Applied Mathematics E-Notes, 21(2021), 478-494 ©
Available free at mirror sites of http://www.math.nthu.edu.tw/∼amen/
ISSN 1607-2510

A Comprehensive Note On Techniques Used For Face Detection And Application*

Vineet Bhatt†, Kailash Chandra Petwal‡, Sunil Kumar§, Sakshi Rajput¶, Manoj‖

Received 16 June 2020

Abstract

Image processing is a technique to apply mathematical operation on a digital image in order to get a better image or to understand other useful information from image data. A methodological study on the significance of image processing and its application in daily life is carried out here. This study paper describes various image processes with techniques and applications such as gender identification, biometric attendance, face detection, feature detection like nose, mouth, eyes, etc. with the help of MATLAB. This paper also deals with some of the face detection techniques, models and feature analysis.

1 Introduction

Natural interaction is a way in which people communicate in day-to-day life through vocal gesture, emotionally, visually, physically and by different types of manipulation in object. In the present time, communicational power, human detection and person identification using technology are becoming more popular and everyone is aware about the development in the technology, especially in the field of human-computer interaction with people. If one wants to interact with technology same as human being, than there is a need for technology to detect our gestures, emotions, physical activities, etc. One such technical way is to interact with technology through face detection, recognition and fingerprint identification. Face detection is a widely used computer technology which detects human faces in digital images. It can also be referred as a psychological process by which people locate faces in a visual scene. The primary goal of face detection, on being given a digital image, is to determine whether or not there are any faces in image, which would otherwise the trivial task for human being. There are some limitations to face detection like variations in scale, location, orientation, pose, facial expression, lightning conditions, and occlusions etc. But this field of face detection has made considerable progress in the past decades thus coming up with various methods to overcome these limitations.

1.1 Literature Survey

We are presenting a glance on the work that has been done in the field of the development of image processing from past 20 years. [1] addressed the issue of evaluating face recognition algorithms using descriptive statistical tools. They used permutation methodology in a Monte Carlo sampling procedure to investigate recognition rate results probability distributions of some algorithms. The results of masterpiece [2] ruled out all extant explanations of prosopagnosia except one that proposed that faces are recognized by a content-specific face processing mechanism. The article [3] came up with the idea of facial mimicry which differentially contributes to recognition of specific facial expressions and this allowed for more refined predictions from embodied cognition theories. [4] discussed that there does appear to be a small, but significant, relationship

*Mathematics Subject Classifications: 68U10; 94A08; 97H60.

†Department of Mathematics, Graphic Era Hill University, Campus Dehradun, Uttarakhand, India
‡Department of Mathematics, H.N.B. Garhwal University, S.R.T. Campus Tehri, Uttarakhand, India
§Department of Mathematics, Graphic Era Hill University, Campus Dehradun, Uttarakhand, India
¶Department of Mathematics, Graphic Era Hill University, Campus Dehradun, Uttarakhand, India
‖Department of Mathematics, Graphic Era Hill University, Campus Dehradun, Uttarakhand, India

between the description measures of accuracy, number of incorrect descriptors, and congruence with that of subsequent identification accuracy. Furthermore, author [4] found certain conditions to strengthen the magnitude of this relationship, including the use of face recognition versus eyewitness identification paradigms and the length of delays between relevant tasks. The [5] presented a review focusing on the evidence, available from experimental investigations of how people recognize faces. Factors affecting recognition are evaluated in terms of how they apply to familiar and unfamiliar faces and categorized according to the nature of their impact. [6] presented conclusions comparing different results, such as: right detections, false positives and false negatives for each color space. According to the results obtained by [6], the most appropriate color spaces for skin color detection are HSV model, and the models YCgCr and YDbDr. [7] showed that photographs are not consistent indicators of facial appearance because they are blind to within-person variability. Crucially, this within-person variability is often very large compared to the differences between people. [8] proposed a generic HFR framework in which both probe and gallery images are separated in terms of nonlinear similarities to a collection of prototype face images. [9] gave us an overview about the research in significant challenges that we face in most video-based applications and present a broad and deep review of how recently proposed methods help in overcoming the difficulties. A new type of algorithm was presented by [10] to improve the fractional differential Tiansi operator that can enhance the edge detection result. [11] presented a literature review on the scientific research on the automatic recognition of personal emotions and later discussed the main limitations, future directions, institutions, and influential figures in this field.

2 Challenges In Face Detection

Challenges reduce accuracy and detection rate in face detection. Researchers who are working in the area of image processing are familiar with some of the challenges which had been discovered in last few decades and figured out few of them. Some of these challenges are given as follows:

(i) Human face having an unusual expression in an image becomes a challenge for face detection.

(ii) Face hidden by an object also hinders the face detection. It is known as face occlusion. It may be hair, hand, hand, glasses, scarf etc.

(iii) Irregular lighting in an image may occur. Some parts have high illumination and other may have very low.

(iv) When there are lots of objects in an image, it also reduces the accuracy of face detection.

(v) Some images have a poor resolution which makes face detection difficult.

(vi) Too many faces in an image are also a challenge for face detection.

(vii) As skin color changes in different geographical locations, it becomes a challenge for face detection.

(viii) When there is too much distance between face and human, detection rate reduces.

(ix) The pose of face with an angle is face orientation. It hinders the process of face detection.

2.1 Mathematical Understanding of A Digital Image

To understanding and learning about an image, first of all one need to convert an image into digital form. A digital image $I[M, N]$ described in a two dimensional discrete space is derived from an analog image $I(x, y)$ in a two dimensional continuous space through a sampling process. The two dimensional digital image is just a matrix of N rows and M columns. The intersection of a row and a column is termed a pixel. The pixels are stored in computer memory as a two dimensional array or matrix of real number. An image also can be defined as a two dimension function $f(x, y)$, where x and y are spatial coordinates, and the amplitude of f at any pair of (x, y) is gray level of the image at that point. A combination of individual 2D images helps in the formation of color images. For monochromatic images, most of the image processing techniques can be extended to color image (3D) by individually processing the three components of image. A digital image is represented as a matrix of order $M \times N$:

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & --- & f(0,N-1) \\ f(1,0) & f(1,1) & --- & f(1,N-1) \\ -- & -- & --- & -- \\ f(M-1,0) & f(M-1,1) & --- & f(M-1,N-1) \end{bmatrix}.$$

The number of gray levels is taken to be power of 2 i.e., 2^k, for some integer $k > 0$. Most of the time digital image takes value such that $0, 1, 2, .., 255$, therefore 256 distinct gray levels.

3 Uses Of Face Detection

In this section, we have discussed some familiar uses of the face detection techniques.

3.1 Gender Identification

We can find gender information from a digital image. The basic gender classification system contains mainly three modules, i.e., preprocessing, features extraction, and classifier. In a pre-processing, basically relevant features are taken out, which are the most potential segment of an image. Feature extraction is done by combing wavelet and Radon transforms. The obtained features are then passed to a powerful supervised learning algorithm using SVM to discriminate male and female [12]using wavelet and Radon transforms together.

3.2 Biometric Attendance and Biometric Pattern Recognition

Human face and human signature represent some of the common biometric patterns that are encountered by our visual systems daily. Here, we present the classification techniques of the above two biometric features. Now-a-days, a lot of interest has been generated in automated face recognition and a number of implementation approaches has been proposed. The major strategies used in face identification are either based on features or on face space, such as Eigen face or Fisher face. Most of the features are extracted from the front view of the face and sometimes from the side view. An automatic face recognition system, using both front and the side view, is more accurate. It takes the advantage of the explicit information available in both the views of the human face. Biometric attendance, one of the use of face detection, is a very useful way to save data without any wastage of paper. Thus, making the process of record-keeping uncomplicated. We take attendance of human beings using their finger prints or face etc. Biometric systems are divided on the basis of medium that are used for authentication. Methods used for identification are Face Recognition, Iris Recognition, Palm Recognition; Voice Recognition, Fingerprint Recognition, ECG signal based recognition method. Different techniques are used to extract the features for recognition [13]. A typical biometric system broadly contains database, feature extractor and classifier.

3.2.1 Feature Extractor

In this, important features are extracted from the training and test images. These features help us in determining real images from fake images.

3.2.2 Database

The extracted features from the training and test data are stored in the database. Input image are stored in the test data-set and the training images are stored in the training data-set.

3.2.3 Classifier

Classifiers categorize the input data into distinct number of classes and decide whether it is real or fake [14].

3.3 Human-Computer Interaction System

This technology is all about the interfaces between users and computers. In the construction of human-computer/ interfaces, a face analysis system is presented and employed. This system is based on three modules, defined as detection, tracking and classification, which are integrated and used to detect, track and classify faces in dynamic environments [15].

3.4 Medical Imaging

Medical imaging is a process by which we get images of the body parts to detect or study diseases, if any. It is developing rapidly due to developments in image processing techniques, including image recognition, analysis and enhancement. Image processing increases the percentage and the amount of tissues detected [16].

3.5 Extracting Facial Feature

With the help of different models, we can detect face as well as facial features such as nose, mouth, eyes, etc. and for this we generally use the RGB technique [17].

3.5.1 Feature Selection

Feature measurement is based on area, angle and distances between some points identified from the front and side views of the human face. Extraction of features from the front view may be performed from the edge images[42]. The template matching may be used for extraction of the eyes from the face image, while features such as nose, lips, chin, etc., may be extracted from the horizontal and vertical edge maps of a human face.

3.5.2 Front Facial Features Extraction

Extraction of front facial components such as eyebrows, eyes, eye points and others are based on correlation technique. Initially the set of eye templates are chosen and facial image $f(i, j)$ is convolved with a set of appropriately chosen templates $T(m, n)$ which is represented by the filter operation $F(i, j) = \sum\sum T(m, n) \times f(i + m, j + n)$. The resultant of the convolution filter represents the position of the eye and for detection of eyes from convolution filter. We can use translation, scale, and rotation invariant affine transformation. After two eye detection, we can easily detect eyebrow positions within a small search region above the eye center. We can also calculate the location of invariant and variant front facial features such as: Distance between left and right iris centers, between two inner eye points, between two outer eye points, and from eye-centre to nose tip are invariant features which do not change with facial expression whereas distance between left iris centre and left eyebrow, between right iris centre and right eyebrow and face width at the nose tip are variant features that change with facial expression.

3.5.3 Face Recognition Using Eigenfaces

The basis of this method is principal component analysis. The main focus of face recognition using eigenfaces is to decompose face images into a set of eigenfaces, which is a set of characteristic feature images. These are the principal components of the original images [43]. When these eigenfaces function as the orthogonal basis vectors of a liner subspace, we call them face space. We project a new face image into the face space and then compare its position in the face space with those of known faces to perform face recognition. This method involves converting training set of facial image patterns into vector of $M \times N \times K$ where $M \times N$ is image size and K is number of training samples. It is important to reduce the dimensionality of $M \times N$ space before attempting face recognition as it is very large and we can do that by using two very common approaches i.e. Principal component analysis (PCA) and linear discriminant analysis (LDA). When we compute the PCA, we represent an $N \times N$ image as a one-dimensional vector of $N * N$ elements, by placing the rows of the image one after another. Later, we compute the covariance matrix of the entire data set. Next we compute the eigenvalues of this covariance matrix. The principal components are the eigenvectors corresponding to the most significant eigenvalue. Principal component analysis is to transform all $M \times N$-dimensional face samples to a single vector such as $x_i(i = 1, 2, ..., k)$ to a single vector $y_i(i = 1, 2, .., n$,i.e., $y_i = v^T x_i$ where face image represent by x_i and weight vectors v represent the scaling. The objective function O calculated by maximization of variance such as $O = max \sum_{i=1}^{n} (y_i - \bar{y})^2$ where $\bar{y} = 1/n \sum_{i=1}^{n} y_i$. The set of weight vectors $v_1, v_2, ..., v_k$ represent the eigenvectors of the covariance matrix computed from the set of sample faces and $k < MN$. To reduce the dimensionality of facial feature we select only a few eigenvectors corresponding to the dominant eigenvalues and thus reduce the dimensionality of facial features. For an accurate face recognition system, the accuracy should be quite high and at the same time the processing time for a test face image should be low. The recognition system should be invariant to the head rotation and translation, and also to the illumination intensity changes. Appendix-1(a),1(b),1(c) and 1(d) are MATLAB code for nose, face, eyes and mouth detection respectively.

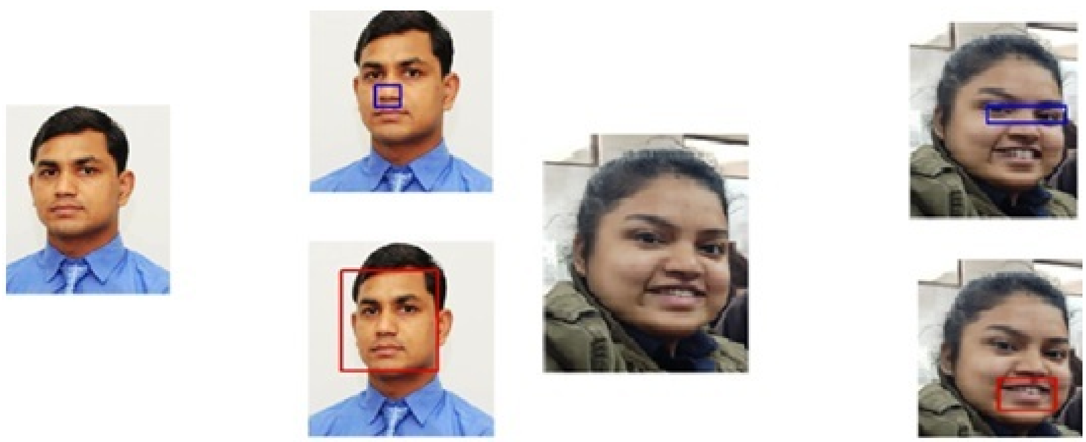

Figure 1: Nose, Face, Eyes and Mouth detection.

3.6 Photography

In photography, face detection plays a vital role as one has to detect the face first and then he/she will be able to capture an image. Photo automatic classification system, which is based on face recognition, is convenient for users to find their own photos. Given a photo of a person, firstly the system indulges in self-deep learning to obtain the feature value of the face and then the system can find all the photos in the given folder. It can also copy the found photos including the sample face to the corresponding person name folder. At the same time, photos that cannot be recognized are also copied to the undetected face folder, so that the user can do a small amount of manual classification [18].

4 Techniques Used For Face Detection

: Face detection can also be regarded as a specific case of object-class detection. In object-class detection, the task is to find the location and size of all objects in an image that belong to a given class. Examples include pedestrians, cars etc.

4.1 Approach Based On Feature

4.1.1 Active Shape Models

Active shape models are statistical models of the shape of objects. They iteratively deform to fit to an example of the object in a new image, developed by [38]. The shapes are constrained by the Point Distribution Model, Statistical Shape Model to vary only in ways seen in a training set of labeled examples. The shape of an object is represented by a set of points that are controlled by the shape model. The Active Shape Model algorithm aims to match the model to a new image. It is classified into three groups i.e. Snakes, Point Distribution Model and deform-able templates.

4.1.2 Active Contour Model-Snakes

Energy minimizing models have a vast history in vision going back at least to Sperlings stereo Model. Snakes are active contour model, used in the domain of image processing to locate contour i.e. parametric curve, which aims to reduce internal energy by moving into a local minimum [19]. The evolution of Snake model is achieved by minimizing an energy function, Esnake, which can be defined as the addition of the three energies.
Esnake = Einternal + Eexternal + Econstraint
These energy terms are defined wisely resulting the final position of the contour will possess a minimum energy. Tracking motion, recognizing object shape, edge detection etc. are some common applications of snakes.

4.1.3 Point Distribution Models

A point distribution model is the collection of those points which describes an object. Point distribution model represent shapes and detect them using active shape model in 2D image. The point distribution model is a powerful shape description technique that may subsequently be used in locating new instances of such shapes in other images [20]. It is most useful for describing features that have well understood general shape, but which cannot be easily described by a rigid model.

4.1.4 Deformable Template

To detect and describe features likes eyes, nose, eyebrows etc. is an important aspect of a face recognition system. Deformable template is the one of the method proposed to detect such features, patialized by a set of guidelines which enables allied information about the expected shape by the features to guide the detection process [21]. The templates are quite adjustable to change its parameters. The deformable template are similar upto some extent to snakes. To extract properties of the image such as peaks, valleys etc. where intensity changes instantly, deformable templates act on three representations of the image along with the image itself. Also, deformable template provides description of features to classify and to match with a database.

4.2 Low Level Analysis

4.2.1 Color Space or Color Models

A variety of color spaces or color models have been suggested and each one of them having a specific color coordinate and each point in the color space representing only a specific color [39],[40][41]. Color images, generated by a digital imaging system, are represented as red, green and blue and are called RGB images. An RGB color image, containing 8-bits of R, G and B pixels, has 2563 or 16,777,216 colors. Other color spaces, similar to RGB, are CMYK, HSV, HIS or LUV, etc.

4.2.2 Uniform Color Space

Uniform color space is defined as the Euclidean color distance between two color points at any part of the color space corresponding to the perceptual difference between the two colors by the human vision system. While in non-uniform color space, two different colors at a distanced apart in one part of the color space do not exhibit the same degree of perceptual difference as two other colors at the same distance apart in other color space. Uniform color images are of great importance in imaging applications. RGB and CMYK color models, being useful in many ways, are in no way similar to the human vision model. One of the major limitations of RGB color space is that it is non-uniform.

4.2.3 Colors and Hues

Color is an attribute of visual perception and hue is that of a human perception. Color can be described by names like red, green, blue, etc. and hue is described as red, green, blue and yellow primary hues or any intermediate combinations of primary hues. Example, a blue car reflects blue hue. Moreover it is an attribute of the human perception. The chromatic component of our perception, hue, can be considered as weak hue or strong hue. The saturation component determines the colorfulness of a color. For example, the color from a single monochromatic source of light, producing colors of a single wavelength only, is highly saturated, while the colors having hues of different wavelengths have less saturation. The gray colors have no hue and are unsaturated. The lightness (L) or intensity (I) or value (V) provides measure of the brightness of the color. The intensity helps human eye to perceive color. Another method of expressing colors in an image is Principal Component Transform (PCT).

4.2.4 Skin-Color Base

As we know that color is an optimum feature of human faces. So it serves many advantages for tracking a face. Color processing is faster as compared to other facial feature. But under certain lightning conditions, color is orientation invariant and it makes motion estimation easier. The reason being that only a translation model is needed for model estimation. Using color to track faces is influenced by many factors like ambient light, object movement etc. and [22] suggested simplest skin color algorithm for detecting skin pixels. Pixels for skin region can be detected using normalized color histogram. When we convert [R, G, B] vector into an [R, G] vector of normalized color provides a fast processing of skin detection. But unfortunately this algorithm failed as there was more skin region like legs, arms etc. Later YC_bCr color space was also introduced. Few researchers noticed that pixels belonging to skin region have similar (band Cr values. So if the values of [Cb, G] fall within the thresholds chosen as [G1, G2] and [Cb1, Cb2], a pixel is classified to have skin tone. But this algorithm has a limitation that the image should be having only face as the skin region [23]. A color predicate in HSV color space to separate skin regions from background has been defined. Although the skin color classification by HIS color space but here the values are hue (h) and saturation (s). Similarly, a pixel is classified to have skin tone, if the values [H,S] fall under the threshold chosen [[H1, S1] and [H2, S2]. Based on RGB, YbCr and HIS color space models, three different face algorithm are available and for implementing these algorithms, there are basically three main steps required: one is classify the skin region in the color space second is apply threshold to mask the skin region and draw bounding box to extract the face image.

4.2.5 RGB Color Model

RGB colors are specified as three primary colors i.e. Red (R), Green (G), and Blue (B). Among other methods of face detection, the normalized RGB has been used because this color space is more suitable for skin detection and facial feature detection[17]. The colors red, green, and blue were taken into consideration because each one corresponds roughly with one of the three types of color-sensitive cones in the human eye. It is one of the most widely used colors paces for processing and storing of digital image data [24]. The main purpose of the RGB color model is to sense, represent, and display images in electronic systems, such as televisions and computers, and also in conventional photography [25]. Appendix-1(e) is the MATLAB code for separating the red, green and blue components of an image (Figure 2).

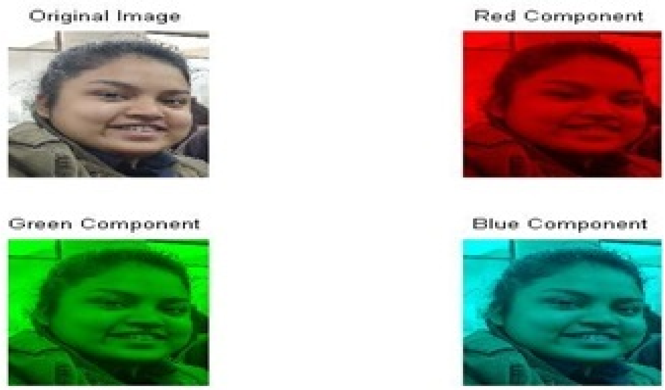

Figure 2: RGB Components of an image.

4.2.6 HSV Color Model

HSV or HSI describes colors as perceived by human beings. HSI (or HSV) stands for hue (H), saturation(S), and intensity (I) or (value V). HSV color model is, basically, cylindrical representation of RGB color model. HSV is specified as hue, saturation and value. In each cylinder, the angle around the central vertical axis corresponds to "hue" as it forms the basic pure color of the image, the distance from the axis corresponds to "saturation" or we can say when white color and black color is mixed with pure color it forms the two different form "tint" and "shade" respectively, and the distance along the axis corresponds to "lightness", "value" or "brightness" as it is providing for an achromatic notion of the intensity of the color or brightness of the color [25]. H varies from 0 -1 in a circular scale, which means H=0 and H=1 represent the same color. S varies from 0 to 1, here 1 represents 100 percent purity of color. H and S scales are partitioned into 100 levels and using H and S, a color histogram is formed. The Hue component describes color in the form of an angle ranging from $[0, 360]$ degrees. The Saturation component determines the amount of color polluted by the white color and ranges from $[0, 1]$. and Value gives us information about the brightness of the color [26].

4.2.7 RGB to HSV Color Space Transformation

The HSV image may be computed from RGB using different forms of transformations. Some of them as are follows: The simplest form of HSV transformation is $H = tan(\frac{3(G-B)}{(R-G)+(R-B)})$, $S = 1 - (\frac{min(R,G,B)}{V})$ where $V = \frac{R+G+B}{3}$. However, the hue (H) becomes undefined when saturation $S = 0$. The most popular form of HSV transformation is shown below, where the r, g, b values are first obtained by normalizing each pixel such that $r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}$ and $b = \frac{B}{R+G+B}$.Accordingly, the H, S and V values can be computed by the equations (1)–(4):

$$V = \max(r, g, b), \tag{1}$$

$$S = \begin{cases} 0 & \text{if } V = 0, \\ V + \frac{min(r,g,b)}{V} & \text{if } V > 0, \end{cases} \tag{2}$$

$$H = \begin{cases} 0 & \text{if } S = 0, \\ \frac{60 \times (g-b)}{S \times V} & \text{if } V = r, \\ 60 \times [2 + \frac{(b-r)}{S \times V}] & \text{if } V = g, \\ 60 \times [4 + \frac{(r-g)}{S \times V}] & \text{if } V = b, \end{cases} \tag{3}$$

and

$$H = H + 360 \text{ if } H < 0. \tag{4}$$

The results we get, after using either of the above transformations, yield reasonably good results.

4.2.8 CIELAB Color Model

$CIE - L^*a^*b^*$ (CIELAB) is specified as the most complete color space by the International Commission on Illumination. The CIELAB color space, which was adopted as an international standard in the 1970s CIE is indeed a perceptually uniform space. In the CIELab color space, the Euclidean distance between two color points corresponds to the perceptual difference between the two colors by the human vision system. This property of the CIELAB color space has made it particularly attractive and useful for color analysis, and the superiority of the CIELAB color space over other color spaces has been demonstrated in many color image applications. For example the CIELAB color space has been successfully used for color clustering. In this approach the color difference in the CIELAB color space has been used in the computation of the dissimilarities between colors and this has been used for color clustering. The color difference formula in the CIELAB color space is used in the computation of the dissimilarities between colors and the formulation of the color membership function. The three coordinates of CIELAB or $CIE - L^*a^*b^*$ represent the lightness of the color such as L^* for the lightness form black (0) to white (100), a^* from green (negative) to red (positive) and b^* from blue (negative) to yellow (positive). The $CIE - L^*a^*b^*$ color space is based on the concept that colors can be considered as combinations of red and yellow, red and blue, green and yellow, and green and blue [31].

4.2.9 RGB to CIELAB Color Space to Transformation

RGB to CIELAB transformation can be performed such as [39, 41]:

$$X = 0.412453R + 0.357580G + 0.180423B$$

$$Y = 0.212671R + 0.715160G + 0.072169B$$

$$Z = 0.019334R + 0.119193G + 0.950227B$$

And $L^*a^*b^*$ is defined as $L^* = 116f(\frac{Y}{Y_n}) - 16$, $a^* = 500f(\frac{X}{X_n}) - f(\frac{Y}{Y_n})$ and $b^* = 200f(\frac{Y}{Y_n}) - f(\frac{Z}{Z_n})$ where

$$f(t) = \begin{cases} t^{\frac{1}{3}} & \text{if } t > 0.008856, \\ 7.787q + \frac{16}{116} & \text{otherwise,} \end{cases}$$

and X_n, Y_n and Z_n are represent the CIE(International Commission on Illumination)-XYZ values of the reference white point [44]. Under illuminant D_{65} with normalization $Y = 100$, the values are $X_n = 95.0489, Y_n = 100$ and $Z_n = 108.8840$, and are obtained by setting $R = G = B = 100(q \in \frac{X}{X_n}, \frac{Y}{Y_n}, \frac{Z}{Z_n})$.

4.2.10 Cyan, Magenta, Yellow and Black (CMYK) Color Model

CMYK color model finds utility in color printers. Most of the output devices use this model. The primary additive colors are red, green and blue. While the primary colors of pigments are magenta, cyan and yellow and the corresponding secondary colors are red, green and blue. The conversion from RGB to CMY is performed as:

$$\begin{pmatrix} C \\ M \\ Y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

where R, G, B represent the normalized color values in the range 0 to 1.

4.2.11 NTSC or YIQ Color Model

In this color model the luminance information Y represents the gray scale information, while hue (I) and saturation (Q) carry the color information. The conversion from RGB to YIQ is

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{pmatrix} - \begin{pmatrix} R \\ G \\ B \end{pmatrix}. \quad (5)$$

The elements of the first row when added become unity and the elements in the second and third row sum to 0. Thus in equation (5) a gray scale image, color components I and Q are zero. The NTSC color space is widely used in television.

4.2.12 YC_bC_r Color Mode

In this model, Y channel is the luminance component of an image, in other words it captures the brightness information of every pixel in an image while C_b and C_r are devoted to capture color information (chrominance). In YC_bC_r space, the distribution of C_b and C_r components of facial color tends to be consistent. It can effectively remove the illumination brightness effect of Y, and it also has a good clustering feature [27]. This color space is used in digital video. The information from RGB to YC_bC_r, is as follows:

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.00 \\ 112.00 & -93.786 & -18.214 \end{pmatrix} \times \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 16 \\ 128 \\ 128 \end{pmatrix}. \quad (6)$$

Equation (6) shows that the three components of YC_bC_r can be easily calculated by linearly combinations of R, G and B components of image. In order to get the skin region, it must satisfy the following equations $-140 \leq C_r \leq 165$ and 140Cb 195 $140 \leq C_b \leq 195$ [28]. Skin color is mainly determined by the darkness or fairness of the skin. The difference in brightness of color mainly determines the Y component rather than C_b and C_r components [29]. In order to get skin regions, some restrictions are made on these two components and Hue. Hence, Skin color can be easily detected by Chrominance and luminance color due to all this property and its simplicity [30]. Refer Appendix-1(f) for the MATLAB code of separation of YC_bC_r components of an image (figure 3).

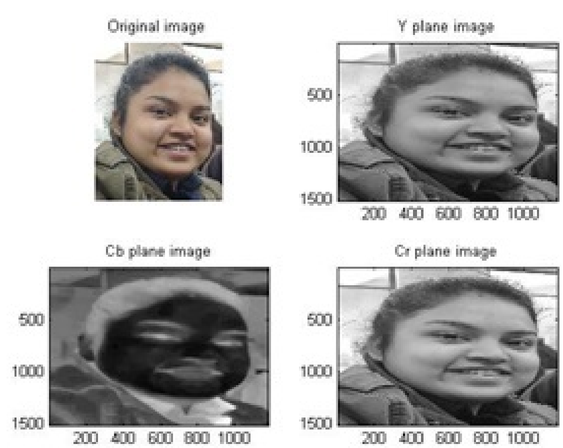

Figure 3: YCbCr components separated in an image.

4.3 Feature Analysis

These algorithms focus on the detection of structural features that even exists if the viewpoint or lighting conditions varies, and then locate them on faces. The main purpose to design these methods is for face localization.

4.3.1 Feature Searching

Viola and Jones presented an approach for object detection which minimizes computation time while achieving high detection accuracy [45]. Viola and Jones proposed a fast and robust method for face detection which is 15 times quicker than existing techniques at the time of release with 95% accuracy [34]. This techniques has a very low false positive rate and but can detect black faces.

4.3.2 Local Binary Pattern

This technique is used to describe the image texture features [35]. It widely uses in the fields of image retrieval, texture examination, face recognition, image segmentation etc. because of high speed computation and rotation invariance but it is insufficient for non-monotonic illumination changes and only used for binary and grey images.

4.3.3 Ada Boost Algorithm

boosting is an ensemble method for improving the model predictions of any given learning algorithm [36]. The Ada boost algorithm was the first practical boosting algorithm. Ada boost is more resistant to over fitting than many machine learning algorithms [37]. A classifier is a hypothesis valued function. Ada boost creates the strong learner (a classifer that is well correlated to the true classifier) by iteratively adding weak learner (a classifier that is only slightly correlated to the true classifier).

5 Comparing Some of the Color Models

5.1 Comparison Between YC_bC_r Color Space and CIELab Color Space Based on Skin Color Segmentation

Skin color is a useful feature for human face detection. In this section, we compare two color spaces YC_bC_r and CIELab. CIELab color space is considered better than other spaces as it can improve the performance of face segmentation under poor or strong lighting conditions. YCbCr color space doesn't depend on luminance. YCbCr can be constructed as weighted sum of RGB values whereas CIELab is perceptual uniform color space. Perceptual uniformity means how two colors differ to see when human beings observe them. And CIELab doesn't depend on device.

Sometimes due to different background and difference in light conditions, advanced technique is less efficient and also not efficient if the user would move.

Performance is decreases if there is variation in illumination, background and user variance [32].

5.2 Comparison Between RGB Color Space and HSV Color Space

RGB color space defines colors as how much red, green and blue is present in it. On the other hand, HSV color space defines colors in terms of Hue, Saturation and Value. In situations where color description plays a vital role, the HSV color model is often preferred over the RGB color model. This is so because the HSV model describes colors similar to how the human eye tends to perceive color. RGB defines color as combination of primary colors, whereas, HSV describes color by using more familiar comparisons, like color, vibrancy and brightness. The color camera used on the robot, uses the RGB model to describe color. After reading the values, camera converts them into HSV values. These values are then used in the code to determine the location of a specific object/color for which the robot is searching [33].

6 Conclusion

Image processing is used to enhance the picture quality, detect people, finding errors, if any, etc. This paper discusses various image processing applications like gender identification, biometric attendance, study of medical images, face feature recognition, photography etc. We have studied some MATLAB codes, also attached with this paper, which are useful for face detection, recognition and application in various fields such as photography. Point Distribution Model represents the mean geometry of a shape and relies on landmark points. Deformable Template provides description of features to classify and to match with a database. We have discussed various color models like RGB that is considered most suitable to detect facial features and skin color, HSV which tells us about the brightness and saturation in image, CIELab describes all the colors that are visible to human eye, CMYK is widely used in color printers, NTSC gives us the gray scale information and YC_bC_r tells us about the luminance and chrominance components of an image because these models help us to study an image with a different viewpoint and also in detecting if there are any defaults in the image or not. Then, we have transformed RGB color model to another, i.e., RGB to HSV, to CIELab, to CMYK, to CMYK, to NTSC and to YC_bC_r, using mathematical transformations and compared YC_bC_r to CIELab and RGB to HSV to better understand the difference between these color models.

Acknowledgment. The author(s) express their heartiest thanks and gratefulness to all those scientists whose masterpieces have been consulted during the preparation of the present research article. We are also thankful and grateful to Professor Dr. Sui Sun Cheng editor in chief AMEN to encourage us and also to give valuable comments for improvement of this article.

References

[1] K. Delac, M. Grgic and S. Grgic, Statistics in face recognition: analyzing probability distributions of PCA, ICA and LDA performance results, In ISPA 2005. Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, (2005), 289–294.

[2] B. C. Duchaine and K. Nakayr, Developmental prosopagnosia: a window to content-specific face processing, Current opinion in neurobiology, 16(2006), 166–173.

[3] L. M. Oberman, P. Winkielman and V. S. Rrchandran, Face to face: Blocking facial mimicry can selectively impair recognition of emotional expressions, Social neuroscience, 2(2007), 167–178.

[4] C. A. Meissner, S. L. Sporer and K. J. Susa, A theoretical review and meta-analysis of the description-identification relationship in memory for faces, European Journal of Cognitive Psychology, 20(2008), 414–455.

[5] R. A. Johnston and A. J. Edmonds, Familiar and unfamiliar face recognition: A review, Memory, 17(2009), 577–596.

[6] J. M. Chaves-Gonzlez, M. A. Vega-Rodrguez, J. A. Gmez-Pulido and J. M. Snchez-Prez, Detecting skin in face recognition systems: A colour spaces study, Digital Signal Processing, 20(2010), 806–823.

[7] R. Jenkins, D. White, X. Van Montfort and A. M. Burton, Variability in photos of the same face, Cognition, 121(2011), 313–323.

[8] B. F. Klare and A. K. Jain, Heterogeneous face recognition using kernel prototype similarities, IEEE transactions on pattern analysis and machine intelligence, 35(2012), 1410–1422.

[9] J. R. Barr, K. W. Bowyer, P. J. Flynn and S. Biswas, Face recognition from video: A review, International Journal of Pattern Recognition and Artificial Intelligence, 26(2012), 1266002.

[10] W. X. Wang, W. S. Li and X. Yu, Fractional differential algorithms for rock fracture images, The Imaging Science Journal, 60(2012), 103–111.

[11] P. Marrero-Fernndez, A. Montoya-Padrn, A. Jaume-i-Cap and J. M. Buades Rubio, Evaluating the research in automatic emotion recognition, IETE Technical Review, 31(2014), 220–232.

[12] R. K. Sangha and P. Rai, An Appearance-Based Gender Classification Using Radon Features, In Data, Engineering and Applications, Springer, Singapore, 2019.

[13] G. Kaur, G. Singh and V. Kumar, A review on biometric recognition, International Journal of Bio-Science and Bio-Technology, 6(2014), 69–76.

[14] M. T. S. K. Kalihal Asst. and J. Kaur, A Review on Different Face Spoof Detection Techniques in Biometric Systems, International Journal of Scientific Research & Engineering Trends, 5(2019), issue 2, 609–611.

[15] J. Ruiz-del-Solar, R. Verschae, P. A. Vallejos and M.Correa, Face analysis for human computer interaction applications, In VISAPP (Special Sessions), (2007), 23–30.

[16] Y. M. Y. Abdallah and T. Alqahtani, Research in Medical Imaging Using Image Processing Techniques, In Medical Imaging-Principles and Applications, IntechOpen, 2019.

[17] R. Subban and R. Mishra, Rule-based face detection in color images using normalized RGB color space-A comparative study, In 2012 IEEE International Conference on Computational Intelligence and Computing Research, (2012), 1–5.

[18] H. Huang and Y. Huang, Photo Automatic Classification System Based on Face Recognition, In Journal of Physics: Conference Series (Vol. 1168, No. 4, p. 042007), 2019.

[19] M. Kass, A. Witkin and D. Terzopoulos, Snakes: Active contour models, International journal of computer vision, 1(1988), 321–331.

[20] A. Kanaujia, Y. Huang and D. Metaxas, Tracking facial features using mixture of point distribution models, In Computer Vision, Graphics and Image Processing, (2006), 492–503.

[21] A. K. Jain, Y. Zhong and M. P. Dubuisson-Jolly, Deformable template models: A review, Signal processing, 71(1998), 109–129.

[22] J. L. Crowley and J. Coutaz, Vision forman machine interaction, Robot Auton Syst., 19(1997), 34758.

[23] R. Kjeldsen and J. Kender, Finding skin in color images, In: Proceeding of the 2nd international conference on automatic face and gesture recognition, (1996), 312–317.

[24] Z. S. Tabatabaie, R. W. Rahmat, N. I. B. Udzir and E. Kheirkhah, A hybrid face detection system using combination of appearance-based and feature-based methods, International Journal of Computer Science and Network Security, 9(2009), 181–185.

[25] Y. Tayal, R. Lamba, and S. Padhee, Automatic face detection using color based segmentation, International Journal of Scientific and Research Publications, 2(2012), 1–7.

[26] H. F. Hashem, Adaptive technique for human face detection using HSV color space and neural networks, In 2009 National Radio Science Conference, (2009), 1–7.

[27] A. Verma, S. A. Raj, A. Midya and J. Chakraborty, Face detection using skin color modeling and geometric feature, In 2014 International Conference on Informatics, Electronics and Vision (2014), 1–6.

[28] J. Fan and H. Wang, A method of face detection based on improved YCBCR skin color model, In 2016 4th International Conference on Machinery, Materials and Computing Technology. Atlantis Press, 2016.

[29] M. M. Hasan, J. M. Thakur and P. Podder, Face Detection in Color Images Using Skin Color, International Journal of Scientific and Engineering Research, 5(2014).

[30] S. Chitra and G. Balakrishnan, Comparative study for two color spaces HSCbCr and YCbCr in skin color detection, Applied Mathematical Sciences, 6(2012), 4229–4238.

[31] R. Khan, K. Zeeshan, A. Muhammad and Q. S. Syed, Static filtered skin detection, International Journal of Computer Science, 9(2012), 257–261.

[32] P. Ganesan, V. Rajini and R. I. Rajkumar, Segmentation and edge detection of color images using CIELAB color space and edge detectors, In INTERACT-2010, (2010), 393–397.

[33] A. Kaur and B. V. Kranthi, Comparison between YCbCr color space and CIELab color space for skin color segmentation, International Journal of Applied Information Systems, 3(2012), 30–33.

[34] H. Zhang, Y. Xie and C. Xu, A classifier training method for face detection based on AdaBoost, In: Proceeding of international conference on transportation, mechanical, and electrical engineering, (2011), pp 731–734.

[35] T. Ahonen, A. Hadid and M. Pietikinen, Face recognition with local binary patterns, In European conference on computer vision, (2001), 469–481.

[36] L. Y. Lang and W. W. Gu, Improved face detection algorithm based on adaboost, In 2009 International Conference on Electronic Computer Technology, (2009), 183–186.

[37] Y. Hou and Q. Peng, Face detection based on adaboost and skin color, In 2009 Second International Symposium on Information Science and Engineering, (2009), 407–410).

[38] T. F. Cootes and C. J. Taylor and D. H. Cooper and J. Graham, Active shape models-their training and application, Computer Vision and Image Understanding, (1995), 3859.

[39] G. Wyszecki and W. S. Stiles, Color Science, 2nd ed., McGraw-Hill, New York, 1982.

[40] B. A. Wandell, Foundations of Vision, Sinauer Associates, Inc., Sunder land, MA, 1995.

[41] H. R. Kang, Color Technology for Electronic Imaging Devices, SPIE Optical Engineering Press, 1997.

[42] S. Y. Lee, Y. K. Ham and R. H. Park, Recognition of human front faces using knowledge based feature extraction and neuro fuzzy algorithm, Pattern Recognition, 29(1996), 1863–1876.

[43] M. Turk and A. Pentland, Eigenfaces for recognition, The Journal of Cognitive Neuroscience, 3(1991), 71–86.

[44] D. A. Broadbent, A critical review of the development of the CIE-1931 RGB color-matching functions, Color Research and Applications, 29(2004), 267–272.

[45] P. Viola and M. Jones, Robust real-time object detection, International journal of computer vision, 4(2001), 4.

7 Appendix-1

(a) Nose Detection MATLAB code:

```
%To detect Nose
clear all
clc
%Detect objects using Viola-Jones Algorithm
%To detect Face
FDetect = vision.CascadeObjectDetector;
%Read the input image
I = imread('C:\Users\computer\Desktop\DSC_0109.JPG');
I1= imresize(I,0.5);
NoseDetect = vision.CascadeObjectDetector('Nose','MergeThreshold',16);
BB=step(NoseDetect,I1);
figure,
imshow(I1); hold on
for i = 1:size(BB,1)
    rectangle('Position',BB(i,:),'LineWidth',4,'LineStyle','-','EdgeColor','b');
end
title('Nose Detection');
hold off;
```

(b) Face Detection MATLAB code:

```
clear all
clc
%Detect objects using Viola-Jones Algorithm
%To detect Face
FDetect = vision.CascadeObjectDetector;
%Read the input image
I = imread('C:\Users\computer\Desktop\DSC_0109.JPG');
I1= imresize(I,0.5);
%Returns Bounding Box values based on number of objects
BB = step(FDetect,I1);
figure,
imshow(I1); hold on
for i = 1:size(BB,1)
    rectangle('Position',BB(i,:),'LineWidth',5,'LineStyle','-','EdgeColor','r');
end
title('Face Detection');
```

(c) Eye Detection MATLAB Code

```
%To detect Eyes
EyeDetect = vision.CascadeObjectDetector('EyePairBig');
%Read the input Image
BB=step(EyeDetect,I1);
figure,imshow(I1);
rectangle('Position',BB,'LineWidth',4,'LineStyle','-','EdgeColor','b');
title('Eyes Detection');
Eyes=imcrop(I1,BB);
figure,imshow(Eyes);
```

(d) Mouth Detection MATLAB Code:

```
clear all
clc
I = imread('C:\Users\computer\Desktop\5.jpg');
I1= imresize(I,0.5);
%To detect Mouth
MouthDetect = vision.CascadeObjectDetector('Mouth','MergeThreshold',100);
BB=step(MouthDetect,I1);
figure,
imshow(I1); hold on
for i = 1:size(BB,1)
 rectangle('Position',BB(i,:),'LineWidth',4,'LineStyle','-','EdgeColor','r');
end
title('Mouth Detection');
hold off;
```

(e) MATLAB Code for extracting RGB components in an image:

```
b=uigetfile('C:\Users\computer\Desktop\5.jpg');
b=imread(b);
[row col dim]=size(b);
Red=b;
Blue=b;
Green=b;
Red(:,:,2:3)=0;
figure;
subplot(2,2,1)
imshow(b)
title ('Original Image');
Green(:,:,1)=0;
Green(:,:,3)=0;
subplot(2,2,2)
imshow(Red);
title ('Red Component');
subplot(2,2,3)
imshow(Green);
title ('Green Component');
Blue(:,:,1)=0;
Green(:,:,2)=0;
subplot(2,2,4)
imshow(Blue);
title ('Blue Component');
```

(f) MATLAB Code for separting Y-Cb-Cr component of an image

```
a=imread('C:\Users\computer\Desktop\5.jpg');
a=double(a);
[row col dim]=size(a);
red=a(:,:,1);
green=a(:,:,2);
blue=a(:,:,3);
Y=16+((65.481*red)+(128.553*green)+(24.966*blue));
Cb=128+((-37.797*red)-(74.203*green)+(112*blue));
Cr=128+((112*red)+(93.786*green)-(18.214*blue));
```

```
figure(1);
subplot(2,2,1)
imshow(uint8(a))
colormap(gray)
title ('Original image');
subplot(2,2,2);
imagesc(Y) %The image function scales image data to the full range of the current colormap and displays
colormap(gray);
title ('Y plane image');
subplot(2,2,3);
imagesc(Cb);
colormap(gray);
title ('Cb plane image');
subplot(2,2,4);
imagesc(Cr);
colormap(gray);
title ('Cr plane image');
```