FACE DETECTION TECHNIQUE BASED ON INTESITY AND SKIN COLOR DISTRIBUTION

Satyanadh Gundimada, Li Tao, and Vijayan Asari

Department of Electrical and Computer Engineering
Old Dominion University, Norfolk, VA 23508
{sgund002@odu.edu, ltaox001@odu.edu, vasari@odu.edu}

ABSTRACT

A rotation invariant human face detection system in color images based on human skin color distribution and intensity is proposed in this paper. Skin color distribution typical to a human face is used as a feature along with the intensity variations to classify the candidate regions into faces and non-faces. The detection process is carried out in YCbCr color space. Sparse Network of Winnows architecture is used to train three networks one for intensity and two for the color distributions for classification of candidate regions. Rotation invariance in detection of faces is achieved by training multiple classifiers, each to detect faces at a particular orientation. The detection process also implements a non linear luminance based lighting compensation method which is very efficient in enhancing and restoring the natural colors into the images which are taken in darker and varying lighting conditions. Experimental results show that the new face detection technique is highly efficient in terms of speed and accuracy in detecting frontal view faces at different orientations in complex environments.

1. INTRODUCTION

Face recognition has gained more prominence in recent years, mainly in applications such as surveillance, security systems and content based video indexing and retrieval. The essential part of any automatic face recognition system is face detection. The goal of any face detection technique is to identify the face regions within a given image. There are about 150 reported face detection techniques proposed in the literature [1] both in gray scale and color. The appearance based algorithms [1] process gray scale images. They rely on extensive training and powerful classification techniques. The classification methods range from neural networks [2], Hidden Markov Models to support vector machines [3]. A similar but much simpler technique called Sparse Network of Winnows (SNoW) [4] has also been implemented for face detection in gray scale images. A typical color based face detection system on the other hand would first do a skin color region extraction on color images based on either pixel based or a combination of pixels and shape based systems in different color spaces [5]. The next step would in general be region merging followed by classification or application of any appearance-based method to classify the skin color regions into faces and non-faces by converting them into gray scale images. Contrary to the conventional procedure, the algorithm proposed in this paper attempts to combine both the operations of skin extraction and face detection into a single process. Instead of using skin extraction as a method to reduce the computational intensity of scanning the entire image, the proposed method uses the skin color information and its distribution on a human face as a feature in addition to the intensity variations. This reduces the number of false positives and also improves the accuracy. SNoW architecture is used for the purpose of classification keeping in view of its computational inexpensiveness. An efficient image enhancement algorithm is applied on the input image to enhance the darker regions within the image to improve the detection process.

The paper is organized as follows. The second section describes the process of image enhancement in color images. Third section describes the architecture of SNoW and also the proposed face detection algorithm, followed by results and conclusion.

2. IMAGE ENHANCEMENT

Luminance dependent nonlinear image enhancement algorithm (LDNE) is used for enhancing the color images captured in extremely dark or non-uniform lighting environments [6]. In this technique, the processing is performed on the intensity/grayscale of the original color image obtained by nonlinearly transforming the RGB spectral bands to the luminance space as:

0-7803-8554-3/04/$20.00 ©2004 IEEE.

$$I(x,y)=\sqrt{R^2+B^2+G^2} \tag{1}$$

Then the luminance image is convolved with a 2-D distribution Gaussian function $H(x,y)$, which is obtained by adding three Gaussian functions (multi-level Gaussian function) with different scales. A Gaussian function is obtained as:

$$G_i(x,y) = K \cdot e^{\left(\frac{-(x^2+y^2)}{c_i^2}\right)} \tag{2}$$

where K determined by

$$\iint K \cdot e^{\left(\frac{-(x^2+y^2)}{c^2}\right)} \cdot dxdy = 1 \tag{3}$$

in which C_i is the Gaussian surround space constant such that C1 = 5, C2 = 20, C3 = 240. H (x, y) is computed as:

$$H(x,y)=\sum_{i=1}^{3} G_i(x,y) \tag{4}$$

Then the filtered image is obtained as:

$$Y(x,y)=I(x,y)*H(x,y) \tag{5}$$

The dynamic range of the output is then nonlinearly compressed by a two step approach: nonlinear compression of the luminance image I(x, y) and the filtered image Y(x, y) by a logarithmic approach; and extraction of the high frequency details from the compressed image. That is,

$$Q(x,y)=\log I(x,y)-\frac{1}{3}\log Y(x,y) \tag{6}$$

Since $Q(x,y)$ is the output in the 'log' domain, we need to apply 'Gain (contrast)' and 'Offset (brightness)' to transfer the output from 'log' domain to display domain.

$$P(x,y)=A\cdot(Q(x,y)+b) \tag{7}$$

where A (Gain) = 150, b (Offset) = 0.6 and both of them are experimentally determined values. Then, a comparison between the original luminance and the enhanced luminance is performed to ensure that the enhanced luminance have no luminance drop for any pixels.

$$S(x,y)=\left(\frac{|P\ (x,y)-I(x,y)|+P\ (x,y)-I(x,y)}{2}\right)+I(x,y) \tag{8}$$

The enhanced color images can be obtained based on the enhanced intensity images through a linear color restoration process described by the following equation.

$$S_j(x,y) = S(x,y)\frac{I_j(x,y)}{I(x,y)}\cdot\lambda \tag{9}$$

where j = r, g, b represent the R, G, B spectral band respectively, and S_r, S_g and S_b are the enhanced RGB values of the enhanced color image. λ is introduced to adjust the color hue of the three spectral bands. The output color images can be further refined with a color saturation and white balance adjustment, which is applied to make the output color looks more natural.

$$S_j(x,y)=S_j(x,y)+(S_a-S_j(x,y))*\kappa_j$$

$$\kappa_1 = 0.15 \quad \kappa_2 = \kappa_3 = 0.3$$

S_a : Average color value of each pixel after enhancement.
$S_j(x,y)$: Enhanced value in each color band.

3. FACE DETECTION

Face detection is performed based on the concept of Sparse Network of Winnows (SNoW) [4]. SNoW is a learning architecture and is used in classifying multiple classes. It is specifically tailored for large scale learning tasks and for domains in which the number of features taking part in decisions is very large. SNoW has been used successfully on a variety of large scale learning tasks in domains like the natural language processing, visual processing etc. In conventional SNoW based face detection, the features used for classification are the active pixel intensities within the candidate region. Pixel intensity at each location (x, y) is expressed in Boolean values as shown in figure 1.

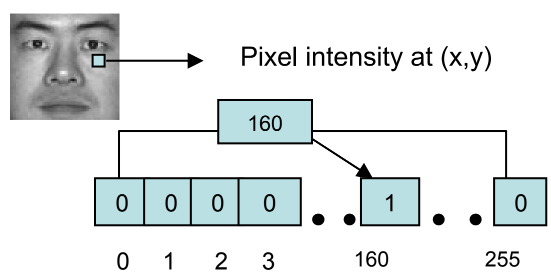

Figure 1. Active pixel representation

3.1 Integrated skin extraction and face detection

Many color spaces have been used in segmenting skin regions in color images. Literature survey shows that the YCbCr color space is one of the successful color spaces in segmenting the skin color accurately, mainly because the chrominance components are almost independent of luminance component in the space. Although skin colors vary from person to person, they tend to get clustered into a compact region in CbCr space. So the skin color distribution information in a human face can be used as an additional feature in classifying the objects into faces and non faces. Chrominance blue component is more prominent [7] around the eyes in a human face and at the same time the Cr component is less. The chrominance red is more prominent in the mouth area. This information which is typical to human face is also added as a feature in classification process. This improves the accuracy of the detection method in color images at the same time reduces the false positives at slight increase in computational complexity. Figure 2 shows the RGB face image transformed to Y, Cb and Cr images respectively. Figure 3 shows the distribution of Cr pixels at the eye region and the mouth region in a human face.

Figure 2. Face image in YCbCr color space

Conventional face detector using SNoW uses online error based training method for updating the sparse network using many gray scale face and non face images. A similar type of online learning method is being used in the algorithm. The novelty in this algorithm is the integration of color information and intensity information for detection of human faces.

150	149	148	148	147	146
151	150	149	148	147	146
152	151	149	148	147	146
154	152	150	149	147	146

160	161	161	160	159	158
160	161	162	163	164	164
159	161	161	162	162	163
159	158	158	157	157	157

Figure 3(a) Cr values around the eye, 3(b) Cr values around the mouth region.

3.2 Training phase

A set of 1500 face and 3000 non-face images collected from different color images from the web and ODU color database, were used for training. The training set contains face images of different skin colors and also includes images taken in varying lighting conditions. For a 25×25 pixel image with intensity range 0-255, SNoW representation is a 625×256 input matrix, but at any time only 625 elements are active. All the nodes at the input representing the input features are connected to the output node by a set of weights. The input is said to be true if the sum of all the weights is greater than a set threshold. The procedure of training is applied to all the three classifiers i.e., intensity, chrominance blue, and chrominance red. All the components are normalized to the range of 0 to 255.

3.3 Candidate face regions and classification

The image is scanned by a window of size of 25×25 at each pixel and sent for classification. The image is sub sampled by a factor of 1.2 each time and then scanned again. This process of sub sampling and scanning by the window is continued for 10 iterations. Summation of weights associated with the input features for each candidate face region in Y, Cb and Cr images of size 25×25 are calculated based on the active pixel locations. Each classifier produces a 1 or 0 based on the comparison between sum of weights and the preset thresholds. A candidate region is said to be face only if all the three classifiers produce a 1. Figure 4 shows the implementation of the proposed face classification technique.

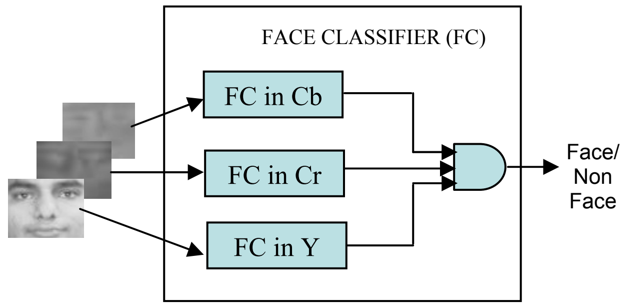

Figure 4. Classification of candidate faces in YCbCr color space

3.4 Rotation invariance

Multiple Face Classifiers (FCs) are trained for the detection of rotated faces at different angles within the image plane. Multiple databases of rotated faces, each for training a particular face classifier are created by rotating the original face database. The face detection system proposed can detect frontal faces and rotated faces up to 45^0 on either side within the image plane. It can also detect faces which are inverted within the image plane. There are six classifiers FC1-FC6. Each classifier can detect faces with approximately 10^0 variation for which it is trained for. Though training time is more, increase in additional computational complexity is very less in the testing stage. Figure 5 gives a picture of how multiple classifiers work to detect faces with different rotations.

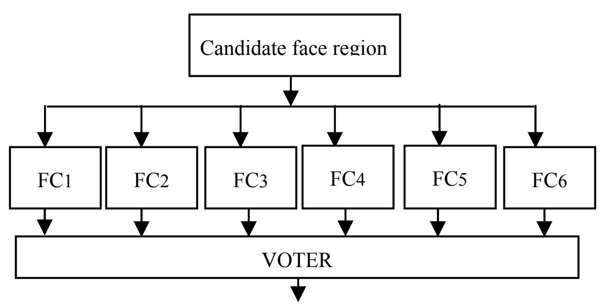

Figure 5. Multiple classifiers for rotation invariance

4. RESULTS

The results obtained by applying the proposed algorithm for detection of face regions in images captured at varying lighting environment are encouraging. The algorithm is tested on 200 frontal face images from the champions [7], ODU databases and 2000 non face samples collected from images containing scenery, texture, surface, desert, and skin patches.

No. of images	Images detected as faces	Detection rate after enhancement	Percentage improvement
200	190	195	2.5%

The rate of increase in detection is reflective of only the number of images that needed enhancement. Had there been more images of darker environments, the detection rate improvement would have been greater because of enhancement. The false negatives were mainly due to out of plane rotation within the face images. One major advantage of using the human face skin color distribution is that there is a reduction in the number of false positives. The algorithm is able to detect faces in varying lighting conditions due to the application of the new nonlinear enhancement technique as a preprocessing step. The face regions appeared at different orientations were also successfully detected by the new technique. Figures 6 and 7 show some of the combined results of image enhancement and face detection applied on various images from web and ODU database.

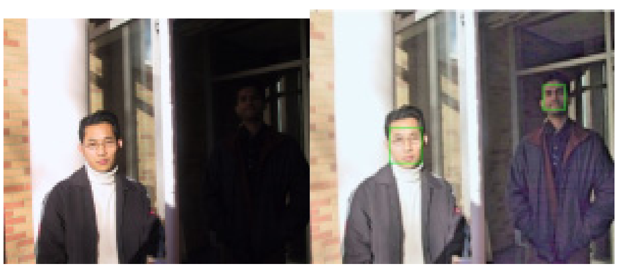

Figure 6. Application of image enhancement and face detection together

5. CONCLUSION

The human skin color and intensity distribution based face detection algorithm presented in the paper has been implemented using the Sparse Network of Winnows learning architecture in YCbCr color space. The algorithm is efficient both in accuracy and computational speed when compared to other leading color based face detection algorithms. Training of the sparse networks to detect human faces irrespective of in-plane or out-of-plane rotation in color images is progressing. Testing the algorithm on various databases and research in the area of finding out the optimum color space which can give better skin color distribution information in human face is also being carried out. Smart scanning of the image for candidate face regions based on the thresholds obtained around the regions can speed up the process of face detection.

6. REFERENCES

[1] M. H. Yang, D.J. Kriegman, N. Ahuja, “Detecting faces in images: A survey,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34-58, Jan. 2002.

[2] H. Rowley, S. Baluja and T. Kanade, "Neural network-based face detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 23-38, January 1998.

[3] E. Osuna, R. Freund, and F. Girosi, “Training support vector machines: An application to face detection*,” Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 130-136, 1997.

[4] M.-H. Yang, D. Roth, and N. Ahuja, “A SNoW-based face detector,” *Advances in Neural Information Processing Systems* 12, S.A. Solla, T. K. Leen, and K.-R. Muller, eds., pp. 855-861, MIT Press, 2000.

[5] Vladimir Vezhnevets, Vassili Sazonov, Alla Andreeva, “A survey on pixel-based skin color detection techniques,” *Graphicon-2003*, Moscow, Russia, September 2003.

[6] L. Tao and V. K. Asari, "Modified luminance based MSR for fast and efficient image enhancement," *Proc. IEEE Int. Workshop on Applied Imagery and Pattern Recognition***,** *AIPR - 2003*, pp. 174-179, October 2003

[7] R. L. Hsu, M. A. Mottaleb, and A. K. Jain, “Face detection in color images,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 696-706, May 2002.

Figure 7. Detected faces with different orientations