November 2019

Enhancing Image Quality of Photographs Taken by Portable Devices by Matching Images to High Quality Reference Images Using Machine Learning and Camera Orientation and Other Image Metadata

Yen-Cheng Chen

Enhancing Image Quality of Photographs Taken by Portable Devices by Matching Images to High Quality Reference Images Using Machine Learning and Camera Orientation and Other Image Metadata

Abstract:

This publication describes systems and techniques to enhance an image quality of a photograph taken by a portable device using a collection of images to find a replacement object. Although the intrinsic image quality of photographs taken by portable devices, such as a smartphone, continues to improve, many pictures still lack detail or have blurriness. Consequently, a photograph, or "original image," can include an object having a poor quality, such as blurriness or an inaccurate reproduction of a texture. To enhance the original image, a machine-learned model detects in the original image the object having poor quality. A collection of images is searched to find a reference image having another version of the object with a superior quality. Metadata for the images can be used to find a matching reference image. Metadata can include positioning, distance to subject, tilt angle, and so forth. An object enhancement module replaces at least a portion of the poor-quality object in the original image with the superior-quality object from the reference image to produce an enhanced image. The image collection with reference images can be stored in the cloud or locally on the portable device. In these manners, the image quality of a photograph taken by a portable device can be enhanced.

Keywords:

portable device, smartphone, camera, photograph, image, image quality, blurry, focus, image collection, replace, substitute, object, building, face, machine learning (ML), neural network, camera position, camera orientation, image metadata

Background:

Many types of portable devices include a camera or image sensor utilized to take a photograph, which is referred to herein generally as an "image." Portable devices include, for example, phones (e.g., smartphones), tablet computers, laptops, convertibles, personal digital assistants (PDAs), smart watches, intelligent glasses, cameras, and so forth. These portable devices are being manufactured with larger sensors, multiple lenses, and image-processing functionality configured to produce images with improved intrinsic quality. Nonetheless, many images still suffer from poor quality, partially because users continue to stretch photographic boundaries. For example, users take pictures in low light, include subjects at different depths of field, and permit movement of the camera sensor. The photographic target may also be moving if the subject is alive.

Consequently, an original image can include an object having a poor quality, such as blurriness, a lack of detail, or an inaccurate reproduction of a texture. Objects include buildings, faces, and so forth. Some issues with poor image quality may be surmountable with more-advanced camera hardware. Other issues with poor image quality may be addressed with longer exposure times or by ensuring that there is no relevant movement between the camera and the photographic subject. Unfortunately, the former, hardware-based approach increases the cost of the portable device. And the latter, human-based approaches inconvenience the user. Therefore, it is desirable to enhance the image quality of photographs taken by portable devices without necessarily requiring upgrades to camera hardware or without requiring users to remain motionless and be limited to still subjects.

Description:

This publication describes systems and techniques for enhancing a photograph taken with a portable device by using a collection of images. The portable device captures an original image in accordance with one or more capturing parameters. Capturing parameters can include location, distance, view angle, time, weather, and so forth. The capturing parameters can be stored in association with the original image as image metadata. A machine-learned (ML) model that is associated with an object enhancement module recognizes one or more objects in the original image. The object enhancement module searches an image collection using the image metadata to find a matching object in a reference image of the image collection. The reference image includes an object that matches an object of the original image in terms of the capture parameters. If a quality of the object in the reference image is superior to a quality of the object in the original image, the object enhancement module performs an object enhancement operation to replace at least a portion of the object in the original image to produce an enhanced image using the matching object of the reference image.

As shown in Fig. 1 (below), an example environment 100 includes an original image 102, an image collection 106, and an object enhancement module 114. The environment 100 further includes at least one reference image 108 and, after an object enhancement operation 110, an enhanced image 112. The original image 102 includes multiple objects 104, such as four objects 104-1, 104-2 104-3, and 104-4. Examples of objects 104 include buildings, natural structures like mountains, and faces of people. The object 104-1b has a poor quality, such as blurriness or a poor reproduction of a texture. The original image 102 is associated with metadata 118, which includes one or more capturing parameters. Examples of metadata 118 are described below with reference to Fig. 2.

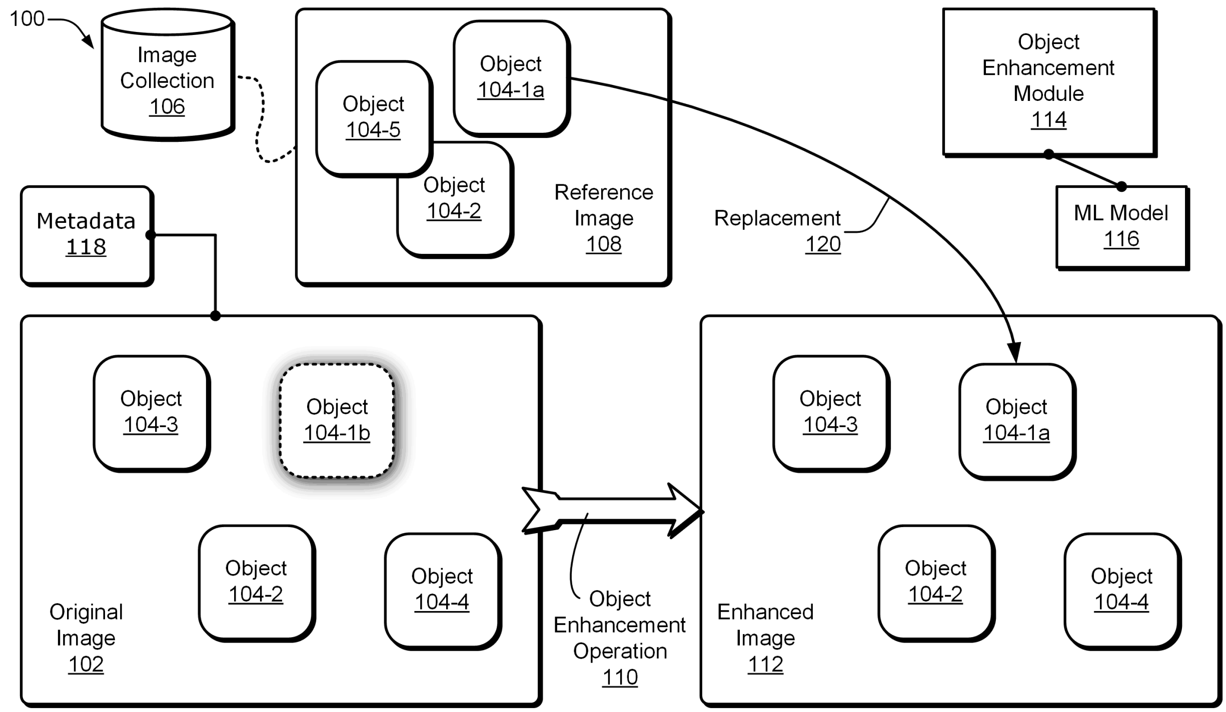

Fig. 1. Environment for enhancing an image using an image collection.

The image collection 106 includes multiple reference images 108, one of which is depicted in Fig. 1. Each reference image 108 includes multiple objects, such as the objects 104-1a, 104-2, and 104-5. In this example, the object 104-1a corresponds to the object 104-1b, but the object 104-1b has an inferior image quality as compared to that of the object 104-1a. To perform an object enhancement operation 110, the object enhancement module 114 analyzes the original image 102 and attempts to replace portions thereof that have a poor image quality. As shown, the object enhancement module 114 includes or is otherwise associated with an ML model 116. The ML model 116 can be realized using, for example, a standard neural-network-based model with corresponding layers to process input features extracted from the images. The ML model may be implemented as a support vector machine (SVM), a recurrent neural network (RNN), a convolutional neural network (CNN), a dense neural network (DNN), one or more heuristics, other machine-learning techniques, a combination thereof, and so forth.

The ML model 116 analyzes the original image 102 to recognize an object 104. If an object is recognized, the ML model 116 can also determine if an image quality thereof is poor. If so, the object enhancement module 114 searches the image collection 106 to find a reference image 108 with a corresponding object 104 with one or more matching capture parameters based on the metadata 118. If an object 104-1 is found that matches the object 104-1b, the object enhancement module 114 determines if an image quality of the matching object 104-1a in the reference image 108 is superior to that of the object 104-1b in the original image 102. In this case, an image quality of the object 104-1a is determined to be superior to that of the object 104-1b in one or more aspects, such as by having clearer edges. To perform the object enhancement operation 110, the object enhancement module 114 performs a replacement 120 of at least one portion of the object 104-1b of the original image 102 with a corresponding, superior portion of the object 104-1a of the reference image 108 to produce the enhanced image 112. The portion that is replaced may be an outline, a texture, a percentage, a shadowed area, a color, and so forth of the object 104-1. An image quality of the object 104-1 is therefore improved in the enhanced image 112 as compared to an image quality of the object 104-1 in the original image 102.

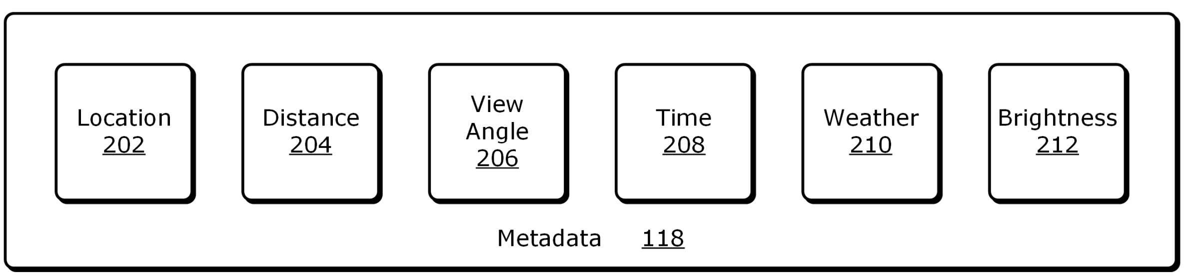

Fig. 2. Metadata of an image being enhanced or referenced.

Fig. 2 depicts an example of the metadata 118 that is associated with an original image 102 and/or a reference image 108. The metadata 118 includes a location 202, a distance 204, a view angle 206, a time 208, weather 210, and brightness 212. Five example types of metadata 118 are depicted, but the metadata 118 for any given image may include more, fewer, or different types of data. The location 202 includes a position of the portable device that took the photograph. The location can be obtained via a global-positioning system (GPS) chip or other hardware that interfaces with a global navigation satellite system (GNSS) (e.g., Galileo). The distance 204 indicates a distance between the portable device and the subject. The distance can be obtained using, for instance, radar or infrared (IR) signaling.

The view angle 206 provides a tilt angle of the portable device during image capture and can be determined using at least one component of an inertial measurement unit (IMU). An IMU can include at least one accelerometer, at least one gyroscope, at least one magnetometer, and so forth. The time 208 reflects a time of day and/or time of year at which the photograph was taken. The weather 210 memorializes atmospheric aspects that can affect a captured image and/or the appearance of a subject thereof, such as cloudiness, fog, an amount of sunshine, a humidity level, and so forth. The weather 210 can be stored as part of the metadata 118 after retrieving weather data from a weather app, a website, or an operating system (OS) application programming interface (API). The brightness 212 indicates a brightness of a photographic subject of the picture.

Fig. 3 illustrates an example public structure enhancement operation 110-1. Public structures can include man-made structures (e.g., a building, bridge, or artistic installation) and natural structures (e.g., a mountain, lake, or landscape). With public structures, the image collection 106 can include a database of photos that are amassed remotely in a server farm from various sources and multiple photographers. In such cases, the object enhancement module 114 can be implemented at least partially using one or more machines of a server farm.

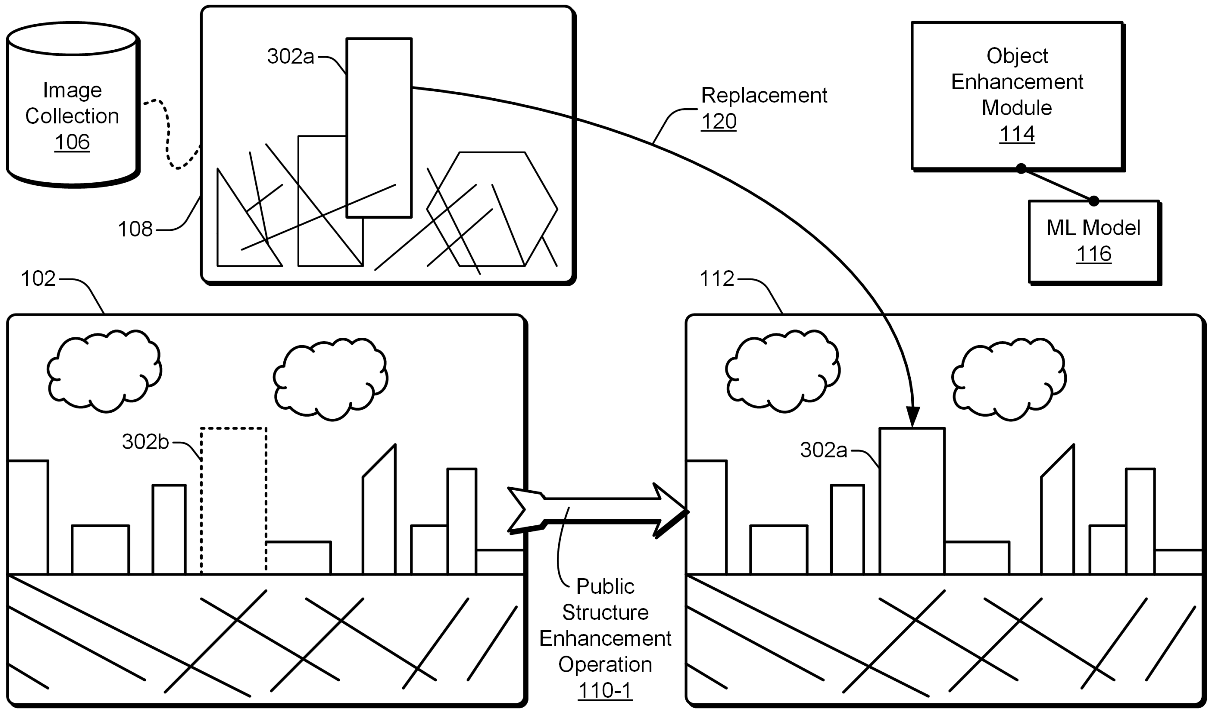

Fig. 3. Image enhancement with public structure replacement.

In Fig. 3, the object of the original image 102 with a poor quality is a building 302b. The building 302b is blurry or does not accurately reflect a texture of the physical structure. Using metadata 118 (e.g., of Fig. 2) for the original image 102, the object enhancement module 114 finds a reference image 108 with a building 302a that has a good quality and matching metadata, such as equivalent weather, a similar location of the photographing devices, and analogous viewing angles. To enhance the original image 102, the object enhancement module 114 replaces at least a portion of the building 302b with some aspect of the building 302a as extracted from the reference image 108 to produce the enhanced image 112. For example, fuzzy details from the building 302b of the original image 102 can be replaced using clearer details from building 302a of the reference image 108 to generate the enhanced image 112.

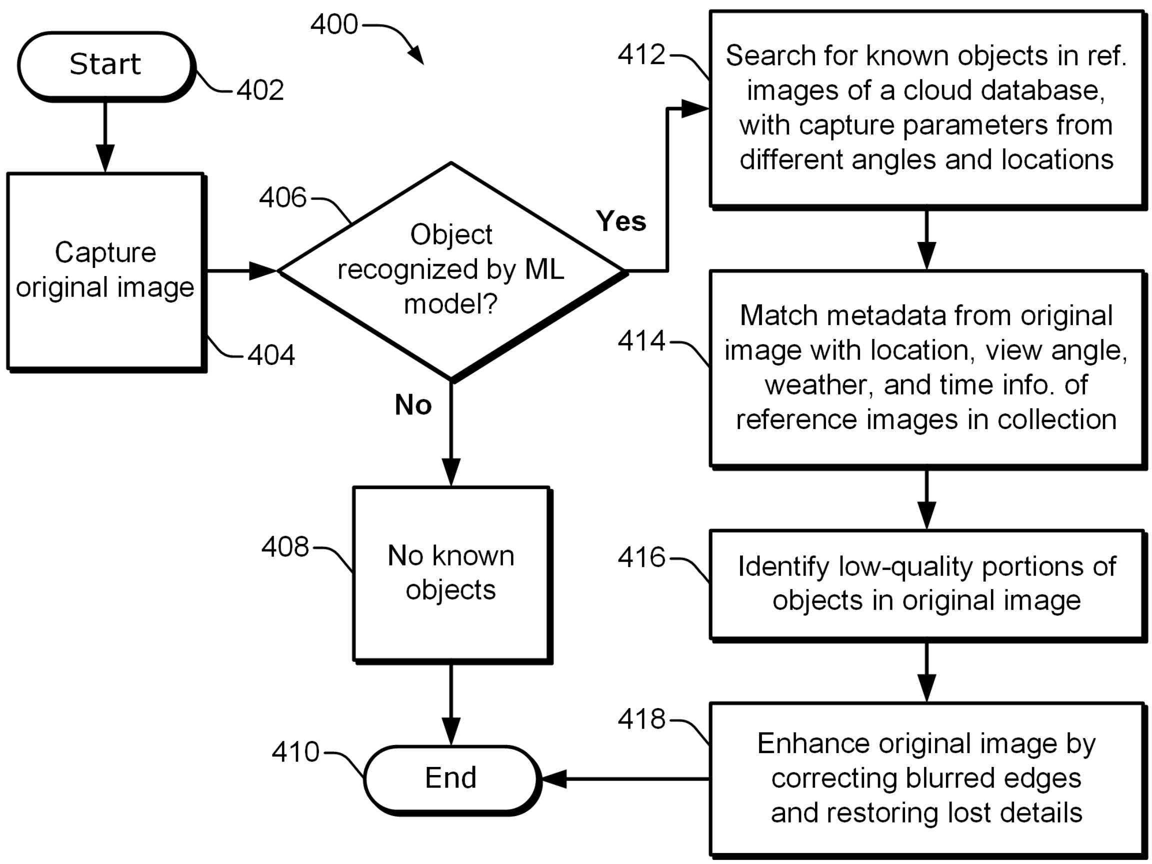

Fig. 4. Flowchart for enhancing an image with a public structure.

Fig. 4 depicts a flowchart 400 of an example process to enhance an image with a public structure. The flowchart 400 starts at 402, and a portable device captures an original image 102 (e.g., of Fig. 1) at 404. At 406, an ML model 116 attempts to recognize one or more objects 104 in the original image 102. If the ML model 116 fails to recognize any objects, an associated object enhancement module 114 determines at 408 that there are no known objects in the original image 102, so the process ends at 410.

On the other hand, if the ML model 116 recognizes at least one object 104-1b (e.g., a building 302b of Fig. 3) in the original image 102 at 406, the process continues at 412. At 412, the ML model 116 searches for known objects 104 in multiple reference images 108 of an image collection 106 that may be in network-accessible storage (e.g., a database in the cloud). The known objects 104 are associated with one or more capture parameters, such as different view angles and locations, using metadata 118 that was obtained when the reference images 108 were captured. At 414, the object enhancement module 114 matches the metadata 118 associated with the original image 102 with metadata—such as location, view angle, weather, and/or time—associated with the reference images 108 in the image collection 106. Alternatively, the matching activity of 414 can be performed before the searching activity of 412. The activities of 412 and 414 jointly determine a reference image 108 including an object 104-1a (e.g., a building 302a) that is recognized as being the same as, or at least similar to, the object 104-1b (e.g., the building 302b) in the original image 102 and that has at least similar metadata to the original image 102.

At 416, the ML model 116 or at least one image-processing algorithm identifies low-quality portions of at least one object 104-1b in the original image 102. At 418, the object enhancement module 114 enhances the original image 102 by replacing the low-quality portions of the object 104-1b in the original image 102 with high-quality portions extracted from the object 104-1a that is present in the reference image 108. For example, the object enhancement module 114 can correct blurred edges and restore lost details. After performing the public structure enhancement operation 110-1, the object enhancement module 114 can end the process at 410.

Fig. 5 illustrates an example face enhancement operation 110-2. Faces can include or exclude head coverings, jewelry, glasses, or hair. For facial enhancement, the image collection 106 can include a database of photos that are accumulated locally in the portable device because of the personal nature of faces. In such cases, the object enhancement module 114 can be implemented at least partially on the portable device, such as by using an onboard neural network engine. In many scenarios, images with one or more faces are taken using a front-facing camera as a “selfie” or are taken using a rear-facing camera as a still portrait of friends or family. However, described strategies are also applicable to photographs that capture people in motion.

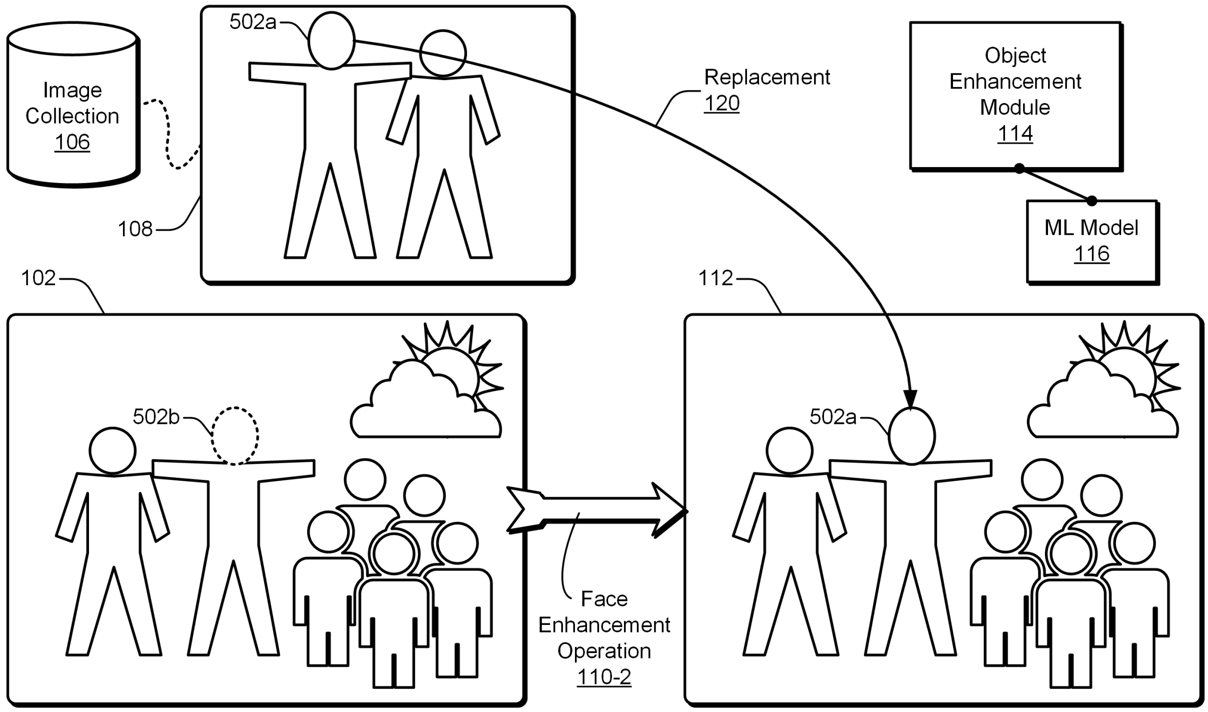

Fig. 5. Image enhancement with face replacement.

In Fig. 5, the object of the original image 102 with a poor quality is a face 502b. The face 502b is blurry, does not accurately reflect a skin tone of the person, or has obscured features. Using metadata 118 (e.g., of Fig. 2) for the original image 102, the object enhancement module 114 finds a reference image 108 with a face 502a that has a good quality and matching metadata, such as similar lighting, an equivalent distance between the camera and the face, and an analogous viewing angle, which may be based on stored measurements provided by an IMU. To enhance the original image 102, the object enhancement module 114 replaces at least a portion of the face 502b with some aspect of the face 502a as extracted from the reference image 108 to produce the enhanced image 112. For example, a noisy skin representation or blurred details of the face 502b of the original image 102 can be replaced using clearer or more accurate details from the face 502a of the reference image 108 to create the enhanced image 112.

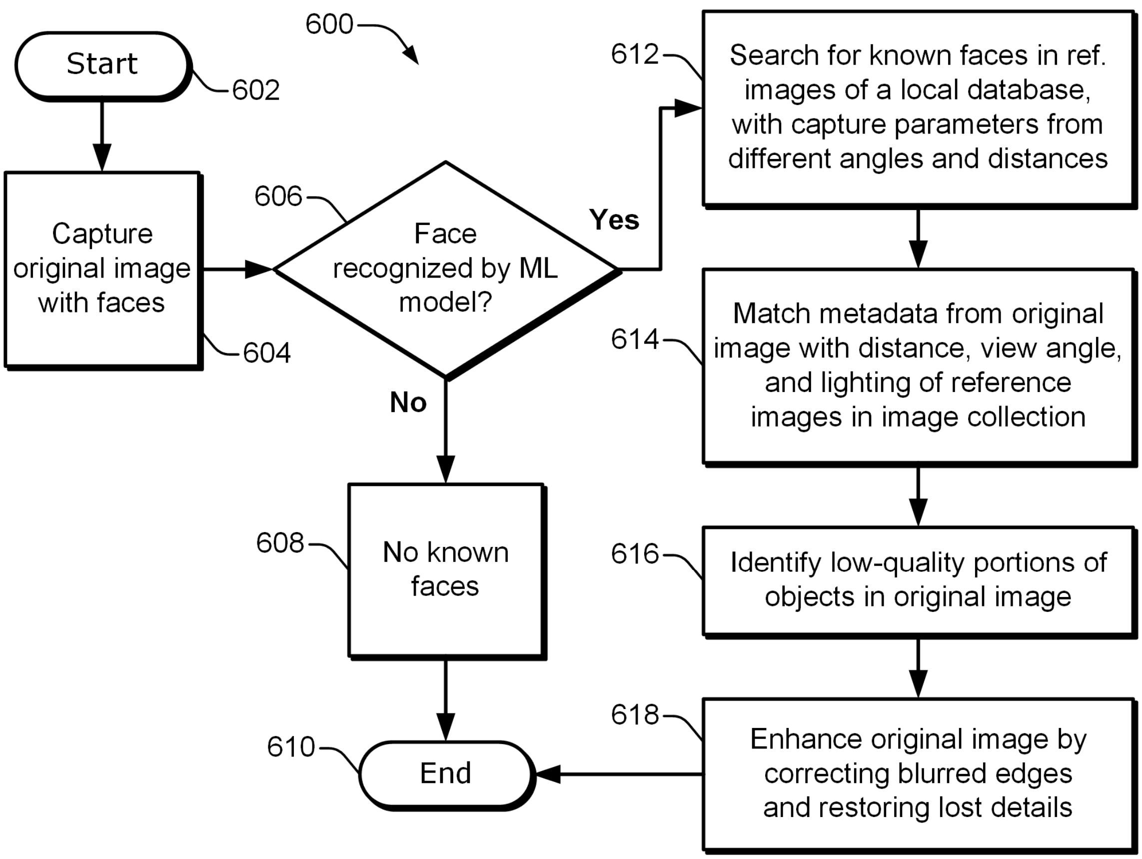

Fig. 6. Flowchart for enhancing an image with faces.

Fig. 6 depicts a flowchart 600 of an example process to enhance an image having at least one face. The flowchart 600 is similar to the flowchart 400 of Fig. 4. However, as shown at 604, the focus is on processing an original image 102 having at least one face 502 (of Fig. 5). At 606, an ML model 116 attempts to recognize one or more faces 502 in the original image 102. If the ML model 116 fails to recognize any faces, an associated object enhancement module 114 determines at 608 that there are no known faces in the original image 102, so the process ends at 610.

On the other hand, if the ML model 116 recognizes, at 606, at least one face 502b in the original image 102, the process continues at 612. At 612, the ML model 116 searches for known faces 502 in multiple reference images 108 of an image collection 106 that is stored in a local database. More generally, although the database may be personal to a user of the portable device, the collection of images may be stored at least partially in remote memory that is accessible via the internet (e.g., "cloud storage"). The known faces 502 are associated with one or more capture parameters, such as different view angles and distances, using metadata 118 that was obtained when the reference images 108 were captured. At 614, the object enhancement module 114 matches the metadata 118 associated with the original image 102 with metadata—such as view angle, lighting, and/or distance—associated with the reference images 108 in the personal image collection 106.

Alternatively, the matching activity of 614 can be performed before the searching activity of 612. The activities of 612 and 614 jointly determine a reference image 108 that includes a face 502a identified as being the same as the face 502b in the original image 102 and that has metadata 118 that is at least similar to that of the original image 102. At 616 and 618, the object enhancement module 114 replaces (e.g., by substitution or merging) the lower-quality aspects of the face 502b of the original image 102 using the higher-quality aspects of the face 502a of the reference image 108 to produce the enhanced image 112 with a superior facial appearance.

Throughout this publication, examples are described in which a computing system (e.g., an end-user device, a server device, a computer, or another type of computing system) may analyze information (e.g., image metadata, image visual data, or facial-recognition data) associated with an end-user, such as the information used to match a face in an original image to a face in a reference image for replacement purposes to produce an enhanced image. Further to the descriptions above, an end-user may be provided with controls allowing the end-user to make an election as to both if and when systems, programs, and/or features described herein may enable collection of information (e.g., image metadata or facial-recognition data), and if or when the end-user sends or receives content or communications to or from a server. The computing system can be configured to only use the information after the computing system receives explicit permission from the end-user of the computing system to use the data. For example, facial recognition or other image-related operations can be performed only on the condition that the end-user has been provided an opportunity to control these and other aspects of the image enhancement features described herein. Further, individual end-users may have constant control over what programs can or cannot do with the information. In addition, information collected may be pre-treated in one or more ways before it is transferred, stored, or otherwise used, so that personally-identifiable information is removed. For example, image metadata may be generalized or otherwise scrubbed prior to uploading an image to a remote server via the internet to remove any user-identifying information or device-identifying information embedded in the data. Thus, the end-user may have control over whether information is collected about the end-user and the end-user's device, and how such information, if collected, may be used by the computing device and/or a remote computing system.

References:

[1] Patent: US8798378B1. Scene classification for place recognition. Priority Date: December 7, 2009.

[2] Cho, Taeg Sang. "Motion Blur Removal from Photographs." PhD diss., Massachusetts Institute of Technology, 2010 (https://www.researchgate.net/publication/279821085_Motion_blur_removal_from_photographs).