiment, the image processing application stored in the memory further directs the processor to warp the segmented object image data to create warped object image data by rescaling the warped object image data to correspond in size to the segmented object image data.

[0012] In a further optional aspect of the embodiment again, the image processing application stored in the memory further directs the processor to inpaint the background image data.

[0013] In another optional aspect of the embodiment again, the image processing application stored in the memory directs the processor to inpaint the background image data by identifying holes in the background image data and performing interpolation to fill the holes.

[0014] In a further additional optional aspect of the embodiment, the memory further contains occluded pixel image data and the image processing application stored in the memory directs the processor to inpaint the background image data by identifying pixels in the occluded pixel image data corresponding to holes identified in the background image data and filling the holes with identified pixels.

[0015] In another additional optional aspect of the embodiment, the image processing application stored in the memory further directs the processor to register the depth map image data to the perspective distortion corrected image data to generate a registered depth map.

[0016] In a still yet further optional aspect of the embodiment, the image processing application stored in the memory further directs the processor to store the perspective distortion corrected image data and a registered depth map in an image file.

[0017] In still yet another optional aspect of the embodiment, the image data and the depth map data are contained within an image file.

[0018] In a still further optional aspect of the embodiment again, the file further includes occluded pixel image data.

[0019] Another embodiment of the invention includes a processor, and memory containing an image processing application, and image data for an image captured from an initial viewpoint and depth map data indicating distances to objects within the scene from the initial viewpoint. In addition, the image processing application stored in the memory directs the processor to: detect a face within the image data and a distance from the initial viewpoint to the face from the depth map data; segment face image data from background image data using the depth map data by: identifying a trimap for the image data using the image data and the depth map data; and generating an alpha matte using the trimap; rerender the face from a synthetic viewpoint by warping the segmented face image data based upon the depth map data to generate warped face image data, where the synthetic viewpoint is a greater distance from the face relative to the distance from the initial viewpoint to the face by: selecting a desired viewpoint distance; projecting the segmented object image data to 3D locations based upon distances to pixels within the segmented object image data contained within the depth map data; re-projecting the 3D locations to new 2D pixel locations based upon the desired viewpoint distance to create warped object image data; and filling holes in the warped object image data. inpaint the background image data; and combine the warped face image data with the background image data to create perspective distortion corrected image data.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] FIGS. 1A – 1H are portraiture images in which perceived perspective distortion can be observed by comparing image pairs taken at 0.3 m and 0.6 m.

[0021] FIG. 2 illustrates an image and a corresponding depth map.

[0022] FIG. 3A is a flow chart illustrating a process for performing automatic depth-assisted perspective distortion correction in accordance with an embodiment of the invention.

[0023] FIG. 3B is a flow chart illustrating a process for performing automatic depth-assisted perspective distortion correction and producing a depth map registered to an image in which perspective distortion has been corrected in accordance with an embodiment of the invention.

[0024] FIG. 4 is a flow chart illustrating a process for segmenting a face from the background of an image in accordance with an embodiment of the invention.

[0025] FIG. 5 is a flow chart illustrating a process for performing a 3D warping of face image data to re-render a face from a different viewpoint in accordance with an embodiment of the invention.

[0026] FIG. 6 is a flow chart illustrating a process for inpainting segmented background image data in accordance with an embodiment of the invention.

[0027] FIG. 7 conceptually illustrates a process for performing automatic depth-assisted perspective distortion correction in accordance with an embodiment of the invention.

[0028] FIGS. 8A and 8B are images taken by a DSLR at 0.6m and 1.0 m respectively.

[0029] FIG. 8C is an image synthesized from light field image data captured by an array camera.

[0030] FIGS. 8D and 8E illustrate perspective distortion corrected version of the image shown in FIG. 8C re-rendered from a viewpoint at a distance of 0.6 m and 1.0 m respectively.

[0031] FIG. 9 conceptually illustrates an image sharing system including image processing systems configured to perform automatic depth-assisted perspective distortion correction in accordance with an embodiment of the invention.

[0032] FIG. 10 conceptually illustrates an image processing system configured to perform automatic depth-assisted perspective distortion correction in accordance with an embodiment of the invention.

DETAILED DESCRIPTION

[0033] Turning now to the drawings, systems and methods for automatic depth-assisted perspective distortion correction in accordance with embodiments of the invention are illustrated. In many embodiments, a face is detected within an image for which a depth map is available. The depth map can be used to segment the face from the background of the image and warp the pixels of the segmented face to rerender the face from a viewpoint at a desired distance greater than the distance from which the camera captured the image of the face. In this way, the perceived perspective distortion in the face can be removed and the rerendered face composited with the image background. In several embodiments, the image background is inpainted to fill any holes created by the segmentation process. In many embodiments, the shifts in pixel locations between the original image and the perspective distortion corrected image are also applied to the original depth map to generate a depth map for the perspective distortion corrected image. While much of the discussion of automatic depth-assisted perspective correction provided below relates to performing depth-assisted perspective correction of faces within images, similar techniques can be applied to apply depth-assisted perspective correction to any of a variety of objects within an image.

Processes for Performing Depth-Assisted Perspective Correction

[0034] A depth map is an image or image channel that contains information relating to the distances from a viewpoint to the surfaces of objects visible within the scene captured by a corresponding (color) image. By way of example, an image and its corresponding depth map are shown in FIG. 2. The image 200 include an image of a human face 202 and the corresponding depth map 204 shows distances to the human face and the background of the image. Various imaging systems exist that are capable of capturing image data from the visible light spectrum and producing a corresponding depth map including (but not limited to) the various array cameras described in U.S. Patent Publication No. 2011/0069189 entitled "Capturing and Processing of Images

Using Monolithic Camera Array with Heterogeneous Imagers”, to Venkataraman et al. The disclosure within U.S. Patent Publication No. 2011/0069189 regarding the implementation of different array camera architectures including monolithic array cameras, non-monolithic array cameras, and arrays of array cameras is hereby incorporated by reference herein in its entirety. The disclosure within U.S. Patent Publication No. 2011/0069189, U.S. Patent Publication No. 2012/0147205 entitled “Systems and Methods for Synthesizing High Resolution Images Using Super-Resolution Processes”, to Lelescu et al., and U.S. Patent 8,619,082 entitled “Systems and Methods for Parallax Detection and Correction in Images Captured Using Array Cameras that Contain Occlusions using Subsets of Images to Perform Depth Estimation” to Ciurea et al. concerning estimating depth and generating a depth map using multiple images of a scene and synthesizing images from different perspectives using depth information is also incorporated by reference herein in its entirety.

[0035] Processes in accordance with many embodiments of the invention use depth maps to perform perspective correction. Objects can be segmented from images and knowledge of the distance of the object from the viewpoint of the camera that acquired the image can be utilized to rerender the object from a viewpoint that is further away. The rerendering can be particularly effective with cameras that have wide fields of view and where images are taken close to the object, which is a typical imaging scenario when taking a self-portrait using a camera on a mobile phone or tablet computing device.

[0036] A process for automatically performing depth-assisted perspective correction of an image using a depth map in accordance with an embodiment of the invention is illustrated in FIG. 3A. The process 300 includes obtaining (302) image data and depth map data, where the image data describes an image captured from a viewpoint and the depth map data describes the distance from the viewpoint to surfaces of objects visible within the image data. In an optional aspect of this embodiment of the invention, the image data and depth map data can be contained within an image file similar to any of the image file formats described in U.S. Patent 8,542,933 entitled “Systems and Methods for Decoding Light Field Image Files” to Venkataraman et al. The disclosure in U.S. Patent 8,542,933 concerning storing image data, depth map data, and auxiliary

data such as (but not limited to) confidence map data and occluded pixel data, and extracting image data, depth map data, and auxiliary data from image files is hereby incorporated by reference in its entirety. In other embodiments, any of a variety of techniques for storing image data and/or depth map data can be utilized as appropriate to the requirements of specific applications.

[0037] In the illustrated embodiment, perspective correction is applied to a face within the image data 302. As noted above, perspective correction can be applied to any of a variety of objects as appropriate to the requirements of a specific application. In order to perform perspective correction, the depth map is used to separate (304) the face image data from background image data. The term face image data is used here to indicate pixels from the original image data corresponding to a face identified from within the image. The background image data refers to pixels that are not part of the face image data. A warp (306) can be applied to the face image data based upon the depth map to rerender the face from a more distant viewpoint, thereby correcting for perspective distortion.

[0038] The background image data is inpainted (308) to fill holes that may exist following segmentation that remain deoccluded following compositing with the rerendered face image data. In an optional aspect of this embodiment of the invention, inpainting can involve a simple interpolation process. In embodiments where the image data and the depth map are generated from light field image data captured by an array camera similar to any of the various array cameras described in U.S. Patent Publication No. 2011/0069189, then inpainting can be performed using image data captured from alternative viewpoints. The process of synthesizing an image using light field image data captured by an array camera can involve selecting a reference viewpoint. Cameras with viewpoints that differ from the reference viewpoint can be said to have alternative views of the scene. Quite often, portions of the background that are not visible in the reference viewpoint are visible in one or more of the alternative viewpoints. Accordingly, pixels occluded by foreground objects in the reference viewpoint that are captured by cameras from alternative viewpoints can be utilized in the inpainting of the background image data of an image from the reference viewpoint following the segmentation of the image data. As disclosed in U.S. Patent 8,542,933, occluded pixels

and their corresponding depths can be included in an image file with image data and depth map data to assist with the rerendering of the image data. The disclosure in U.S. Patent 8,542,933 of utilizing occluded pixel data in the rerendering of image data from different viewpoints is hereby incorporated by reference herein in its entirety. Although specific inpainting processes are described herein, any of a variety of inpainting processes can be utilized as appropriate to the requirements of specific applications in accordance with embodiments of the invention.

[0039] The rerendered face image data and the inpainted background image data can be composited (310) to generate a new image in which perspective distortion is eliminated. In an optional aspect of this embodiment of the invention, a user can control the distance from which to rerender the face image data. In another optional aspect of this embodiment of the invention, the distance from which to rerender the face can be predetermined and/or automatically determined.

[0040] In an optional aspect of this embodiment of the invention, the pixel shifts that result from the 3D warping applied to the face image data are applied to the depth map to generate a depth map registered with respect to the composited image. In another optional aspect of this embodiment of the invention, the depth values in the depth map are also modified to reflect the increase in the distance between the new viewpoint and the surfaces in the perspective distortion corrected image. In a further optional aspect of this embodiment of the invention, holes within the warped depth map can be filled using an inpainting process. As noted above, inpainting can be performed using interpolation and/or using depth information from occluded background pixels that become disoccluded by the warping of the face image data and for which depth information is available. A process for correcting perspective distortion in images using a depth map and for registering the depth map to the perspective distortion corrected image in accordance with an embodiment of the invention is illustrated in FIG. 3B. The process 330 is similar to the process 300 shown in FIG. 3A with the addition that the depth map is adjusted (332) to register the depth map with respect to perspective distortion corrected image. Accordingly, corrected image data and corresponding depth map image data (334) are output.

[0041] Although specific processes are described above for rerendering faces to correct for perceived perspective distortion with reference to FIGS. 3A and 3B, any of a variety of processes can be utilized to segment foreground objects from images, rerender the segmented objects from alternate viewpoints including (but not limited to) more distant viewpoints, and composite the rerendered objects with the background of the original image in accordance with embodiments of the invention. Processes for segmenting foreground objects such as (but not limited to) faces using depth maps in accordance with embodiments of the invention are discussed further below.

SEGMENTATION BASED UPON DEPTH

[0042] Perspective distortion correction in accordance with many embodiments of the invention involve segmentation of foreground objects from the backgrounds of images. In a number of embodiments, the process of segmentation can leverage distance information contained within a depth map to separate foreground and background objects. In several embodiments, an object detection technique is used to identify the presence of a specific type of object (e.g. a face). Many embodiments of the invention use a cascade of classifiers based approach to object detection in which a decision tree of weak classifiers trained to detect a specific type of object is applied in stages to different regions within an image until the candidate region is rejected or all of the stages are passed. The use of a cascade of classifiers based approach to detect faces is disclosed in P. Viola, M. Jones, *Robust Real-time Object Detection*, IJCV 2001, the disclosure of which is hereby incorporated by reference in its entirety. In other embodiments, any of a variety of object detection techniques can be utilized including other feature based approaches, template matching approaches, bag-of-words models and/or any other of a variety of techniques appropriate to the requirements of specific applications. Once an object is detected, the distance to the object can be utilized to construct a trimap. The term trimap can be utilized to determine a map that categorizes pixels and/or blocks of pixels as being either part of the detected (foreground) object,

part of the background, or of unknown region. A trimap can be utilized to generate an alpha matte. An alpha matte α_z can be determined by modeling and observed image I_z as a convex combination of a foreground image F_z and a background image B_z :

$$I_z = F_z \alpha_z + (1 - \alpha_z) B_z$$

where α_z can be any value in $[0, 1]$.

[0043] The foreground pixels in the trimap can be assigned $\alpha_z = 1$, the background pixels assigned $\alpha_z = 0$, and a value for α_z in the range $[0, 1]$ determined for unknown pixels using any appropriate technique including, but not limited to, Laplacian color matting. Segmentation can then be performed by applying a threshold to the alpha matte. Processes for generating trimaps, alpha mattes, and performing segmentation using depth information in accordance with embodiments of the invention are described further in U.S. Provisional Patent Application Serial No. 61/949,999, entitled "Depth Regularization and Semiautomatic Matting Using RGB-D Images", filed March 7, 2014, the relevant disclosure from which is incorporated by reference herein in its entirety.

[0044] A process for depth-assisted segmentation of faces from image backgrounds in accordance with an embodiment of the invention is illustrated in FIG. 4. The process 400 includes obtaining (402) image data and corresponding depth map data. A face detector can be applied (404) to the image data and/or depth map data to detect the presence of a face within the image. In addition to defining a region of the image containing a face, the face detector can utilize the depth map data to determine (406) a distance and/or a range of distances corresponding to the distance from the viewpoint of the camera that captured the image to the detected face. The distance and/or range of distances can be utilized in the identification of pixels corresponding to the detected face and the background. In an optional aspect of this embodiment of the invention, a trimap is created (408) by classifying pixels according to the determined (406) distance and/or range of distances. In a further optional aspect of this embodiment of the invention, the process of creating a trimap can also involve the consideration of confidence scores related to the depth measurements in the depth map in the

assignment of pixels to regions in the trimap. Where an occlusion map is also available (i.e. a map that shows regions of the image that are occluded in the images captured by one or more cameras in the array of cameras used to capture the image data), the occlusion map can also be used in the generation of the trimap. As noted above, the trimap can be used as an input to a matting process in which an alpha map is generated, and the alpha matte used to segment (410) the face image data from the background image data (412).

[0045] Although specific processes for segmenting foreground objects from image backgrounds are described above with respect to FIG. 4, any of a variety of processes that can be utilized to segment foreground and background image data can be utilized as appropriate to the requirements of specific applications in accordance with embodiments of the invention. Processes for warping image data for segmented foreground objects to remove perspective distortion and for inpainting segmented background image data in accordance with various embodiments of the invention are discussed further below.

WARPING IMAGE DATA TO RERENDER FROM MORE DISTANT VIEWPOINT

[0046] Warping image data to rerender a foreground object from a more distant viewpoint based upon a depth map can be considered to be a process involving rerendering based upon an explicit geometry. In many embodiments, the warping process is a 3D warping process that leverages the fact the depth information provided in the depth map for each of the foreground pixels to rerender the pixels from a new (different) point of view by projecting the pixels from the original image to their 3D locations in space and re-projecting the pixels to the appropriate locations for the new viewpoint.

[0047] A process for warping foreground image data based upon a depth map in accordance with an embodiment of the invention is illustrated in FIG. 5. The process 500 includes determining (502) the distance of a new viewpoint. While much of the discussion above assumes that the new viewpoint is at a greater distance than the original viewpoint, processes in accordance with embodiments of the invention can select a distance that is closer than the original viewpoint. In an optional aspect of this

embodiment of the invention, the distance of the new viewpoint is predetermined. In a further optional aspect of this embodiment of the invention, the distance can be selected by an automated process trained using a machine learning algorithm based upon features such as (but not limited to) the size and/or ratios of different facial features in the facial image data and/or by user input.

[0048] The pixels in the face image data are projected (504) to 3D locations based upon the distances to each pixel found within the depth map data. The 3D locations of the pixels can then be used to re-project (506) the pixels to new 2D locations based upon the distance to the new viewpoint. Where the re-projection process creates holes in the warped face image data, then the holes are filled (508) and/or other artifacts in the warped face image data can be corrected for using appropriate forms of filtering. The holes can be filled using any infilling process appropriate to the requirements of specific applications including, but not limited to, increasing the size of pixels and/or using any of a variety of appropriate interpolation processes. In an optional aspect of this embodiment of the invention, the resulting warped face image data is rescaled to correspond in size to the original face image data to aid with compositing the warped face image data with the background image data.

[0049] Although specific processes are described above for warping object image data to rerender an object from a different viewpoint, any of a variety of processes that utilize explicit geometry can be utilized as appropriate to the requirements of specific applications in accordance with embodiments of the invention. Furthermore, processes similar to those above can be used to warp foreground image data to viewpoints that are shifted along the optical axis and/or perpendicular to the optical axis that can utilize occluded pixel data. The inpainting of background image data to aid compositing with warped object image data in accordance with various embodiments of the invention is discussed further below.

INPAINTING BACKGROUND IMAGE DATA

[0050] Distortion of a foreground object to correct for perspective distortion can result in portions of the background of a scene becoming unoccluded when the rendered foreground is composited with the image background. The unoccluded portions of the

background can manifest as holes (absences of pixels) in the composited image. Inpainting can be performed to prevent holes from manifesting themselves in the composited image. The specific inpainting process utilized largely depends upon the requirements of a specific application. Any of a variety of interpolation processes can be utilized. Where occluded pixel image data is available (i.e. image data concerning portions of the background that are occluded in the original image by the foreground object), then the occluded pixel image data can be used to perform inpainting. The provision of occluded pixel data in image files containing image data and depth map data is described in detail in U.S. Patent 8,542,933, the relevant portions of which are incorporated by reference above.

[0051] A process for performing inpainting of a segmented background in accordance with an embodiment of the invention is illustrated in FIG. 6. The process 600 includes identifying (602) holes in the background image data resulting from segmentation to remove a foreground object. Where occluded pixel image data is available, holes for which pixel data is available in the occluded pixel image data are identified (604) and inserted (606) into the background image data to fill the relevant holes. Where holes remain and/or in the absence of occluded pixel image data, an interpolation process can be applied (608) to fill the remaining holes.

[0052] Although specific inpainting processes are described above with reference to FIG. 6, any of a variety of inpainting processes can be utilized as appropriate to the requirements of specific applications in accordance with embodiments of the invention.

PERSPECTIVE DISTORTION CORRECTION IMAGE PROCESSING PIPELINES

[0053] A variety of processes that can be utilized to perform perspective distortion correction in accordance with embodiments of the invention are described above. A process in which segmentation is performed by generating a trimap and in which background inpainting is performed in accordance with an embodiment of the invention is conceptually illustrated in FIG. 7. The process receives as inputs an image 700 and a corresponding depth map 702. Any of the processes described above can be utilized to create a trimap based upon the depth map 702 and/or the image 700 and to segment the face image data 706 from the background image data 708. A 3D warp can be

applied to the face image data 706 based upon the depths of the pixels within the face image data and a desired viewpoint from which to rerender the face to produce warped face image data 710 and the background image data is inpainted 712. The warped face image data 710 and the inpainted background image data 712 are then composited to produce a perspective distortion corrected image 714. In an optional aspect of this embodiment of the invention, the perspective distortion corrected image and/or a depth map corresponding to the perspective distortion corrected image can be inserted as layers into the file containing the original image data and/or depth map data. In a further optional aspect of this embodiment of the invention, the perspective distortion corrected image and/or a depth map corresponding to the perspective distortion corrected image are stored in a separate file.

[0054] The application of the image processing pipeline in FIG. 7 to an image synthesized using light field image data captured by an array camera is illustrated in FIGS. 8A – 8E. FIGS. 8A and 8B are images captured by a Digital SLR camera at 0.6 m and 1.0 m. FIG. 8C is an image synthesized from an array camera from light field image data captured at 0.3 m. Using observed disparity between the images captured by the cameras in the array camera, a corresponding depth map was generated for the image shown in FIG. 8C and used to perform depth-assisted perspective distortion correction using an image processing pipeline similar to the image processing pipeline described above with respect to FIG. 7. A perspective distortion corrected image synthesized using a viewpoint of 0.6 m is shown in FIG. 8D. A perspective distortion corrected image synthesized using a viewpoint of 1.0 m is shown in FIG. 8E.

[0055] Although a specific image processing pipeline is illustrated in FIG. 7, as can readily be appreciated from the above discussion any of a variety of processes can be combined together to form a depth-assisted perspective distortion correcting image processing pipeline as appropriate to the requirements of specific applications in accordance with embodiments of the invention.

SYSTEMS FOR PERFORMING PERSPECTIVE DISTORTION CORRECTION

[0056] Systems for performing perspective distortion correction in accordance with embodiments of the invention can be located anywhere within a network. As noted above, the techniques for performing perspective distortion correction are particularly effective with respect to removing perspective distortion from images captured by mobile phones and mobile computing devices. Accordingly, the mobile devices that captured the images can themselves be configured via software as image processing systems capable of performing perspective distortion correction on captured images either automatically and/or at the direction of the user. In a number of embodiments, images can be shared over networks using a number of different communication techniques including (but not limited to) email, SMS, an image sharing service (e.g. the Facebook service provided by Facebook, Inc. of Menlo Park, California) and/or a synchronized content management system (i.e. a server that automatically coordinates the copying of images from one device to another). In many embodiments, a server that receives an image shared by a mobile computing device that captured the image can be configured by software to perform perspective distortion correction. As can readily be appreciated, almost any computing device within a network can be configured to perform perspective distortion correction as appropriate to the requirements of specific applications in accordance with embodiments of the invention.

[0057] Computing devices configured to share images and perform perspective distortion correction in accordance with an embodiment of the invention are illustrated in FIG. 9. The image sharing system 90 includes a number of mobile phones 92 configured to capture image data and corresponding depth map data. Although mobile phones are illustrated, one can readily appreciate that any computing device (including computing devices that do not possess cameras) can be utilized to process and share images in accordance with embodiments of the invention. The mobile phones 92 can upload the image data and depth map data to an image processing server 94 via a network 96. In an optional aspect of this embodiment of the invention, the image processing server 94 can share uploaded images automatically and/or in accordance with instructions provided by one of the mobile phones 92.

[0058] Where an image includes perspective distortion, any of the computing devices within the image sharing system 90 can potentially be utilized to perform perspective distortion correction. In an optional aspect of this embodiment of the invention, the mobile phones 92 include software that configures the mobile phone to perform perspective distortion correction on images captured, stored, and/or received by the mobile phone. In an optional aspect of this embodiment of the invention, the image processing server can perform perspective distortion correction on images received from mobile phones. In a further optional aspect of this embodiment of the invention, the computing devices can share the original images, the perspective distortion corrected images, the corresponding depth maps and/or intermediary work product such as (but not limited to) the alpha matte used to extract the object that was warped to perform perspective distortion correction.

[0059] An image processing system that can be utilized to automatically perform depth-assisted perspective distortion correction in accordance with an embodiment of the invention is illustrated in FIG. 10. The image processing system 1000 includes a processor 1010, memory 1020, a network interface 1030, an (optional) display interface 1040, and an (optional) camera interface 1050. Servers typically will not require a display interface or a camera interface and many computing devices that include displays do not include camera interfaces. The memory 1020 includes an image processing application 1060, image and depth map data 1070, (optional) occluded pixel data 1075 and optics calibration information 1080. In an optional aspect of this embodiment of the invention, the optics calibration data can be utilized to create characteristic lens distortions in re-projected images to simulate the optical systems of specific cameras. The image processing application can configure the processor to perform depth-assisted perspective distortion correction using any of the techniques described above to synthesize a perspective distortion corrected image from a more distant viewpoint.

[0060] Although specific image sharing systems and image processing systems are described above with reference to FIGS. 9 and 10, any of a variety of computing devices can be utilized to acquire, share, and/or perform perspective distortion correction on images using depth maps as appropriate to the requirements of specific applications in accordance with embodiments of the invention.

[0061] While the above description contains many specific embodiments of the invention, these should not be construed as limitations on the scope of the invention, but rather as an example of one embodiment thereof. Accordingly, the scope of the invention should be determined not by the embodiments illustrated, but by the appended claims and their equivalents.

WHAT IS CLAIMED IS:

1. An image processing system, comprising:
a processor;
memory containing an image processing application, and image data for an image captured from an initial viewpoint and depth map data indicating distances to objects within the scene from the initial viewpoint;
wherein the image processing application stored in the memory directs the processor to:
detect an object within the image data and a distance from the initial viewpoint to the object from the depth map data;
segment object image data from background image data using the depth map data;
rerender the object from a synthetic viewpoint by warping the segmented object image data based upon the depth map data to generate warped object image data, where the synthetic viewpoint is a greater distance from the object relative to the distance from the initial viewpoint to the object; and
combine the warped object image data with the background image data to create perspective distortion corrected image data.
2. The image processing system of claim 1, wherein the object detected within the image data is a face.
3. The image processing system of claim 2, wherein the image processing application stored in the memory directs the processor to segment object image data from background image data using the depth map data by identifying a trimap for the image data using the image data and the depth map data.
4. The image processing system of claim 3, wherein the trimap comprises pixels within the image data identified as: foreground pixels; background pixels; or unknown pixels.

5. The image processing system of claim 3, wherein the image processing application stored in the memory directs the processor to segment object image data from background image data using the depth map data by segmenting the object image data by generating an alpha matte using the trimap.

6. The image processing system of claim 1, wherein the image processing application stored in the memory directs the processor to warp the segmented object image data to create warped object image data by:

selecting a desired viewpoint distance;

projecting the segmented object image data to 3D locations based upon distances to pixels within the segmented object image data contained within the depth map data;

re-projecting the 3D locations to new 2D pixel locations based upon the desired viewpoint distance to create warped object image data; and

filling holes in the warped object image data.

7. The image processing system of claim 6, wherein the image processing application stored in the memory further directs the processor to warp the segmented object image data to create warped object image data by rescaling the warped object image data to correspond in size to the segmented object image data.

8. The image processing system of claim 1, wherein the image processing application stored in the memory further directs the processor to inpaint the background image data.

9. The image processing system of claim 1, wherein the image processing application stored in the memory directs the processor to inpaint the background image data by identifying holes in the background image data and performing interpolation to fill the holes.

10. The image processing system of claim 1, wherein the memory further contains occluded pixel image data and the image processing application stored in the memory directs the processor to inpaint the background image data by identifying pixels in the occluded pixel image data corresponding to holes identified in the background image data and filling the holes with identified pixels.

11. The image process system of claim 1, wherein the image processing application stored in the memory further directs the processor to register the depth map image data to the perspective distortion corrected image data to generate a registered depth map.

12. The image process system of claim 11, the image processing application stored in the memory further directs the processor to store the perspective distortion corrected image data and a registered depth map in an image file.

13. The image processing system of claim 1, wherein the image data and the depth map data are contained within an image file.

14. The image processing system of claim 13, wherein the file further comprises occluded pixel image data.

15. An image processing system, comprising:
a processor;
memory containing an image processing application, and image data for an image captured from an initial viewpoint and depth map data indicating distances to objects within the scene from the initial viewpoint;
wherein the image processing application stored in the memory directs the processor to:
detect a face within the image data and a distance from the initial viewpoint to the face from the depth map data;

segment face image data from background image data using the depth map data by:

- identifying a trimap for the image data using the image data and the depth map data; and

- generating an alpha matte using the trimap;

rerender the face from a synthetic viewpoint by warping the segmented face image data based upon the depth map data to generate warped face image data, where the synthetic viewpoint is a greater distance from the face relative to the distance from the initial viewpoint to the face by:

- selecting a desired viewpoint distance;

- projecting the segmented object image data to 3D locations based upon distances to pixels within the segmented object image data contained within the depth map data;

- re-projecting the 3D locations to new 2D pixel locations based upon the desired viewpoint distance to create warped object image data; and

- filling holes in the warped object image data.

inpaint the background image data; and

combine the warped face image data with the background image data to create perspective distortion corrected image data.

1/11

FIG. 1A



FIG. 1B

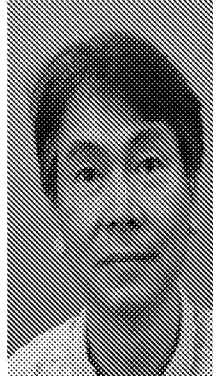


FIG. 1C

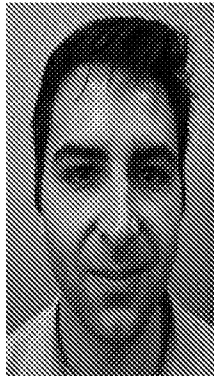


FIG. 1D

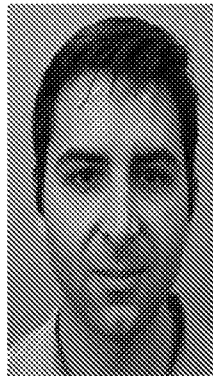


FIG. 1E



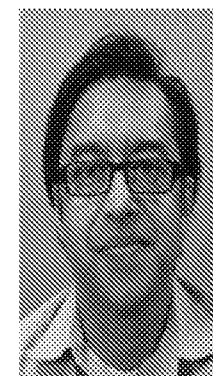
FIG. 1F



FIG. 1G



FIG. 1H





3/11

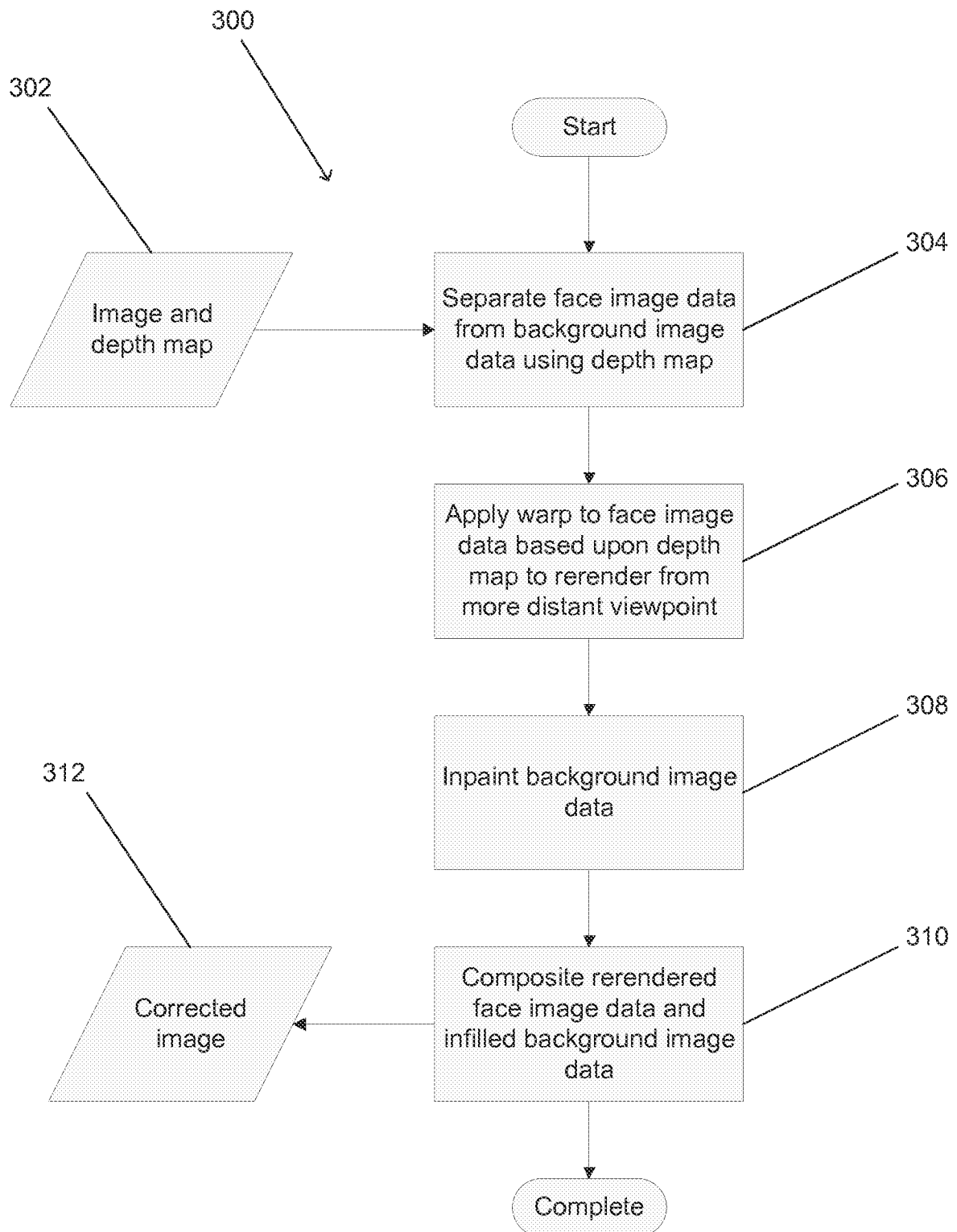
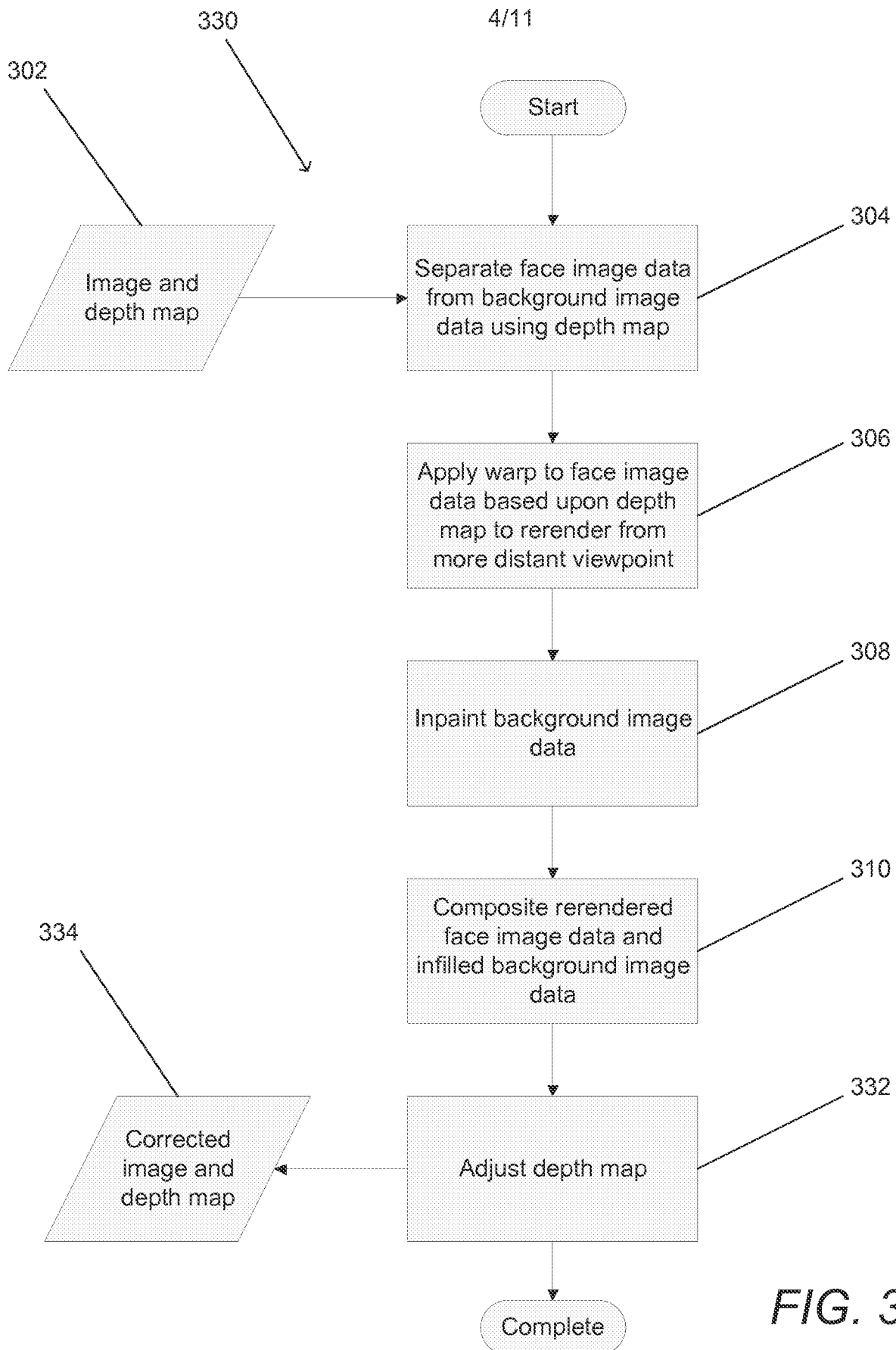
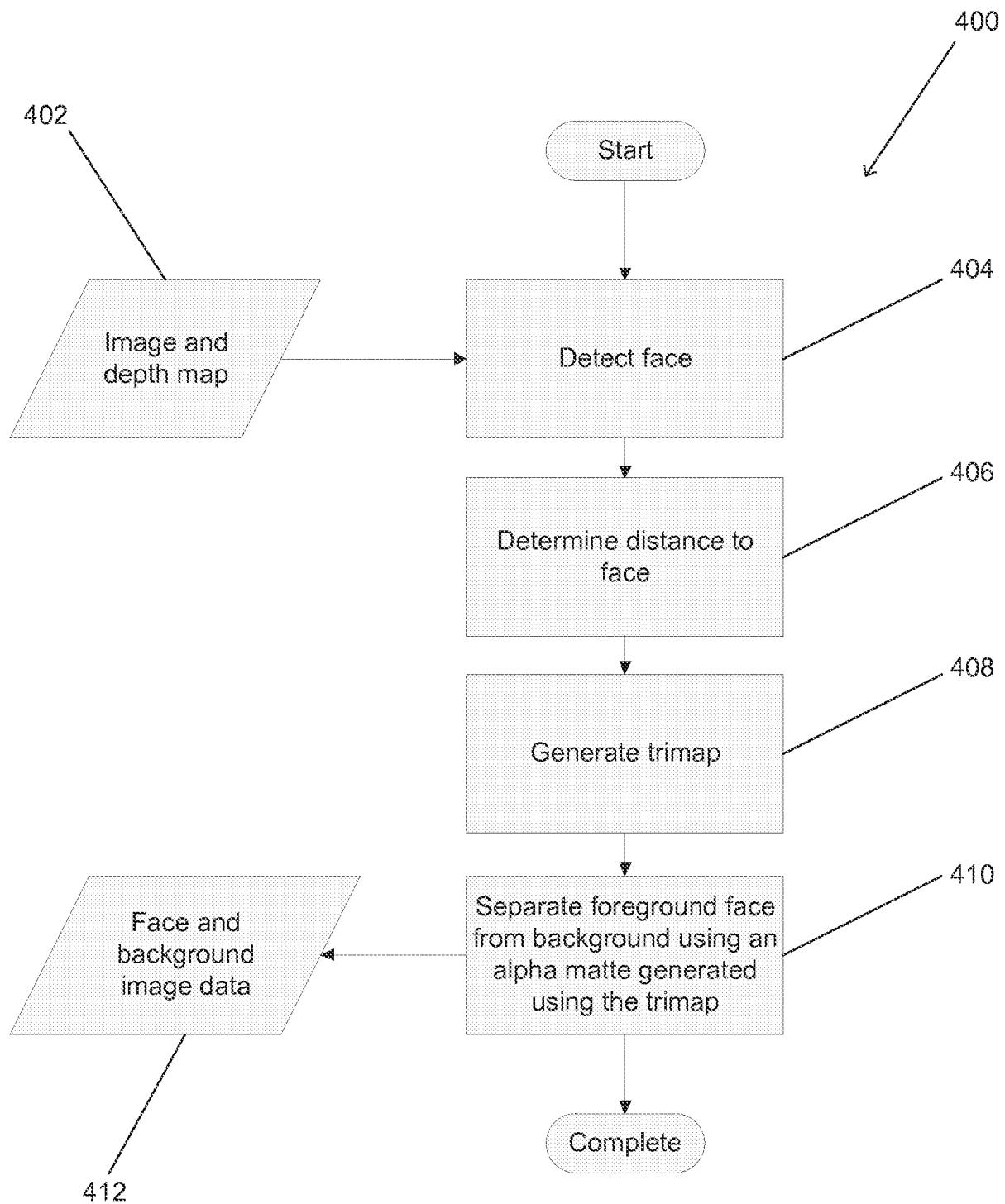


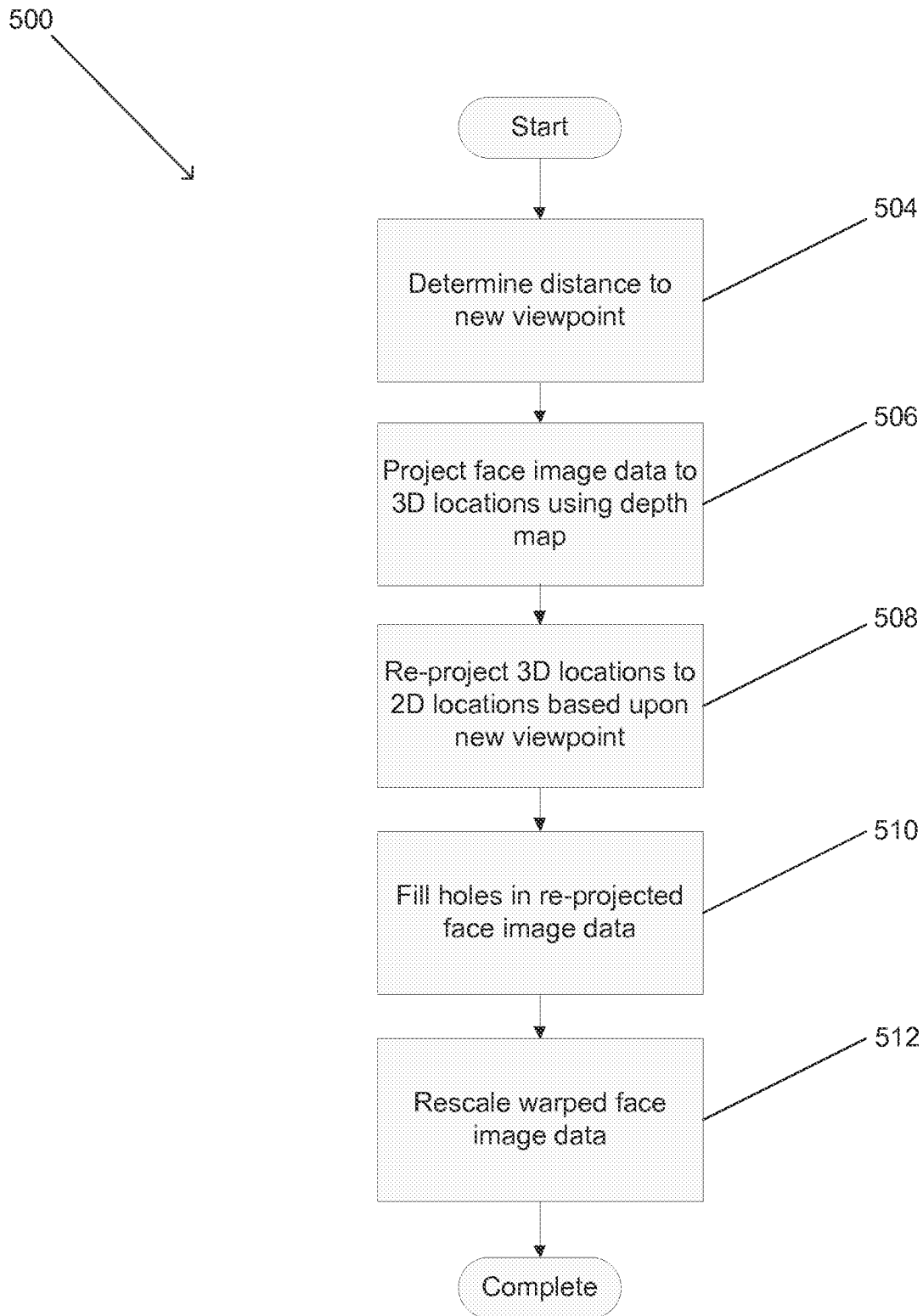
FIG. 3A

**FIG. 3B**

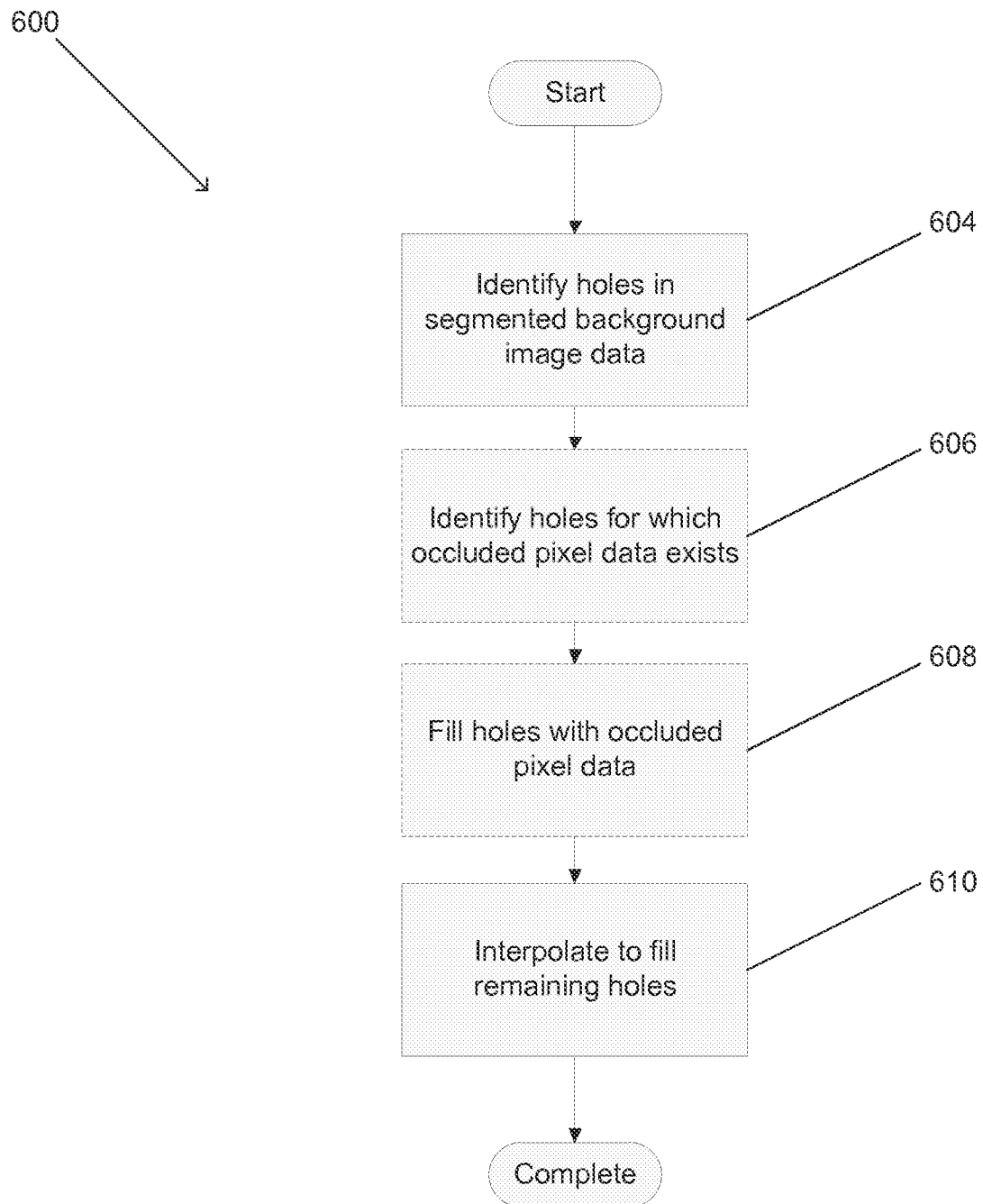
5/11

**FIG. 4**

6/11

**FIG. 5**

7/11

**FIG. 6**

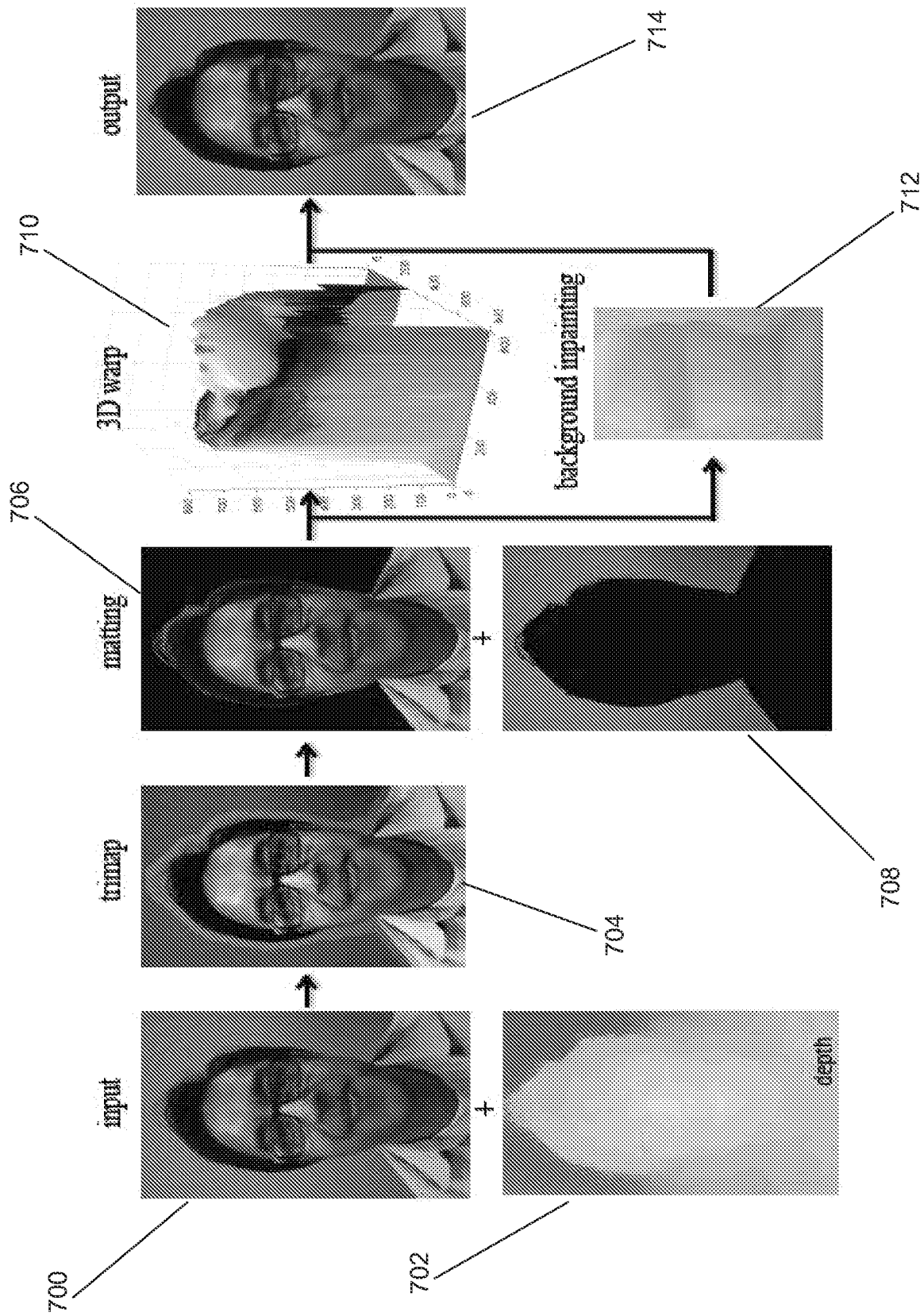
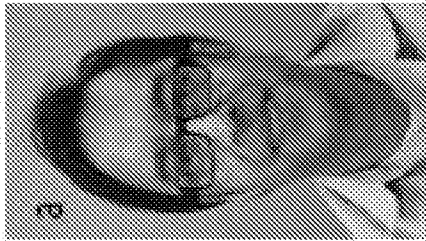


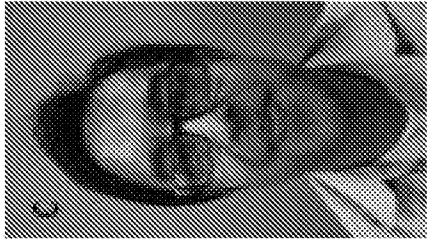
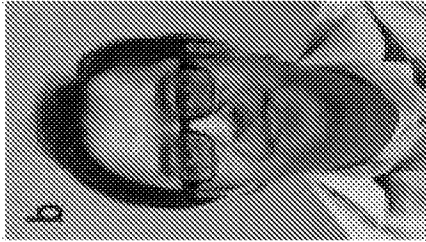
FIG. 7

FIG. 8A FIG. 8B

0.6m



1.0m



input (0.3m)

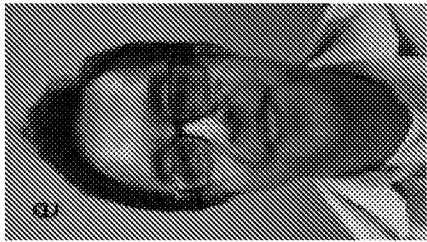
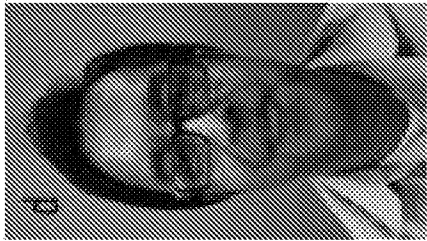


FIG. 8C FIG. 8D FIG. 8E

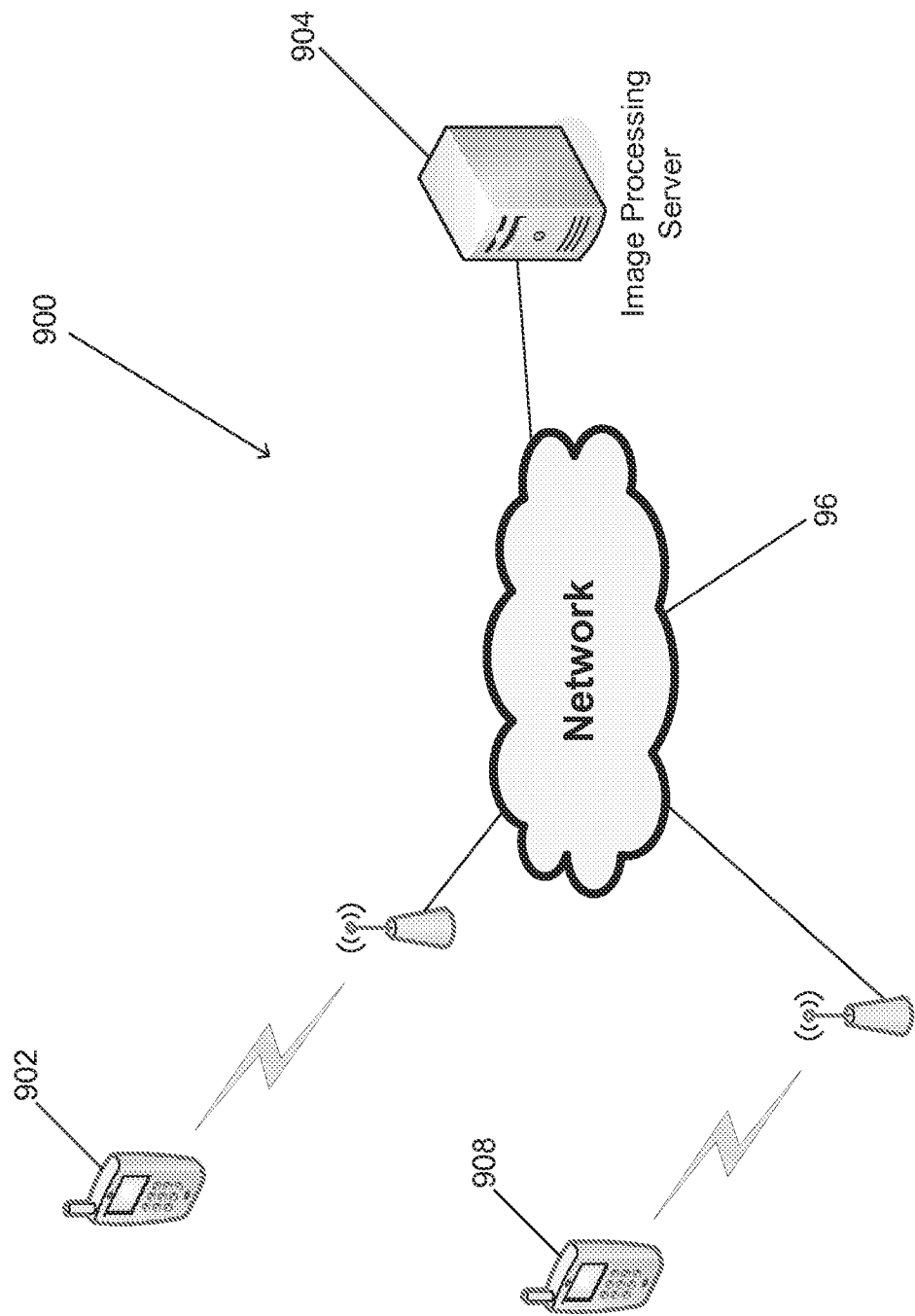


FIG. 9

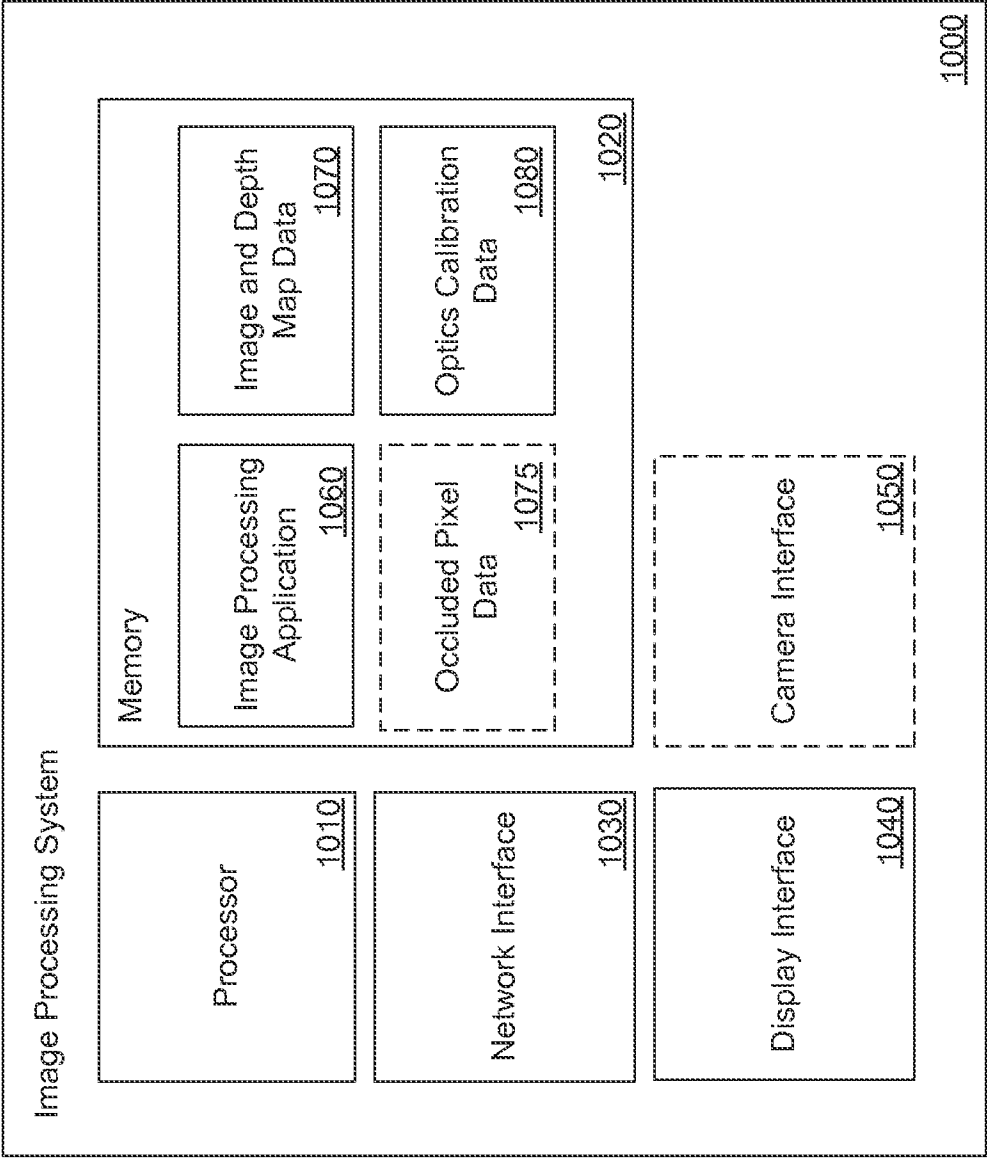


FIG. 10