The Thirteenth International Symposium on Artificial Life and Robotics 2008(AROB 13th '08),
B-Con Plaza, Beppu, Oita, Japan, January 31-February 2, 2008

Data Fusion for Skin Detection

Jamal Ahmad Dargham[1] Ali Chekima[1] Sigeru Omatu[2]*
[1]*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)

[2]*Computer and Systems Sciences, Graduate School of Engineering,*
Osaka Prefecture University, Sakai City, 599-8531, Japan
(omatu@cs.osakafu-u.ac.jp)

Abstract: In this paper, two methods of data fusion to improve the performance of skin detection were tested. The first method fuses two chrominance components from the same colour space while the second method fuses the outputs of two skin detection methods each based on a different colour space. The colour spaces used are the normalised rgb colour space, referred to here as pixel intensity normalisation and a new method of obtaining the r, g, and b components of the normalised rgb colour space, called maximum intensity normalisation. The MLP neural network and histogram thresholding were used for skin detection. It was found out that fusion of two chrominance components gives lower skin detection error than a single chrominance component regardless of the database or the colour space for both skin detection methods. In addition, the fusion of the outputs of two skin detection methods further reduces the skin detection error

Keyword: Skin Detection, Data Fusion, Intensity Normalisation, MLP Neural Networks, Histogram Thresholding

I. INTRODUCTION

Skin detection has been used as a primitive in a wide variety of human-related images processing systems such as face recognition1,2,3 (Albiol et al., 2001b; Brand and Mason 2001; Wang and Sung, 1999), lips reading4 (Lievin and Luthon, 2000), Hand recognition and tracking5 (Yin et al. 2001), and pornography filtering6 (Fleck et al, 1996). A skin detector should divide an image into two distinct classes one representing skin regions and the other non-skin regions. Since human-skin does not have a particular geometrical shape, the only attributes that can be used for skin detection are texture and colour. Most skin detection systems used by researchers either use a single colour system or a single skin detection technique. However, recently data fusion has been reported for skin detection. According to Steinberg7, data fusion is "a process of combining data or information to estimate or predict entity states". Thus, the important issues that need to be addressed in data fusion are: What are the features to be fused and how to fuse them? Aureli et.al.[8] presented a framework based on the employment of the fuzzy integral for fusing the outputs of three skin detector. They argued that the performance of their method subsumes the performance of more simple fusion operators. In this paper, we present two methods of data fusion to improve the performance of skin detection. The first method fuses two chrominance components from the same colour space while the second method fuses the outputs of two skin detection methods each based on a different colour space

II. THE DATABASES

Two databases were used. The first database is called the In-house database, while the second database is called the WWW database.

2.1 The In-House 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. Thus, this database has a total of 180 images. These images were all taken indoor with a single digital camera under normal lighting conditions and with uniform background. This database will be used to study the effects of scaling and facial expressions on skin and lips detection. In addition it is used to investigate the effects of minor occlusion with glasses on facial skin detection.

2.2 The WWW Database

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. This database is used to study the effects of images taken under unknown conditions on the performance of the face detection system in comparison with images taken under controlled environment.

III. 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 8 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.

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

VI. FUSING TWO CHROMINANCE COMPONENTS

To evaluate the performance of data fusion on skin detection the histogram thresholding and MLP neural networks were used.

4.1 Histogram Thresholding

When a single chrominance component is used for skin detection then the value of the pixel O(x,y) in the output image will be set to skin if the value of the pixel P(x,y) in the input image for the chrominance component C is below the threshold T^C otherwise it will be set to non-skin as illustrated in Equation 3.

$$O(x,y) = \begin{cases} \text{Skin} & \text{if } P(x,y) \geq T^C \\ \text{Non-skin} & \text{otherwise} \end{cases} \quad (3)$$

However, when two chrominance components are fused, the value of the pixel O(x,y) in the output image will be set to skin if the value of the pixel P(x,y) in the input image for the chrominance component C_1 is below the threshold T^{c_1} and if the value of the pixel P(x,y) in the input image for the chrominance component C_2 is below the threshold T^{c_2} otherwise it will be set to non-skin as illustrated in Equation 4.

$$O(x,y) = \begin{cases} Skin & if\ P(x,y) \geq T^{C_1}\ and\ P(x,y) \geq T^{C_2} \\ Non-Skin & Otherwise \end{cases} \quad (4)$$

Thus, the two chrominance components are fused by the simple and operator. The following chrominance components r, g, b, r/g, r/b which have been used by researchers as well as r-b, r-g, r.b, r.g which, to our knowledge, have not been used by other researchers for skin detection were used as well as the fusion of r and r/b, r and r-b, r and r-g, r and r.g, as well as r-b and r-g. Figures 1 and 2 show the percentage skin detection error when a single chrominance component is used as well as when two chrominance components are used for skin detection for the In-house and WWW databases respectively. As can be seen from Figures 1 and 2, the fusion of two chrominance components gives lower skin detection error regardless of the database and the colour system.

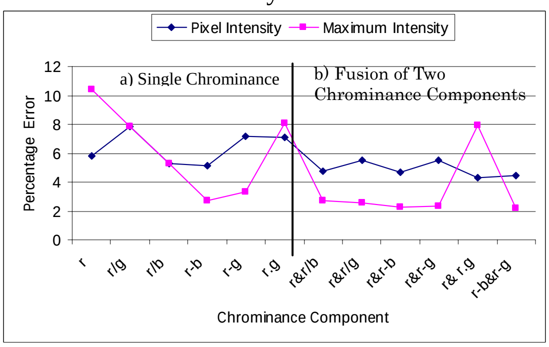

Fig. 1. Percentage skin detection error on the In-house database when (a) a single chrominance component is used and (b) when two chrominance components are used

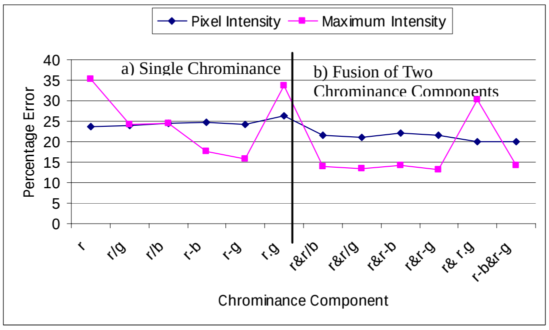

Fig. 2. Percentage skin detection error on the WWW database when (a) a single chrominance component is used and (b) when two chrominance components are used

4.2 Neural Network

The chrominance component or the combination of chrominance components that gave the lowest skin detection error when using the histogram thresholding will be used as the input to a multi-layer perceptron (MLP). The number of neurons in the input layer is 9 and 18 for a single and two chrominance components respectively. The number of neurons in the hidden layer was obtained by using a modified network growing technique. A single neuron decoded as 1 for skin and 0 for non-skin was used in the output layer. The logsig transfer function was used for all neurons. Table 1 shows the network structures for single and double chrominance components for both databases and for both intensity normalisation methods.

Table 1. Network Structures and the Corresponding Chrominance Components Used for Skin Detection With Two Chrominance Components

Database	Normalisation Method			
	Maximum Intensity		Pixel Intensity	
	Chrominance Component	Network Structure	Chrominance Component	Network Structure
In-House	r-b	9-4-1	r-b	9-15-1
	r-b and r-g	18-3-1	r-b and r-g	18-8-1
WWW	r-g	9-35-1	r.g	9-16-1
	r and r-g	18-35-1	r and r.g	18-19-1

For each network structure, 30 networks were trained. Forty percent of the skin pixels were used

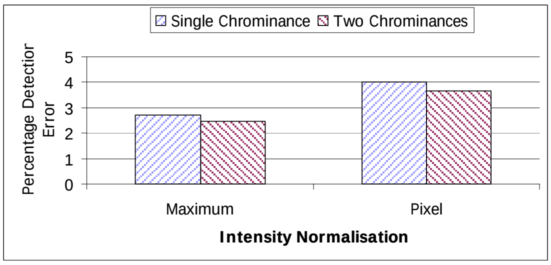

Fig. 3. Percentage skin detection error on the In-House database

for training and the other forty percent were used for validation. The networks were trained using the Levenberg-Marquardt training algorithm. The trained networks were used to segment all the images in each database and the average skin detection error for the 30 networks for each network structure for each database and for each intensity normalisation method was calculated. Figures 3 and 4 show the percentage skin detection error for the In-house and WWW databases respectively. As can be seen from Figures 3 and 4, the fusion of two chrominance components gives lower skin detection error regardless of the database or the colour system used.

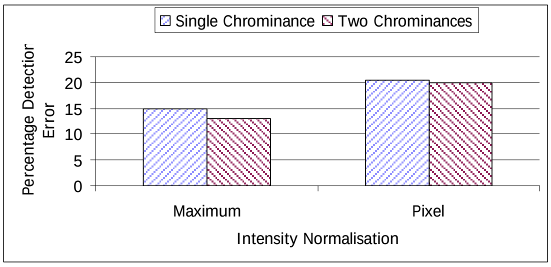

Fig. 4. Percentage skin detection error on the WWW database

V. FUSING THE OUTPUTS OF TWO SKIN DETECTION SYSTEMS

In the previous section, two chrominance components from the same colour space were fused. In this section, the outputs of two skin detection systems each using a different colour space are fused together using the AND operator as shown in Figure 5.

5.1 Histogram Thresholding

For the histogram thresholding method, the skin detection systems use the combination of two chrominance components from the same colour system for skin detection. Table 2 shows the percentage skin detection error when a single colour system is used for skin detection and when their outputs are fused together. As can be seen from Table 2, fusing the output of the two skin detector systems gives lower skin detection error for both databases.

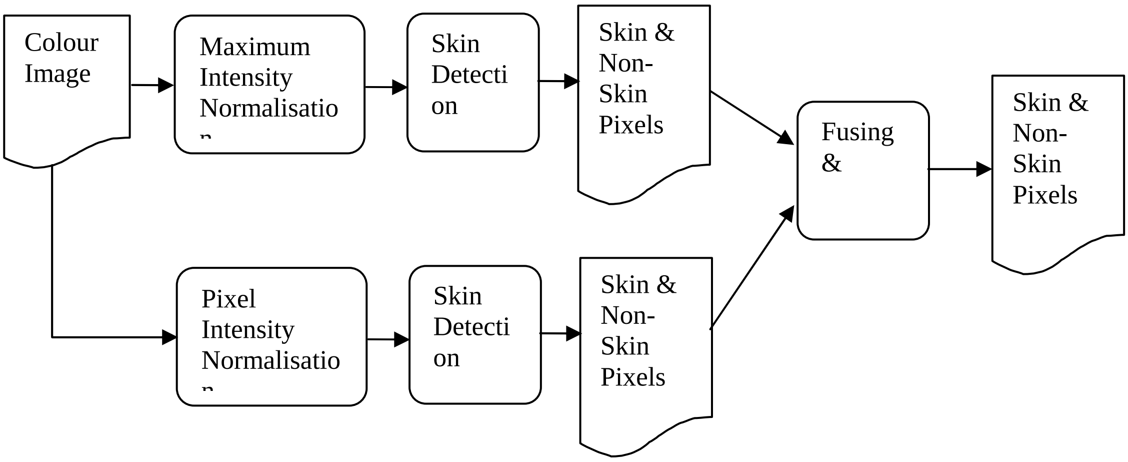

Fig. 5. Fusing the outputs of two skin detection methods

Table 2. Fusing the outputs of two skin detection systems each using different colour systems for the In-house database (top) and WWW database (bottom)

Skin Detection Method	Chrominance	Percentage Skin Detection Error
Pixel	r & r-b	4.7165
Maximum	r & r-b	2.2896
Combination		**2.1988**

Skin Detection Method	Chrominance	Percentage Skin Detection Error
Pixel	r & r.g	19.8786
Maximum	r & r-g	**13.1661**
Combination		**11.42**

5.2 Neural Networks

For the neural networks, the outputs of the networks that gave the lowest skin detection error are fused together. As can be seen from Table 3, the fusion of the outputs of two neural networks each using a different colour space give lower skin detection error than either of them used alone regardless of the database.

VI CONCLUSIONS

In this paper, two methods of data fusion to improve the performance of skin detection were tested. The first method fuses two chrominance components from the same colour space while the second method fuses the outputs of two skin detection methods each based on a different colour space. The colour spaces used are the normalized rgb and a modified version of it. The simple AND operator was used to fuse either two chrominance components from the same colour space or the outputs of two skin detection systems each using a different colour space. It was found out that the fusion with a simple operator improves the performance of the system

Table 3. Fusing the outputs of two skin detection systems each using different colour systems for the In-house and WWW databases

Skin Detection Method	Database	
	WWW	In-House
Pixel	19.4405	3.3793
Maximum	12.0505	2.3443
Combination	11.7195	2.246

REFERENCES

[1] Albiol, A., Torres, L., Delp, E.J. 2001. An Unsupervised Color Image Segmentation Algorithm for Face Detection Applications. Proceedings of the International Conference on Image Processing, Tessaloniki. 681 –684.

[2] Brand, J.D., Mason, J.S.D. 2001. Skin Probability Map and Its Use in Face Detection. Proceedings of the International Conference on Image Processing, Thessaloniki 1034 –1037.

[3] Wang Jian-Gang and Sung Eric. 1999. Frontal-view face detection and facial feature extraction using color and morphological operation. Pattern Recognition Letters, 20(10):1053-1068.

[4] Liévin M., Luthon Franck. 2000. A Hierarchical Segmentation Algorithm for Face Analysis Application for Lipreading. Proceedings of the IEEE International Conference on Multimedia and Expo, 1085-1088

[5] Yin Xiaoming, Dong Guo and Ming Xie, 2001. Hand image segmentation using color and RCE neural network, Robotics and Autonomous Systems, 34(4):235-250

[6] Fleck Margaret M., Forsyth David A., Bregler Chris. 1996. Finding Naked People. Proceeding of the 4th European Conference on Computer Vision, Cambridge.593-602.

[7] Steinberg A.N. (2001) Data Fusion Systems Engineering. IEEE Aerospace and Electronic Systems Magazine 16(6): 7-14

[8] Aureli S-F, Rodrigo V, Aitor O. (2007) Fuzzy fusion for skin detection. Fuzzy Sets and Systems 158 325 – 336