April 2021

Predictive Modeling Techniques for Adaptive Face Retouching

Wei Hong

Shengyi Lin

Li Wei

Predictive Modeling Techniques for Adaptive Face Retouching

Abstract:

This publication describes predictive modeling techniques for adaptive face retouching according to a predictive model on a computing device. The adaptive face retouching enables preservation of important features that may depict diversity, persona, and uniqueness in images where retouching is applied by a computer. In the techniques, an algorithm may be implemented to balance image attributes from a variety of sources and categories. The algorithm may further reduce unnecessary, redundant processing on the computing device by storing computed attributes within a facial attribute database. As such, diverse facial features may be maintained while allowing adequate face retouching without requiring overuse of computing resources.

Keywords:

Face, retouching, photo editing, facial corrections, strength prediction model, face attributes, exposure time, camera, image, photograph, facial recognition, temporal filter, machine learning, artificial intelligence, convolutional neural network

Background:

The utilization of a retouching technique can improve the quality of images on computing devices (e.g., phones). As an example, a face retouching technique can remove, add, or improve facial features (e.g., blemishes, wrinkles, hair, feature shapes). Sometimes, a face retouching process can unintentionally diminish the value of underlying facial features. For instance, high-value features that highlight diversity, persona, or uniqueness may be treasured by photographers and subjects alike. While retouching may seek to remove some irregularities or distortion, important features that provide that distinctiveness may be unintentionally removed. As an example, smoothing techniques may remove blemishes, facial hair, stubble, and other unique characteristics. That is, when smoothing, retouching, and enhancement techniques are broadly applied, high-value features that underscore diversity, persona, or uniqueness may become lost.

Description:

Consider, for example, that a user would like to preserve facial-hair stubble without being required to select such preservation. This disclosure provides data-driven machine learning techniques to solve this and other problems. As an example, a machine-learned model may be used to predict the preference of users and adjust the strength of face retouching accordingly. A simple technique may use an adaptive strength based on the gradient-based sharpness of face regions. An improved technique may use a machine-learned model to predict the appropriate strength of the face retouching. The predictive model can be any machine-learned model (e.g., regression, support-vector machine, decision tree, neural network, convolution). As illustrated in Figure 1, image retouching may be based on a camera image or another image source and may result in an output, which may be formatted as a JPEG by the computer. The predictive model may be trained to provide a desirable strength based on the inputs.

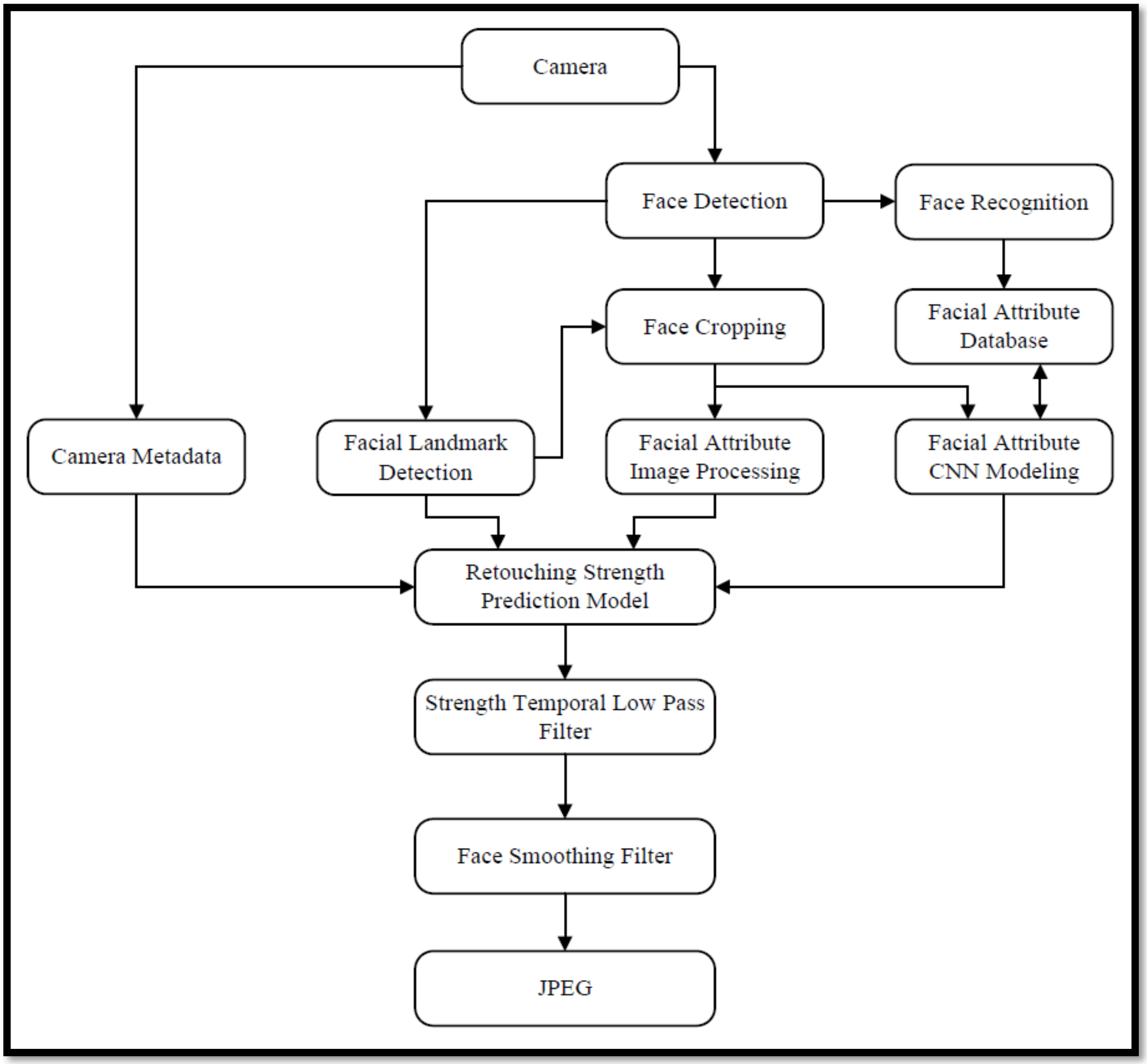

Figure 1

The inputs to the predictive model that defines the retouching strength may be defined within one of four categories. As an example, the camera or image may define camera metadata. The camera metadata may include manufacturer name, model number, image orientation, version numbers, date and time information, YCbCr positioning, compression information, resolution, exposure time, F-number, biases, aperture indications, flash information, focal lengths, and various other information related to the image. Such information may be encoded and provided to the predictive model.

As another example, input category, facial landmark detection information may be provided to the prediction model. As an example, the facial landmark detection information may include the size of a subject's face within the image. Other facial landmark detection information may include facial orientation information, and movement information. Additional facial landmark information may be provided by facial landmark detection processing functions.

Unintelligent image processing may provide inputs that fall within a third example input category, facial attribute image processing. As an example, gradient-based sharpness detection may indicate the overall sharpness of the image or face within the image. The image may be cropped by the algorithm based on the facial landmark and detection functions to provide an indication of characteristics of only the face within the image instead of the image as a whole. The image processing within this category may further provide brightness, contrast, and other quality indications of the facial region.

Lastly, artificial intelligence may provide inputs that fall within a fourth and final example input category. As an example, the image or cropped facial region may be subjected to layers of a convolutional neural network. The convolutional neural network may be trained to indicate features of the face, including facial hair, stubbles, freckles, acne, scars, age, hardware (e.g., glasses, nose rings), other diversity, persona, or uniqueness information and other image-related information (e.g., distortion, blur). As an example, the convolutional neural network (CNN) may output a confidence factor or floating-point number corresponding to the likelihood of any of the trained features. That is, an image with facial stubble may output a 0.99 in the stubble category, while a person with a blemish or beauty mark may output a 0.50 in the freckle category. These indications, along with the others, may impact the strength output of the prediction model and enable stubble, a beauty mark, or other valuable uniqueness characteristics to be maintained with retouching.

The artificial intelligence calculations associated with the fourth and final input category may have high computational loads. The face attribute database may store familiar faces and associated attributes that are fed as the fourth example input category to the prediction model. As an example, an unknown face may be detected as having a similarity index to other faces in the face attribute database less than a threshold. The CNN may be instantiated to perform recognition of features discussed above. As features are detected, the features may be saved in the face attribute database such that instantiation of the CNN is not performed for known faces. In another aspect, one CNN may output confidence indications for each of the categories (e.g., stubble, freckles, scars). In another aspect, the fourth input category may be based on CNNs for each associated feature. As an example, one CNN may be used to determine freckles, while another is used to determine stubble. As such, each CNN may be engaged or ignored based on the stored features within the face attribute database. For instance, the CNNs may determine respective confidence indications for stubble and face blurriness based on the cropped facial image.

To reduce the number of CNN operating on each image, the CNN for each confidence indication (e.g., stubble, face blurriness) may be categorized as long-term or short-term. Long-term confidence indications may be stored within the face attribute database along with the respective face. Short-term indications may be processed for every image received. As an example, confidence indications sent to the prediction model may correspond with stubble and facial blurriness. If the face within the image is recognized and the confidence indication for stubble (e.g., long-term indications) has been previously computed or stored, the confidence indications may be retrieved from the facial attribute database and sent to the prediction model. For short-term indications (e.g., facial blurriness indications), the confidence indication may be computed by the respective CNN and sent to the prediction model. Additional categories or inputs to the predictive model are contemplated by this disclosure.

Drastic changes in the strength determined by the predictive model may lead to inconsistent results. As such, a temporal low-pass filter may smooth changes in the strength determination provided by the prediction model. The temporal low-pass filter can avoid retouching strength fluctuations across camera sessions and avoid abrupt frame-to-frame changes in the filtering strength to produce more consistent results. A face-smoothing filter may be applied to the original image or portion of the image containing the face to retouch the face according to the determined strength.

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 implement 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.

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 collection of user information (e.g., information about a user's appearance, social groups, persona, diversity attributes, uniqueness, social 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: US20190130166A1. Method and Apparatus for Image Processing, and Computer-Readable Storage Medium. Priority Date: October 31, 2017.

[2] Patent Publication: US20120162478A1. Digital Photographing Apparatus and Control Method Thereof. Priority Date: December 22, 2010.

[3] Patent Publication: US20200151964A1. Scalable Real-Time Face Beautification of Video Images. Priority Date: September 11, 2015.

[4] Patent Publication: WO2005096213A1. Face Recognition System and Method. Priority Date: March 5, 2004.