March 2021

Image Deblurring According to Facially Recognized Locations Within the Image

Wei Hong

Li Wei

Xuhui Jia

Marius Renn

Yukun Zhu

Image Deblurring According to Facially Recognized Locations Within the Image

Abstract:

This publication describes techniques for image deblurring according to a facially recognized locations within the image. An algorithm may use facial detection and recognition to selectively sharpen aspects of faces within an image and the surrounding area associated with the facial detection. In one or more aspects, the selectivity of sharpening improves the computational load and other aspects of image provision to improve overall computer function, power consumption, and user experience. Individual faces within the image may be cropped or thumbnailed, providing portions of the image that include the faces. Counterpart images associated with the individual faces may be found within a database having a repository of sharp features associated with the counterpart images. As such, the features may be integrated with the blurred faces of the original image to sharpen an image output.

Keywords:

Face, deblurring, camera, YUV, homography, landmarks, canonical, convolutional neural network (CNN), encoder, decoder, distortion, sharp, blur, image, photograph, machine-learned model

Background:

Image blur and distortion can ruin even the most skillfully taken photographs. As mobile devices drive photographic mechanisms into ever-smaller footprints, smaller sensors, apertures, and motion reducers are challenged to mitigate blur and distortion. To compound the issue, as apertures decrease, exposure time increases, resulting in increased blur and distortion similar to photos of active scenes. Image sharpening and deblurring implementations can attempt to correct distortion with the consumption of expensive computational resources. As pixel density and image size increase, the cost to enhance entire images becomes prohibitive, even though photographers may only desire particular features in the image to become clear.

Description:

This publication describes techniques for image deblurring according to facially recognized locations within the image. The described techniques may be supported by various computers that may be portable or stationary. A computer can include various sensors, processors, computer-readable media, and components for taking pictures. Some of the additional components may include lenses, apertures, shutters, and flashes for producing photographic images. A computer-readable medium may include various memory types, including static random-access memory (RAM), dynamic RAM, non-volatile RAM, read-only memory (ROM), and other available memory technologies. The computer-readable medium may include instructions in computer-readable form, for instance, a program defining instructions operable to implement the teachings of this disclosure. The instructions may be of any implementation and may include field-programmable gate arrays (FPGA), machine code, assembly code, higher-order code (e.g., RUBY), or various combinations thereof. The processors may execute the instructions to follow a combination of steps and executions as provided in this disclosure. Other implements or variations of the processors may execute the instructions to follow a combination of steps and executions as provided in this disclosure.

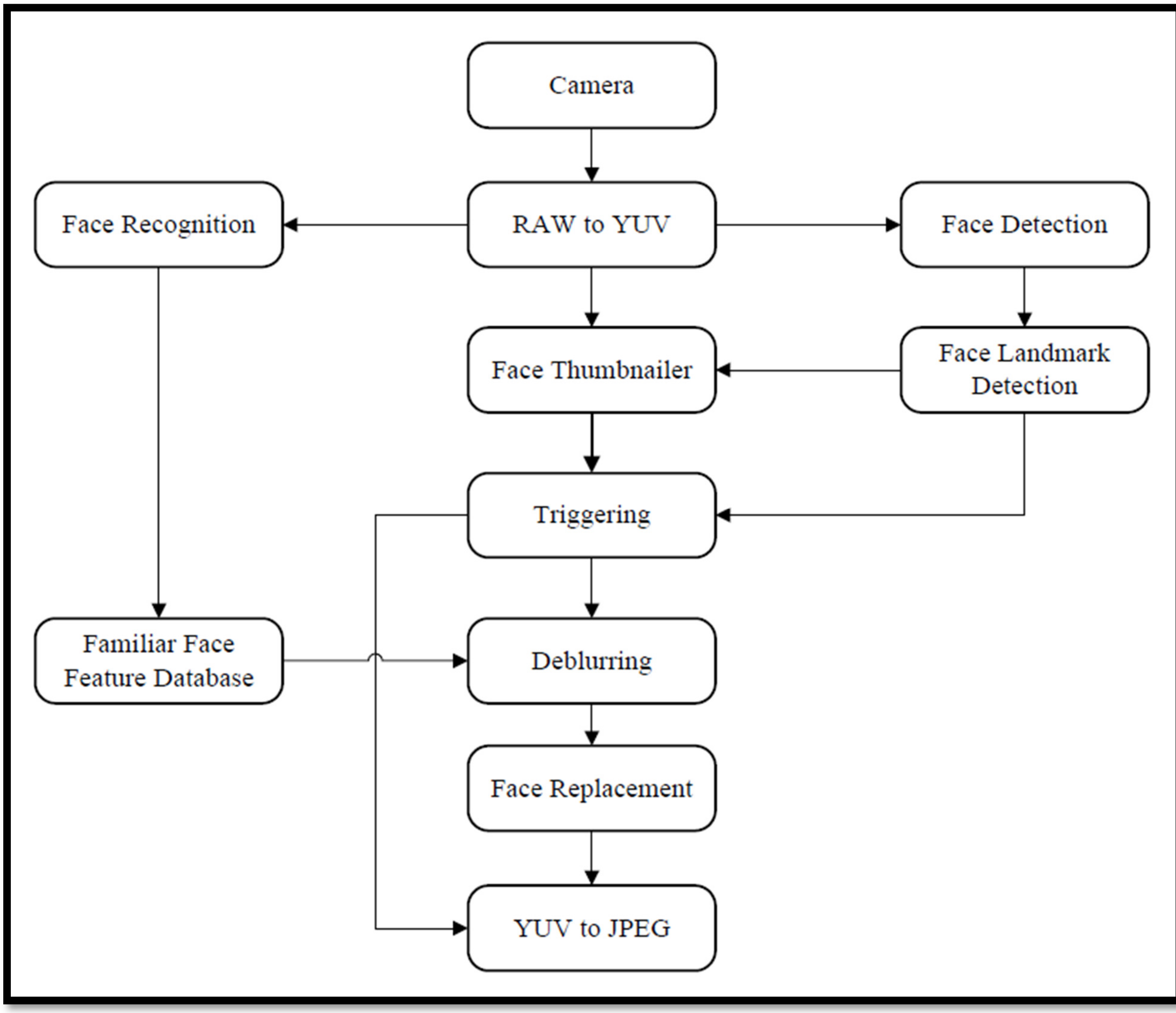

Figure 1

Figure 1 depicts a post-processing algorithm for an image provided by a camera, or other means (e.g., digital device), using computational photography and machine learning techniques. Figure 1, or portions of Figure 1, may be considered a top-level diagram of a face-deblurring pipeline. Because deblurring can be a computationally burdensome and time-consuming process, thumbnails or cropped portions of the image may be selected to decrease computation time while adding value to images provided. That is, photographers may appreciate crisp and clear facial content over background image elements. Sharpening features of faces within an image may provide an adequate balance between computational loads and image clarity. Faces may be more likely to move during picture-taking, and the clarity of facial features may be more desirable than those features in the background. These clarity tools may be applied to various picture modes, including portrait mode and night mode.

A camera may capture a raw image after a shutter of the camera is closed. The image may be processed into a YUV format or another format for further processing along the pipeline. The image may be processed to detect faces and also recognize those faces according to a familiar-face database. In one aspect, the recognized faces may undergo a facial-landmark-detection function that detects landmarks associated with the faces. In another, the facial recognition function may associate faces recognized within the image with faces from the familiar-face database and provide associated facial features.

Based on the detected faces, thumbnails are generated to show one or more of the face landmarks detected. A homography transform between the face landmarks and a set of canonical landmarks may be generated and applied for one or more faces in the image, generating aspect-specific thumbnails for one or more of the detected faces. The homography transform may adjust, align, and resize the thumbnail.

Sharpening and other image enhancements may be performed on one or more of the detected faces. The algorithm may include a triggering function configured to receive one or more of the thumbnails in the pipeline. The triggering function determines whether the thumbnail requires sharpening, deblurring, or other post-processing. The triggering function may be a trained classifier based on a variety of inputs. For example, the classifier may receive the image, sharpness metrics derived from the thumbnail, movement of the face, camera movement, exposure time, autofocus status, metadata, and other data available to the digital device. If the classifier determines that deblurring should be performed, the individual thumbnail or a set of blurry, classified thumbnails is sent to the deblurring or sharpening function. Sharp or undistorted images may bypass the sharpening function to conserve computation, battery, and memory resources.

The thumbnails may be sharpened by an encoder-decoder pipeline depicted in Figure 2.

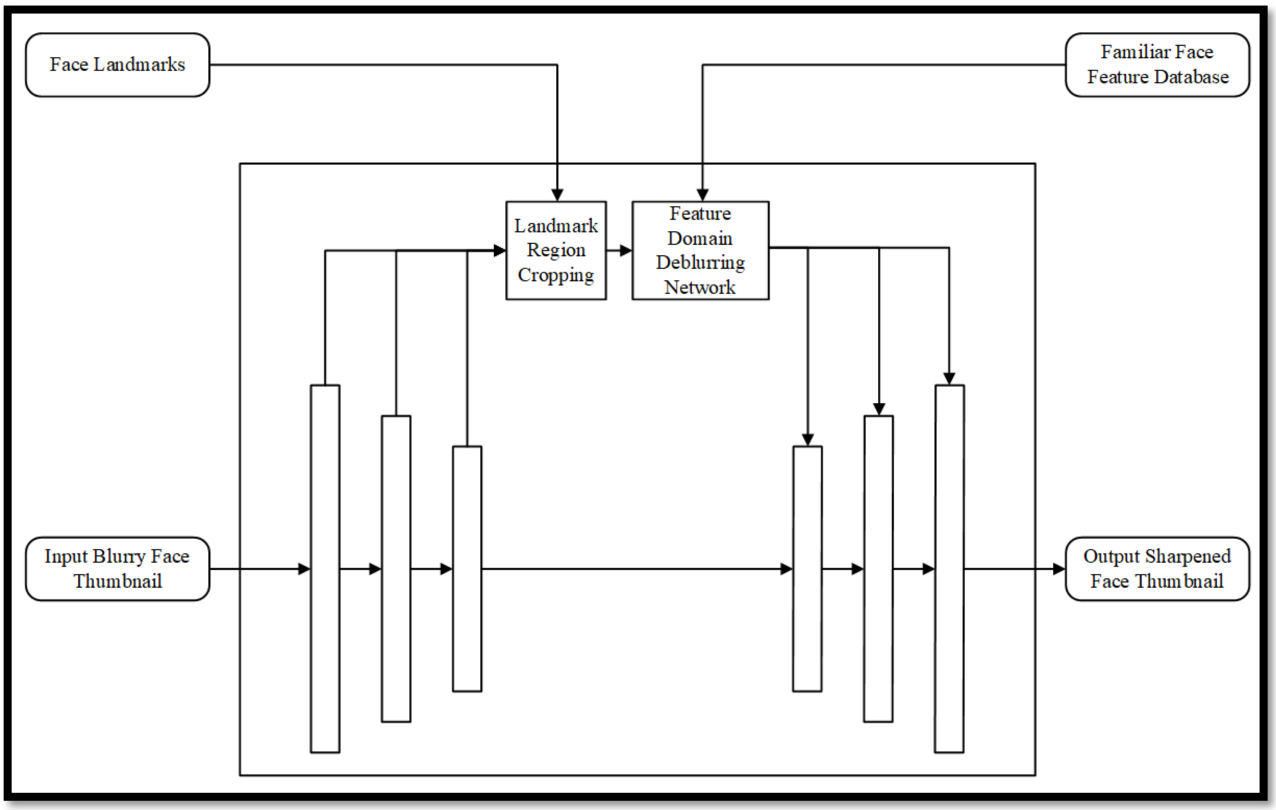

Figure 2

The encoder-decoder pipeline may include one or more convolutional neural networks trained to sharpen the thumbnails or portions of the thumbnail. The convolutional neural network may have a cascade of encoding layers and decoding layers. The encoding layers may define weighted features (e.g., shapes, edges, contours) that are associated with thumbnails of faces. The face landmarks are provided to the sharpening function to influence individual cropping of features (e.g., eyes, nose, mouth) within the thumbnail. For example, landmark-region cropping may provide an indication of the size or shape of facial features, landmarks, or the like and draw the attention of the network to the particular locations of the thumbnail related to the landmark.

Through a combination of the features gleaned from layers of the convolutional neural network and landmarks determined within the thumbnail, features isolated from the particular layers may be cropped using the landmarks and provided to the feature-domain deblurring network. As an example, thumbnails that contain particular features filtered by the layers are cropped based on location and provided to the feature-domain deblurring network.

A familiar-face-feature database may provide crisper or clearer features or landmarks for faces recognized within the thumbnail to enhance the distorted features. Sharp thumbnails associated with a particular face may be combined with the cropped landmark features in the feature-domain deblurring network. As an example, a neural network may combine the sharp feature of the recognized face in the thumbnail with the blurred feature cropped from the thumbnail to deblur or sharpen the particular feature.

Facial recognition may provide an identity of the face. If an image is received of a face that is sharp, the face and the facial landmarks detected may be added to the familiar-face database and used to sharpen distorted facial features of other images provided. For instance, the encoder portion of Figure 2 may be used to provide face features for the identified thumbnail. As an example, the familiar-face-feature database may include the thumbnails of images and the associated facial features. That is, if a sharp face is provided as a thumbnail of an image and the triggering function decides not to trigger deblurring because the face is already sharp, the encoder may encode the thumbnail and save the sharp face features along with the thumbnail. The image may be stored in the database with the encoded features.

To provide sharpening, the similarity of the face provided in the thumbnail may drive the selection of the features retrieved from the database. A similarity index may be determined by defining the similarity between the face provided in the thumbnail and those of the database. If none of the images compared for similarity exceed a threshold or match the image provided in the thumbnail, a general sharp face-feature dataset may be used, or one of many general sharp face-feature datasets that are most similar to the face in the thumbnail may be used. In one aspect, the feature-domain deblurring network receives the distorted or blurry face features from the encoder layers and the sharp face-features from the database to sharpen or deblurred the face, the face features, or a combination thereof. The deblurred face features are then used in the decoder to produce the deblurred face thumbnail. The deblurred face features may be concatenated with the layers or integrated in another fashion.

In one aspect, the deblurred face thumbnail may be transformed back to the location and size of the original face by the inverse homography transform used to generate the thumbnail. The transformed deblurred face replaces the original blurred face. A smooth transition around the face region is created to avoid a harsh boundary between the deblurred face and the background.

Further to the above descriptions, a user may be provided with controls allowing the user to make an election as to both if and when systems, applications, and/or features described herein may enable the collection of user information (e.g., facial features, appearance information, social interactions and activities, profession, a user's preferences, a user's current location), and if the user is sent content and/or communications from a server. In addition, certain data may be treated in one or more ways before it is stored and/or used so that personally identifiable information is removed. For example, a user's identity may be treated so that no personally identifiable information can be determined for the user. In another example, a user's geographic location may be generalized where location information is obtained, such as to a city, ZIP code, or state level, so that a particular location of a user cannot be determined. Thus, the user may have control over what information is collected about the user, how that information is used, and what information is provided to the user.

References:

[1] Patent Publication: US 20170351932 A1. Method, Apparatus and Computer Program Product for Blur Estimation. Priority Date: December 19, 2014.

[2] Patent Publication: US 20170213112 A1. Utilizing Deep Learning for Automatic Digital Image Segmentation and Stylization. Priority Date: January 25, 2016.

[3] Patent Publication: CN 111738230 A. Face Recognition Method, Face Recognition Device and Electronic Equipment. Priority Date: August 05, 2020.

[4] Patent Publication: US 20130242129 A1. Method and Device for Recovering a Digital Image from a Sequence of Observed Digital Images. Priority Date: September 28, 2010.