Global Journal of Computer Science and Technology
Volume 12 Issue 3 Version 1.0 Fabruary 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Two-Handed Sign Language Recognition for Bangla Character Using Normalized Cross Correlation

By Kaushik Deb, Helena Parvin Mony & Sujan Chowdhury
Chittagong University of Engineering and Technology,Chittagong, Bangladesh

Abstract - Sign language detection and recognition (SLDR) using computer vision is a very challenging task. In respect to Bangladesh, sign language users are around 2.4 million [1]. In this paper, we try to focus for communicating with those users by computer vision. In this respect, an efficient method is propose consists of two basic steps: (a) refinement and (b) recognition. Initially in refinement, a Red-Green-Blue (RGB) color model is adopted to select heuristically threshold value for detecting candidate regions (i.e. hand and wrist band sign regions). After the candidate regions are obtained by applying color segmentation, then procedures for refining the candidate region are followed by using two different color wrist band regions and filtering. Finally, statistically based template matching technique is used for recognition of hand sign regions. Various hand sign images are used to test the proposed method and results are presented to provide its effectiveness.

Keywords : RGB color model, Color segmentation,Template matching.

GJCST Classification: I.2.9

TWO-HANDED SIGN LANGUAGE RECOGNITION FOR BANGLA CHARACTER USING NORMALIZED CROSS CORRELATION

Strictly as per the compliance and regulations of:

Two-Handed Sign Language Recognition for Bangla Character Using Normalized Cross Correlation

Kaushik Deb [α], Muhammad Ibrahim Khan [α], Helena Parvin Mony [σ] & Sujan Chowdhury [α]

Abstract - Sign language detection and recognition (SLDR) using computer vision is a very challenging task. In respect to Bangladesh, sign language users are around 2.4 million [1]. In this paper, we try to focus for communicating with those users by computer vision. In this respect, an efficient method is propose consists of two basic steps: (a) refinement and (b) recognition. Initially in refinement, a Red-Green-Blue (RGB) color model is adopted to select heuristically threshold value for detecting candidate regions (i.e. hand and wrist band sign regions). After the candidate regions are obtained by applying color segmentation, then procedures for refining the candidate region are followed by using two different color wrist band regions and filtering. Finally, statistically based template matching technique is used for recognition of hand sign regions. Various hand sign images are used to test the proposed method and results are presented to provide its effectiveness.

Keywords : RGB color model, Color segmentation, Template matching.

I. Introduction

According to sociolinguistic survey deaf, dumb and sign language users are neglected by society. In time to time, the scenario becomes changed. In the modern world, they are also treated as imaginative, creative and as intelligent as any other normal human being. But their disabilities is the main obstacle to lead a normal social life. According to study, instead of verbal communication, the deaf and dumb people use sign language, which is a visual form of communication including the combination of hand shapes, orientation and movement of the hands, arms or the body, and facial expressions. Sign language is the organized collection of gestures. Gestures are usually understood as hand and body movement which can pass information from one to another. In this paper we work with two hand gesture. Many researchers have attempted to recognize sign language through various techniques. However none of them have ventured into the area of Bangla sign language. However, Bangla-sign language users community is the largest community among the language based minority communities in Bangladesh. We focused our research on deaf and dump that live in Bangladesh. This paper propose a framework for Bangla SLDR. The basic concept involves the use of wearing two different color wrist bands by the disabled person who makes the sign and will interact with the system by gesturing in the view of the camera. Shape and position information about the hand will be gathered using detection of skin and wrist band color. The SLDR task is quite challenging from sign images due to viewpoint changes, uneven illumination condition during image acquisition. As far as detection and recognition of hand sign region are concerned, researchers have found various methods of hand sign detection and recognition. For example, glove based recognition strategy had developed in Pakistan sign language, described in [4]. Authors describe the use of statistical template matching for gesture recognition in Boltay Haath. In [2], the model describes the development of a video-based continuous sign language recognition system. The system is based on continuous density Hidden Markov Models (HMM) with one model for each sign which is developed for German sign language. According to [3], the one-state transitions of the English language are projected into shape space for tracking and model prediction using a HMM like approach. In [5], accomplished a computer based method to generate American sign language gestures from natural speech and display them through a 2D animated character. The domain is chosen as the sequential generation of gestures of letters in a text, corresponding to the speech. In the second segment of the work, we have implemented a system to generate American Sign Language gestures from speech and display them by a 2D Virtual Human.

However, to the best of our knowledge all previous work on Bangla hand sign recognition is done only on one hand. In this respect, the main emphasis of this paper is to develop computerized sign language recognition system for two-handed Bangla sign language. In this regard, ten (10) different hand sign created by various people and matched it with our dataset and get better result and ensure its effectiveness.

Author [α] : Faculty Member, Department of Computer Science and Engineering, Chittagong University of Engineering and Technology, Chittagong, Bangladesh
Author [σ] : Department of Computer Science and Engineering, Chittagong University of Engineering and Technology, Chittagong, Bangladesh

II. Description of Two- Handed Sign

The sign language used by the deaf community in Bangladesh is called Bangla Sign Language Anthology (BSLA) controlled by Centre for Disability Development (CDD). Like other sign languages of the world, it is a rich and complex visual-spatial language, with a vocabulary and syntax of its own. It is different from other sign languages and, of course, from spoken language such as English. It includes hand shapes and movements, facial expression and body movements to express meaning, and can be used to express a full range of meaning. Various types of two handed Bangla sign language images are shown in Figure 1. All signs have two different color wrist bands, one is red and another is swan. Each sign of these images consist of four different colors: red, swan, skin color and black (background color). Color is an identifiable feature and wrist band color is an invariant manner is of main objectives of the proposed color based hand sign extraction method.

Fig.1. Two -handed sign images.

III. The Proposed Framework

The following block diagram shows the working approach towards the solution of the stated problem. In order to fulfill our desired goal, the proposed framework divides the whole process into two basic steps: a) detection and b) recognition.

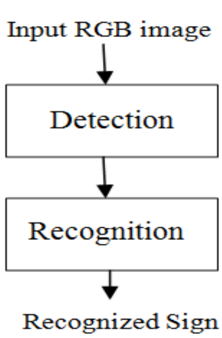

Fig.2. The proposed hand sign recognition framework.

IV. Detection Algorithm Architecture

According to our proposed framework first portion is detection module as shown in Figure 3. Detection means find out proper candidate region that will be used as an input in algorithmic implementation of recognition. The propose detection module consists of three major steps i.e. color segmentation, centroid calculation and localizing the edge of two wrist bands. The output of the initial step will be the binary image. In second step, center is calculated to detect the edge point. Finally, candidate region is extracted by removing forearm pixel.

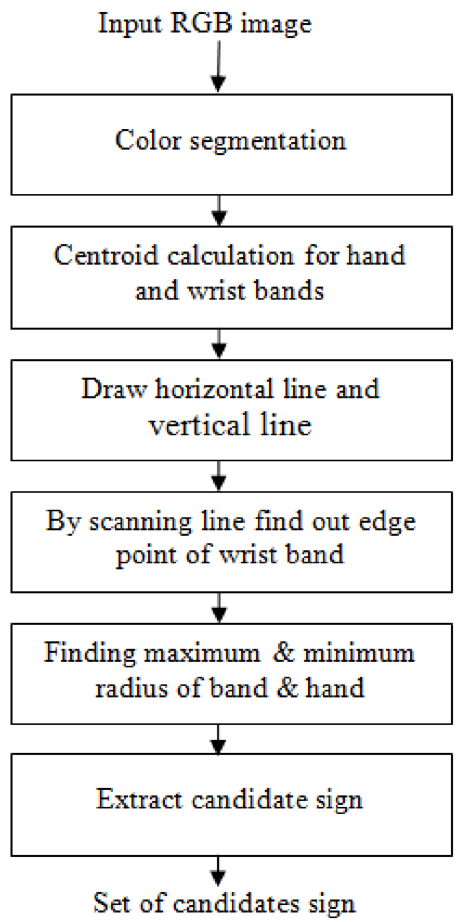

Fig.3. The proposed hand sign detection algorithm.

a) Color segmentation

Color is a distinctive feature of image. In this proposed method, candidate region is detected based on its color properties. For detecting hand and two wrist bands region a threshold values are obtained using heuristical approach on different hand images. The result of this step is two binary masks (where 0 means background and 1 means possible sign) which represent candidate locations.

Pixels belongs to hand sign if

$$(I_{i,j}^{R} > .40) \,\&\, (I_{i,j}^{G} > .36) \,\&\, (I_{i,j}^{B} > .30) \qquad (1)$$

Pixels belongs of red band sign if

$$(I_{i,j}^{R} > .40) \,\&\, (I_{i,j}^{G} > .007) \,\&\, (I_{i,j}^{B} > .02) \qquad (2)$$

And pixels belongs of swan band sign if

$$(I_{i,j}^{R} > .01) \,\&\, (I_{i,j}^{G} > .57) \,\&\, (I_{i,j}^{B} > .50) \qquad (3)$$

Where $I_{i,j}^{R}, I_{i,j}^{G}, I_{i,j}^{B}$ pixels are belongs to red, green and blue component of input image I with coordinates (i, j). A hand sign image and its color segmentation results are shown in Fig. 4.

Fig.4. A hand sign image (left) and its color segmentation results (right) using an RGB color model.

In color segmentation, we get three different regions, such as skin color region and two wrist band color region. The skin color is discrete and it is separated by wrist band color. In order to obtain the full desired binary image we have to make an OR masking operation. Obtained binary image still contain noises and that is not ideal. These noise means forearm pixels of hand.

As there is no difference between the color ranges of a skin pixel of the hand and a skin pixel of the forearm, our next approach is to remove forearm pixel from the position information. We can obtain it using centroid calculation and localizing the wrist band which is described in later part. In first step, calculate the centroid of band and hand and in next step try to find the edge point, which will help us to remove the noise and obtain the candidate image for further decision.

b) Centroids calculation

By averaging the position of the pixels that is detected earlier, it is possible to calculate the centroid of both the hand and two wrist bands.

The following equation used to find the hand & band centroid is as follows:

$$C_{hand,band} = \frac{1}{|N|}\sum_{i=1}^{N} X \qquad (4)$$

Where N is total number of pixels

c) Localizing the edge of two wrist band

In this step we try to find out the edge of the two wrist bands. It was considered that if the distance and angle of the edges of the wrist band relative to the hand centroid, the forearm skin pixels could be removed by comparing their distances and angles with them.

The equation used for vector joining for two centroid is

$$C_{dif} = (x_{dif}, y_{dif}) = C_{hand} - C_{band} \qquad (5)$$

The yaw angle of the hand is :

$$\theta_{hand} = \tan^{-1}(y_{diff} / x_{diff}) \qquad (6)$$

Horizontal & Vertical Line through center of band are found by this equation

$$P_1(s_1) = C_{band} + s_1 \begin{bmatrix} \cos\left(\theta_{hand} + \frac{\pi}{2}\right) \\ \sin\left(\theta_{hand} + \frac{\pi}{2}\right) \end{bmatrix} \qquad (7)$$

Where ($-35 < s_1 < 35$)

For each s_1 count the number of wrist band pixels $n(s_1)$ along the line:

$$P_2(S_1, S_2) = P_1 + S_2 \begin{bmatrix} \cos(\theta_{hand}) \\ \sin(\theta_{hand}) \end{bmatrix} \qquad (8)$$

Where ($-50 < s_2 < 50$)

The edges points of the band left and right are equal to $P_1(s_1)$ when $n(s_1)$ falls below a certain threshold. In this way, find the edge point of two wrist band as well as maximum & minimum radius of band & hand.

$$r_{band} = \max\left(\left|band_{left} - C_{band}\right|, \left|band_{right} - C_{band}\right|\right) \quad (9)$$

In this way, removing the forearm and band pixels and extract the candidate sign as shown in Figure 5. During this step, main geometrical properties of sign candidate such as area, is computed. And also combination of pixel position and priority based information is used to remove erroneous detected pixels.

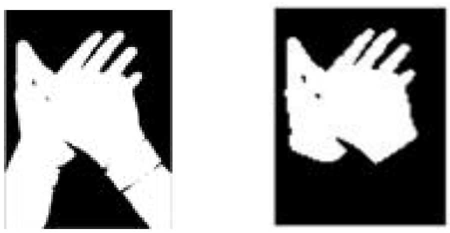

Fig.5. A binary sign image with forearm pixel (a), candidate detected by proposed method (b) after removing forearm pixel.

V. Sign Reconized Algorithm

Next step of proposed framework is recognizing the hand sign after extracting the candidate regions from detection step. In this paper, template matching with the use of normalizes cross-correlation to perform recognition for different types of hand sign. The proposed recognition scheme is shown in Figure 6.

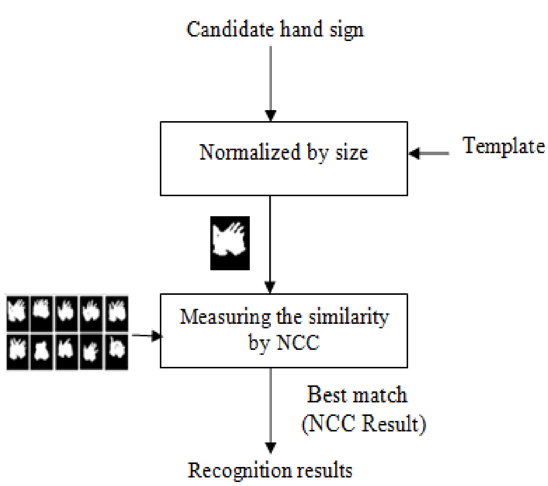

Fig.6. The proposed hand sign recognition scheme.

Before recognition scheme, the candidate hand signs are normalized. Normalization is to refine the hand signs into a block containing no extra white spaces (i.e. pixels) in all four sides of the hand sign. Then each hand sign is fit to equal size. Fitting approach is necessary for template matching. For matching the candidate hand sign with the template, input candidate sign images must be equal sized with the template sign. Here the sign are fit to 96×72. For recognition of hand sign, we use the database where containing 10 prototypes for the hand signs with the size of 96×72. The database formed is shown in Figure 7.

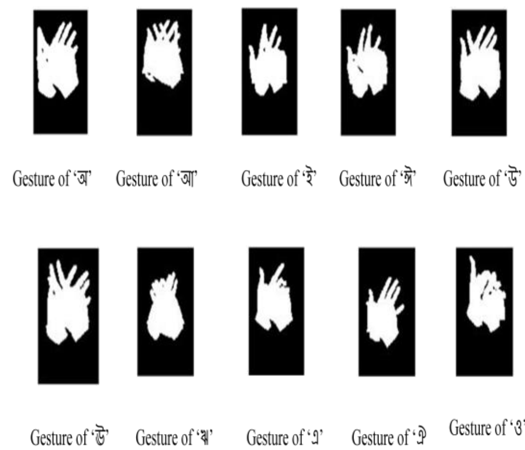

Fig.7. Template (10 prototypes for the hand sign images with Bangla meaning) used for pattern matching.

To measure the similarity and find the best match, a statistical method correlation is used [7]. Correlation is an effective technique for image recognition. This method measures the correlation coefficient between a number of known images with the same sized unknown images or part of an image with the highest correlation coefficient between the images producing the best match.

Let t be a template image and $\bar{t}$ its average of binary image. Let us assume that c is a candidate hand sign image, having the same size of the template and let $\bar{c}$ be its average of binary image. We use the normalized cross-correlation function between the image pair and define in the discrete case as follows.

$$Ncc_{ct} = \frac{\sum_{x=0}^{m-1}\sum_{y=0}^{n-1}(c-\bar{c})(t-\bar{t})}{\sqrt{\sum_{x=0}^{m-1}\sum_{y=0}^{n-1}(c-\bar{c})^2(t-\bar{t})^2}}$$

Where Ncc_{ct} is the correlation coefficient. The candidate hand sign recognition process is based on the value of the correlation coefficient. If the value of the correlation coefficient exceeds a threshold set by the user, then the similarity measure is large enough and input hand sign can be assumed to present. Finally, a red box on the target hand sign is plotted as shown in Figure 8.

Illustration of hand sign recognition is portrayed in Figure 8.

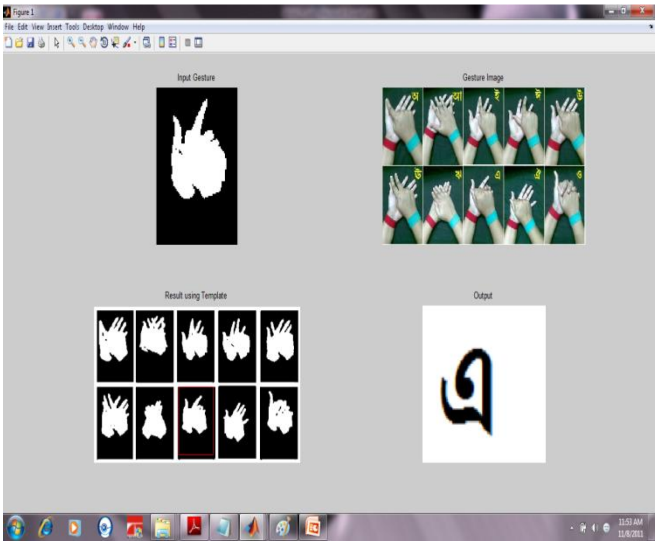

Fig.8. Illustration of hand sign recognition: (a) extracted candidate binary hand sign image, (b) gesture image, (c) and (d) candidate hand signs measuring the similarity by NCC and recognized gesture sign converted to text, respectively.

VI. Result Discussion

All experiments were done on core-i3 2.53 GHz with 2 GB RAM under MATLAB environment. In the experiments, 80 images were employed and the size of the images is 320×240 pixels. Some images are shown in Figure 7. The hand sign detection and recognition results, and average computational cost for hand sign recognition are shown in Tables 1, 2, and 3.

Table 1 : Detection results.

Input Image	Detected hand sign	Success rate (%)
80	78	97.5

Table 2 : Recognition results.

Stage	Success rate (%)
Recognition of hand sign	96

Table 3 : Average computational cost.

Computational Time of Refinement Step		Computational Time of Recognition Step	
Stage	Avg. Time(s)	Stage	Avg. Time(s)
Color segmentation	.39	Normalization	.21
Remove forearm pixel	.27	Plotted output	1.92
-	-	Convert to text	.37
	.66		2.50

VII. Conclusions

In this paper, we adopt a new method for automatic recognition process of sign language to develop the quality of life of these disable people. Initially in refinement, a Red-Green-Blue (RGB) color model is adopted to select heuristically threshold value for detecting candidate regions (i.e. hand and wrist band sign regions). After the candidate regions are obtained by applying color segmentation, then procedures for refining the candidate region are followed by using two different color wrist band regions and filtering. And also combination of pixel position and priority based information was used to remove any erroneous detected pixels. Finally, we focus in this paper on the conducting an experiment using template matching technique with the use of normalized cross-correlation to perform for recognition of hand sign regions.

While conducting the experiments, different illumination conditions and varied distances between hand sign and camera often occurred. In such cases, confirmed the result is very effective when the proposed method is used. However, the proposed method is sensitive with different back ground condition and rotated hand sign images. We leave these issues for consideration in future studies.

References Références Referencias

1. “Bangla Sign Language Anthology: CDD's JOURNEY”, Centre for Disability in Development (CDD), Bangladesh, 2009.
2. Siti Noraini Sulaiman, Mohd Fauzi Alias, Nor Ashidi Mat Isa and Mohamad Faizal Abd Rahman, “An Expert Image Processing System on Template Matching”, School of Electrical & Electronic Engineering, Universiti Sains Malaysia, Malaysia, July 2007.
3. Aleem Khalid Alvi, M. Yousuf Bin Azhar, Mehmood Usman, Suleman Mumtaz, Sameer Rafiq, Razi Ur Rehman,IsrarAhmed, “Pakistan Sign Language Recognition Using Statistical Template Matching”, World Academy of Science, Engineering and Technology, Page 52-54, 2005.

4. Ryo Takei, "A New Grey-Scale Template Image Matching Algorithm Using the Cross-Sectional Histogram Correlation Method", Dynax Corporation Fuchu-City, Fuchu-cho, Tokyo, Japan, 2003.
5. Bowden & Sarhadi, "Building temporal models for gesture recognition", in proceedings British Machine Vision Conference, pages 32-41, 2000.
6. Bauer & Hienz, "Relevant feature for video-based continuous sign language Recognition", Department of Technical Computer Science, Aachen University of Technology, Aachen, Germany, 2000.
7. Ray Lockton, "Hand Gesture Recognition Using Computer Vision", Balliol College, 4th year project Oxford University.