Advanced Hybrid Color Space Normalization for Human Face
Extraction and Detection

Jayakrishna.V¹ Akhila G.P.² Shafeena Basheer³
¹,²Faculty ³PG Student
¹,³Amal Jyothi College of Engineering, Kanjirappally
²UKF College of Engineering &Technology,Paripally

Abstract— This paper presents a new color space normalization (CSN) technique for enhancing the discriminating power of color space along with the principal component analysis (PCA) for the face recognition process. The common RGB technique is not suitable for the characterizing of the skin color due to the presence of luminance factor. In the YCbCr color space, the luminance information is contained in Y component, and the chrominance information is in Cb and Cr. Therefore, the luminance information can be easily de-embedded. Different color spaces have different discriminating power, in this paper, eye can be perfectly detected by using YcbCr color space and the mouth regions can be perfectly detected by using the YIQ color space. Then PCA is used to express the large 1-D vector of pixels constructed from 2-D facial image into the compact principal components of the feature space. Each face image may be represented as a weighted sum (feature vector) of the eigenfaces, which are stored in a 1D array. PCA allows us to compute a linear transformation that maps data from a high dimensional space to a lower dimensional space. It covers standard deviation, covariance, eigenvectors and eigenvalues. Face recognition is obtained by PCA without much loss of information. Experiments using different databases by varying the facial expressions (open/closed eyes, smiling/not smiling) show that the proposed method by combining color space discrimination and PCA can improve face recognition to a great extend.

Keywords: Color space, Color Space normalization, Color Model, Principal Component Analysis

I. INTRODUCTION

Face recognition is a very active research area as evidenced by the large number of publications in the journals and conferences of computer vision and pattern recognition. Recently scientists suggest that researchers should concentrate on "face recognition problems that are harder, as defined by the image sets in the experiments and the performance by a control algorithm" rather than work on problems that have already been solved. There is a Principal Component Analysis (PCA) algorithm that has been optimized for large scale problems. This project presents a method that applies color configurations in the YIQ and the YCbCr color spaces to improve face recognition performance. Color provides an important clue or useful feature for object detection, tracking and recognition, image (or video) segmentation, indexing and retrieval etc. Different color spaces (or color models) possess different characteristics and are suitable for different visual tasks. For instance, the HSV color space and the YCbCr color space are effective for face detection, while the modified $L^a*b^v$ color space is useful for image segmentation. As a result, when applying color information, we should first choose an appropriate color space, and such a choice is very important for achieving the best result for a specific visual task.

The RGB color space is a fundamental and widely used color space, and other color spaces (or color models) are usually defined by transformations of the RGB color space. The transformations involved are either linear or nonlinear. The color spaces generated via the nonlinear transformations (of the RGB color space), such as the HSV and $L^a*b^v$ color spaces, generally associate with the human vision system, while the color spaces determined by the linear transformations, such as the YUV and YIQ color spaces usually associate with color display of some hardware (eg: television and color monitors) for adapting to human color-response characteristics.

Although color has been demonstrated helpful for face detection and tracking, some past research suggest color appears to confer no significant face recognition advantage beyond the luminance information. Recent research efforts, however, reveal that color may provide useful information for face recognition. The experimental results show that the principal component analysis (PCA) method using color information can improve the recognition rate compared to the same method using only luminance information. The results reveal that color cues do play a role in face recognition and their contribution becomes evident when shape cues are degraded. The results further demonstrate that color cues can significantly improve recognition performance compared with intensity-based features for copying with low-resolution face images. Other research findings also demonstrate the effectiveness of color face recognition. Different color spaces derived from different transformations of the RGB color space revealed different face recognition performance. The YUV color space, for example, is shown more effective than the RGB color space. The YOQ color configuration (a hybrid color space), where the Y and Q color components are from the YIQ color space and the Cr color component is from the YCbCr color space, is more powerful than the RGB, HSV and $L^a*b^v$ color spaces. Another two hybrid color spaces, RIQ, RQC are demonstrated effective recently. Some color spaces generated by evolution algorithms and discriminant models also turn out to be very powerful. Current research findings showed that some linear color spaces, which are derived by linear transformations from the RGB color space, perform much better those derived by nonlinear transformations from the RGB color space. We therefore focus on linear color spaces in this paper. Rather than searching for a more effective color space as the previous research, we try to explore general ways for enhancing the
This paper assesses the performance of different color spaces using a large scale database. The assessment results reveal that some color spaces such as RGB, XYZ, HSV and L*a*b* color spaces are relatively weak whereas the other color spaces, such as I1I2I3, YUV, YIQ and LSLM color spaces are relatively powerful in achieving good face recognition performance. What characteristics make the I1I2I3, YUV, YIQ and LSLM color spaces more powerful than the RGB and XYZ color spaces for face recognition? By analyzing the transformation matrices of the I1I2I3, YUV, YIQ and LSLM color spaces, we find out that these matrices all share a common characteristic: the sum of the elements in the second and third rows of the transformation matrix are both zero. The RGB and XYZ color spaces, however, do not have such a property. Inspired by the finding of the difference of the transformation matrices between the weak and the powerful color spaces, we present the concept of color space normalization (CSN) and develop two CSN techniques. These CSN techniques normalize any color space that is derived by a linear transformation of the RGB color space, so that the normalized color space possesses the same properties as the powerful color spaces do, i.e., the sums of the elements in the second and third rows of the transformation matrix are both zero. The proposed two techniques are demonstrated to be very effective: the normalized RGB and XYZ color spaces are as powerful as or even more powerful than I1I2I3, YUV, YIQ and LSLM color space recognition.

The proposed CSN techniques, which are capable of converting weak color spaces into powerful ones, provide us with more flexibility for color space selection for specific pattern recognition tasks. Previous color space selection is limited to set of conventional color spaces or their hybrids. Specifically, we choose a powerful color space by experiments from the two set of hybrid color spaces that are generated by choosing some color components from the conventional color spaces. The weak color spaces are simply left behind unsatisfactory performance. The proposed color space normalization techniques, however, can convert the weak color spaces into powerful ones, and these normalized color spaces form a new set of color spaces, from which we might find a more effective color space for a specific recognition task. The three sets of color spaces are illustrated in the Fig. 1.

![Fig. 1: Illustration of three sets of color spaces.](image)

Detection of faces is a crucial step in the identification applications. Most face recognition algorithms assume that the face location is known. Similarly, face tracking algorithms often assume the initial face location is known. Face detection can be viewed as a two-class classification problem. Therefore, some techniques developed for face recognition. The remainder of the paper is organized as follows. Section 2 outlines some conventional color spaces. Section 3 presents the concept of color space normalization (CSN). In section 4, the proposed CSN techniques are assessed, and the problem of why the proposed CSN techniques can improve the face verification and recognition performance is addressed. Section 5 describes the face detection algorithm. Section 6 presents the detection results of our algorithm on several face databases. Finally the conclusions and the future works are specified.

## II. CONVENTIONAL COLOR SPACES

### A. Color Spaces

The RGB color space is a fundamental and commonly used color space. Other Color spaces can be calculated from the RGB color space by means of either linear or nonlinear transformations. It is apparent that every color space derived by the linear transformation of the RGB color space is uniquely determined by the associated transformation matrix. In the following, we review five color spaces derived from the RGB color space via linear transformations\[(1)\].

The XYZ color space was derived from a series of experiments in the study of the human perception by the International Commission on Illumination (CIE) in 1931. The transformation from the RGB color space to the XYZ is as follows:

\[
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix} =
\begin{bmatrix}
0.607 & 0.174 & 0.201 \\
0.299 & 0.587 & 0.114 \\
0.000 & 0.066 & 1.117
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

(1)

The I1I2I3 color space was obtained through the decorrelation of the RGB color components using K-L transform by Ohta et al. in 1980. The transformation from the RGB color space to the I1I2I3 color space is as follows:

\[
\begin{bmatrix}
I_1 \\
I_2 \\
I_3
\end{bmatrix} =
\begin{bmatrix}
1/3 & 1/3 & 1/3 \\
1/2 & 0 & -1/2 \\
-1/2 & 1 & -1/2
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

(2)

The YUV color space is defined in terms of one luminance (Y) and two chrominance components (U and V), and is used in the PAL (Phase Alternating Line), NTSC (National Television System Committee), and SECAM (Sequential Couleur a memoire) composite color video standards. The transformation from the RGB to the YUV color space is as follows:

\[
\begin{bmatrix}
Y \\
U \\
V
\end{bmatrix} =
\begin{bmatrix}
0.2990 & 0.5870 & 0.1140 \\
-0.1471 & -0.2888 & 0.4359 \\
0.6148 & -0.5148 & -0.1000
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

(3)

The YIQ color space was formerly used in the National Television System Committee (NTSC) television standard. The YIQ system, which is intended to take advantage of human color response characteristics, and can be derived from the corresponding RGB space as follows:

\[
\begin{bmatrix}
Y \\
I \\
Q
\end{bmatrix} =
\begin{bmatrix}
0.2990 & 0.5870 & 0.1140 \\
0.5957 & -0.2744 & -0.3213 \\
0.2115 & -0.5226 & 0.3111
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

(4)
The I and Q components in the YIQ color space are obtained via clockwise rotation (33 degree) of the U and V color components in the YUV color space.

The LSLM color space is a linear transformation of the RGB color space based on the opponent signals of the cones: black-white, red-green and yellow-blue. The LSLM color space is defined as follows:

\[
\begin{bmatrix}
L \\ S \\ LM
\end{bmatrix} =
\begin{bmatrix}
0.209 & 0.715 & 0.076 \\
0.209 & 0.715 & -0.924 \\
3.148 & -2.799 & -0.349
\end{bmatrix}
\begin{bmatrix}
R \\ G \\ B
\end{bmatrix}
\]

(5)

B. Goals and Discussions

The work concentrates mainly on finding out a good and effective method for face recognition and face extraction from a series of color spaces available and to make the effective combination of them in order to get a better result. The paper proposes a way to validate the performances of the effectiveness of the various normalization techniques for face recognition and identification in color images. Categorizing face detection methods based on the representation used reveals that detection algorithms using holistic representations have the advantage of finding small faces or faces in poor-quality images, while those using geometrical facial features provide a good solution for detecting faces in different poses. A combination of holistic and feature-based approaches is a promising approach to face detection as well as face recognition. Motion and skin-tone color are useful cues for face detection. However, the color-based approaches face difficulties in robustly detecting skin colors in the presence of complex background and different lighting conditions. We propose a face detection algorithm that is able to handle a wide range of variations in static color images, based on a lighting compensation technique and a nonlinear color transformation. Our approach models skin color using a parametric ellipse in a two-dimensional transformed color space and extracts facial features by constructing feature maps for the eyes, mouth, and face boundary[4].

III. COLOR SPACE NORMALIZATION TECHNIQUES

A. Concept and Techniques

Different color spaces usually display different discriminating power, and our experiments on a large scale face recognition data base problem reveal that some color spaces, such as the RGB and XYZ color spaces, are relatively weak, where as other color spaces, such as the I1I2I3, YUV, YIQ and LSLM color spaces, are relatively powerful. What characteristics make the I1I2I3, YUV, YIQ and LSLM color spaces more powerful than RGB and XYZ color spaces for recognition? By analyzing the transformation matrices of the I1I2I3, YUV, YIQ and LSLM color spaces, we find out that these matrices all share a common characteristic: the sums of the elements in the second and third rows of the transformation matrix are both zero. The RGB and XYZ color spaces, however, do not have such a property.

The transformation matrix of the RGB color space is an identity matrix:

\[
\begin{bmatrix}
R \\ G \\ B
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
R \\ G \\ B
\end{bmatrix}
\]

(6)

The CSN techniques normalize any color space that is derived by a linear transformation of the RGB color space, so that the normalized color space possesses the same property as the powerful color spaces do, i.e., the sums of the elements in the second and third rows of the transformation matrix are both zero.

B. Within-color-component normalization

To achieve the goal that the sums of the elements in the second and third rows of the color space transformation matrix are zero, the within-color-component normalization technique works by directly removing the means of the second and third row vectors, respectively. Let \( C_1 \), \( C_2 \) and \( C_3 \) be the three color components derived by the following linear transformation of RGB color space:

\[
\begin{bmatrix}
C_1 \\ C_2 \\ C_3
\end{bmatrix} =
\begin{bmatrix}
A_{11} & A_{12} & A_{13} \\
A_{21} & A_{22} & A_{23} \\
A_{31} & A_{32} & A_{33}
\end{bmatrix}
\begin{bmatrix}
R \\ G \\ B
\end{bmatrix}
\]

(7)

The mean of the second row vector of the transformation matrix \( A \) is \( m_2 = (a_{21} + a_{22} + a_{23})/3 \) and the mean of the third row vector is \( m_3 = (a_{31} + a_{32} + a_{33})/3 \). Removing \( m_2 \) from the second row vector and \( m_3 \) from the third row vector, we obtain a normalized transformation matrix \( A_1 \), which determine the normalized color space: \( C = C_1C_3 \).

\[
\begin{bmatrix}
C_1 \\ C_2 \\ C_3
\end{bmatrix} =
\begin{bmatrix}
A_{11} & A_{12} & A_{13} \\
A_{21} & A_{22} & A_{23} \\
A_{31} & A_{32} & A_{33}
\end{bmatrix}
\begin{bmatrix}
R \\ G \\ B
\end{bmatrix} - \begin{bmatrix} m_2 \end{bmatrix}
\]

(8)

The within-color-component normalization technique is named color space normalization1 (CSN-1). The normalized RGB color space using CSN-1 is

\[
\begin{bmatrix}
R \\ G \\ B
\end{bmatrix} =
\begin{bmatrix} 1 & 0 & 0 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3
\end{bmatrix}
\begin{bmatrix}
R \\ G \\ B
\end{bmatrix}
\]

(9)

The normalized XYZ color space using CSN-1 is

\[
\begin{bmatrix}
\tilde{X} \\ \tilde{Y} \\ \tilde{Z}
\end{bmatrix} =
\begin{bmatrix}
0.6070 & 0.2470 & 0.9200 \\ -0.0343 & 0.2537 & -0.2193 \\ -0.3940 & 0.3280 & 0.7220
\end{bmatrix}
\begin{bmatrix}
R \\ G \\ B
\end{bmatrix}
\]

(10)

C. Across-color-component normalization

To make the sums of these elements in the second and third rows of the color space transformation matrix zero, the across-color-component normalization technique works in the following way. The original three row vectors of the color space transformation matrix are first used to generate two zero-mean row vectors via a linear combination. A new color space transformation matrix is then obtained by replacing the second and third row vectors of the original transformation matrix with the generated two zero-mean row vectors. The linear combination of the three row vectors of the original color space transformation matrix \( A \) may be written as follows:

\[
\xi = K_1A_1 + K_2A_2 + K_3A_3 = \\
(\Sigma_{i=1}^{3} k_i )a_{1i} , (\Sigma_{i=1}^{3} k_i )a_{2i} , (\Sigma_{i=1}^{3} k_i )a_{3i}
\]

(11)

Let the sum of the elements of this linear combination vector \( \xi \) (row vector) be zero, i.e.
Face detection or image retrieval is carried out based on the spatial arrangement of pixel gray values on Markov random fields and Markov chains, making use of frontal faces in real-time applications. Face detectors based on training examples and are designed primarily to locate approaches require a large number of face and nonface template matching, Hough transform, motion extraction, and information theory, geometrical modeling, (deformable) These approaches utilize techniques such as principal

\[ \sum_{i=1}^{3} k_i a_{i1} + \sum_{j=1}^{3} k_j a_{j2} = \sum_{i=1}^{3} k_i + a_{i1} \sum_{j=1}^{3} k_j + a_{j2} = [s_1, s_2, s_3][k_1, k_2, k_3]^T = 0 \] 

Where \( S_i = \sum_{i=1}^{3} a_{ij} \), obviously, \( s_i \) is the sum of the elements of the \( i \) th row vector of the color space transformation matrix \( A \).

The previous equations show that the linear combination coefficient vector \( [k_1, k_2, k_3]^T \) can be chosen as the basis vectors of the null space of \( [s_1, s_2, s_3] \). Since this null space is two-dimensional, it has only two basis vectors. Let the two basis vectors be \( k_i = [k_{i1}, k_{i2}, k_{i3}] \).

The normalized color space transformation matrix is defined as follows:

\[ \check{A}_{II} = \begin{bmatrix} A_{11} & \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{bmatrix} \] 

which determines the following normalized color space \( C_1, C_2, C_3 \):

\[ \begin{bmatrix} \check{C}_1 \\ \check{C}_2 \\ \check{C}_3 \end{bmatrix} = \check{A}_{II} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \] 

D. Face Detection/Retrieval System

The face detection or image retrieval is carried out based on the segmentation method as follows:

(a) A significant scene change is detected in a video footage.
(b) A sampling point is scanned on the beginning frame of a new scene.
(c) If the color at a sampled point is within a color window, segmentation is carried out for a number of errors and for a few sets of the weights on the HSV components.
(d) The segmented image is made binary, which is then checked in some requirements as face.
(e) Then, the pattern is correlated with an input face pattern.
(f) The segmented image with the largest correlation is output as the face for the frame, where it also is possible to detect multiple faces.
(g) Segmented face images are displayed according to their correlation values.

The scene change is detected by evaluating the difference between the neighboring frames.

IV. FACE EXTRACTION USING COLOR SPACE NORMALIZATION

Various approaches to face extraction are discussed. There are also recent surveys on face detection. These approaches utilize techniques such as principal component analysis neural networks, machine learning, information theory, geometrical modeling, (deformable) template matching, Hough transform, motion extraction, and color analysis. The neural network-based and view-based approaches require a large number of face and nonface training examples and are designed primarily to locate frontal faces in gray-scale images. Facial templates and Hough transform were incorporated to detect grayscale frontal faces in real-time applications. Face detectors based on Markov random fields and Markov chains, make use of the spatial arrangement of pixel gray values.[7]

Categorizing face detection methods based on the representation used reveals that detection algorithms using holistic representations have the advantage of finding small faces or faces in poor-quality images, while those using geometrical facial features provide a good solution for detecting faces in different poses. A combination of holistic and feature-based approaches is a promising approach to face detection as well as face recognition[2]. Motion and skin-tone color are useful cues for face detection. However, the color-based approaches face difficulties in robustly detecting skin colors in the presence of complex background and different lighting conditions. We propose a face detection algorithm that is able to handle a wide range of variations in static color images, based on a lighting compensation technique and a nonlinear color transformation. Our approach models skin color using a parametric ellipse in a two-dimensional transformed color space and extracts facial features by constructing feature maps for the eyes, mouth, and face boundary.

Research on face detection in images and its related areas has extensively been made in recent years especially in the fields of image processing and computer vision. The previous algorithms are aimed at detecting or recognizing the face in image. The detection is required to be in real time in computer vision, possibly at the sacrifice of reliability for each frame but not for a sequence of frame images. In the field of multimedia, on the other hand, the focus has been on not just its detection or recognition but also identification of faces, people, or some specific objects in video images or video footages. Satoh et al., for example, tried to retrieve the name from the face or the face from the name using the video, video caption and the transcripts. Since the segmentation accuracy affects to the identification and the images may be available in a limited duration of time, several improvements have been reported. They combine temporal segmentation or tracking with spatial segmentation or adopt manual segmentation. Long et al., for example, presented a method that uses three consecutive frames to take into account motion and user interaction when automatic detection fails. That may also be the case of retrieving some visual information from video footages, where accuracy may also be crucial.

A. Face Extraction Algorithm

The face extraction algorithm contains two major modules: (1) Face localization for finding face candidates and (2) facial feature detection for verifying detected face candidates. The algorithm first estimates and corrects the color bias based on a lighting compensation technique. The detected red, green, and blue color components are then nonlinearly transformed in the YCbCr color space. The skin-tone pixels are detected using an elliptical skin model in the transformed space. The parametric ellipse corresponds to contours of constant Mahalanobis distance under the assumption of Gaussian distribution of skin tone color. The detected skin-tone pixels are iteratively segmented using local color variance into connected components which are then grouped into face candidates based on both the spatial arrangement of these components and the similarity of their color. The size of a face candidate can range from \( 13 \times 13 \) pixels to about three fourths of the input image size. The
facial feature detection module rejects face candidate region that do not contain any facial features such as eyes, mouth, and face boundary [4].

B. Lighting compensation and skin tone detection

The appearance of the skin-tone color depends on the lighting conditions. We introduce a lighting compensation technique that uses “reference white” to normalize the color appearance. We regard pixels with the top 5 percent of the luma (nonlinear gamma-corrected luminance) values in the image as the reference white only if the number of these pixels is sufficiently large (>100). The R, G, and B components of a color are adjusted so that the average gray value of these reference white pixels is linearly scaled to 255. The image is not changed if a sufficient number of reference white pixels are not detected or the average color is similar to skin tone. This assumption is reasonable not only because an image contains “real white” pixels in some regions of interest (such as eye regions), but also because the dominant bias color always appears as “real white”. With lighting compensation, our algorithm detects fewer non face pixels and more skin-tone facial pixels.

The appearance of the skin-tone color depends on the lighting conditions. We introduce a lighting compensation technique that uses “reference white” to normalize the color appearance.

Modeling skin technique requires choosing an appropriate color space and identifying a cluster associated with skin color in this space. It has been observed that the normalized red-green (rg) space is not the best choice for face detection. Based on Terrillon et al.’s comparison of nine different color spaces for face detection, the tint-saturation-luma (TSL) space provides the best results for two kinds of Gaussian density models (unimodal and a mixture of Gaussians). We adopt the YCbCr space since it is perceptually uniform, is widely used in video compression standards (e.g., MPEG and JPEG), and it is similar to the TSL space in terms of the separation of luminance and chrominance as well as the compactness of the skin cluster. Many research studies assume that the chrominance components of the skin-tone color are independent of the luminance component. However, in practice, the skin-tone color is non-linearly dependent on luminance. The luma dependency of skin-tone color in different color spaces is based on skin patches (853,571 pixels) collected from nine subjects (137 images) in the Heinrich-Hertz-Institute (HHI) image database. Detecting skin tone based on the cluster of training samples in the CbCr subspace, results in many false positives. Face detection based on the cluster in the (Cb/Y)-(Cr/Y) subspace, results in many false negatives. Therefore, we non-linearly transform the YCbCr color space to make the skin cluster luma-independent. This is done by fitting piecewise linear boundaries to the skin cluster. The transformed space, enables a robust detection of dark and light skin tone colors. More skin-tone pixels with low and high luma are detected in the transformed subspace than in the CbCr subspace.

C. Localization of facial features

Among the various facial features, eyes and mouth are the prominent features for recognition and estimation of 3D head pose. Most approaches for eye localization are template-based. However, we directly locate eyes, mouth and face boundary based on their feature maps derived from both the luma and chroma of an image. We consider only the area covered by a face mask that is built by enclosing the grouped skin-tone regions with a pseudo convex hull.

D. Eye map

We first build two separate eye maps, one from the chrominance components and the other from the luminance component. These two maps are then combined into a single eye map. The eye map from the chroma is based on the observation that high Cb and low Cr values are found around the eyes. It is constructed by

\[ \text{EyeMap}_C = \frac{1}{3} \left( (C_b)^2 + \left( C_r \right)^2 + \left( \frac{C_b}{C_r} \right) \right) \] (15)

where \( C_b \) and \( C_r \) are defined as

\[ C_b = (Y - 0.2681 \times C_r) \]

The eye map from the chroma is based on the observation that high Cb and low Cr values are found around the eyes. It is constructed by

\[ \text{EyeMap}_L = \frac{Y(x,y) \times \text{EyeMap}(x,y)}{Y(x,y) \times \text{EyeMap}(x,y) + 1} \] (16)

where the gray-scale dilation \( \oplus \) and erosion \( \ominus \) operations on a function \( f: \mathbb{R}^2 \rightarrow \mathbb{R} \) using a structuring function \( g: \mathbb{G} \subset \mathbb{R}^2 \rightarrow \mathbb{R} \) are defined. The eye map from the chroma is enhanced by histogram equalization and then combined with the eye map from the luma by an AND (multiplication) operation.

The resulting eye map is then dilated, masked, and normalized to brighten both the eyes and suppress other facial areas, as shown in Fig. 3. The locations of the eye candidates are

![Diagram](image-url)
initially estimated from the pyramid decomposition of the eye map and then refined using iterative thresholding and binary morphological closing on the eye map.

![Diagram](Image)

**Fig. 3:** Eye map construction

### E. Mouth Map

The color of mouth region contains stronger red component and weaker blue component than the other facial regions. Rather than the usual procedure of getting the mouth map using the YCbCr color space we have used the YIQ color space. Hence the, the Q component is greater than I in the mouth region. We further notice that the mouth has relatively low response in the Q/I feature, but it has a high response in Q^2. We construct the mouth map as follows:

\[
\text{MOUTHMAP} = Q^2(Q^2 - \eta((Q/I))^2)
\]  

(17)

Where both Q^2 and Q/I are normalized to the range [0, 255], and \( \eta \) is the number of pixels within the face mask, FG. The parameter \( \eta \) is estimated as a ratio of the average Q^2 to the average Q/I. Fig. 4 shows the construction of the mouth map for the subject in Fig. 4.

![Diagram](Image)

**Fig. 4:** Mouth map construction

Overall selection of a structuring element depends upon the geometric shapes you are attempting to extract from the image data. For example, if you are dealing with biological or medical images, which contain few straight lines or sharp angles, a circular structuring element is an appropriate choice[3]. When extracting shapes from geographic aerial images of a city, a square or rectangular element will allow you to extract angular features from the image. While most examples in this chapter use simple structuring elements, you may need to create several different elements or different rotations of a singular element in order to extract the desired shapes from your image. For example, if you wish to extract the rectangular roads from an aerial image, the initial rectangular element will need to be rotated a number of ways to account for multiple orientations of the roads within the image. The size of the structuring element depends upon what features you wish to extract from the image. Larger structuring elements preserve larger features while smaller elements preserve the finer details of image features. The following table shows how to easily create simple disk-shaped, square, rectangle, diagonal and custom structuring elements using IDL[13]. The visual representations of the structures, shown in the right-hand column, indicate that the shape of each binary structuring element is defined by foreground pixels having a value of one.

IDL Code for Structuring Element Shapes. Disk-Shaped Structuring Element. Use SHIFT in conjunction with DIST to create the disk shape.

\[
\text{radius} = 3
\]

\[
\text{strucElem} = \text{SHIFT}(\text{DIST}(2*\text{radius} + 1), \text{radius}, \text{radius})
\]

**V. IMPLEMENTATION AND EXPERIMENTAL RESULTS**

The algorithm is evaluated using several face image databases, including family and news photo collections. Face databases designed for face recognition, usually contain gray-scale mugshot-style images and, therefore, in our opinion, are not suitable for evaluating face detection algorithms. Most of the commonly used databases for face detection, including the Carnegie Mellon University (CMU) database, contain gray-scale images only[11]. Therefore, we have constructed our database for face detection from JPEG2000, the World–Wide–Web– and personal photo collections. These color images have been taken under varying lighting conditions and with complex backgrounds. Further, these images contain multiple faces with variations in color, position, scale, orientation, 3D pose, and facial expression.

Our algorithm can detect faces of different sizes with a wide range of facial variations in an image. Further, the algorithm can detect both dark skin-tone and bright skin-tone because of the nonlinear transformation of the CbCr color space. All the algorithmic parameters demonstrate that our algorithm can successfully detect dark skin faces. Figures show the results for subjects with some facial variations (e.g., closed eyes or open mouth) and for those who are wearing glasses. Our algorithm can detect nonformal faces as long as the eyes and mouth are visible in half-profile views. Face can also be detected in the presence of facial hair. A summary of the detection results (including the number of false positives, detection rates, and average CPU time for processing an image) on the HHI JPEG2000 image database and the champion database. The database contains 106 images, each of size 640 × 480 pixels. Subjects in the database belong to several racial groups and the lighting conditions (including overhead lights and side lights) change from one image to another. Further, these images contain frontal, near-frontal, half-profile, and profile face views of different sizes. A detected face is a correct detection if the detected locations of the eyes, the mouth, and the ellipse bounding a human face are found with a small amount of tolerance, otherwise it is called a false positive. The detection rate is computed by the ratio of the number of correct detections in a gallery to that of all human faces in the gallery. The detection rate on the database after
the first two stages (before facial feature extraction) is ~ 97 percent for all the poses[8].

A. Output Snap Shots of Different phases

Fig. 5: RGB image

Fig. 6: Extracted R components

Fig. 7: Extracted G component

Fig. 8: Extracted B components

Fig. 9: The YC\_b C\_r color space

Fig. 10: Extracted Y components
Advanced Hybrid Color Space Normalization for Human Face Extraction and Detection

Fig. 11: Extracted $C_b$ components

Fig. 12: Extracted $C_r$ components

Fig. 13: Eye map after Histogram equalization

Fig. 14: Eye map luminance

Fig. 15: Final Eye Map

Fig. 16: YIQ Color Space
Fig. 17: Extracted I component

Fig. 18: Extracted Q component

Fig. 19: Mouth Map after Histogram Equalization

Fig. 20: Combined Eye and mouth map

Fig. 21: Masked imag after combining eye and mouth map

Fig. 22: Face Segmentation
VI. CONCLUSION

This paper presents the concept of color space normalization (CSN) and two CSN techniques for enhancing the discriminating power of color spaces for face recognition. Our experimental results reveal that some color spaces, like RGB and XYZ are, relatively weak for recognition, whereas other color spaces such as I1I2I3, YUV, YIQ and LSLM, are relatively powerful. The proposed CSN techniques are applied to the RGB and XYZ color spaces, the three hybrid color spaces XGB, YRB and ZRG which are generated by configuring the components from the RGB and XYZ color spaces, and the 10 randomly generated color spaces. All experimental results demonstrated the effectiveness of the proposed CSN techniques.

To address the problem of why the CSN techniques can improve the face recognition performance of weak color spaces, we perform the correlation analysis on color component images corresponding to different color spaces and show that the proposed CSN techniques can significantly reduce the correlation between color component images and thus can enhance the discriminating power of the concatenated color component images.

Finally, it should be pointed out that the focus of this paper is on validating the effectiveness of the color space normalization techniques for color images based face recognition. We only use a basic face feature extraction method. If using and combining more complicated feature extraction methods, we can achieve state-of-the-art database verification results based on the normalized color spaces.

We have presented a face detection algorithm for color images using a skin-tone color model and facial features. Our method first corrects the color bias by a lighting compensation technique that automatically estimates the reference white pixels. We overcome the difficulty of detecting the low-luma and high-luma skin tones by applying a nonlinear transform to the YCr and YIQ color spaces. Our method detects skin regions over the entire image and then generates face candidates based on the spatial arrangement of these skin patches[2]. Our algorithm constructs eye, mouth, and boundary maps to verify the face candidates. Detection results on several photo collections have been presented.

- Future Enhancements

Additionally, we can further improve the verification rates of color spaces once the z-score normalization technique is applied. Our future goal is to design a system that detects faces and facial features, allows users to edit detected faces, and use these detected facial features as indices for identification and retrieval from image and video databases.

ACKNOWLEDGMENT

We would like to acknowledge the contributions from Department of Computer Science, Amal Jyothi College of Engineering. We are indebted to our guide, Dr. George Varkey, for providing the project directions and also in the preparation of this paper. We would also like to thank Mr. Akas G Kamal and Ms. Bindu S Moni, for providing the necessary instructions for carrying out the experiments.

REFERENCES

[6] Robust Histogram Construction from Color Invariants for Object Recognition Theo Gevers, Member, IEEE, and Harro Stokman
[7] Guo Dong, Member, IEEE, and Ming Xie, Member, IEEE, Color Clustering and Learning for Image Segmentation Based on Neural Networks.


