

# Face Detection based Neural Networks using Robust Skin Color Segmentation

*Aamer Mohamed, Ying Weng, Jianmin Jiang and Stan Ipson,*

School of Informatics, University of Bradford, UK  
{ a.s.s.mohamed, Y.Weng, j.jiang1,S.S.Ipson}@Bradford.ac.uk

## ABSTRACT

This paper proposes a robust schema for face detection system via Gaussian mixture model to segment image based on skin color. After skin and non skin face candidates' selection, features are extracted directly from discrete cosine transform (DCT) coefficients computed from these candidates. Moreover, the back-propagation neural networks are used to train and classify faces based on DCT feature coefficients in Cb and Cr color spaces. This schema utilizes the skin color information, which is the main feature of face detection. DCT feature values of faces, representing the data set of skin/non-skin face candidates obtained from Gaussian mixture model are fed into the back-propagation neural networks to classify whether the original image includes a face or not. Experimental results shows that the proposed schema is reliable for face detection, and pattern features are detected and classified accurately by the back-propagation neural networks

*Index Terms*— Face detection, feature extraction, DCT, Neural Networks

## 1. INTRODUCTION

Face detection has received much more attention in recent years. It is the first step in many applications such as face recognition, facial expression analysis, content based image retrieval, surveillance system and intelligent human computer interaction. Therefore, the performance of these systems depends on the efficiency of face detection. The comprehensive survey on face detection has been given out [1, 12]. These approaches utilize techniques such as Neural Networks, Principle Component Analysis (PCA), Machine Learning, and Support Vector Machine. Many several face detection algorithms that are use neural networks, and support vector machine [3, 5, and 6]. Face detection is an easy task for a human, but still far way for the computer to be easy task. Why? One of the most main reason arise is the high variability among the faces. Face detection according to various approaches are classified into four categories (i) knowledge based method (ii) feature based method (iii) template matching method (iv) Appearance based method. The objectives of this research are to develop better normalization method and also aim

to improve the segmentation that will assist and speed up detecting faces from images. And also to implement a classifier face based on neural networks for face detection. This paper presents Gaussian model applied on samples of face image in order to obtain the likelihood skin belonging to the face regions. Then applying DCT coefficient on the gray image obtained from the previous stage in order to get features from DCT Domain. The algorithm works directly on the DCT coefficient parameters, moreover color information is used as the main detection clues , a skin color model is created in the level of YCbCr color space .The reason for choosing chrominance blue and chrominance red (Cb and Cr) color space instead of YCbCr in order to eliminate the affect of illumination changes by using Y component. Classification using only pixel chrominance Cb and Cr (pure color), skin segmentation may become more robust to lighting variations if pixel luminance is discarded, moreover to narrow the search and speed up the calculation for detecting the skin face regions. The new algorithms introduced combines two methods to perform fast and accurate face detection system, which are rely on a feature based methods and image based methods, the feature based method used a pre-processor of the image based method and guides the search of image based methods using neural networks that examine the face candidate regions instead of performing huge search in every part of the test image [5]. Hwei proposed extraction regions of skin can be either pixel-based [7] or region based [8, 1]. The system in this paper propose techniques presents skin segmentation applied using the Gaussian mixture model and the likelihood will be the gray-scale image whose values represent the likelihood of the pixel belonging to the skin face region. 2D Discrete Cosine Transform (DCT) for each sub block image is computed and features vector are formed from the DCT coefficients; where DCT can be as signature useful for recognitions tasks such as facial expression recognitions.[9, 10]. The rest of the paper is organized as follows, in section 2 presents brief description of face detection in images, section 3 gives how face detection algorithms implemented, section 4 presents a conclusion and future works.

## 2. FACE DETECTION IN IMAGES

Techniques for face detection in image are classified into four categories –

## 2.1. Knowledge Based Method

It dependence on using the rules about human facial feature .It is easy to come up with simple rules to describe the features of a face and their relationships. But the difficulty of it is how to translate human knowledge into well known rules in order to detect faces in different poses. For example, a face often appears in an image with two eyes that are symmetric to each other, a nose, and a mouth. If try to define detailed rules then there may be a large number of faces stratifying the rules. Few rules are unable to describe the face exactly. This approach is good for frontal face image

## 2.2. Image Based Method

In this approach, there is a predefined standard face pattern is used to match with the segments in the image to determine whether they are faces or not. It uses training algorithms to classify regions into face or non-face classes. Image-based techniques depends on multi-resolution window scanning to detect faces, so these techniques have high detection rates but slower than the feature-based techniques. Eigen-faces and neural networks are examples of image-based techniques. This approach has advantage of being simple to implement, but it cannot effectively deal with variation in scale, pose and shape.

## 2.3. Feature Based Method

This approach depends on extraction of facial features that are not affected by variations in lighting conditions, pose, and other factors. These methods classified according to the extracted features [1]. Feature-based techniques depend on feature derivation and analysis to gain the required knowledge about faces. Features may be skin color, face shape, or facial features like eyes, nose, etc.... Feature based methods are preferred for real time systems where the multi-resolution window scanning used by image based methods are not applicable. Human skin color is an effective feature used to detect faces, although different people have different skin color, several studies have shown that the basic difference based on their intensity rather than their chrominance. Texture of human faces has a special texture that used to separate them from different objects. Facial Features method depends on detecting features of the face. Some users use the edges to detect the features of the face, and then grouping the edges. Some others use the blobs and the streaks instead of edges. For example, the face model consists of two dark blobs and three light blobs to represent eyes, cheekbones, and nose. The model uses streaks to represent the outlines of the faces like, eyebrows, and lips .Multiple Features methods use several combined facial features to locate or detect faces. First, find the face by using features like skin color, size and shape and then

verifying these candidates using detailed features such as eyebrows, nose, and hair.

## 2.4. Template Based Method

Template matching methods use the correlation between pattern in the input image and stored standard patterns of a whole face /non face features to determine the presence of a face or non face features. if the window contains a pattern which is close to the target pattern, then the window is judged as containing a face.

## 3. FACE DETECTION ALGORITHMS

Information of skin color in a color image is a very popular and useful technique for face detection. The obvious advantage of this method is simplicity of skin detection rules that leads to construction of a very rapid classifier. We can use color information as a feature to identify a person's face in an image because human faces have a special color distribution that differs significantly, although not entirely from those of the background objects. Previous studies have found that pixels belonging to skin region exhibit similar chrominance components within and across different human races. In the YCbCr color space, chrominance components are represented by Cb and Cr values. In [13], thus, skin color model can be deriving from these values. By using threshold techniques, skin color pixels are identified by the presence of a certain set of Cb and Cr values which corresponding to the respective ranges of  $R_{Cb}$  and  $R_{Cr}$  values of skin color. Otherwise, the pixel classified as non-skin color. The system being design divided into three main categories, pre-processing, segmentation, and classification using neural networks.

### 3.1. Pre-processing

In fact, processing skin color is faster than other facial features and proven to be an effective feature in many application; different people have different skin color, while several studies shown that the difference lies mostly between their intensity not in their chrominance color itself [1]. Literature survey shows that YCbCr color space is one of the successful color spaces in segmenting skin color accurately, among which Y is the luminance,  $C_b$  and  $C_r$  are the chrominance components. Choosing the suitable color space to model skin color is very important in order to a void variation of lighting condition. Collect a dataset of the skin face and non skin face by cropping or cutting manually the image skin face and non-skin face to get two classes of face and non-face. Before training, each of the training dataset image samples are intensity normalized using the equations (1) and (2), in order to be uniform and to be applied in the next segmentation stage.

$$I^- = \frac{1}{N} \sum_{i=1}^N I_i \quad (1)$$

$$I_i'' = (I_i - I^-) + 128 \quad (2)$$

Where N is the number of pixels in the training sample,  $I^-$  is the average gray value.  $I_i''$  is the gray value of  $i^{\text{th}}$  pixel, and  $I_i$  is the normalized gray value.

### 3.2. Segmentation

Skin color information and the accuracy of skin color segmentation are very important for many face detection researches [6]. Many skin detection methods ignore the luminance component of the color space, in order to achieve independent model of the differences on the skin appearance that may arise from the differences of human races. In this paper we convert the color space of the image from RGB to YCbCr color space. Moreover, skin pre-process were used to avoid the exhaustive search for faces, and also to reduce the space dimension. A skin color was modelled by collecting color samples from different face images. A total of 97200 skin face pixel sample from 100 color face images were used to determine the color distribution of human skin in chrominance blue Cb and chrominance red Cr color. The skin sampled filtered using a low-pass filter in order to reduce or remove the noise. The process containing the following steps:- In order to apply the color distribution for skin color of different people. Obtain the means and covariance matrix of both color space Cb and Cr as in equation (4).

Figure.1.a shows that the color distribution of these skin sample color space components cluster in small area. Moreover, the color histogram indicates that the skin color distributions of different people are clustered in the chromatic color space.

A maximum likelihood detection scheme applied in which the image location with highest likelihood chooses as the face region, fitting the skin likelihood of gray-scale image which is represent the likelihood of pixels belongs to skin face. The likelihood for skin face region denoted as skin color space  $x$  given its class  $k$ , where  $k \in \{\text{skin, non-skin}\}$ , it can be calculated by:-

$$p(x/k) = EXP[-0.5(x - \mu)^T C^{-1}(x - \mu)] \quad (3)$$

Where  $x = (Cb, Cr)^T$

And  $\mu$  is the mean of both Cb and Cr and C represented the covariance matrix of the two dimensional Gaussian distribution and N is the total number of the pixels samples in the face regions. Figure 1.b shows fitting skin color using Gaussian distribution, therefore the color can be represented using a Gaussian distribution model  $G=(M,C)$  where:-

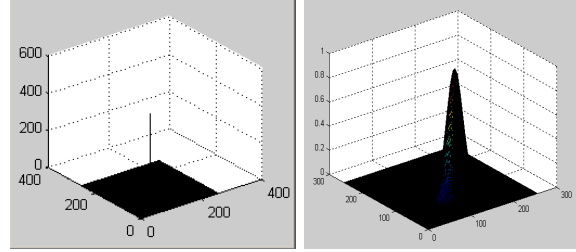
$$M = (C\bar{b}, C\bar{r})$$

$$C\bar{b} = \frac{1}{N} \sum_{i=1}^N Cb_i \quad (4)$$

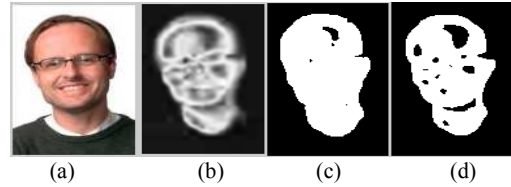
$$C\bar{r} = \frac{1}{N} \sum_{i=1}^N Cr_i$$

$$C = \begin{pmatrix} \sigma_{CbCb} & \sigma_{CbCr} \\ \sigma_{CrCb} & \sigma_{CrCr} \end{pmatrix}$$

Where  $C\bar{b}$  and  $C\bar{r}$  are the Gaussian Mean of both color spaces distribution Cb and Cr respectively.



**Figure 1:** (a) Color distribution of different skin color people, (b) Skin color fitting into Gaussian distribution



**Figure 2:** (a) Original face image (b) Likelihood skin region (c) Gray skin region (d) binary skin region

In this case the data set image is frontal face image, moreover, the detected region and the skin mainly corresponds to the skin face region and the skin regions are brighter than other part of the image. In order to segment skin region from the rest of image we apply threshold process, the likelihood skin region is shown as in Figure 2.(a) original test image, (b) likelihood skin region, (c) candidate gray-scale skin region image, and (d) binary skin regions.

### 3.3. Feature Extraction

Discrete cosine transform (DCT) is used widely in much application and mainly used in the compressed data domain, the formula of DCT and its inverse can be found in [14]. In addition it forms the basis well known JPEG image compression format. Jiang et.al introduced simple low cost and fast algorithms that extract dominant color feature directly from DCT rather than in the pixel domain [11]. The extracted DCT Coefficient can be used as type of a signature of which might be useful for recognition task, such as facial expression recognition [9]. The proposed technique calculates the 2D-DCT for each cropped skin gray-scale block coming out of the previous stage. A matrix of  $1 \times N$  coefficients obtained of both Cb and Cr color space components within the processed image block, these coefficient values is taken to construct the feature vector. Empirically, the upper left corner of

the 2D-DCT matrix contains the most important values; because these correspond to low-frequency coefficient do contain more energy. The upper left corner coefficient is called DC and it's corresponding to average light intensity of the block. The others are called AC, and those coefficients provide useful information about the texture detail in the block. For each block of  $C_b$  and  $C_r$ , we arrange DCT coefficient features by zig zag order, selecting the DC and the first three AC's as a set of  $1 \times N$  vector coefficients feature from both chromatic  $C_b$  and  $C_r$ , and then merging the features to be homogenize in order for neural networks to understand and converge as shown in Fig.3.

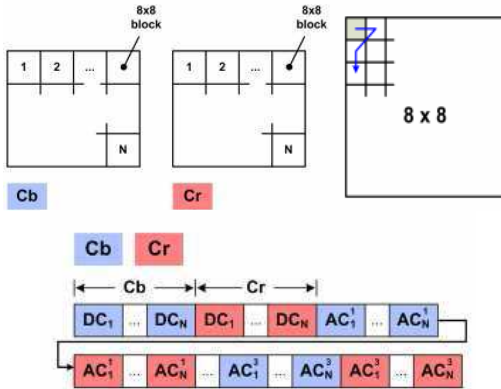


Figure 3: Feature extraction from DCT Coefficients

### 3.4. Back-propagation Neural Networks

Neural networks have been applied in many pattern recognition problems like object recognition. There is many image based face detection algorithms using neural networks, the most successful system was introduced by Rowley et al [3]. The neural network used in this paper is back-propagation neural networks and the reason of choosing that because of the simplicity and its capability in supervised pattern matching. The structure of the neural network defines three layers as shown in figure.4, the input layer can be fed by a vector of  $1 \times N$  DCT coefficient feature vector of each image either face or non face image, the hidden layers has  $n$  neurons, and the output layer is a single neuron which is 0.9 if the face is presented and 0.1 otherwise. The neural networks is trained using the extracted DCT coefficient feature vectors of face skin color candidate obtained from the segmentation stage, which are the DC and the first three AC's ordered by zig zag operation as shown in figure 3. One feature vector obtained from all coefficients image blocks, in order for the back-propagation neural network to learn and classify each feature vector as output value 0.9 for a face and 0.1 for non-face.

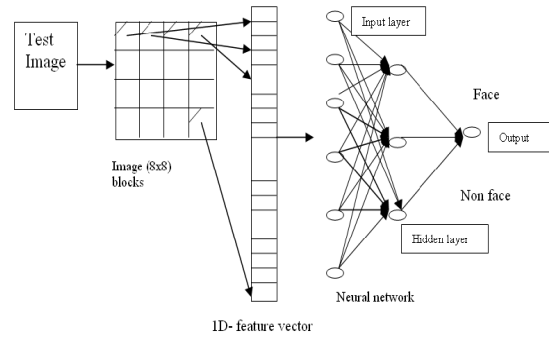
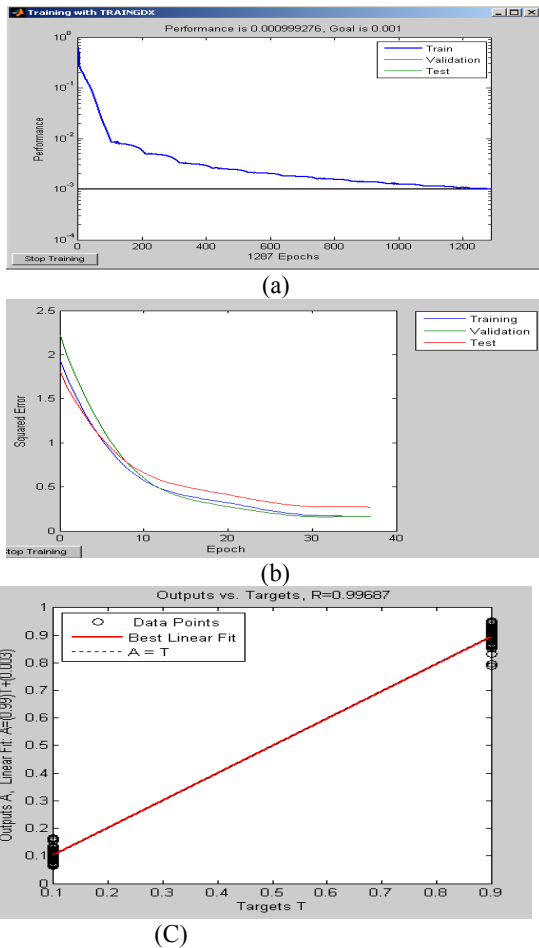


Figure 4: Neural network system

### 3.5. Experimental Results

In this section we show a set of experimental results to presents the performance of the proposed system, the experimented was implemented using Matlab Version 7.2 on the Intel Pentium (4) 2.80Ghz 1.00GB of RAM and Windows XP operating system .This results of experiment applied on the unknown input test image containing a face or non-face as shown in figure.5. Starting with sliding overlapping window  $18 \times 27$ , by overlap scanning the window across the test image, further, different overlap parameters are used 1, 2 up to the half pixels of window size, 9 pixels was the half of the window it might be maximum overlap, then each part of the unknown test image is scanned using slid window and extracted the DCT features and feed it to the trained neural networks of the dataset of images. Therefore the neural networks tested with the trained neural networks and classify it to see if the part containing a face or non face. The experiment results shows that our face detection system is promising results, that the neural network able to detect and classify pattern features accurately under different overlap sliding scan window over the unknown input test image. The converge response of training dataset shows accurate and excellent face and non face classification as in Fig.5a , In Fig.5b the result is reasonable, since the test set error and the validation set error have similar characteristics. The next step is to perform some analysis of the network response, by putting the entire data set through the network (training, validation and testing) to perform a linear regression between the network outputs and the corresponding targets as shown in Fig.5c, according to the excellent response of the Back-propagation neural networks with the target desired, the classification performance provides a comprehensive excellent picture of the classification performance of the classifier.



**Figure 5:** (a) Training Performance (b) NN's trained with training, testing and validation (c) Linear regression

#### 4. CONCLUSION AND FUTURE WORKS

This paper proposes a new algorithm for face detection in the compressed domain, which extracts DCT coefficient vector features after segmenting a face skin candidate by Gaussian mixture model in both Cb and Cr colour spaces, along with the back-propagation neural networks classifier. The problem is divided into three stages: pre-processing, segmentation, and classification using the back-propagation neural networks. This schema has been tested on a dataset of upright frontal colour face images from the Internet and achieved excellent detection rate compared with [13]. This schema will be improved for face detection of compressed images in further research such as face image retrieval based on skin color, and face detection using other classifier methods such as Support Vector Machine (SVM) or Adaboost classifier. Moreover, the schema proposed here can be used as first step in face recognition system [5].

#### 5. REFERENCES

[1] Y. Ming-Hsuan, J.D. Kriegman and N. Ahuja, "Detecting faces in images: a survey," Pattern Analysis and Machine

Intelligence, IEEE Transactions on, vol. 24, pp. 34-58, 2002

[2] H. Rein-Lien, M. Abdel-Mottaleb, and A.K Jain, "Face detection in color images," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 24, pp. 696-706, 2002

[3] H. Rowle . Baluja .A, S, and T. Kanade, "Neural network-based face detection," in Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 20, pp. 23-38, 1998.

[4] F. Smach, Atri. M.J Miteran and M. Abid " Design of a Neural Networks Classifier for Face Detection" in Journal of computer Science 2(3):pp. 257-260,2006

[5] Lamiaa Mostafa, Sharif Abdelazeem " Face Detection Based on Skin Color Using Neural Networks" in GVIP 05 Conference, pp19-21 , CICC, Cairo, Egypt, Dec 2006.

[6] Hwei-Jen Lin, Shu-Yi Wang, Shwu-Huey, and Yang -Ta-Kao " Face Detection Based on Skin Color Segmentation and Neural Network" IEEE Transactions on, Volume: 2, pp1144- -1149, ISBN: 0-7803-9422-4

[7] V. Vezhnevets, V Sazonov, and A. Andreeva, "A survey on pixel-based skin color detection techniques", in Proc. Graphicon-2003

[8] Kruppa .H, M. Bauer .A, and Schiele .B, "Skin patch detection in real world images", in Proc. of the DAGM-Symposium, pp. 109--116, (2002)

[9] L. Ma, Y. Xiao, K. Khorasani, and R. K. Ward, "A new facial expression recognition technique using 2D DCT and k-means algorithm", in Proc. International Conference on Image Processing, pp. 1269-1272, Oct.2004.

[10] L. Ma and K. Khorasani, "Facial expression recognition using constructive feed-forward neural networks", IEEE Transactions on Systems, Man and Cybernetics, Part B, Vol. 34, No. 3 pp.1588 - 1595, June 2004.

[11] J.Jiang, Y.Weng, and P.LI "Dominant colour extraction in DCT Domain" in Image and Vision computing 24 , published by Elsevier B.V.pp.1269-277, 2006.

[12] E. Hjelmås, and B.K Low, "Face detection: a survey", Computer Vision and Image Understanding, Vol. 83, No. 3, pp. 236-274, Sep. 2001.

[13] Aamer .S.S.Mohamed, Ying Weng, Stan Ipson, Jianmin Jiang "Face Detection based on Skin Color in Image by Neural Networks"; Accepted by IEEE International Conference on Intelligent and Advanced Systems ICISA 2007