CV: Facial recognition and emoji generation

Image Processing:

It refers to performing certain methods/processes on image in order to get enhanced image or to get certain useful information from it.

It is a type of signal processing which uses image as the input and output can be an image or other characteristic features via an algorithm.

Preprocessing steps are:

1. Read image
2. Resize image
3. Remove noise
4. Segmentation
5. Morphology

Classification Techniques:

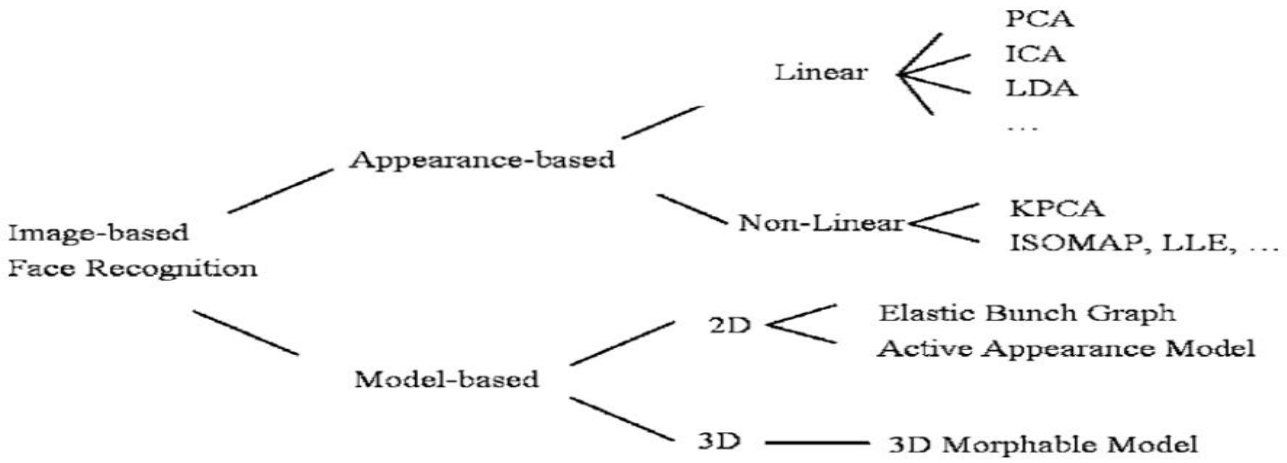

Image classification is the task of categorizing and assigning labels to groups of pixels or vectors within an image dependent on particular rules. The categorization law can be applied through one or multiple spectral or textural characterizations.

Face Detection Techniques:

Viola-Jones Algorithm:

Despite being an outdated framework, Viola-Jones is quite powerful, and its application has proven to be exceptionally notable in real-time face detection. This algorithm is painfully slow to train but can detect faces in real-time with impressive speed. Given an image (this algorithm works on grayscale image), the algorithm looks at many smaller subregions and tries to find a face by looking for specific features in each subregion. It needs to check many different positions and scales because an image can contain many faces of various sizes. Viola and Jones used Haar-like features to detect faces in this algorithm.

R-CNN Algorithm:

R-CNN creates bounding boxes, or regions, using selective search. Selective search looks at the image through windows of different sizes, and for each size it tries to group together adjacent pixels by texture, color, or intensity to identify objects. Generate a set of regions for bounding boxes. Run the images in the bounding boxes through a pre-trained neural network and finally an SVM to see what object the image in the box is. Run the box through a linear regression model to output tighter coordinates for the box once the object has been classified.

Holistic Matching (eigenfaces):

In this approach, complete face region is taken into account as input data into face catching system. One of the best examples of holistic methods are Eigenfaces, PCA, Linear Discriminant Analysis and independent component analysis etc.

Eigenfaces are made by extracting characteristic features from the faces. The input images are normalized to line up the eyes and mouths. Then they are resized so that they have the same size. Eigenfaces can now be extracted from the image data by using a mathematical tool called PCA. Now each image will be represented as a vector of weights. System is now ready to accept queries. The weight of the incoming unknown image is found and then compared to the weights of already present images in the system. If the input image's weight is over a given threshold it is considered to be unidentified. The identification of the input image is done by finding the image in the database whose weights are the closest to the weights of the input image. The image in the database with the closest weight will be returned as a hit to the user.

Feature-Based:

Detection: Identify the Interest Point

Description: The local appearance around each feature point is described in some way that is (ideally) invariant under changes in illumination, translation, scale, and in-plane rotation. We typically end up with a descriptor vector for each feature point.

Matching: Descriptors are compared across the images, to identify similar features. For two images we may get a set of pairs $(Xi, Yi) \leftrightarrow (Xi`, Yi`)$, where (Xi, Yi) is a feature in one image and $(Xi`, Yi`)$ its matching feature in the other image.

Modeling techniques:

The modelling of facial image starts by creating synthetic images with different types of faces and facial features measurement. In fact, there are various types of human faces in terms of shape, size and facial measurements. There are a few researches related with the study of human face and mostly addressing that there are commonly 5 types of human face shapes.

The synthetic face is a matrix with a dimension of 360x240 thus there are a total of 86400 data for a single image. The data indicate total pixels of the synthetic image. Inside the matrix data, it consists of three different variables declared with separate values.

The synthetic models will be tested under three parameters, as adding noise, manipulating the pixel intensity and changing the size of the frontal image. The first process is manipulating the intensity difference between object and background image. Next, the object size is varied. Then, the process conducted is adding noise to the input models. The Gaussian noise is added with the mean = 0 but with different variance. Then, the intensity difference between object and background is varied.

Synthetic models have been developed to represent facial images. The model portrays five different types of face shape and based on the face features as eyes, nose and lips. Additionally, the scene parameters of facial images – intensity, size and noise can be varied using the synthetic model. These synthetic models are useful as they can serve as input for performance analysis of facial recognition system. In total, the synthetic model contains 2500 images from 5 face types with variation of 10 levels of intensity differences, 10 levels of object size and 5 levels of measurement noise. The number of images could be increase with additional levels of variation accordingly.

Conclusion:

Preferred facial detection/recognition technique:

Hybrid Method:

This uses a combination of both holistic and feature extraction methods. Generally, 3D Images are used in these methods. The image of a face is caught in 3D, to note the curves of the eye sockets, or the shapes of the chin or forehead. Even a face in profile would serve because the system uses depth, and an axis of measurement, which gives it enough information to construct a full face. The 3D system includes Detection, Position, Measurement, Representation and Matching.

Detection - Capturing a face by scanning a photograph or photographing a person's face in real time.

Position - Determining the location, size and angle of the head.

Measurement - Assigning measurements to each curve of the face to make a template.

Representation - Converting the template into a numerical representation of the face.

Matching - Comparing the received data with faces in the database. The 3D image which is to be compared with an existing 3D image, needs to have no alterations.