Component-based Face Detection in Colour Images

Jamal Ahmad Dargham, Ali Chekima
School of Engineering and Information Technology,University Malaysia Sabah,
Locked Bag 2073, Teluk Likas,88999 Kota Kinabalu, Sabah, Malaysia
{jamalad, Chekima}@webmail.ums.edu.my

Abstract: **Face detection is an important process in many applications such as face recognition, person identification and tracking, and access control. The technique used for face detection depends on how a face is modelled. In this paper, a face is defined as a skin region and a lips region that meet certain geometrical criteria. Thus, the face detection system has three main components: a skin detection module, a lips detection module, and a face verification module. The Multi-layer perceptron (MLP) neural networks was used for the skin and lips detection modules. In order to test the face detection system, two databases were created. The images in the first database, called In-house, were taken under controlled environment while those in the second database, called WWW, were collected from the World Wide Web and as such have no restriction on lighting, head pose or background. The system achieved a correct detection rate of 87 and 80 percent on the In-house and WWW databases respectively**

Keywords: Face Detection; MLP Neural Network; Component-Based Detection

I. INTRODUCTION

Face detection is defined as the ability to ascertain whether a face is present in the input image and if a face is present to returns the location of the face. The technique used for face detection depends on how the face is modelled. In [6] face-processing systems are classified into four categories: template-based, feature-based, appearance-based, and knowledge-based systems. However, these four categories can be grouped into two categories: global and component. In the global category the whole face is modelled either using template-based or appearance-based approaches. While in the second category, the face components and their relationships are used to model the face.

A face detection system proposed by [5] that combines a lips detection neural network (LDNN) and skin distinction neural network (SDNN). The lips are first detected by the LDNN and then the presence of a face is validated by the SDNN if the area around the lips is skin. The LDNN is a multi-layer feed forward neural network trained with the Back propagation algorithm. The network has 300 input neurons, 10 hidden neurons, and one output neuron coded as 1 for lips and 0 otherwise. The size of the lips is restricted to between 5 to 15 pixels vertically and 10 to 30 pixels horizontally. The SDNN is similar to that of the LDNN. However, the search area for the SDNN is limited to a box centred on the centre of the lips and extend 20 pixels vertically and 40 pixels horizontally. In [4] a face detection system was proposed. The proposed system has two modules: a face localisation module and a facial features detector module. The face localisation module generates potential face candidates in the image based on the chrominance and luminance of the colour image as well as the size of the potential face. The facial feature detector directly locates the eyes, mouth, and face boundary based on measurement derived from the colour-space components of the image. The system was tested on two image databases containing images in different poses from frontal to profile. The performance of the system is found to be dependent of the face pose as well as the occlusion of the facial features. Thus, the detection rate varied from around 75% to 91%.

A trainable, component-based face detection system for frontal and near-frontal images in gray images was presented by [3]. The system consists of a two-level hierarchy of Support Vector Machine classifiers. In the first level fourteen SVM classifiers independently detect the facial components. A single SVM classifier at the second level checks if the geometrical configuration of the detected components in the image matches a geometrical model of a face

II. THE DATABASES

Two databases were developed. The first database is called the In-house database while the second database is called the WWW database. The In-house database comprises 15 subjects both males and females from the various races in Malaysia namely Chinese, Malay, Indian, and Indigenous people. Each subject has 12 images showing the frontal facial image of the subject at three distances from the camera. These distances are called scaling factors. Scale factor 1 represents a distance of 36 cm between the camera and the subject while scale factor 2 and 3 represent a distance of 72 cm and 108 cm respectively. For each scale factor, images for three facial expressions namely neutral, smiling, and laughing were taken as well as with glasses with neutral expression only. These images were all taken indoor with a single digital camera under normal lighting conditions and with uniform background.

The WWW database has 45 images of Asian subjects collected from the World Wide Web. The subjects represent males and females of different ages. Some of the images were taken indoor while others were taken outdoor with varying backgrounds and lighting conditions. In addition, the pose of the face in the image varied widely from frontal to near portrait. The cameras used for taking these images as well the image processing techniques applied to them are unknown. The only restriction on these images is that they must show a face of an Asian person.

III. THE PROPOSED SYSTEM

The proposed system labels any region in a colour image as a face if that region contains a skin region and a lips region that meet certain geometrical constraints. Fig. 1 shows a block diagram of the proposed face detection system. As can be seen from Fig. 1, skin detection is carried out first on the colour image. Then post-processing operations such as erosion and dilation are carried out on the detected skin pixels followed by connected component labelling to group skin pixels into skin regions. Next, size filtering is performed on these regions to remove objects that are either smaller or larger than a particular size. The skin regions remaining in the image after the size filtering are considered as face candidates. The size of the smallest and largest skin region to be considered as a face candidate is selected as a ratio of the skin region to the size of the image. The values for these ratios were obtained by trials and errors. Then lips detection is carried out only on the face candidates and thus reducing the search space. Then post-processing steps similar to those performed on skin detection are carried out followed by size filtering to determine the lips candidates. The size of the smallest and largest lips region to be considered as a lips candidate is selected as a ratio of the lips region to the size of the image. A face is then assumed to exist when a face candidate and a lips candidate meet certain conditions that relate the ratio of the face size to the lips size as well as the orientation between the face candidate and the lips candidate. Unlike the systems described in this paper, our system does not put any limitation on the image size or on the size of the facial components.

A. Colour Transformation

For an image having M by N pixels, the r, g, and b component of the normalised rgb colour scheme are obtained from the normal RGB colour scheme as given by Equation 1.

$$\begin{cases} r(x,y) = \dfrac{R(x,y)}{R(x,y)+G(x,y)+B(x,y)} \\ g(x,y) = \dfrac{G(x,y)}{R(x,y)+G(x,y)+B(x,y)} \\ b(x,y) = \dfrac{B(x,y)}{R(x,y)+G(x,y)+B(x,y)} \end{cases} \quad (1)$$

The set of equations given in Equation 1 perform pixel-by-pixel normalization of the RGB components of the RGB colour scheme. Thus, we call this method pixel intensity normalisation. One of the problems of this method is that it breaks down when R+G+B = 0. Thus, we propose to normalise the RGB components by the maximum value of (R+G+B) over the entire image, we call this method maximum intensity normalisation. Thus, the set of equations in Equation 1 become as expressed in Equation 2.

$$\begin{cases} r(x,y) = \dfrac{R(x,y)}{Max(R+G+B)} \\ g(x,y) = \dfrac{G(x,y)}{Max(R+G+B)} \\ b(x,y) = \dfrac{B(x,y)}{Max(R+G+B)} \end{cases} \quad (2)$$

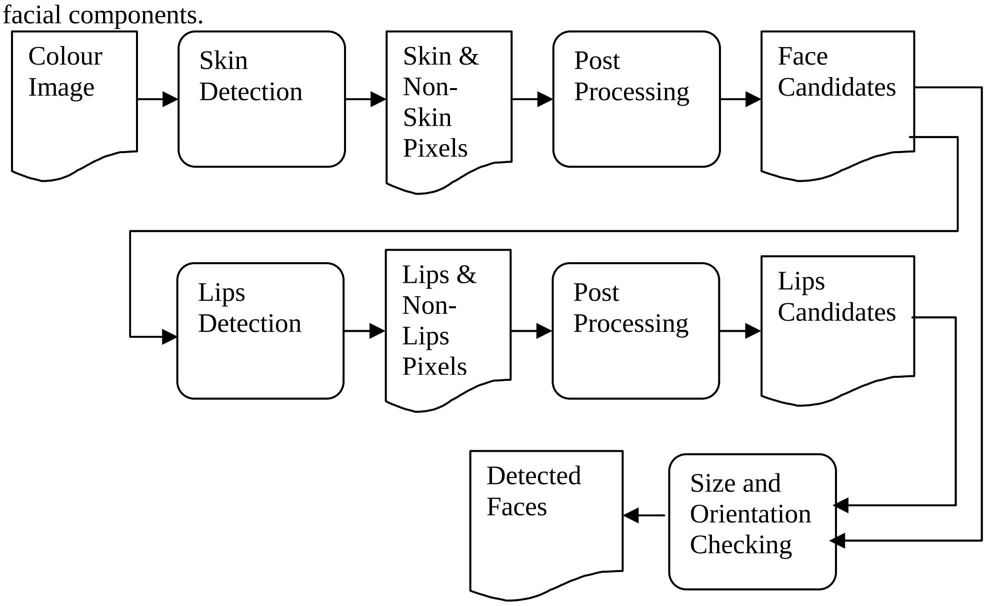

Figure 1 Block Diagram of the Proposed Face Detection System

B. Skin and Lips Detection

The Multi-layer Perceptron neural networks (MLP) were used for the skin and lips detection modules. The network structures for the skin detection and lips detection networks for each database and for each color transformation are given in Tables 1and 2 respectively. The number of the neurons in the hidden layer was determined by using network-growing technique. The networks were trained using the Levenberg-Marquardt training algorithm. For detailed descriptions see [1 and 2] Fig. 2 shows the process of finding the face candidates for an image from the WWW database while Fig. 3 shows the process of finding the lips candidates for the same image.

TABLE I. NETWORK STRUCTURE FOR THE MLP NEURAL NETWORKS USED FOR SKIN DETECTION

Database	Maximum Intensity		Pixel Intensity	
	Chrominance Components	*Network Structure*	*Chrominance Components*	*Network Structure*
In-House	r-b & r-g	18-3-1	r-b & r-g	18-8-1
WWW	r & r-g	18-35-1	r & r.g	18-19-1

TABLE II. NETWORK STRUCTURE FOR THE MLP NEURAL NETWORKS USED FOR LIPS DETECTION

Database	Maximum Intensity		Pixel Intensity	
	Chrominance Components	*Network Structure*	*Chrominance Components*	*Network Structure*
In-House	r/g & r+b-2g	2-3-1	r/g & r+b-2g	2-4-1
WWW	r+b-6g & r-g	2-5-1	g & r/g	2-1-1

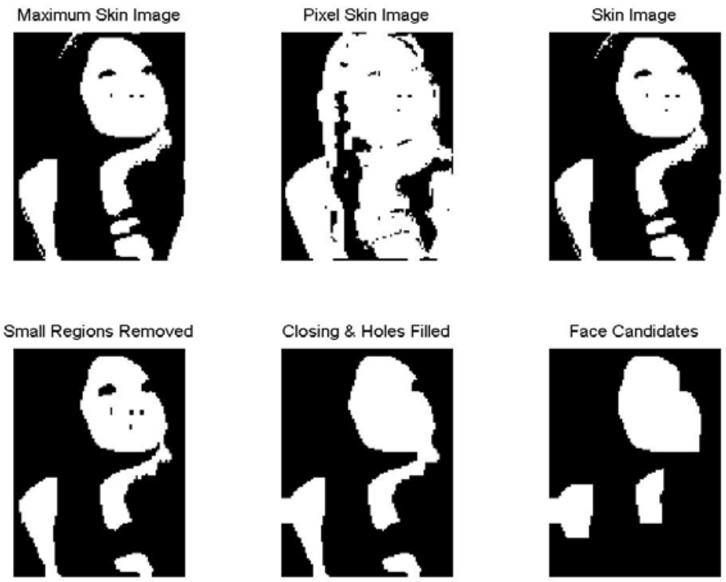

Figure 2. The Process of Finding the Face Candidates

C. Results

As has been mentioned earlier, a face is detected only when a face candidate and a lips candidate meet certain conditions. Thus, as shown in Fig. 4a, when the detected face by the system contains both the correct facial skin and the correct lips then a correct detection is given. However, if either the lips or the facial skin is incorrectly classified, as shown in Fig. 4b, this is termed as false acceptance. Finally, false rejection is defined as when either the facial skin or the lips is not detected when a face exists in an image as shown in Fig. 4c. The performance of the system on both databases is given in Table 3.

TABLE III PERFORMANCE OF THE FACE DETECTION SYSTEM ON BOTH DATABASES

Database	%Correct	FAR	FRR
In-House	87	6.67	6.67
WWW	80	8.89	11.11

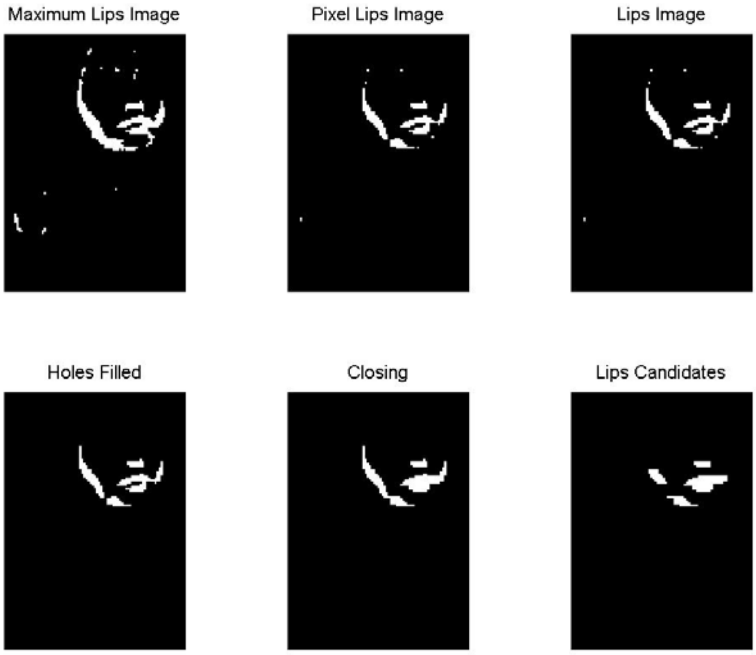

Figure 3. The Process of Finding the Lips Candidates

a b c

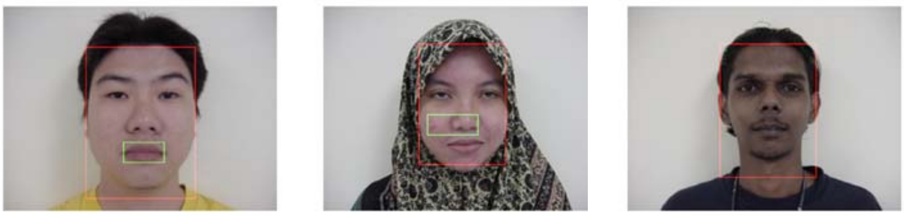

Figure 4 a) Correct Detection, b) False Acceptance, c) False Rejection

IV. CONCLUSIONS

In this paper, a component-based face detection system was presented. The system has two modules: skin detection and a lips detection. The MLP neural networks were used for both skin and lips detections using only the skin and lips colors. Post-processing was carried on the detected skin and lips pixels to detect potential face and lips candidates. If a face candidate and a lips candidate meet certain geometrical conditions then a face is detected. The system was tested on two databases with a correct detection rate of 87 and 80 percent respectively

ACKNOWLEDGMENT

The Authors would like to acknowledge the financial assistance given by the Ministry of Science and Technology, Malaysia under the E-science research grant 03-01-10-SF0045

REFERENCES

[1] Dargham j., Chekima, A.. Omatu, S. Lips detection by use of neural networks. To appears in the Proceedings of the The Thirteenth International Symposium on Artificial Life and Robotics (AROB 13th ’08), Beppu, Oita, Japan, 2008

[2] Dargham j., Chekima, A.. Skin Detection using MLP Neural Networks. Proceedings of the 9th International Conference on Mechatronics Technology, ICMT’05. Kuala Lumpur, Malaysia, December 2005.

[3] Heisele B., Serre T., Pontil M., and Poggio T.. Component-based face detection. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, volume 1, pages 657–662, Hawaii, 2001.

[4] Hsu, R.L., Abdel-Mottaleb, M., and Jain, A. K.. Face Detection In Color Images, IEEE Transaction on Pattern Analysis and Machine Intelligence, 24(5): 696-706, 2002

[5] Mitsukura, Y. Fukumi, E. Akamatsu, N.. A Design of Face Detection System by Using Lip Detection Neural Network and Skin Distinction Neural Network. Proceedings of the International Conference on Systems, Man and Cybernetics, 2789 –2793, 2000

[6] Yang M. H., Roth D., and Ahuja N.. A SNoW-Based Face Detector, Advances in Neural Information Processing Systems. S. A. Solla, T.K. Leen and K.-R. Muller (eds), pp. 855-861, MIT Press. 2000