IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 1, March, 2013
ISSN: 2320 - 8791
www.ijreat.org

Identify To Detect Face Based On Skin Color Using Neural Networks

V.Vinothraj[1], Sangeetha.V[2] , Sumithra devi.S[3]

[1]Assistant Professor, Department of Computer Science and Engineering
VRS college of engineering and technology arasur-607107, India.

[2,3]UG Scholar, Department of Computer Science and Engineering
VRS college of engineering and technology arasur-607107, India.

ABSTRACT

Face detection is one of the challenging problems in image processing. A novel face detection system is presented in this paper. The system combines two algorithms for face detection to achieve better detection rates. The two algorithms are skin detection and neural networks. In the first module of the system a skin color model based on normalized RGB color space is built and used to detect skin regions. The detected skin regions are the face candidate regions. In the second module of the system, the neural network is created and trained with training set of faces and non-faces. The network used is a two layer feed-forward network. There are two modifications for the classical use of neural networks in face detection. First, the neural network tests only the face candidate regions for faces, thus the search space is reduced. Second, the window size used by the neural network in scanning the input image is adaptive and depends on the size of the face candidate region. This enables the face detection system to detect faces with any size. In experiments on images having upright frontal faces with any background our system has achieved high detection rates and low false positives.

1.INTRODUCTION:

Face detection has attracted many researchers because it has a wide area of applications. Given an image, face detection involves localizing all faces - if any - in this image. Face detection is the first step in any automated system that solves problems such as: face recognition, face tracking, and facial expression recognition. Several face detection systems have been introduced[8].

Detection rate and the number of false positives are important factors in evaluating face detection systems. Detection rate is the ratio between the number of faces correctly detected by the system and the actual number of faces in the image. While false positive is the number of regions claimed to be faces by the system but they are not. The number of false positives depends mainly on the number of regions that have been tested for faces that is why the false positive in most systems is represented as a number not ratio Techniques of face detection can be classified into two main categories: feature-based and image-based [1]. All face detection techniques requires a priori knowledge of the face. 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. On the other hand image based techniques treat face detection as a general pattern recognition problem. 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. Eigenfaces [6] and neural networks [3] are examples of image-based techniques. The new algorithm introduced in this paper combines two methods to perform fast and accurate face detection. The system combines a feature based method with an image based method. The feature-based method is used a preprocessor of the image based method. Feature-based method guides the search of the image-based method so that the multi-resolution window scanning in all image locations is avoided. The first module of the system is a feature-based technique using a skin detector based on explicitly defined skin regions in normalized RGB color space. The regions where skin is detected are considered as face candidate regions. The second module is an image-based technique using a neural network; the neural network only examines the face candidate regions instead of performing exhaustive search in every part of the test image as in previous neural networks systems thus the search space is reduced. Also the window size used by the neural network is adaptive and determined by the size of the face candidate regions. The remainder of the paper is organized as follows :Section (2) focuses on skin detection. Section (3)emphasizes on the use of neural networks in face detection. In Section (4) we describe in detail the new system for face detection and present some results. Finally, conclusions and future work are in Section (5).

2. SKIN DETECTION:

Skin color is one of the most important features in the human face. Feature-based face detection techniques may use skin color information to detect faces in color images having complex background [2]. The skin detector detects whether certain regions in a color image represent human skin or not.

It must define certain decision rules to discriminate between skin and non-skin pixels. To build these rules, a human skin model must be built. Several skin color modeling methods have been introduced [9].Skin color modeling methods can be classified into three main categories: explicitly defined skin regions, nonparametric skin distribution modeling and parametric skin distribution modeling. In explicitly defined skin regions, both the color space and the decision rules – the skin region boundaries - are found empirically. The main advantage of explicitly defined skin regions method is the simplicity of the classification rules and its speed. On the other hand its main challenge is the empirical choice of a proper color space and adequate decision rules. In nonparametric methods a skin color distribution is established using the training data without derivation of the explicit model of the skin color. The result of these methods is sometimes referred to as construction of skin probability map. Non-parametric methods are fast in building the skin model and classification but it requires large storage space to represent the training skin samples.

The last method of building a decision rule is using parametric skin distribution modelling where an explicit skin distribution model is obtained. These methods can be fast also it can generalize the results but it very slow in building the model and in detecting the skin regions.The selection of the color space that will be used in modelling skin color is very important; it is well known that different people have different skin color appearance, but these differences lie mostly in the color intensity not in the color itself. That is why many skin detection methods drop the luminance component of the color space. Dropping the luminance component achieves two important goals; first the model will be independent of the differences in skin appearance that may arise from the difference in human race, or the difference in the lighting of the image; second the color space dimensions will be reduced so the calculations would be easier. There are lots of color spaces that have been used in early work of skin detection, such as RGB, normalized RGB, YCbCr, HIS and TSL [9]. Although RGB color space is one of the most used color spaces for processing and storing digital images, it is not widely used in skin detection algorithms because the chrominance and luminance components are mixed. Normalized RGB and YCbCr are often used by skin detection techniques. Some work has been done to compare different skin color space performance in skin detection problems [5]. The conclusion was that normalized color space yields the best skin detection results. Normalized RGB color space is obtained from RGB using simple normalization:

$$r = \frac{R}{R+G+B}$$

$$g = \frac{G}{R+G+B} \qquad (1)$$

$$b = \frac{B}{R+G+B}$$

The three normalized components *r*, *g* and *b* are called pure colors; they contain no information about the luminance. Also it can be deduced from the above equations that the sum of the three components is always equal to one, so it is enough to use only two components *r* and *g* to completely describe the skin color space .In our work, we build a skin color model based on explicitly defined skin regions in normalized RGB color space. Normalized RGB color space was chosen in our work for many reasons; first, it contains no information about luminance which yields a more general skin color model; also it has only two components which helps to speed up the calculations; also the transformations from RGB color space into normalized RGB color space is done using simple and fast transformations. The main reason for using explicitly defined skin regions in building the skin detector is its speed in detecting the skin regions.

To build the model, we collected samples of human skin from different races like Africans, Americans, and Arabs.... etc. For every pixel in the skin samples the values of *r* and *g* are calculated, and then the mean and standard deviation of *r* and *g* in all skin samples are calculated. The mean values of the red and green colors are denoted as μr and μg respectively. While the standard deviations of the red and green colors are denoted as σr and σg respectively.

After the normalized RGB skin color model is built, it can be used for skin detection. The first step in skin detection is pixel-based skin detection, where the skin detector tests every pixel of the input image and computes its normalized red value *r* and normalized green value *g*. If *r* and *g* values of the pixels satisfy the inequalities (2), then this pixel is considered skin. The value of α determines how accurate the skin detector will be and its value is to be determined experimentally.

$$\mu_r - \alpha\sigma_r < r < \mu_r + \alpha\sigma_r$$

$$(2)$$

$$\mu_g - \alpha\sigma_g < g < \mu_g + \alpha\sigma_g$$

The output of the pixel-based skin detector is a binary mask that contains ones in the skin regions and zeros in non-skin regions. The detected skin regions may be discontinuous; this discontinuity may be due to lighting effects that leads to missed skin pixels or due to the presence of non-skin face features like the eyes and brows. In order to make the detected skin regions continuous, region-based skin detection is applied to the output of the pixel-based skin detection. In region-based skin detection, after the binary mask is formed it is subject to median filtering of size 15*15 to make the skin regions continuous. All detected skin regions are considered as face candidates. Figure 1 clarifies the steps of skin detection. The input image where skin regions is to be determined is shown in Figure 1.a. Figure 1.b shows the output binary mask of the pixel-based skin detector, while Figure 1.c shows the output of the region based.

Figure 1.a. Skin Detector Input

Figure 1.b. Pixel-based Skin Detector Output

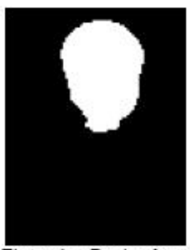

Figure 1.c. Region-based Skin Detector Output

3. NEURAL NETWORKS:

Neural networks have been applied in many pattern recognition problems like optical character recognition and object recognition. Instead of following a set of predefined human designed rules, the neural network in neural networks based face detection systems learns the underlying rules from a given samples. There are many image-based face detection algorithms using neural networks, the most successful system was introduced by Rowley *et al.* [3]. In neural networks based face detection techniques, the structure of the network is chosen first. The structure defines how many layers the network will have, the size of each layer, the number of inputs of the network and the value of the output for faces and non-faces. Then the network is trained using samples of faces and non-faces. To test an input image for faces, most of image-based approaches apply a window scanning technique to detect faces .The window has a fixed predetermined size and moves with certain step till it scans all parts of the input image. Each time the output is computed and if it is above certain threshold the window is classified as face. These systems can detect faces with size near to the window size. To be able to detect faces with different scales, an image pyramid is formed by successively resizing the input image. Then each level in the image pyramid is scanned by the moving window. Image-based techniques make exhaustive search for faces in all levels in image pyramid, so it may be computationally expensive. The neural network used in this paper is a two layer, feed forward neural network. A multilayer network is used as it can learn a mapping of any complexity not only linearly separable. This type of networks is used in many applications such as: speech recognition, hand written character recognition and object detection. The input layer of the network is 408 neurons. The size of the hidden layer is 26 neurons The input layer size was chosen as follows; at our work the smallest face size was chosen to have width of 30 pixels and height of 18 pixels. The ratio between the width and the height of the face obey the golden ratio of human faces which is approximately 1.6 [1]. If we input every pixel of this region, we will need 540 input neurons. But since the human face is oval in shape, we do not need all of the pixels in the rectangular 30*18 window. We need only pixels that represent human facial features. So using an oval mask which resembles the shape of the human face, only 408 pixels in the region marked by the oval mask is used as input to the neural network. The hidden layer size was chosen experimentally.

The transfer function used by the neural network is the sigmoid bipolar transfer function. In the training phase, the training image is displayed to the user and the user manually crops some regions that represent faces and some other regions as non-face samples. Since choosing a representative set of non-face samples could be hard ,bootstrapping is done where the false detections are added to the training set as training progresses. Before using the training sample in training it is subject to preprocessing. First, each training sample will be re-sized into size 30*18 this size is equal to the size of the smallest face that can be detected by the neural network, and then the training sample is converted into gray scale images. It is important to note that in the training and testing, the images are converted into gray scale before being input to the neural network. This is done because at this level no color information is needed because only regions that have been detected as skin regions are presented to the network. So using color information at this part would be redundant. After converting into gray scale the next step in pre-processing will be multiplying the gray scale sample by the oval mask to eliminate any background pixels that were cropped into the training sample. Multiplying the sample by the oval mask ensures that we do not wrongly introduce any unwanted background pixels to the neural network. The last step in the pre-processing will be histogram equalization. Histogram equalization enhances the contrast of an intensity image; it compensates changes in illumination in different images. Figure 2 shows the pre-processing steps that will be applied to training samples. Manually cropped face image is shown in Figure 2.a. After the face image is selected it is resized into the windows size –30*18 – and the result are shown in Figure 2.b. In Figure 2.c the gray scale image of the resized face is shown.

After the face sample is converted into gray scale it is multiplied by the oval mask – shown in Figure 4.a – and the output after multiplying the face sample by the mask is show in Figure 2.d. The output of the last step in the pre-processing – the histogram equalization – is shown in Figure 2.e.


2.a 2.b 2.c 2.d 2.e
Figure 2: Pre-processing Steps for the Neural Networks Training Face Samples

In the testing phase, only regions marked as skin region by the skin detector are presented for the neural network.It is important to note that in the testing, the images are subject also to the above pre-processing steps before applying to the neural network

4. SYSTEM DESCRIPTION AND RESULTS:

A practical system is implemented using MATLAB 6.5.In the first step the system is initialized to be able to detect faces. To initialize the skin detection module, skin samples are prepared. We used over 30 samples – 6000 pixels - of manually cropped human skin images. These samples are used to build normalized RGB skin color model. The histogram of the model is shown in Figure 3.It is clear from the figure that the *r* and *g* components of the skin color occupy a small cluster in the color histogram, so skin and non-skin regions can be separated using normalized RGB skin color model.

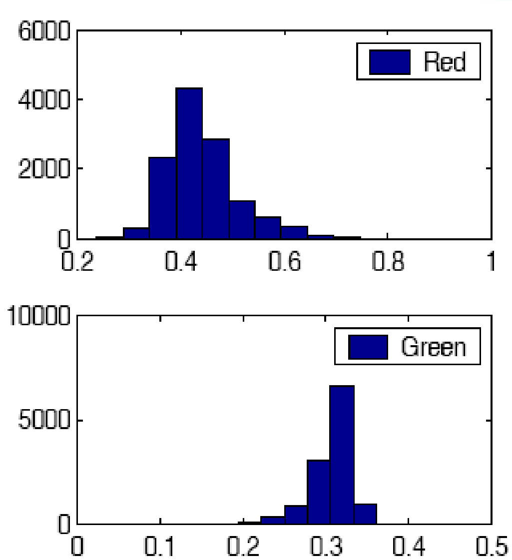

Figure 3. Histogram of Normalized RGB Skin Color Model μr =0.4408, μg =0.3070, σr =0.0713 and σg =0.0271

To initialize the neural network, training images areprepared. The training set we used was composed of 70 color images with complex background. Each of the training set images contains at least one face. The training samples are manually cropped from the training set. Before using the training sample in training it is subject to pre-processing. The training samples are composed of 92 face images and 167 non-face images. Figure 4 shows the oval mask used and some examples of face and non-face samples after the pre-processing


Figure 4.a. Oval Mask


Figure 4.b. Face Samples


Figure 4.c. Non-Face Samples

The neural network was trained to produce an output of +1.0 for faces, and -1.0 for non-faces. To detect faces in a test image, first the skin detector detects all skin regions in that image. Only the face candidate regions are tested for faces using the neural network. We assume that each face candidate region will contain maximum of one face. There are three possibilities for the face candidate region; first the region contain any other skin features like hand, legs or any other background pixels having skin color but not faces as indicated in Figure 5.a; second a face fully occupies the face candidate region as shown in Figure 5.b; third a face occupies a small part of the region and the rest is background pixels or other skin features like neck as shown in Figure 5.c.

In the first and second case, if the region is resized to the window size and tested by the neural network. It could be classified as non-face or face directly by comparing the output to the threshold. In the third case, the problem is that the face does not fully occupy the face candidate region, so the neural network will not be able to detect the face. The solution for this problem is to make window scanning in the face candidate regions. Since our goal is to detect all faces regardless of their scale, and it is expected that the face will occupy large part of the face candidate region, then the window size will be dependant on the size of the face candidate region itself. If the width and height of the face candidate region is *W* and *H*, respectively, the following window sizes will be used in scanning: W*H, W/2*H/2 and W/4*H/4. So even if the face occupies quarter of the face candidate region it will be detected. In each face candidate region, if more than one window gives output greater than the neural network threshold, only the window which gives the maximum output will be detected as a face. In reporting the results, a correct detection is counted if the region contains at least both eyes and the mouth, other than that it is reported as false detection [7].The test set is composed of 70 images. The images were tested by the skin detector; the best skin region segmentation was achieved using α equals 1.5. After that the face candidate regions are tested by the neural network. Multiple threshold values were tried. The results are shown in Figure 6. The best detection ratio was 91.43% and it was achieved using threshold equals 0. The number of false positives at this case was only 7.

Input Image Binary Mask Output Gray Image

Figure 5.a

Figure 5.b

Figure 5.c

Figure 5: Examples of the three cases of skin detector output

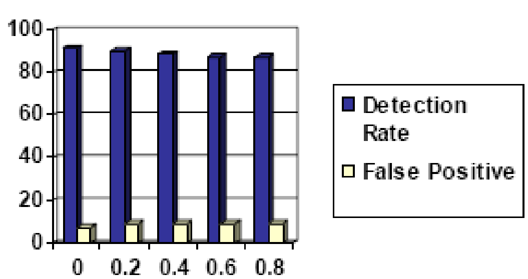

Figure 6: Detection Rate and False Positive Using Skin Detector with α = 1.5 and Neural Network with Adaptive Window Scanning

Figure 7 shows some results of the proposed system using α with value 1.5 and the neural network threshold equals 0. In Figure 7.a the face candidate region contains a complete face only and when tested by the neural network gives an output of 0.9892. In Figure 7.b the neck and the hair are selected by the skin detector to be part of the face candidate region. Also in Figure 7.c part of the background is included in the face candidate region. Because of the window scanning technique used by the neural network the face is detected with an output equals to 0.9999 in both cases. In Figure 7.d the skin detector fails to detect all face pixels due to the moustache. In addition, the neural network was not accurate because there are other features like the moustache and the glasses that were not included in any of the training images. The network detects a face with output equals 0.9986 but this is counted as a false positive because the region does not include the mouth and eyes so it is not considered as a complete face. Figure 7.e shows an image with more than one face. The skin detector detects two face candidate regions, and each region is tested with the neural network. The first region output was 0.9965 and the second region output was 0.8717. Input Image Binary Mask Output

Input Image Binary Mask Output Image

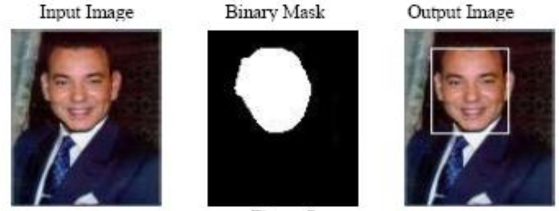

Figure 7.a

Figure 7.b

Figure 7.c

Figure 7.d

Figure 7.e

5. CONCLUSIONS AND FUTURE WORK:

A new system for face detection was presented in this paper. The new system was designed to detect upright frontal faces in color images with simple or complex background. There is no required a priori knowledge of the number of faces or the size of the faces to be able to detect the faces in a given image. The system has acceptable results regarding the detection rate, false positives and average time needed to detect a face. As a future work, the system proposed in this paper can be used as a pre-processor for face recognition systems as it gives high detection rate, and the false positives will be rejected by the face validation process of the recognition system. Also the system can be simply modified to detect faces with in-plane rotation using the method described by Rowley et al. [4]. The first step would be detecting the face candidate regions, and then a router network can be used to detect the rotation angle of the face in the face candidate region. After that the face candidate region will be de-rotated by the orientation angle detected by the router network. The last step would be a detector network that will verify the existence of a face in the de-rotated face candidate region.

6. References

[1] E. Hjelmas, Boon Kee Low, "Face Detection: A Survey", 2001.

[2] S. Kannumuri and A.N. Rajagopalan, "Human Face Detection in Cluttered Color Images Using Skin Color and Edge Information", Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP'02), (Ahmedabad, India), 2002.

[3] H. A. Rowley, S. Baluja, and T. Kanade, "Neural Network Based Face Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, 20, January, pp. 23-38, 1998.

[4] H. A. Rowley, S. Baluja, and T. Kanade, "Rotation Invariant Neural Network-based Face Detection", In Proc. IEEE Intl. Conf. on Computer Vision and Pattern Recognition, pp. 38–44, 1998.