

Face and Hand Shape Segmentation Using Statistical Skin Detection for Sign Language Recognition

Bahare Jalilian^{1,*}, Abdolah Chalechale²

¹Department of Computer Engineering, Kermanshah Science and Research Branch, Islamic Azad University, Kermanshah, Iran

²Department of Computer Engineering, Razi University, Kermanshah, Iran

*Corresponding Author: baharejalilian@gmail.com

Copyright © 2013 Horizon Research Publishing All rights reserved.

Abstract An accurate face and hand segmentation is the first and important step in sign language recognition systems. In this paper, we propose a method for face and hand segmentation that helps to build a better vision based sign language recognition system. The method proposed is based on YCbCr color space, single Gaussian model, Bayes rule and morphology operations. It detects regions of face and hands in complex background and non-uniform illumination. This method tested on 700 posture images of the sign language that are performed with one hand or both hands. Experimental results show that our method has achieved a good performance for images with complex background.

Keywords Hand Segmentation, Face Segmentation, Ycbcr Color Space, Single Gaussian Model, Bayes Rule

1. Introduction

The ability to postures recognition is necessary for human machine interface. Recently, there is a significant amount of research on vision-based hand sign recognition. Image pre-processing is necessary for image enhancement and gaining good results. The first step of a posture recognition process is face and hand detection from the background. Since the following steps of recognition process strongly rely on the influence of segmentation, face/hand segmentation is the key important step in posture recognition process. Segmentation and detection of hand and face reduce processing time and increase precision of recognizing postures in sign language recognition systems. One of the major difficulties in vision based approaches is segmentation of face and hand from complex background with non-uniform illumination.

Some of the posture recognition systems can only operate in constrained environments using colored gloves [1,2], uniform or fixed backgrounds [3]. For detecting hand there are various algorithms including skin color based algorithms. Stergiopoulou and Papamarkos [4] proposed YCbCr segmentation method for hand gesture recognition,

while the background of images should be clear, simple and uniform. YCbCr color model is used in order to improve the detection accuracy [5]. HSV color space can be used to extract the skin-like region by estimating the parameter values for skin pigment [6]. In [1] color glove is used for input gestures and HSI color space is used for the segmentation process. Nine skin color classifiers (Bayesian classifier, three linear classifiers, three single Gaussian classifiers, a Gaussian mixture classifier and multilayer perceptron classifier) were compared in [7] in different color space. In [8], Y.P. Lew segmented the skin color in normalized RGB color space and modeled the distribution by Gaussian model. A combination of HSI and YCbCr color space can be used for segmentation [9]. Comparison of two chrominance models (single Gaussian model and Gaussian mixture model) and nine different chrominance spaces has been studied in [10]. All of these methods get rid of the background information which can split image in reversed side to enhance the performance.

Color is an important feature of human skin. Using skin-color as a feature has several advantages for hand and face detection. Color processing is much faster than processing other hand features. Under certain lighting conditions, color is orientation invariant. This property makes motion estimation much easier because only a translation model is needed for motion estimation. Detection human skin using color as a feature has several problems such as different cameras produce significantly different color values even for the same person under the same lighting conditions and skin color differs from person to person and changing illumination condition. In order to use color as a feature for segmentation, we have to solve these problems. The properties of skin color can be characterized by Gaussian distribution [11]. The single Gaussian model is one of the simplest models to model the distribution of the certain objects which is widely used in computer vision and pattern recognition.

In this paper a method for detection and segmentation of face and hand in color images with complex backgrounds is described. The algorithm begins with the modeling of skin and non-skin color using a database of skin and non-skin

pixels respectively. A single Gaussian model is used to estimate the underlying density function. Then, Bayes rule is used to calculate the Skin Probability Image (SPI) from an input color image. Finally, SPI is converted to a binary image and morphology operation is performed to improve the result of segmentation. Skin color is a simple but powerful pixel based feature. Also skin color analysis is robust to changes in scale, resolution.

The rest of the paper is organized as follows. A detail description of the proposed face and hand segmentation algorithm is presented in Section 2. In Section 3, we exhibit our experimental results with discussion. Finally, the conclusions and further work are presented in Section 4.

2. Materials and Methods

Image Pre-processing is necessary for image enhancement and getting good results. The proposed algorithm introduces a method for segmenting face and hands using YCbCr color space, single Gaussian model, Bayes rule and morphology operators. In particular, our proposed method consists of five main modules. The block diagram of the proposed segmentation method is shown in (Figure1).

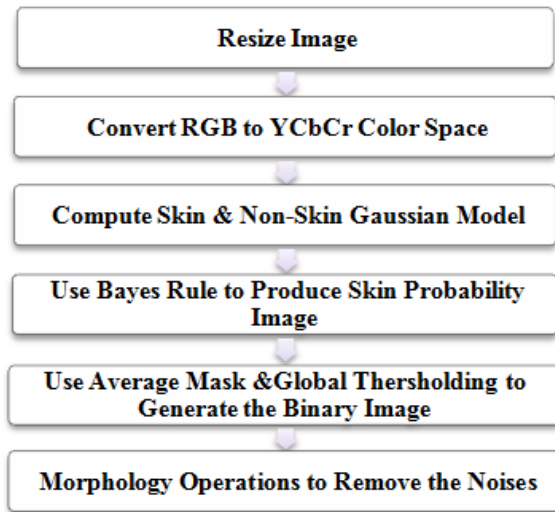


Figure 1. structure of the proposed method

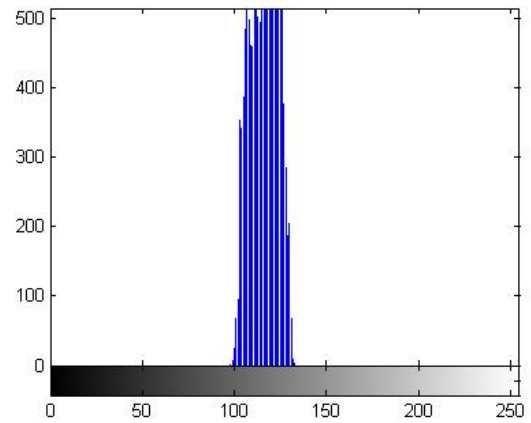
2.1. YCbCr Color Space Conversion

One important factor that should be considered while building a statistical model for color is the choice of a right color space. For most images, the RGB color space is considered as the default color space. In order to convert into other color spaces, we can apply linear or non-linear transformation on the RGB components. In this algorithm, the input RGB image is converted in to YCbCr images due to the fact that RGB color space is more sensitive to different light conditions so we need to transform the RGB values in to YCbCr. The color space transformation is assumed to decrease the overlap between skin and non-skin pixels to

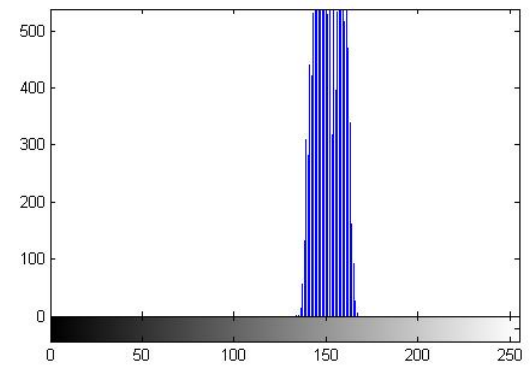
classify skin-pixel and to provide robust parameter against varying illumination conditions. RGB values can be transformed to YCbCr color space using (1):

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 16 \\ 128 \\ 128 \end{pmatrix} + \begin{pmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1)$$

If only the chrominance component is used, segmentation of skin colored regions becomes powerful in this process. Therefore, the variations of luminance component are eliminated as much as possible by choosing the CbCr plane (chrominance components) of the YCbCr color space to build the model. Research has shown that skin color is clustered in a small region of the chrominance space [4], as shown in (Figure 2).



(a)



(b)

Figure 2. Histogram of (a) Cb component and (b) Cr component.

2.2. Single Gaussian Model

The skin color distribution in CbCr plane is modeled as a single Gaussian model. According to section 2.1 the reason for using a single Gaussian model is the localization of skin color to a small area in the CbCr chrominance space. A database of labeled skin pixels is used to train the Gaussian

model. The mean and the covariance of the database characterize the model. The images containing human skin pixels as well as non-skin pixels are collected. The skin pixels from these images are carefully cropped out to form a set of training images.

Let $c = [Cb Cr]^T$ denote the chrominance vector of an input pixel. Then the probability that the given pixel lies in the skin distribution is given by (2):

$$p(c / skin) = \frac{\exp\left[-\frac{1}{2}(c - \mu_s)^T \Sigma_s^{-1} (c - \mu_s)\right]}{2\pi\sqrt{|\Sigma_s|}} \quad (2)$$

where μ_s and Σ_s represent the mean vector and the covariance matrix of the training pixels, respectively. Thus the mean and the covariance have to be estimated from the training data to characterize the skin color distribution as illustrated by (3) and (4). In these equations, n is number of samples in training set.

$$\mu_s = \frac{1}{n} \sum_{j=1}^n c_j \quad (3)$$

$$\Sigma_s = \frac{1}{n-1} \sum_{j=1}^n (c_j - \mu_s)(c_j - \mu_s)^T \quad (4)$$

This model is used to obtain the Skin Probability Image of an input color image. Once the skin color is modeled using a single Gaussian, it can be used to calculate the probability of an input pixel representing skin, i.e. $p(skin/c)$, where c is the input color value. $p(c/skin)$ is again used to compute the required probability $p(skin/c)$. To compute this probability, the Bayes rule is used [12]:

$$p(skin / c) = \frac{p(c / skin)P(skin)}{p(c / skin).P(skin) + p(c / non - skin).P(non - skin)} \quad (5)$$

To calculate the probability, $p(skin/c)$, for each input pixel, The probabilities $p(skin)$ and $p(non-skin)$ can be estimated from skin and non-skin image in the training database [12]. In this study, for training set we assumed that all pixels are belong to the skin or non-skin clusters. Hence, we used:

$$p(skin) = p(non - skin) = 0.5 \quad (6)$$

$$p(skin / c) = \frac{p(c / skin)}{p(c / skin) + p(c / non - skin)} \quad (7)$$

To compute the probability $p(c/non-skin)$ a Gaussian model similar to skin pixels is built for non-skin pixels also which is called the non-skin or the background model. Thus, the two conditional probabilities and the above ratio are computed pixel-by-pixel to give the probability of each pixel representing skin given its chrominance vector c . These results are in a gray level image where the gray value at a pixel indicates the probability of that pixel representing skin. This is called the Skin Probability Image given by (8):

$$SPI(i, j) = a.p(skin / c_{ij}) \quad (8)$$

Where a is a proper scaling factor and c_{ij} is the chrominance value of pixel (i, j) . Here a is chosen to be 255 so that the highest probability value results in a gray level of 255 in the Skin Probability Image. Then gray image obtained is converted into a binary image by global thresholding. In order that enhancement of global thresholding before thresholding, the gray image (the Skin Probability Image) is smoothed by Average mask.

2.3. Morphology Operations

The binary image obtained in the previous section may contain white pixels at non-skin regions (background) where the background color resembles the color of skin or black pixels at hand region. These noises may be caused due to bad lighting conditions or existing pixels similar to skin pixels in those regions. In order to detect the hand and the face clearly, it has been further implemented morphology operations, to fill up the black pixels on the segmented hand and white pixels on background. There are two operations involve namely dilation and erosion. Firstly, dilation operation is performed. Dilation adds pixels to fill up any missing pixels in hand region. Secondly, erosion operation is performed. Erosion removes any white pixels which do not belong to the hand region. This stage is performed to improve the result of hand segmentation.

(Figure 3) shows the images obtained after applying each of steps in the proposed segmentation algorithm. (Figure 3(a)) is original image; (Figure 3(b)) is Skin Probability Image. (Figure 3(c)) is the binary image after thresholding. (Figure 3(d)) is the hand detected after morphology operation.

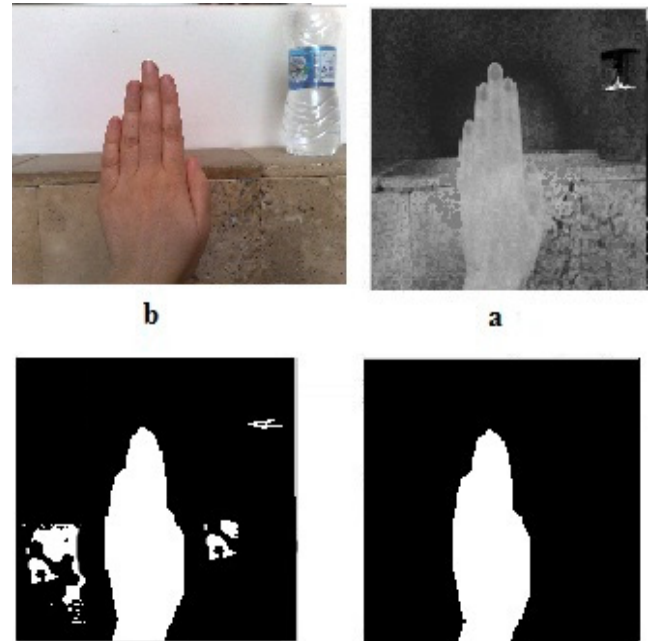


Figure 3. (a): original image, (b): skin probability image, (c): segmented hand after thresholding (d): the hand detected after morphology operations

3. Results and Discussion

In order to evaluate the segmentation algorithm's performance, we have applied it on 700 posture images of the sign language that are performed with one hand or both hands. These images are taken in different and complex environments, under several daylight and illumination conditions, as well as different persons are used to perform signs. The pixels similar to skin pixels are used in shooting background for more challenging. Some of these posture images are shown in (Figure 4).

For the skin modeling a training set of skin images are built. These images were collected by cropping off skin regions in 95 images containing hands of different persons that were created manually. Some of the skin images from

the skin database are shown in (Figure 5). Besides, a training database containing 95 non-skin images are used for the background modeling. For obtaining more accuracy, we added to the database the pixels similar to skin pixels. Some of non-skin images from the background database are shown in (Figure 6).

The mean and the covariance of the skin database specify the skin model. Similarly, the mean and the covariance of the background database characterize the non-skin model. As discussed in Section 2.1, the skin color is localized to a small region in the CbCr chrominance space. The single Gaussian skin and non-skin models and Bayesian rule are used to calculate the Skin Probability Image for an input test image.

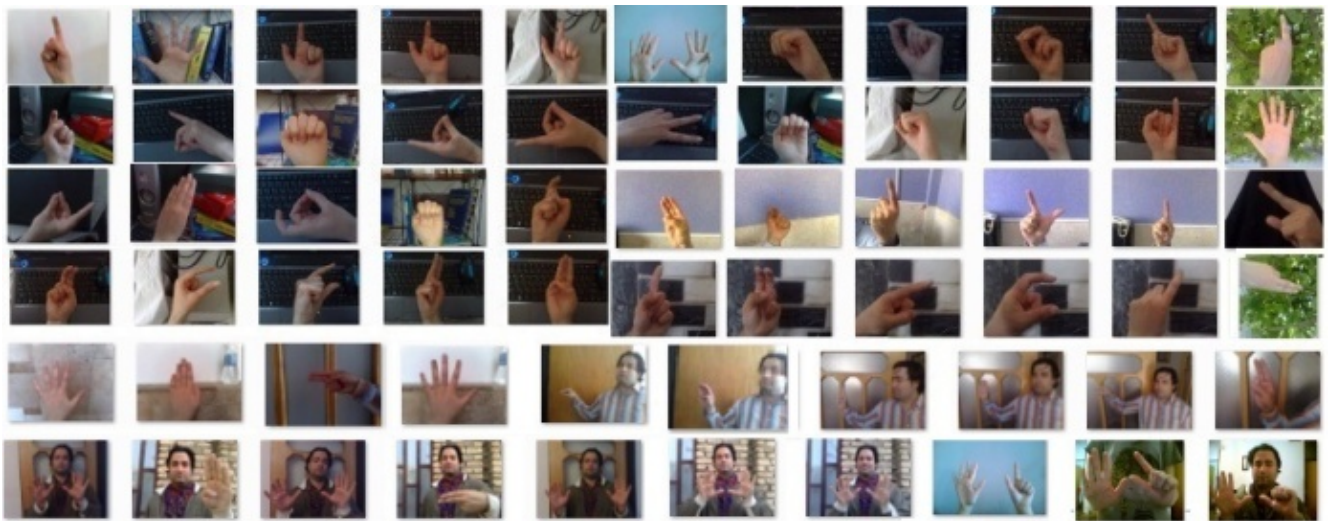


Figure 4. some of postures images of the sign language with complex background and changing illumination.

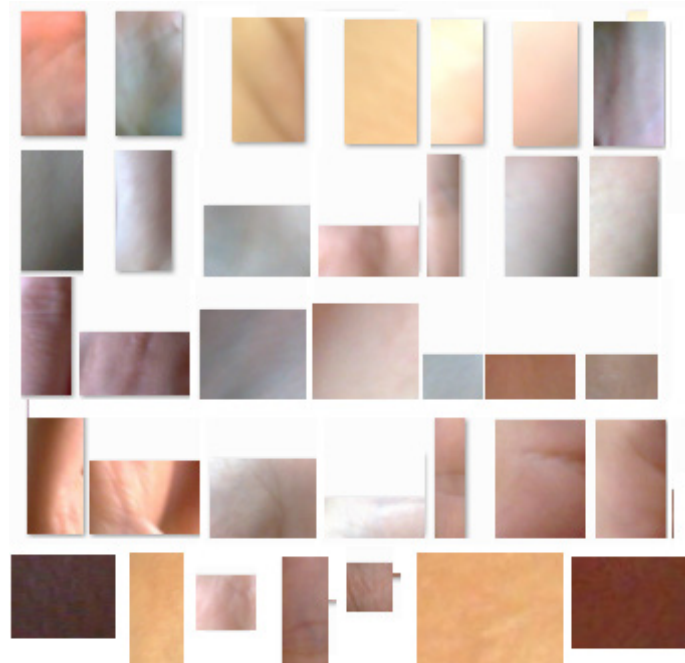


Figure 5. some of skin images from the skin database

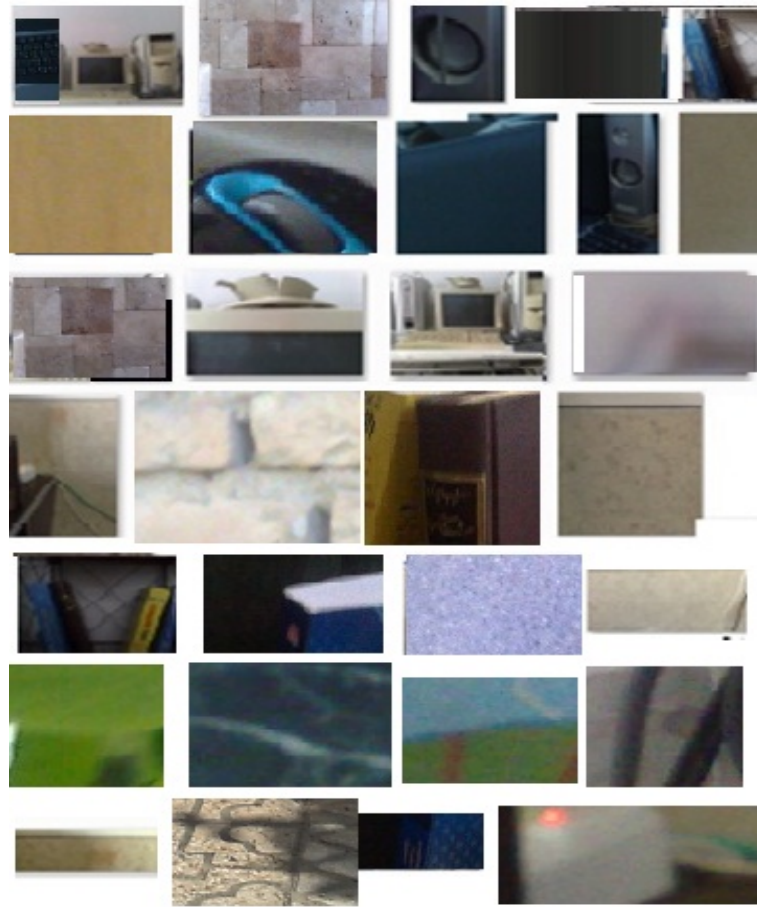


Figure 6. some of non-Skin images from the non-Skin Database

The results obtained after applying the proposed algorithm are shown in (Figure 7). Original images with different background and various skin color are shown in (Figure 7(a,b,c,d)). (Figure 7(e,f,g,h)) are Single Probability Images. After thresholding, binary segmented images are produced, see (Figure 7(i,j,k,l)). To detect face and hand clearly, it has been further implemented the erosion and dilation operations, to fill up the black pixels on the segmented image and white pixels on background, as shown in (Figure 7(m,n,o,p)).

To evaluate the segmentation, this paper utilizes true positive rate (TPR), and false negative rate (FPR) for the images such as shown in (Figure 7(g,h,I)) using equations (8) and (9), respectively:

$$TPR = \frac{TP}{TP + FN} \times 100 \tag{8}$$

$$FPR = \frac{FP}{FP + TN} \times 100 \tag{9}$$

Where *TP* is the True Positive, *FN* is the False Negative and *TN* is the True Negative values. The segmentation results are listed in Table I. It is observed that our method provides better results than other skin color based method. The proposed method has high *TPR* and low *FPR*. Hands and face under exposure of changing illumination condition are well segmented using our algorithm as shown in (Figure 7).

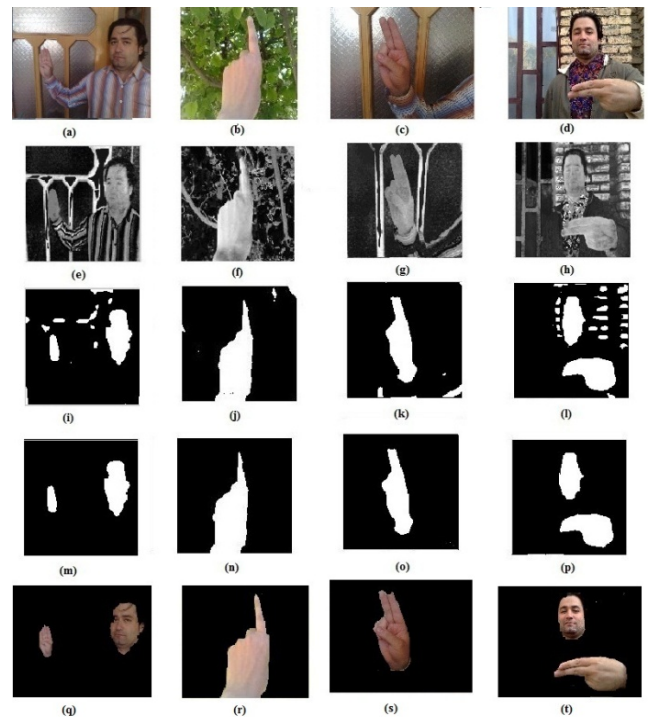


Figure 7a,b,c,d. original images; e,f,g,h: skin probability images; i,j,k,l: segmented hands; m,n,o,p: hands detected after morphology operations; q, r, s, t: final results.

Table 1. Performance of different skin detectors reported by the authors

Method	TPR	FPR
Single Gaussian in I-Q[13]	90%	33.3%
Single Gaussian in CbCr[13]	90%	33.4%
Elliptical boundary model in CIE-xy[13]	90%	20.9%
Gaussian Mixture models in RGB [14]	80%	9.5%
Modified ellipse model[15]	93.1%	11.9%
Histogram model[15]	94%	12.4%
Proposed Method	95%	4.2%

3. Conclusion

In this paper, a segmentation method using single Gaussian model in CbCr color space, Bayes rule, a suitable thresholding and morphology operations is proposed for face and hand detection phase in sign language recognition systems. In this method, a model for human skin color distribution is built using a database of labeled skin pixels. A single Gaussian model has been trained using the same database of skin pixels. Likewise, a single Gaussian model for the non-skin background pixels is built. Then, Bayes rule is used to calculate the Skin Probability Image (SPI) that shows probability of an input pixel representing skin. Next, SPI is smoothed by Average mask and is converted into binary image by global thresholding. Finally, morphology operations are applied on binary image to improve the result of segmentation phase.

Comparing to other algorithms in the domain of skin color modeling and Gaussian distribution by using YCbCr and RGB color space, the performance of our algorithm is better because it can efficiently remove most of the non-skin colored background objects. Experimental results show that the proposed method can achieve a relatively high TPR and low FPR. In the future work, we intend to use the results of proposed hand and face segmentation in a sign language recognition system.

REFERENCES

- [1] M. Maraqa and R. Abu-Zaiter, Recognition of Arabic Sign Language (ArSL) Using Recurrent Neural Networks, *Journal of Intelligent Learning Systems and Applications*, vol.4, No.1, 41-52, 2012.
- [2] L. Lamberti and F. Camastra, Real-time Hand Gesture Recognition Using A Color Glove, Springer-Verlag Berlin Heidelberg, Vol.6978, No.1, 365-373, 2011.
- [3] R. Grzeszczuk, G. Bradski, M. Chu and J. Bouguet, Stereo Based Gesture Recognition Invariant to 3D Pose and Lighting, *Computer Vision and Pattern Recognition, IEEE Computer Society Conference*, 18-26, 2000.
- [4] E. Stergiopoulou and N. Papamarkos, Hand Gesture Recognition Using a Neural Network Shape Fitting Technique, *Elsevier Engineering Applications of Artificial Intelligence*, Vol.22, No.8, 1141-1158, 2009.
- [5] C. Chuqing and L. Ruifeng, Real-Time Hand Posture Recognition Using Haar-like and Topological Feature, *International Conference on Machine Vision and Human-machine Interface*, 683 – 687, 2010.
- [6] M.M. Hasan and P.K. Mishra, HSV Brightness Factor Matching for Gesture Recognition System, *International Journal of Image Processing*, Vol.4, No.5, 456-467, 2011.
- [7] S.L. Phung, Skin Segmentation Using Color Pixel Classification: Analysis and Comparison, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol.27, No.1, 148-154, 2005.
- [8] Y.P. Lew, A.R. Ramli and S.Y. Koay, A Hand Segmentation Scheme Using Clustering Technique in Homogeneous Background, *Student Conference on Research and Development Proceedings*, 305-358, 2002.
- [9] H.C.V. Lakshmi and S. Patikulakarni, Segmentation Algorithm for Multiple Face Detection in Color Images with Skin Tone Regions Using Color Spaces and Edge Detection Techniques, *International Conference on Signal Acquisition and Processing*, 162-166, 2010.
- [10] J.C. Terrillon, M.N. Shirazi and H. Fukamachi, Comparative Performance of Different Skin Chrominance Models and Chrominance Spaces for the Automatic Detection of Human Faces in Color Images, *Automatic Face and Gesture Recognition, Fourth IEEE International Conference on Automatic Face and Recognition*, 54-61, 2000.
- [11] M. H. Yang and N. Ahuja, Gaussian Mixture Model for Human Skin Color and Its Applications in Image and Video Databases, *Storage and Retrieval for Image and Video Databases*, Vol.3656, No.7, 1999.
- [12] V. Vezhnevets, V. Sazonov and A. Andreeva, A Survey on Pixel-Based Skin Color Detection Techniques, *GRAPHICON3*, 85–92, 2003.
- [13] J. Y Lee and S.I. Yoo, An Elliptical Boundary Model for Skin Color Detection. In *proceedings of the International Conference on Imaging Science, Systems and Technology*, 2002.
- [14] M.J. Jones and J.M. Rehg, Statistical Color Models with Application to Skin Detection, *International Journal of Computer Vision*, Vol.46, No.1, 81-96, 2002.
- [15] H.k Tang and Z.Q Feng, Hand's Skin Detection Based on Ellipse Clustering, *International Symposium on Computer Science and Computational Technology*, 758-761, 2008.