



(51) International Patent Classification:

G06K 9/00 (2006.01) G06K 9/46 (2006.01)
G06K 9/20 (2006.01)

(21) International Application Number:

PCT/US2017/034339

(22) International Filing Date:

24 May 2017 (24.05.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/340,894 24 May 2016 (24.05.2016) US

(72) Inventor; and

(71) Applicant: MARTIN, Brian K. [US/US]; 296 Concord
Road, Billerica, Massachusetts 01821 (US).

(74) Agent: RENNER, W. Karl et al.; FISH & RICHARDSON
P.C., P.O. Box 1022, Minneapolis, Minnesota 55440-1022
(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, DJ, 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, KH, KN, KP, KR,
KW, 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,

(54) Title: DOCUMENT IMAGE QUALITY ASSESSMENT

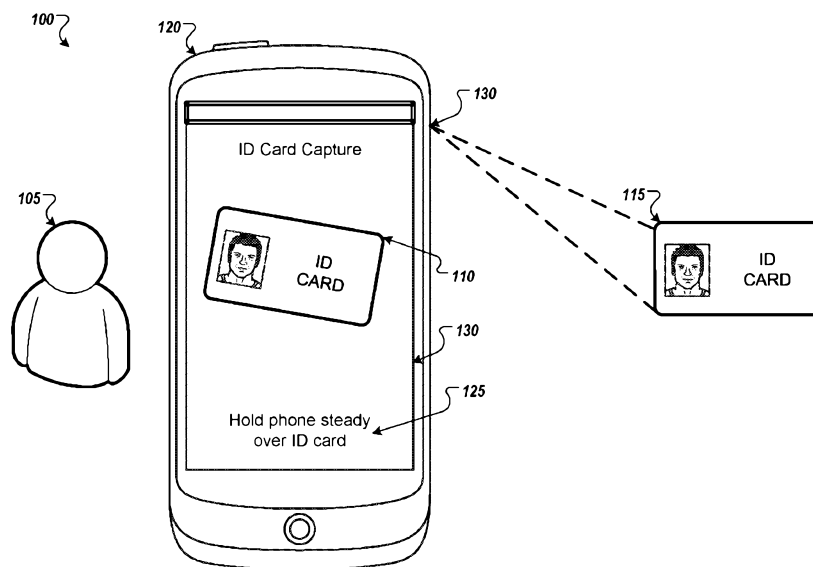


FIG. 1

(57) Abstract: Methods, systems, and apparatus, including computer programs encoded on a computer storage medium, for assessing document image quality are disclosed. In one aspect, a method includes the actions of receiving one or more images of a document. The actions further include detecting an object that has a shape that is similar to an expected shape of the document. The actions further include determining a focus score that reflects a sharpness of the first image. The actions further include determining a border fit score that reflects a likelihood that the object corresponds to the document. The actions further include determining an expected shape score that reflects a correlation between the shape of the object and an expected shape of the document. The actions further include determining a quality score that reflects a quality of the first image. The actions further include determining whether to perform document authentication on the first image.

UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

DOCUMENT IMAGE QUALITY ASSESSMENT

CROSS REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to U.S. Application No. 62/340,894, filed on May 24, 2016, which is incorporated by reference.

5

FIELD

[0002] This specification relates to document authentication.

BACKGROUND

10 **[0003]** Identification ("ID") documents play a critical role in today's society. One example of an ID document is an ID card. ID documents are used on a daily basis to prove identity, to verify age, to access a secure area, to evidence driving privileges, to cash a check, and so on. Airplane passengers are required to show an ID document during check in, security screening, and prior
15 to boarding their flight. In addition, because we live in an ever-evolving cashless society, ID documents are used to make payments, access an automated teller machine (ATM), debit an account, make a payment, and the like.

20

SUMMARY

[0004] The subject matter of this application relates to automated document authentication (e.g., ensuring a document is authentic by examination of the document's visual characteristics). In some implementations, authenticating a document accurately depends on having properly captured digital images.
25 These images may be scanned on a flat glass image plan similar to a flatbed scanner. Doing this ensures correctly focused, illuminated and scaled image of the document.

[0005] The system described below is able to extend document authentication abilities to uncontrolled capture (e.g., acquiring an image from
30 an un-calibrated configuration). An example would be to take an image of a

driver's license with a cell phone and use that image to authenticate the driver's license.

[0006] The system is able to quantify the 'goodness' of a digital image of a document for the purpose of both providing automatic corrections to the image and authenticating the document. This will enable a mobile phone to be used to capture images of documents, and the usability of those images to be evaluated on the phone to either provide feedback during the capture, or to simply accept or reject the image.

[0007] A method for assessing the quality of a captured document prior to performing document authentication includes the following actions. One or more digital images are acquired of a document. The rectangular shaped document is detected with the borders of the document defined for the first image. The quality of the first image is quantified using (i) the sharpness of focus of the document information, (ii) the confidence of the fit of the borders of the document, and (iii) the deviation of the document shape from the expected shape of the document. If the quality of any metric or combination of metrics in the action where the quality of the first image is quantified are below thresholds, the image is rejected and the other images are considered. For each additional image, the quality using the three features above (e.g., (i) the sharpness of focus of the document information, (ii) the confidence of the fit of the borders of the document, and (iii) the deviation of the document shape from the expected shape of the document) and using the values or some combination of them from the earlier quantification of the quality of the first image, an image is retained as a candidate for document authentication if the image is of sufficient quality. The image is sent to the document authentication system. If no image was above minimum thresholds, the capture process fails. In some implementations, the image is retained as a candidate for document authentication is the one with a highest quality score.

[0008] In some implementations, user feedback on failures can be provided (e.g., out of focus measure). Feedback to the capture device can be provided

(e.g., focus the camera). The results of the quality metrics can be used to select the candidate image from multiple shape candidates in a single image. The resulting image can go to a pre-processing step prior to document authentication to correct the perspective distortion of the camera.

5 **[0009]** According to an innovative aspect of the subject matter described in this application, a method for assessing document image quality includes the actions of receiving one or more images of a document; detecting, in a first image of the one or more images, an object that has a shape that is similar to an expected shape of the document; determining a focus score that reflects a sharpness of the first image; determining a border fit score that reflects a
10 likelihood that the object corresponds to the document; determining an expected shape score that reflects a correlation between the shape of the object and an expected shape of the document; based on the focus score, the border fit score, and the expected shape score, determining a quality score
15 that reflects a quality of the first image; determining whether the quality score satisfies a quality score threshold; and based on determining whether the quality score satisfies the quality score threshold, determining whether to perform document authentication on the first image.

[0010] These and other implementations can each optionally include one or
20 more of the following features. The action of determining whether the quality score satisfies a quality score threshold includes determining that the quality score satisfies the quality score threshold. The action of determining whether to perform document authentication on the first image includes determining to perform document authentication on the first image based on determining that
25 the quality score satisfies the quality score threshold. The action of determining whether the quality score satisfies a quality score threshold includes determining that the quality score does not satisfy the quality score threshold. The action of determining whether to perform document authentication on the first image includes determining not to perform document
30 authentication on the first image based on determining that the quality score does not satisfy the quality score threshold. The actions further include, in

response to determining not to perform document authentication on the first image, determining a quality score for a second image of the one or more images.

[0011] The actions further include determining whether the focus score satisfies a focus score threshold. The action of determining whether to perform document authentication on the first image is further based on determining whether the focus score satisfies the focus score threshold. The actions further include determining whether the border fit score satisfies a border fit score threshold. The action of determining whether to perform document authentication on the first image is further based on determining whether the border fit score satisfies the border fit score threshold. The actions further include determining whether the expected shape score satisfies an expected shape score threshold. The action of determining whether to perform document authentication on the first image is further based on determining whether the expected shape score satisfies the expected shape score threshold. The focus score reflects a sharpness of only the object.

[0012] The document is an identification document. The one or more images of the document are frames of a video. The actions further include determining, for each of the one or more images of the document, a quality score; and determining that the quality score of the first image is greater than the quality scores of each of the one or more images other than the first image. The action of determining whether the quality score of the first image satisfies a quality score threshold includes determining that the quality score of the first image satisfies the quality score threshold. The action of determining whether to perform document authentication on the first image includes determining to perform document authentication on the first image based on determining that the quality score of the first image satisfies the quality score threshold and based on determining that the quality score of the first image is greater than the quality scores of each of the one or more images other than the first image. The action of determining whether the quality score satisfies a quality score threshold includes determining that the quality score does not satisfy the

quality score threshold. Based on determining that the quality score does not satisfy the quality score threshold the actions further include providing, to the user, feedback for improving quality of an image of the document.

[0013] Other implementations of this aspect include corresponding systems, apparatus, and computer programs recorded on computer storage devices, each configured to perform the operations of the methods.

[0014] The subject matter described in this application may have one or more of the following advantages. A system may perform document authentication only on images of documents that are of sufficient quality.

10 Limiting document authentication to only specific images may decrease computing resources needed for document authentication, such as processing power and battery power.

[0015] The details of one or more implementations of the subject matter described in this specification are set forth in the accompanying drawings and the description below. Other features, aspects, and advantages of the subject

15 matter will become apparent from the description, the drawings, and the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] FIG. 1 illustrates an example of a user capturing an image of an identification card using a mobile device.

20

[0017] FIG. 2 illustrates an example system for assessing the quality of an image of a document.

[0018] FIG. 3 illustrates an example process for assessing the quality of an image of a document.

[0019] FIG. 4 illustrates an example video capture from a laptop computer of a business card with thin lines to illustrate a recommended capture area and bold lines that highlight the detected business card.

25

[0020] FIG. 5 illustrates, in the top image, an example video capture of a business card with example line candidates that are the thin lines, and, in the bottom image, an example diagnostic output of the edges detected and used for locating the business card.

5 **[0021]** FIG. 6 illustrates an example output rectangle with linear skew correction applied to make a standard shape business card (or driver's license card).

[0022] FIG. 7 illustrates an example of a computing device and a mobile computing device.

10

DETAILED DESCRIPTION

[0023] FIG. 1 illustrates an example of a user 105 capturing an image 110 of an identification card 115 using a mobile device 120. Briefly, and as described in more detail below, the user 120 is attempting to capture an image of sufficient quality so that an application or entity can authenticate the
15 identification card 115. To authenticate the identification card 115, the image 110 should be in focus, have a shape that approximately matches the expected shape of the identification card 115, and have a sufficiently clear border.

[0024] In the example system 100 of FIG. 1, the user 105 may use the
20 mobile device 120 to execute an application or visit a website that requests an image of the user's identification card 115. The user 105 may be unable to hold the mobile device 120 perfectly parallel to the identification card 115 for a sufficient period. Therefore, the captured image 110 may be slightly askew.

[0025] The application or website may present instructions 125 to hold the
25 mobile device 120 parallel and steady over the identification card 115. While the user 105 holds the mobile device 120 over the identification card, a camera 130 of the may capture video. The mobile device 120 may provide the video to a server. The server may analyze the video frame by frame. The server

may identify a particular frame that has an image of the identification card that is of sufficient quality for authentication purposes. The server executes a process to identify an image of the identification card 115 that will be described in more detail below with respect to FIGS. 2 and 3.

5 **[0026]** The server identifies a particular frame of the video and provides data to the mobile device 120 that the server captured an image of the identification card 115 of sufficient quality. The mobile device 120 may provide feedback to the user 105 through the display 130. For example, the mobile device 120 may provide on the display 130 text that indicates that the server identified an
10 image of the identification card 115 of sufficient quality. As another example, the mobile device 120 may display a graphical indicator that the server identified an image of the identification card 115 of sufficient quality. The graphical item may be a green border around the image of the identification card 115. With an image of the identification card 115 of sufficient quality, the
15 server may perform document authentication on the image or provide the image to another server for document authentication.

[0027] FIG. 2 illustrates an example system 200 for assessing the quality of an image of a document. Briefly, and as described in more detail below, the system 200 processes images of a document 205. In processing the images
20 of the document 205, the system 200 identifies an image that is of sufficient quality to perform document authentication on the document 205. The system 200 may be implemented on one or more computing devices. For example, the system 200 may be implemented on a mobile phone. As another example, portions of the system 200 such as the video capturer 210 may be
25 implemented on a mobile phone while the other portions may be implemented on a server.

[0028] In more detail and as illustrated in FIG. 2, a video capturer 210 captures video of a document 205. The document 205 may be any type of document. For example, the document 205 may be an identification document
30 (e.g., driver's license), a passport, a contract, or any other similar type of

document that may require authentication. The video capturer 210 may begin to capture video when the system 200 receives input from a user or when the system 200 executes an instruction to begin capturing video.

5 **[0029]** The video capturer 210 may provide the captured video to the frame sorter 215. The frame sorter 215 may be configured to divide the captured video into individual frames. The frame sorter 215 may continuously divide the captured video as the frame sorter 215 receives video. In some implementations, the frame sorter 215 may divide a video clip into individual frames.

10 **[0030]** The frame sorter 215 provides video frames to the shape detector 220. The shape detector 220 is configured to detect a border of rectangular-like shapes in each of the video frames. In some implementations, the shape detector is configured to detect borders of different shapes depending on the expected shape of the document 205. For example, the expected shape may
15 be a four-sided shape with ratios that are similar to a credit card or driver's license. The expected shape may be a four-sided shape with ratios that are similar to a passport. The expected shape may be a circular shape. The shape detector 220 may detect contrast differences between pixels to identify a likely border of the document 205.

20 **[0031]** The shape detector 220 provides data identifying the likely border of the document 205 and the corresponding image to the focus scorer 225, the expected shape scorer 230, and the border fit scorer 235. The focus scorer 225 is configured to calculate a focus score that reflects a focus level of the image. In some implementations, the focus scorer calculates a focus score for
25 the portion of the image within the likely border of the document 205. For example, the focus score may be 0.9 for a sharp image of the document 205. The focus score may be 0.5 for a less sharp image of the document 205.

[0032] The expected shape scorer 230 may be configured to calculate an expected shape score of the likely border of the document 205 compared to an

expected shape of the document 205. For example, the expected shape score may be 0.9 if the likely border has corners that are within two degrees of ninety degrees and a ratio of a long side to a short side that is within five percent of the expected ratio. The expected shape score may be 0.5 if the likely border
5 has corners that are between ten and fifteen degrees of ninety degrees and a ratio of a long side to a short side that is between fifteen and twenty percent of the expected ratio. In some implementations, the system 200 may receive input from a user or another computing device indicating the expected document. The system 200 may then retrieve an expected shape of the
10 document. For example, the system 200 may receive input indicating that the expected document is a driver's license. The system may then retrieve an expected shape and ratio of a driver's license.

[0033] The border fit scorer 235 may be configured to calculate a border fit score that reflects a likelihood that the identified border is the actual border of
15 the document 205. The border fit score may be based on a contrast between pixels located along the identified border. More pixels with higher contrast along the identified border may indicate a higher confidence that the identified border is the actual border of the document. For example, the border fit score may be 0.88 if the border fit scorer 235 determines that the identified border is
20 likely the actual order of the document 205. The border fit score may be 0.4 if the border fit scorer 235 determines that the identified border is less likely the actual border of the document 205.

[0034] The focus scorer 225, the expected shape scorer 230, and the border fit scorer 235 may provide their respective scores to the image quality scorer
25 240. The image quality scorer 240 may combine the focus score, the expected shape score, and the border fit score to calculate a quality score that reflects the overall quality of the video frame as it relates to the quality of the image of the document 205. In some implementations, the image quality scorer 240 may sum or average the focus score, the expected shape score,
30 and the border fit score. In some implementations, the image quality scorer 240 may compute a quality score by weighting the focus score, the expected

shape score, and the border fit score differently. For example, the image quality scorer 240 may weigh the expected shape score twice as much as the border fit score and the focus score. The image quality scorer 240 may weigh each score differently depending on the expected document. For example,
5 focus may be more important for a contract and expected shape may be more important for a driver's license.

[0035] The image quality scorer 240 provides the quality score, the focus score, the expected shape score, and the border fit score to the quality score comparer 245. In some implementations, the quality score comparer 245
10 compares the quality score to a quality score threshold from the threshold storage 250. If the quality score satisfies a quality score threshold, then quality score comparer 245 provides the corresponding image of the document 205 to an authentication module. For example, if the quality score is 0.85 and the quality score threshold is 0.80, then the quality score comparer 245 may
15 provide the corresponding image to an authentication module. If the quality score does not satisfy the quality score threshold, then the quality score comparer 245 continues comparing additional quality scores received from the image quality scorer 240. If the quality score comparer 245 provides an image of the document to the authentication module, then the quality score comparer
20 245 may provide a signal to the video capturer 210 to cease video capture.

[0036] In some implementations, the quality score comparer 245 may compare each of the quality score, the focus score, the expected shape score, and the border fit score to a respective threshold. If each score satisfies its respective threshold, then the quality score comparer 245 provides the image
25 of the document 205 to the authentication module. If one of the scores does not satisfy its respective threshold, then the quality score comparer 245 may continue comparing scores of additional images from the image quality scorer 240.

[0037] In some implementations, the video capturer 210 may capture a fixed
30 number of image frames of the document 205. In this instance, the system

200 may process all the frames. The quality score comparer 245 may provide the image with the highest quality score to the authentication module. The quality score compare 245 may only provide the image with the highest quality score to the authentication module if the quality score and, if necessary, the focus score, the expected shape score, and the border fit score satisfy the thresholds in the thresholds storage 250. If none of the images have a sufficient quality score, then the quality score comparer 245 may provide data indicating that the user should attempt to capture a better image of the document 205.

10 **[0038]** In some implementations, the quality score comparer 245 may provide feedback to the user that is specific to the scores that did not satisfy their respective thresholds. For example, if the focus score does not satisfy the focus score threshold, the quality score comparer 245 may provide feedback to the user to try to improve the focus. If the expected shape score and the border fit score both do not satisfy their respective thresholds, then the quality score comparer 24 may provide feedback to the user to hold the mobile device parallel to the document and to place the document 205 on a surface with a contrasting background.

[0039] FIG. 3 illustrates an example process 300 for assessing the quality of an image of a document. In general, the process 300 determines a quality score that reflects the quality of the document captured in the image. The quality score is based on the focus of the document, a confidence that the process identified the document border, and the deviation of the identified shape of the document to the expected shape of the document. The process 300 will be described as being performed by a computer system comprising one or more computers, for example, the system 100 as shown in FIG. 1 or system 200 as shown in FIG. 2.

[0040] The system received one or more images of a document (310). In some implementations, the document is an identification document (e.g., driver's license, passport, etc.). In some implementations, the document is

contract or other legal document. In some implementations, the images of the document are frames of a video

[0041] The system detects, in a first image of the one or more images, an object that has a shape that is similar to an expected shape of the document (320). In some implementations, the system detects contrast between pixels within the image. Areas where neighboring pixels have greater contrast may be areas that are more likely to correspond to edges of the document.

[0042] The system determines a focus score that reflects a sharpness of the first image (330). In some implementations, the focus score only reflects a sharpness of the portion of the image identified as the object.

[0043] The system determines a border fit score that reflects a likelihood that the object corresponds to the document (340). The border fit score may reflect a likelihood that the identified border of the object is the border of the document. In some implementations, the border fit score may indicate a higher likelihood that the identified border of the object is the border of the document if the contrast between pixels along the border is greater than other areas of the image.

[0044] The system determines an expected shape score that reflects a correlation between the shape of the object and an expected shape of the document (350). In some implementations, the correlation score may be greater if the expected shape of the document has right angle corners and the shape of the object has nearly right angle corners. In some implementations, the correlation score may be greater if the expected shape of the document has a long side to short side ratio and the same ratio of the object is near the same value.

[0045] The system, based on the focus score, the border fit score, and the expected shape score, determines a quality score that reflects a quality of the first image (360). In some implementations, the system sums or averages the

focus score, the border fit score, and the expected shape score. In some implementations, the system determines the quality score by weighing each of the focus score, the border fit score, and the expected shape score. The system may weight each score differently depending on the type of document or on the type of device or both. The type of device may also cause the system to weight the scores differently. For example, the system may weigh the border fit score differently depending on the system capturing the images with a mobile phone camera or a webcam.

[0046] The system determines whether the quality score satisfies a quality score threshold (370). In some implementations, the system determines whether the focus score satisfies a focus score threshold. In some implementations, the system determines whether the border fit score satisfies a border fit focus score. In some implementations, the system determines whether the expected shape score satisfies an expected shape score threshold.

[0047] The system, based on determining whether the quality score satisfies the quality score threshold, determines whether to perform document authentication on the first image (380). In some implementations, the system performs document authentication on the image if the quality score satisfies a quality score threshold. In some implementations, the system performs document authentication on the image if the quality score, the border fit score, the expected shape score, and the focus score each satisfy their respective thresholds. In some implementations, the system performs document authentication on the image if any combination of the quality score, the border fit score, the expected shape score, and the focus score each satisfy their respective thresholds.

[0048] In some implementations, the system performs stages 310 to 380 for each of the one or more images of the document. In this instance, the system selects the image with the highest quality score for document authentication. The system may only select the image with the highest quality score if the

quality score and any other scores satisfy their respective thresholds. In some implementations, none of the images may be of sufficient quality. In this instance, the system may not perform document authentication on any of the images. In some implementations, the system may not perform document authentication on any of the images if the system is processing images for more than a particular period of time. For example, if the system processes images for one minute, then the system may not perform document authentication on any of the images. In this instance, the system may provide feedback to the user to attempt a different configuration for capturing images of the document.

[0049] In some implementations, the thresholds may be device dependent. For example, if the system captures the images with a mobile camera phone, then the thresholds may be different than if the system captures the images with a webcam.

[0050] In more detail, the process described in this application is designed to take real time video input and detect driver's license cards in the field of view. The detected licenses are segmented from the image and returned in at in a standard rectangular format at the requested resolution along with quality metrics to determine the suitability for further processing.

[0051] The method may be customized for grayscale landscape frame buffers with height of three-hundred pixels. Any other format image will be converted to grayscale and resampled to minimum image dimension of three-hundred pixels. Any size input video frame can be used, but it should be able to support the desired resolution of the output. If it cannot meet the desired output resolution, the output will be up-sampled and the resolution quality metric will show a low score.

[0052] The method may expect a card to be presented into the middle of the video image with approximately one sixth of the area being the background border. This can be seen in FIG. 4 that illustrates a sample video capture (e.g.,

from laptop computer and of business card). The lines that are perpendicular to the edges of the figure show the recommended capture area with one sixth border and the bold border (e.g., marginal quality) highlights of the detected card.

- 5 **[0053]** Each video frame that is captured is down sampled, converted to grayscale, and edges in the vicinity of the one sixth boarder are detected to find line candidates, as shown in FIG. 5. The two edge candidates, that likely correspond to the most accurate edges, for each of the four boarders are used to construct rectangle candidates. These rectangle candidates are scored on
10 their shape to pick the most likely card candidate (e.g., large cards with aspect ratio 3.370 to 2.125).

- [0054]** In some implementations, since the method is edge based, when the card has other high contrast edges in the print or when the card is against a background of similar color, the may perform additional processing to find the
15 correct line candidates to construct the rectangle. If an accurate set of lines or if an accurate rectangle cannot be detected, the capture process will continue to the next video frame.

- [0055]** The likeliest rectangle candidate (e.g., a convex quadrilateral at this point) is chosen and then mapped to a rectangle (e.g., 3.370 x 2.125) and re-
20 rendered at the requested resolution to that shape. The rectangle fit may be only coarsely aligned to the card as illustrated in FIG. 6. In some implementations, the actions include fine tuning the alignment of the rectangle at higher resolutions.

- [0056]** The actions include analyzing the rectangular sample for quality. The
25 metrics calculated include: Focus, Exposure, Color Saturation, and Resolution. In some implementations, these are returned on a 0-100 scale (e.g., 100 indicating a highest score). These quality scores along with the rectangle confidence score are used as an overall capture score.

[0057] After one capture is complete, the software can display the detected rectangle on the video feed and color code the rectangle to reflect the capture score (e.g., red=poor, yellow=marginal, green=good).

[0058] The capture method continues and captures another frame, detects
5 the rectangle and scores the result. If the new capture is likely better than the old capture, the new result is kept. If it is likely worse, then the method continues to capture. If a likely good capture is detected after a particular amount of time (e.g., three seconds), the capture is complete. If a likely good capture is not detected after the particular amount of time (e.g., three
10 seconds), the previous likely best result is erased and another capture attempt is started. After a specified number of capture attempts (e.g., nine seconds worth) the method determine that the capture has not completed successfully.

[0059] In some implementations, if a likely good capture is detected in the first particular amount of time (e.g. three seconds), it may be beneficial to store
15 that result but take another attempt (e.g., a three second attempt) to determine if the result is better. After acquiring a likely good quality image, the rectangular output can be used for further processing such as fine tuning alignment, optical character recognition, driver's license state detection, color (white balance) correction, illumination correction, etc. In some
20 implementations, a device, such as a laptop computer executes the method at a rate of ten millisecond per frame and may run at near real time on a mobile device.

[0060] FIG. 7 shows an example of a computing device 700 and a mobile computing device 750 that can be used to implement the techniques described
25 here. The computing device 700 is intended to represent various forms of digital computers, such as laptops, desktops, workstations, personal digital assistants, servers, blade servers, mainframes, and other appropriate computers. The mobile computing device 750 is intended to represent various forms of mobile devices, such as personal digital assistants, cellular
30 telephones, smart-phones, and other similar computing devices. The

components shown here, their connections and relationships, and their functions, are meant to be examples only, and are not meant to be limiting.

[0061] The computing device 700 includes a processor 702, a memory 704, a storage device 706, a high-speed interface 708 connecting to the memory 704 and multiple high-speed expansion ports 710, and a low-speed interface 712 connecting to a low-speed expansion port 714 and the storage device 706. Each of the processor 702, the memory 704, the storage device 706, the high-speed interface 708, the high-speed expansion ports 710, and the low-speed interface 712, are interconnected using various busses, and may be mounted on a common motherboard or in other manners as appropriate. The processor 702 can process instructions for execution within the computing device 700, including instructions stored in the memory 704 or on the storage device 706 to display graphical information for a GUI on an external input/output device, such as a display 716 coupled to the high-speed interface 708. In other implementations, multiple processors and/or multiple buses may be used, as appropriate, along with multiple memories and types of memory. Also, multiple computing devices may be connected, with each device providing portions of the necessary operations (e.g., as a server bank, a group of blade servers, or a multi-processor system).

[0062] The memory 704 stores information within the computing device 700. In some implementations, the memory 704 is a volatile memory unit or units. In some implementations, the memory 704 is a non-volatile memory unit or units. The memory 704 may also be another form of computer-readable medium, such as a magnetic or optical disk.

[0063] The storage device 706 is capable of providing mass storage for the computing device 700. In some implementations, the storage device 706 may be or contain a computer-readable medium, such as a floppy disk device, a hard disk device, an optical disk device, or a tape device, a flash memory or other similar solid state memory device, or an array of devices, including devices in a storage area network or other configurations. Instructions can be

stored in an information carrier. The instructions, when executed by one or more processing devices (for example, processor 702), perform one or more methods, such as those described above. The instructions can also be stored by one or more storage devices such as computer- or machine-readable
5 mediums (for example, the memory 704, the storage device 706, or memory on the processor 702).

[0064] The high-speed interface 708 manages bandwidth-intensive operations for the computing device 700, while the low-speed interface 712 manages lower bandwidth-intensive operations. Such allocation of functions is
10 an example only. In some implementations, the high-speed interface 708 is coupled to the memory 704, the display 716 (e.g., through a graphics processor or accelerator), and to the high-speed expansion ports 710, which may accept various expansion cards. In the implementation, the low-speed interface 712 is coupled to the storage device 706 and the low-speed
15 expansion port 714. The low-speed expansion port 714, which may include various communication ports (e.g., USB, Bluetooth, Ethernet, wireless Ethernet) may be coupled to one or more input/output devices, such as a keyboard, a pointing device, a scanner, or a networking device such as a switch or router, e.g., through a network adapter.

[0065] The computing device 700 may be implemented in a number of different forms, as shown in the figure. For example, it may be implemented as a standard server 720, or multiple times in a group of such servers. In addition, it may be implemented in a personal computer such as a laptop computer 722. It may also be implemented as part of a rack server system
25 724. Alternatively, components from the computing device 700 may be combined with other components in a mobile device, such as a mobile computing device 750. Each of such devices may contain one or more of the computing device 700 and the mobile computing device 750, and an entire system may be made up of multiple computing devices communicating with
30 each other.

[0066] The mobile computing device 750 includes a processor 752, a memory 764, an input/output device such as a display 754, a communication interface 766, and a transceiver 768, among other components. The mobile computing device 750 may also be provided with a storage device, such as a micro-drive or other device, to provide additional storage. Each of the processor 752, the memory 764, the display 754, the communication interface 766, and the transceiver 768, are interconnected using various buses, and several of the components may be mounted on a common motherboard or in other manners as appropriate.

[0067] The processor 752 can execute instructions within the mobile computing device 750, including instructions stored in the memory 764. The processor 752 may be implemented as a chipset of chips that include separate and multiple analog and digital processors. The processor 752 may provide, for example, for coordination of the other components of the mobile computing device 750, such as control of user interfaces, applications run by the mobile computing device 750, and wireless communication by the mobile computing device 750.

[0068] The processor 752 may communicate with a user through a control interface 758 and a display interface 756 coupled to the display 754. The display 754 may be, for example, a TFT (Thin-Film-Transistor Liquid Crystal Display) display or an OLED (Organic Light Emitting Diode) display, or other appropriate display technology. The display interface 756 may comprise appropriate circuitry for driving the display 754 to present graphical and other information to a user. The control interface 758 may receive commands from a user and convert them for submission to the processor 752. In addition, an external interface 762 may provide communication with the processor 752, so as to enable near area communication of the mobile computing device 750 with other devices. The external interface 762 may provide, for example, for wired communication in some implementations, or for wireless communication in other implementations, and multiple interfaces may also be used.

[0069] The memory 764 stores information within the mobile computing device 750. The memory 764 can be implemented as one or more of a computer-readable medium or media, a volatile memory unit or units, or a non-volatile memory unit or units. An expansion memory 774 may also be
5 provided and connected to the mobile computing device 750 through an expansion interface 772, which may include, for example, a SIMM (Single In Line Memory Module) card interface. The expansion memory 774 may provide extra storage space for the mobile computing device 750, or may also store applications or other information for the mobile computing device 750.
10 Specifically, the expansion memory 774 may include instructions to carry out or supplement the processes described above, and may include secure information also. Thus, for example, the expansion memory 774 may be provide as a security module for the mobile computing device 750, and may be programmed with instructions that permit secure use of the mobile computing
15 device 750. In addition, secure applications may be provided via the SIMM cards, along with additional information, such as placing identifying information on the SIMM card in a non-hackable manner.

[0070] The memory may include, for example, flash memory and/or NVRAM memory (non-volatile random access memory), as discussed below. In some
20 implementations, instructions are stored in an information carrier. that the instructions, when executed by one or more processing devices (for example, processor 752), perform one or more methods, such as those described above. The instructions can also be stored by one or more storage devices, such as one or more computer- or machine-readable mediums (for example,
25 the memory 764, the expansion memory 774, or memory on the processor 752). In some implementations, the instructions can be received in a propagated signal, for example, over the transceiver 768 or the external interface 762.

[0071] The mobile computing device 750 may communicate wirelessly
30 through the communication interface 766, which may include digital signal processing circuitry where necessary. The communication interface 766 may

provide for communications under various modes or protocols, such as GSM voice calls (Global System for Mobile communications), SMS (Short Message Service), EMS (Enhanced Messaging Service), or MMS messaging (Multimedia Messaging Service), CDMA (code division multiple access),
5 TDMA (time division multiple access), PDC (Personal Digital Cellular), WCDMA (Wideband Code Division Multiple Access), CDMA2000, or GPRS (General Packet Radio Service), among others. Such communication may occur, for example, through the transceiver 768 using a radio-frequency. In addition, short-range communication may occur, such as using a Bluetooth,
10 WiFi, or other such transceiver. In addition, a GPS (Global Positioning System) receiver module 770 may provide additional navigation- and location-related wireless data to the mobile computing device 750, which may be used as appropriate by applications running on the mobile computing device 750.

[0072] The mobile computing device 750 may also communicate audibly
15 using an audio codec 760, which may receive spoken information from a user and convert it to usable digital information. The audio codec 760 may likewise generate audible sound for a user, such as through a speaker, e.g., in a handset of the mobile computing device 750. Such sound may include sound from voice telephone calls, may include recorded sound (e.g., voice
20 messages, music files, etc.) and may also include sound generated by applications operating on the mobile computing device 750.

[0073] The mobile computing device 750 may be implemented in a number of different forms, as shown in the figure. For example, it may be implemented as a cellular telephone 780. It may also be implemented as part of a smart-
25 phone 582, personal digital assistant, or other similar mobile device.

[0074] Various implementations of the systems and techniques described here can be realized in digital electronic circuitry, integrated circuitry, specially designed ASICs (application specific integrated circuits), computer hardware, firmware, software, and/or combinations thereof. These various
30 implementations can include implementation in one or more computer

programs that are executable and/or interpretable on a programmable system including at least one programmable processor, which may be special or general purpose, coupled to receive data and instructions from, and to transmit data and instructions to, a storage system, at least one input device, and at
5 least one output device.

[0075] These computer programs (also known as programs, software, software applications or code) include machine instructions for a programmable processor, and can be implemented in a high-level procedural and/or object-oriented programming language, and/or in assembly/machine
10 language. As used herein, the terms machine-readable medium and computer-readable medium refer to any computer program product, apparatus and/or device (e.g., magnetic discs, optical disks, memory, Programmable Logic Devices (PLDs)) used to provide machine instructions and/or data to a programmable processor, including a machine-readable medium that receives
15 machine instructions as a machine-readable signal. The term machine-readable signal refers to any signal used to provide machine instructions and/or data to a programmable processor.

[0076] To provide for interaction with a user, the systems and techniques described here can be implemented on a computer having a display device
20 (e.g., a CRT (cathode ray tube) or LCD (liquid crystal display) monitor) for displaying information to the user and a keyboard and a pointing device (e.g., a mouse or a trackball) by which the user can provide input to the computer. Other kinds of devices can be used to provide for interaction with a user as well; for example, feedback provided to the user can be any form of sensory
25 feedback (e.g., visual feedback, auditory feedback, or tactile feedback); and input from the user can be received in any form, including acoustic, speech, or tactile input.

[0077] The systems and techniques described here can be implemented in a computing system that includes a back end component (e.g., as a data
30 server), or that includes a middleware component (e.g., an application server),

or that includes a front end component (e.g., a client computer having a graphical user interface or a Web browser through which a user can interact with an implementation of the systems and techniques described here), or any combination of such back end, middleware, or front end components. The components of the system can be interconnected by any form or medium of digital data communication (e.g., a communication network). Examples of communication networks include a local area network (LAN), a wide area network (WAN), and the Internet.

[0078] The computing system can include clients and servers. A client and server are generally remote from each other and typically interact through a communication network. The relationship of client and server arises by virtue of computer programs running on the respective computers and having a client-server relationship to each other.

[0079] Although a few implementations have been described in detail above, other modifications are possible. For example, while a client application is described as accessing the delegate(s), in other implementations the delegate(s) may be employed by other applications implemented by one or more processors, such as an application executing on one or more servers. In addition, the logic flows depicted in the figures do not require the particular order shown, or sequential order, to achieve desirable results. In addition, other actions may be provided, or actions may be eliminated, from the described flows, and other components may be added to, or removed from, the described systems. Accordingly, other implementations are within the scope of the following claims.

CLAIMS

1. A computer-implemented method comprising:
 - receiving one or more images of a document;
 - 5 detecting, in a first image of the one or more images, an object that has a shape that is similar to an expected shape of the document;
 - determining a focus score that reflects a sharpness of the first image;
 - determining a border fit score that reflects a likelihood that the object corresponds to the document;
 - 10 determining an expected shape score that reflects a correlation between the shape of the object and an expected shape of the document;
 - based on the focus score, the border fit score, and the expected shape score, determining a quality score that reflects a quality of the first image;
 - determining whether the quality score satisfies a quality score threshold;
 - 15 and
 - based on determining whether the quality score satisfies the quality score threshold, determining whether to perform document authentication on the first image.
- 20 2. The method of claim 1, wherein:
 - determining whether the quality score satisfies a quality score threshold comprises determining that the quality score satisfies the quality score threshold, and
 - determining whether to perform document authentication on the first
 - 25 image comprises determining to perform document authentication on the first image based on determining that the quality score satisfies the quality score threshold.
3. The method of claim 1, wherein:
 - 30 determining whether the quality score satisfies a quality score threshold comprises determining that the quality score does not satisfy the quality score threshold, and

determining whether to perform document authentication on the first image comprises determining not to perform document authentication on the first image based on determining that the quality score does not satisfy the quality score threshold.

5

4. The method of claim 3, comprising:
in response to determining not to perform document authentication on the first image, determining a quality score for a second image of the one or more images.

10

5. The method of claim 1, comprising:
determining whether the focus score satisfies a focus score threshold,
wherein determining whether to perform document authentication on the first image is further based on determining whether the focus score satisfies
15 the focus score threshold.

6. The method of claim 1, comprising:
determining whether the border fit score satisfies a border fit score threshold,
20 wherein determining whether to perform document authentication on the first image is further based on determining whether the border fit score satisfies the border fit score threshold.

7. The method of claim 1, comprising:
25 determining whether the expected shape score satisfies an expected shape score threshold,
wherein determining whether to perform document authentication on the first image is further based on determining whether the expected shape score satisfies the expected shape score threshold.

30

8. The method of claim 1, wherein the focus score reflects a sharpness of only the object.

9. The method of claim 1, wherein the document is an identification document.
10. The method of claim 1, wherein the one or more images of the document are frames of a video.
11. The method of claim 1, comprising:
determining, for each of the one or more images of the document, a quality score; and
10 determining that the quality score of the first image is greater than the quality scores of each of the one or more images other than the first image, wherein determining whether the quality score of the first image satisfies a quality score threshold comprises determining that the quality score of the first image satisfies the quality score threshold, and
15 wherein determining whether to perform document authentication on the first image comprises determining to perform document authentication on the first image based on determining that the quality score of the first image satisfies the quality score threshold and based on determining that the quality score of the first image is greater than the quality scores of each of the one or
20 more images other than the first image.
12. The method of claim 1, wherein:
determining whether the quality score satisfies a quality score threshold comprises determining that the quality score does not satisfy the quality score
25 threshold,
based on determining that the quality score does not satisfy the quality score threshold the method further comprises providing, to the user, feedback for improving quality of an image of the document.
- 30 13. A system comprising:
one or more computers; and
one or more storage devices storing instructions that are operable,

when executed by the one or more computers, to cause the one or more computers to perform operations comprising:

receiving one or more images of a document;

detecting, in a first image of the one or more images, an object

5 that has a shape that is similar to an expected shape of the document;

determining a focus score that reflects a sharpness of the first image;

determining a border fit score that reflects a likelihood that the object corresponds to the document;

10 determining an expected shape score that reflects a correlation between the shape of the object and an expected shape of the document;

based on the focus score, the border fit score, and the expected shape score, determining a quality score that reflects a quality of the first image;

15 determining whether the quality score satisfies a quality score threshold; and

based on determining whether the quality score satisfies the quality score threshold, determining whether to perform document authentication on the first image.

20

14. The system of claim 13, wherein:

determining whether the quality score satisfies a quality score threshold comprises determining that the quality score satisfies the quality score threshold, and

25 determining whether to perform document authentication on the first image comprises determining to perform document authentication on the first image based on determining that the quality score satisfies the quality score threshold.

30 15. The system of claim 13, wherein:

determining whether the quality score satisfies a quality score threshold comprises determining that the quality score does not satisfy the quality score threshold, and

determining whether to perform document authentication on the first image comprises determining not to perform document authentication on the first image based on determining that the quality score does not satisfy the quality score threshold.

5

16. The system of claim 13, wherein the operations further comprise:
determining whether the focus score satisfies a focus score threshold,
wherein determining whether to perform document authentication on the first image is further based on determining whether the focus score satisfies
10 the focus score threshold.

17. The system of claim 13, wherein the operations further comprise:
determining whether the border fit score satisfies a border fit score
threshold,

15 wherein determining whether to perform document authentication on the first image is further based on determining whether the border fit score satisfies the border fit score threshold.

18. The system of claim 13, wherein the operations further comprise:
20 determining whether the expected shape score satisfies an expected shape score threshold,
wherein determining whether to perform document authentication on the first image is further based on determining whether the expected shape score satisfies the expected shape score threshold.

25

19. The system of claim 13, wherein the operations further comprise:
determining, for each of the one or more images of the document, a quality score; and
determining that the quality score of the first image is greater than the
30 quality scores of each of the one or more images other than the first image,
wherein determining whether the quality score of the first image satisfies a quality score threshold comprises determining that the quality score of the first image satisfies the quality score threshold, and

wherein determining whether to perform document authentication on the first image comprises determining to perform document authentication on the first image based on determining that the quality score of the first image satisfies the quality score threshold and based on determining that the quality score of the first image is greater than the quality scores of each of the one or more images other than the first image.

20. A non-transitory computer-readable medium storing software comprising instructions executable by one or more computers which, upon such execution, cause the one or more computers to perform operations comprising:

- receiving one or more images of a document;
- detecting, in a first image of the one or more images, an object that has a shape that is similar to an expected shape of the document;
- 15 determining a focus score that reflects a sharpness of the first image;
- determining a border fit score that reflects a likelihood that the object corresponds to the document;
- determining an expected shape score that reflects a correlation between the shape of the object and an expected shape of the document;
- 20 based on the focus score, the border fit score, and the expected shape score, determining a quality score that reflects a quality of the first image;
- determining whether the quality score satisfies a quality score threshold;
- and
- based on determining whether the quality score satisfies the quality score threshold, determining whether to perform document authentication on the first image.

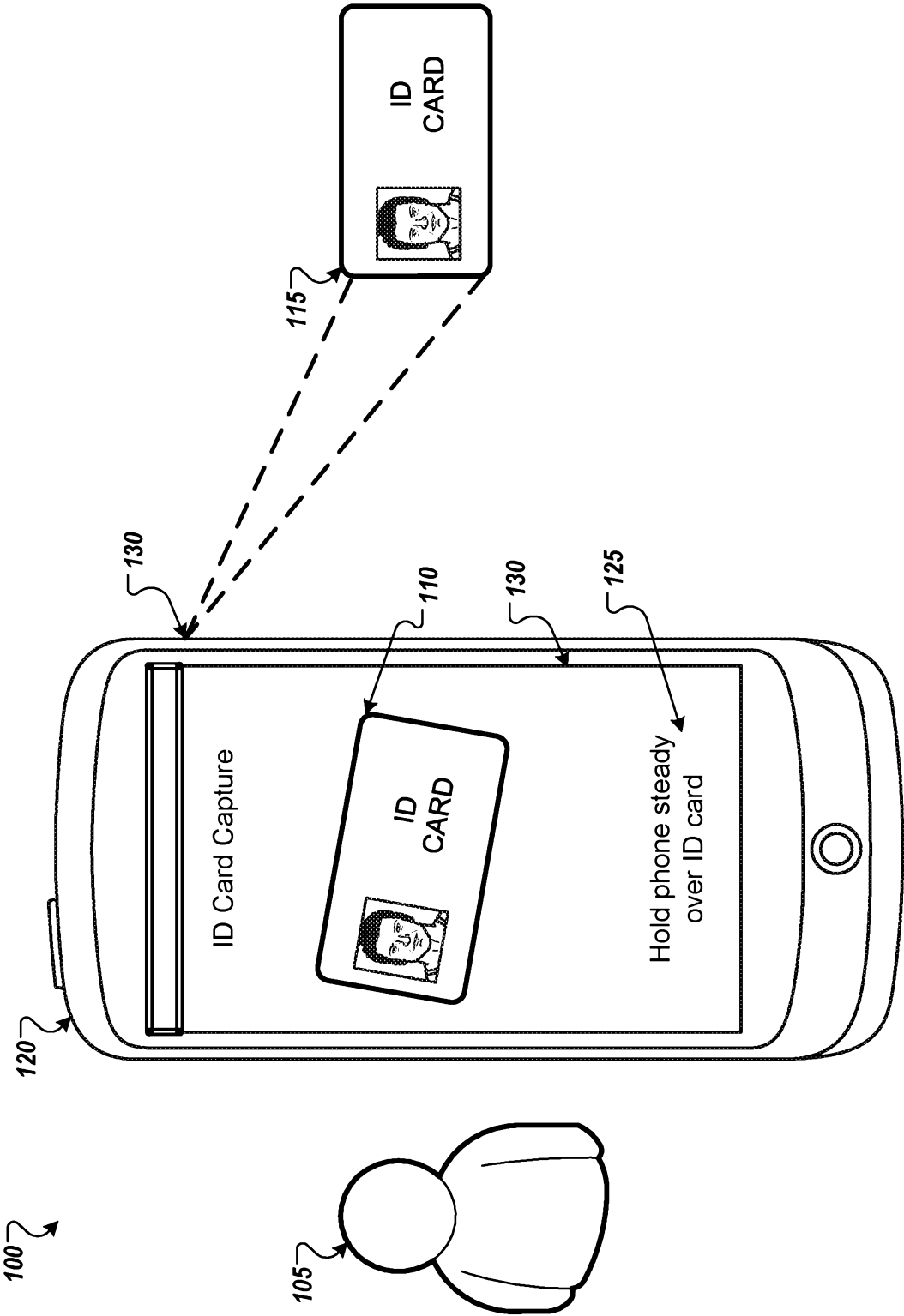


FIG. 1

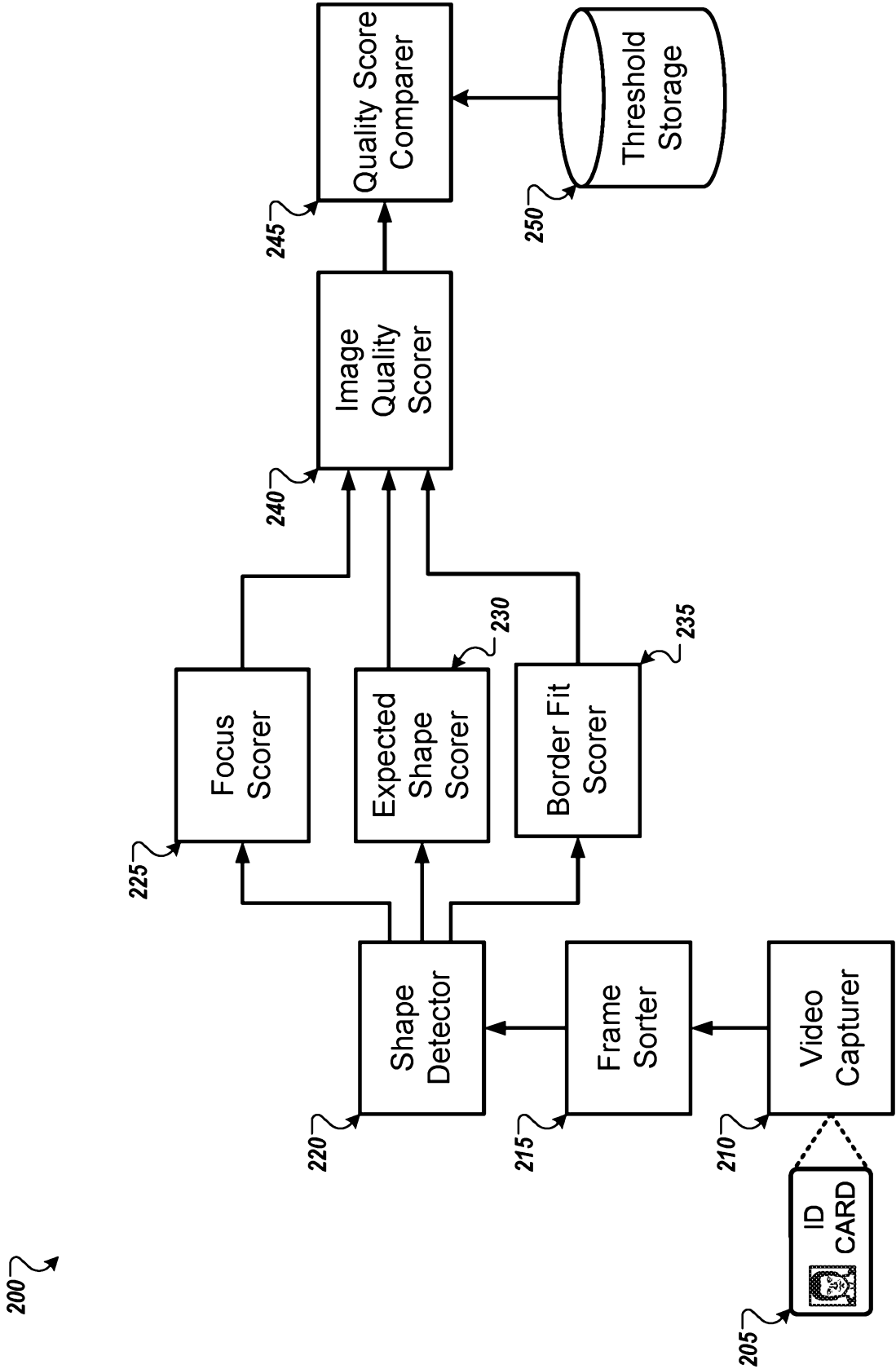


FIG. 2

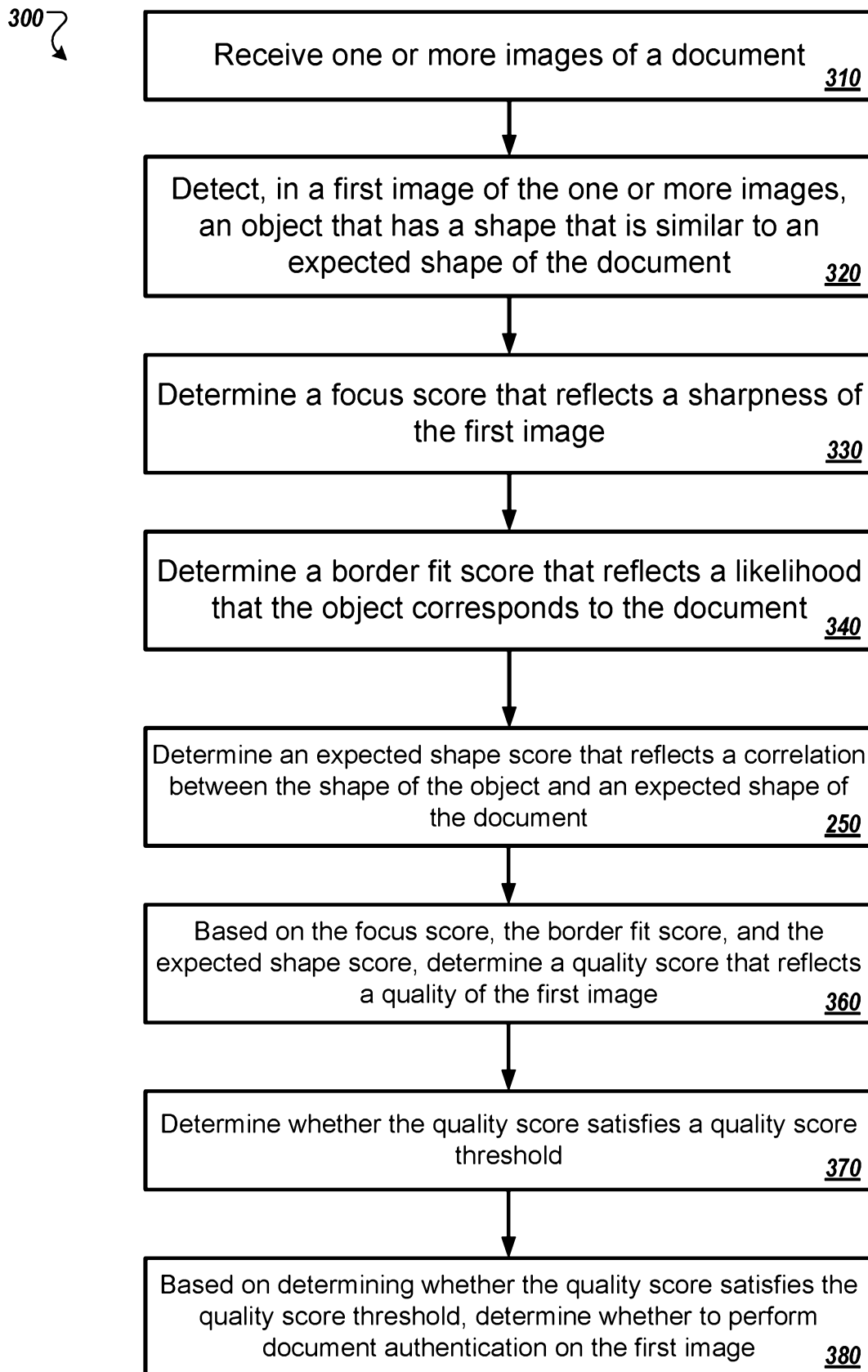


FIG. 3

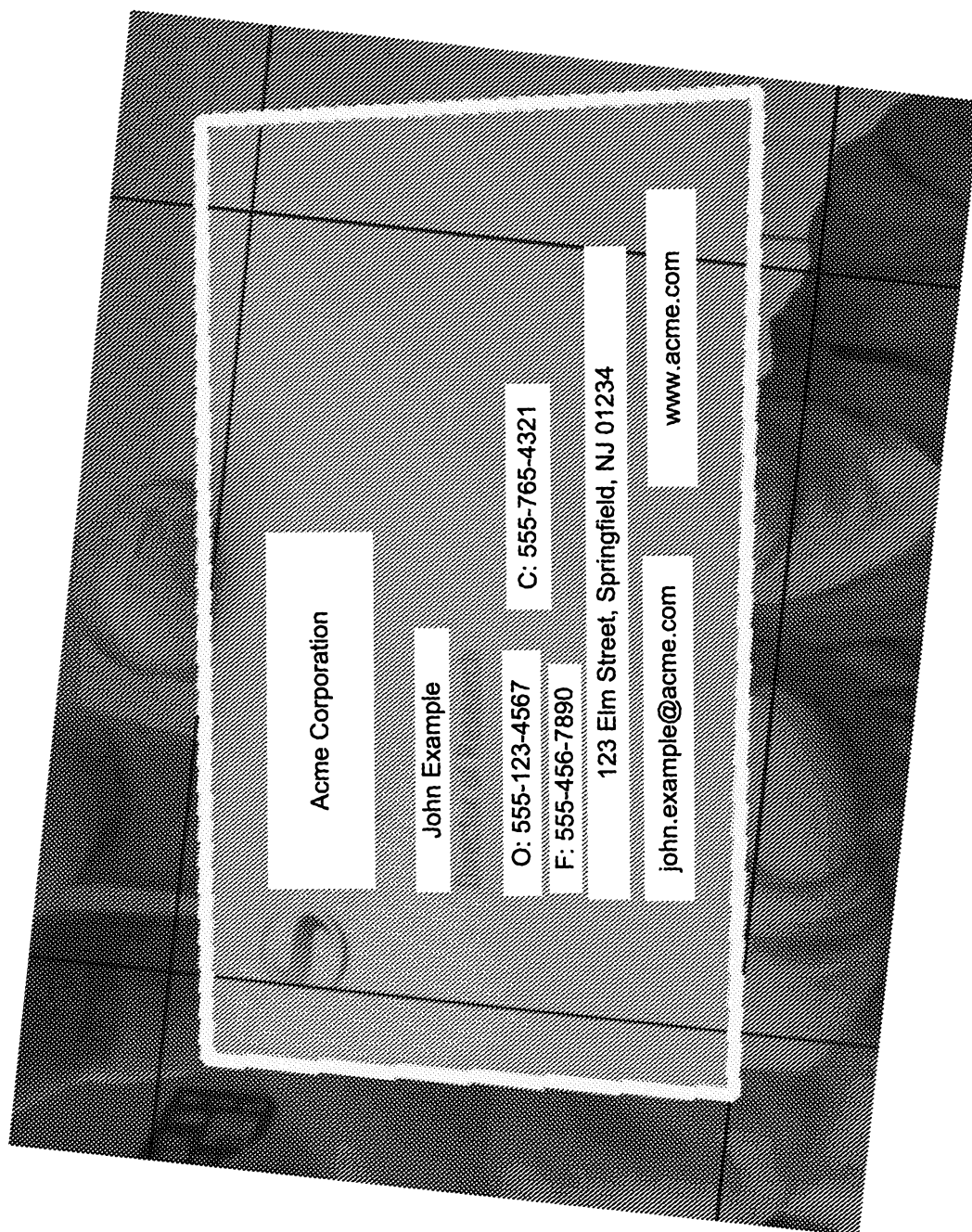


FIG. 4

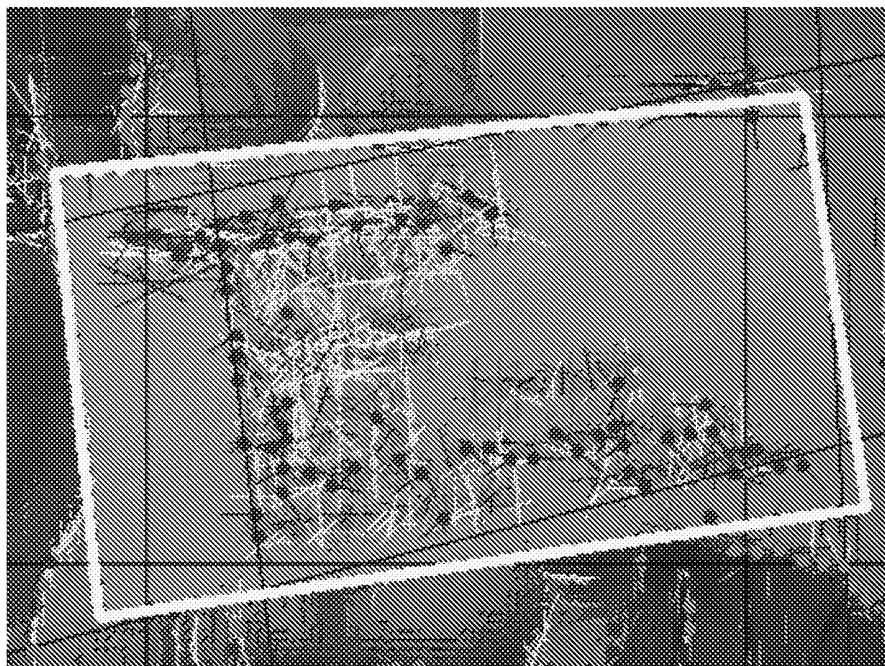
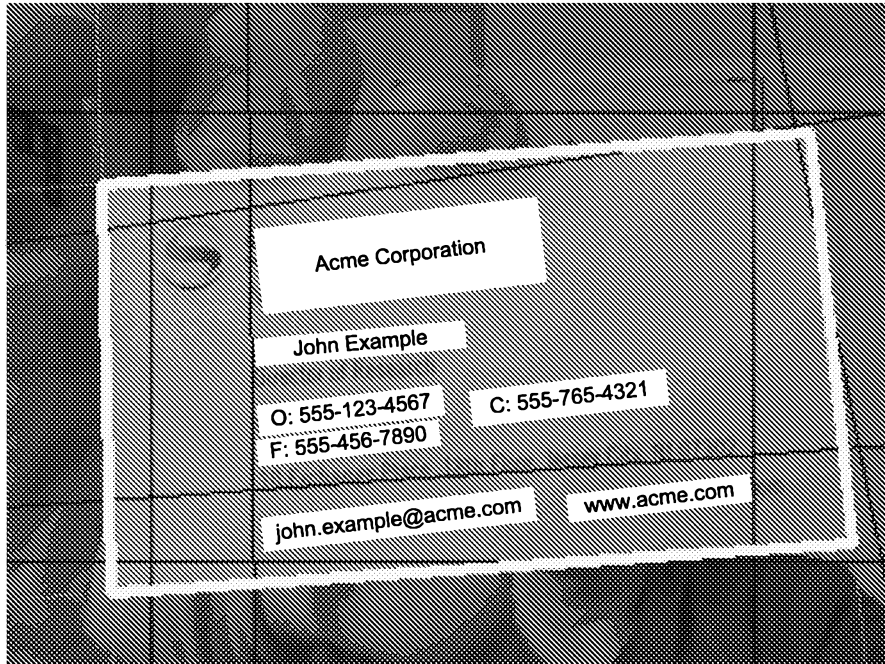


FIG. 5

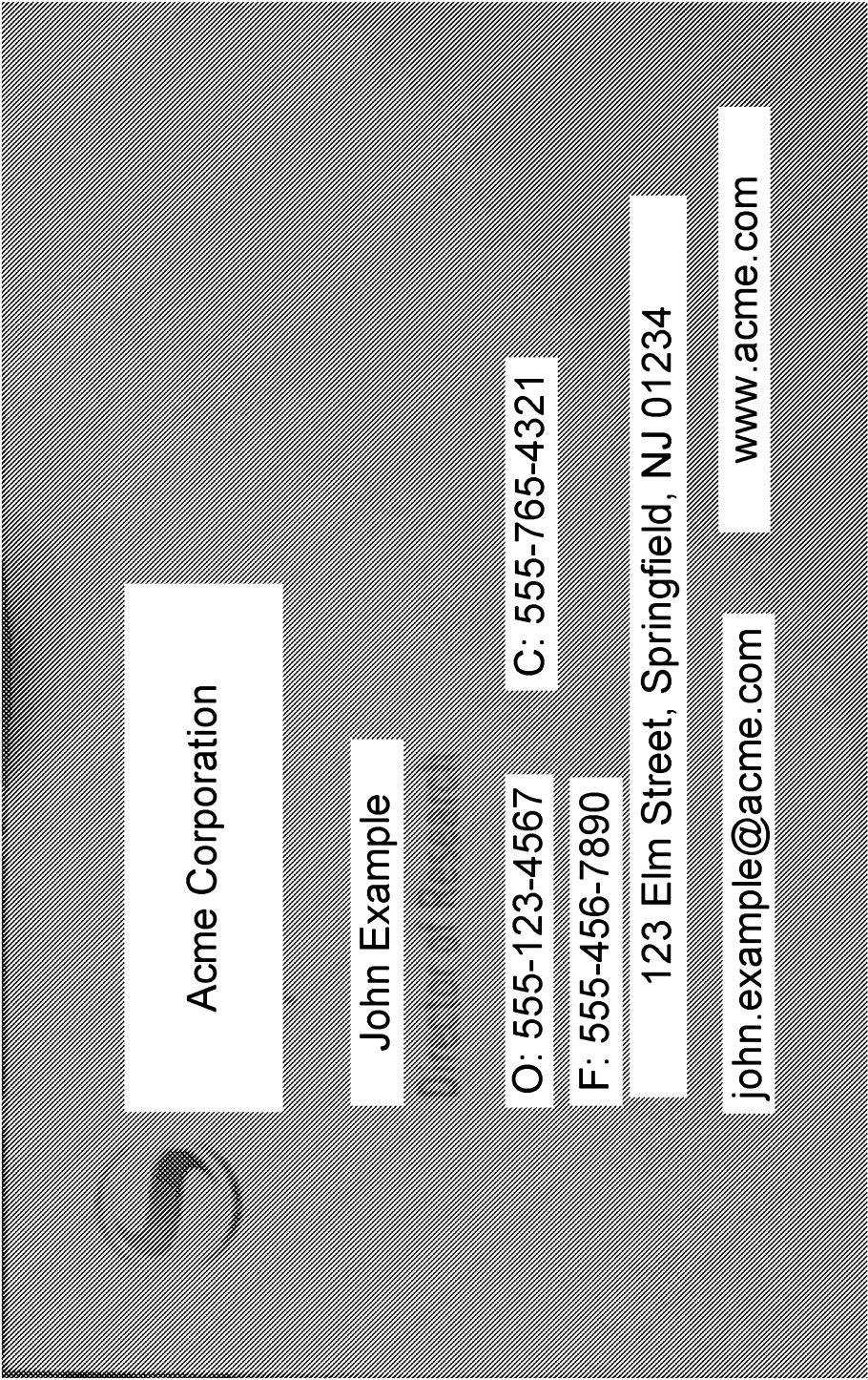


FIG. 6

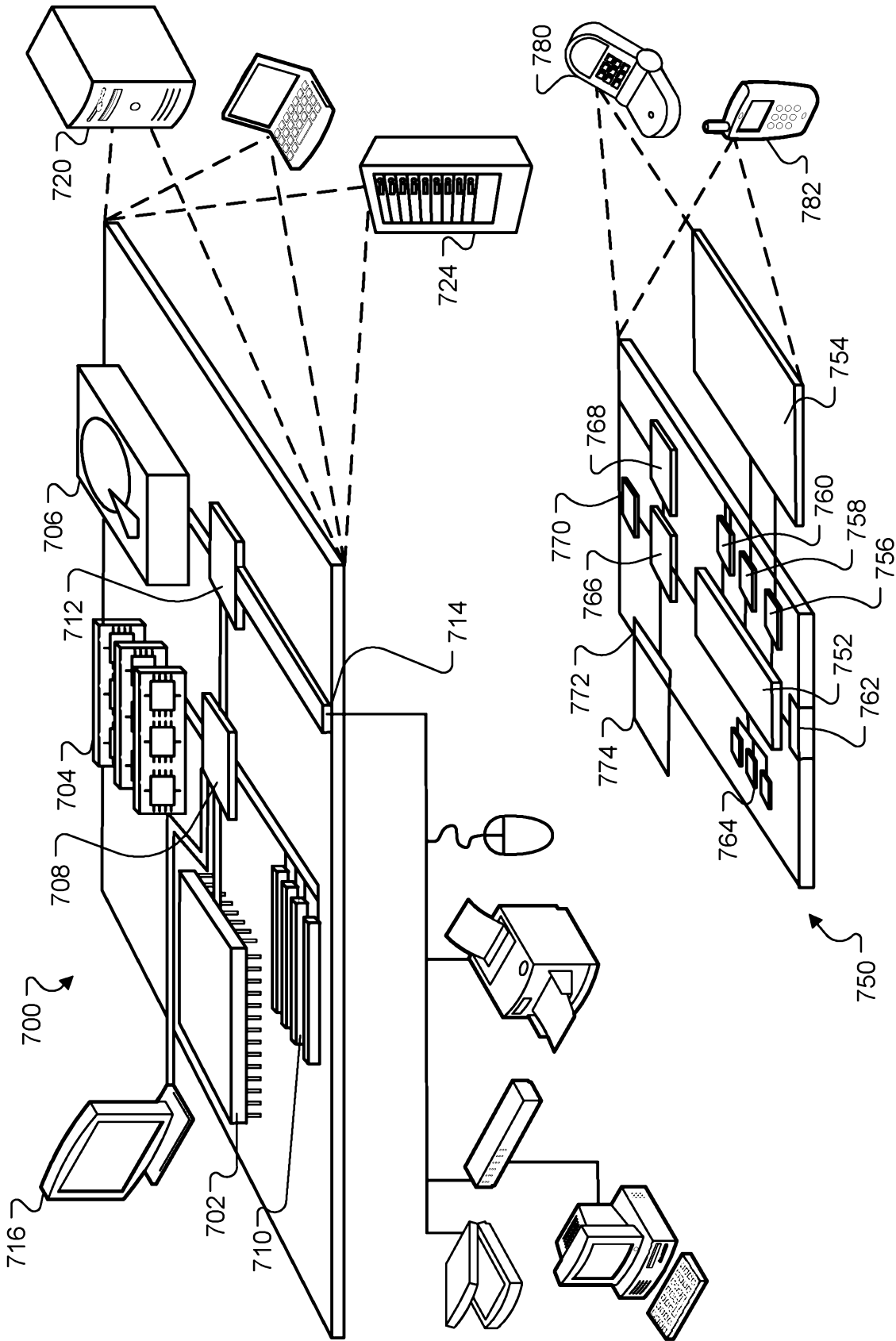


FIG. 7