Front.Comput.Sci.
DOI

RESEARCH ARTICLE

A Real-Time Hand-Signs Segmentation And Classification System Using Fuzzy Rule Based RGB Model And Grid-Pattern Analysis

Muhammad Aminur Rahaman (✉)[1], Mahmood Jasim [1], Md. Haider Ali [1], Tao Zhang [2], and Md. Hasanuzzaman [1]

1 Department of Computer Science and Engineering, University of Dhaka, Dhaka-1000, Bangladesh
2 Department of Automation, School of Information Science & Engineering, Tsinghua University, Beijing 100084, China

Abstract Hand-signs segmentation and classification in cluttered and dynamic background is extremely challenging, especially when there are variation of illumination in the environment. This paper presents a robust solution for hand-signs segmentation and classification using Fuzzy rule based RGB model and grid-pattern analysis in real-time. The system uses Haar-classifiers to initialize region of interest (ROI) by detecting predefined hand-signs. The proposed Fuzzy rule based RGB (FRB-RGB) model is used to segment human skin-color from the ROI, in which each skin-color pixel is classified according to fuzzy rules. The system extracts target binary hand-signs based on segmented skin-color pixels with specific motions. The use of FRB-RGB model with Haar-classifier and motion queue makes the system robust in extracting target binary hand-signs only from the ROI instead of whole image in different challenging environment. To remove noise, morphological closing operation is done on the binary image and then Gaussian smoothing filter is used to smooth the outer boundary of the target binary hand-signs. The system applies rotation followed by clipping and normalization on the binary image to make it rotation and scale invariant. A 5×5 Window-Grid Mask (WGM) is applied on the normalized binary image to generate Grid-Pattern which is represented by Window-Grid Vector (WGV). Then a feature vector is formed based on the WGV, Area and Mean height of the target hand-signs. The system classifies the hand-signs by calculating the maximum similarity between test and pre-trained feature vectors using the proposed simple mathematical analysis based classifier. The system is tested in six different challenging environments for 46 hand-signs in Bengali sign Language (BdSL) and 10 numeral hand-signs in Chinese sign language (ChSL), achieving mean accuracy of 95.67% for BdSL and 96.57% for ChSL with the computational cost of 8.01 milliseconds per frame.

Keywords Hand-sign, Haar-classifiers, Fuzzy Rule Based RGB (FRB-RGB) Model, Window-Grid Vector (WGV), Bengali Sign Language (BdSL), Chinese Sign Language (ChSL).

1 Introduction

Computer vision based researches of automatic hand-signs classification using simple camera based image, rather than expensive devices like Microsoft Kinect, different types of gloves or sensors, are increasing rapidly to achieve naturality, simplicity, availability and accessibility in the present world. Hand-signs recognition researches were used previously in virtual reality [1], human–computer interface (HCI)

Received March 07, 2017; accepted September 22, 2017

E-mail: aminur.wg@gmail.com

technologies [2], remote controller [3], and touch screen [4, 5]. Recently hand-signs classification system is used to assist storytelling for children [6]. Hand-signs classification has attracted extensive research within computer vision communities for sign language recognition [7–21]. Hand-signs are the basic unit of the sign languages. So, successful sign language recognition depends on the recognition or classification of hand-signs successfully.

Bengali language is one of the 10 most spoken languages where 233 million people speak in Bengali worldwide among which 215 million are native speakers [22]. About 16 million sign people are living in Bangladesh which include speaking and hearing impaired [18]. Not a large number of researches have been carried out in Bengali Sign Language Recognition (BdSLR). So hand-sign segmentation and classification system (HSSCS) in BdSL is an important issue for speech or hearing impaired people to communicate with other people and themselves. BdSL is structurally different from sign languages of other countries [23]. In 1994, Bangladesh Sign Language Committee [24], introduced a Bengali sign language dictionary which uses 36 (6 vowels and 30 consonants) two-handed Bengali sign alphabets from 51 Bengali written alphabets according to pronunciation. The system uses these 6 vowels (অ(a), আ(aa), ই(i), উ(u), এ(e), ও(o)) and 30 consonants (ক(ka), খ(kha), গ(ga), ঘ(gha), চ(cha), ছ(chha), জ(ja), ঝ(jha), ট(toa), ঠ(thhoa), ড(doa), ঢ(dhho-a), ত(ta), থ(tha), দ(da), ধ(dha),ণ(na), প(pa), ফ(pha), ব(ba), ভ(bha), ম(ma), য়(wa), র(ra), ল(la), স(sa), হ(ha), ড়(rha), ং(ng), ঃ(hh)) of BdSL alphabet signs and 10 BdSL numeral signs (০(zero), ১(one), ২(two), ৩(three), ৪(four), ৫(five), ৬(six), ৭(seven), ৮(eight), ৯(nine)) as sample dataset for training and testing process. We have also selected the 10 hand-signs of ChSL numerals (0, 1, 2, 3, 4, 5, 6, 7, 8 and 9) for classification by the proposed system to prove that the proposed system will be able to recognize any hand-sign of any sign language.

Many researchers have attempted to develop sign language recognition system in different countries. Kulkarni et al. [7] developed an American Sign Language (ASL) recognition system using appearance-based method. British Sign Language (BSL) recognition system was developed by Ong et al. [8] using multi-class classifier based on Sequential Pattern Trees. Naoum et al. [9] developed an Arabic sign language (ArSL) recognition system using KNN classifier. German Sign Language (GSL) recognition system was developed by Bauer et al. [10] based on Hidden Markov Models (HMM). Japanese Sign Language (JSL) recognition system was developed by Sako et al. [11] and Australian Sign Language (AuSL) recognition system was developed by Holden et al. [12] using HMM. Zhang et al. [13] proposed a novel approach for multi-modality ASL recognition. In Bangladesh, Deb et al. [25] presented a two-handed BdSL alphabet recognition system using template matching with a success rate of 96% accuracy. They used two different color wrist bands for hand segmentation. Karmokar et al. [26] used Neural Network Ensemble (NNE) in their system to recognize BdSL involving different numbers of individual NNs in negative correlation learning (NCL) with the recognition accuracy of 93%. They have used only one-handed static BdSL and it is not a real-time system. Jasim et al. [14] developed a real-time system for recognizing static and dynamic hand gestures. KNN algorithm is used to recognize the static hand-sign and Principle Directional Features (PDF) is used to recognize the dynamic hand gestures in the system. The system was tested by the ChSL numeral dataset and BdSL numeral dataset with good accuracy. More recently, notable researches on BdSLR [18–21] are done for limited sign. But there are no complete solution for HSSCS in BdSL. Our previously developed system [16] used contour vector of hand-sign instead of segmenting skin-color to recognize hand-signs. But that system could not recognize among the hand-signs where outer contour are similar but inner shapes are different. In this paper, we have proposed a rotation, translation and scale invariant feature vector denoted by Window-Grid Vector (WGV) which includes not only outer contour but also the inner shape to overcome this limitation. Another limitation was that the system sometimes failed to recognize the hand-signs, if it was occluded by other 3D objects in cluttered background. To overcome this limitation, we have proposed FRB-RGB model for skin-color segmentation.

Skin-color is an efficient component for hand-sign segmentation. Many researches [27–29] have been done on skin-color segmentation in an optimal color space such as RGB, HSV, YCbCr and YUB which are basically traditional skin-color segmentation methods. This paper presents a fuzzy rule based non-parametric

RGB skin-color model (denoted by 'FRB-RGB') for skin-color segmentation which minimizes the computational cost for hand-signs segmentation in cluttered background with illumination variation. The RGB system was used to represent skin-colors [30–32], where a fuzzy model was used to recognize human faces based on skin-color. The proposed method classifies each skin-color pixel according to a set of fuzzy rules from the ROI containing hand-signs using RGB color spaces. Combination of Haar-classifiers [33, 34] to detect hand-signs from the captured images which was described in details in the previous system [14, 15] with the FRB-RGB model, enables the system to segment skin-color robustly only from the ROI instead of whole image and which reduces the computational cost significantly.

Proper selection of classifier is another challenge in hand-sign classification. Many researchers have used traditional classifier such as KNN [9, 14, 15, 35], SVM-KNN [36], HMM [10, 12, 37, 38], ANN [39] in their systems. We have proposed simple mathematical analysis based classifier instead of traditional classifiers that classifies the hand-signs accurately in real-time using WGV analysis.

The challenges of hand-signs recognition or classification include background and illumination variation, feature extraction for training and recognition, environmental noise reduction, signer dependency, high computational cost and proper selection of recognition algorithm [15]. To overcome these challenges, this paper proposes a computer vision-based signer independent, natural and real-time hand-signs segmentation and classification system (HSSCS) using fuzzy rule based RGB (FRB-RGB) model and WGV analysis. The main objective of the proposed system is the minimization of the computational cost with keeping the higher success rate for recognizing the hand-signs of any sign language in cluttered and dynamic background with illumination variation environments.

The contributions of the proposed system are in three areas- (1). Hand-signs detection and segmentation using a robust Fuzzy rule based RGB (FRB-RGB) model combining with Haar-classifiers and motion queue in cluttered and dynamic background environment with illumination variation; (2). Rotation, translation and scale invariant feature vector (WGV) is proposed instead of row vector used in previous system [15], in which WGV includes not only outer contour of previous system [16] but also inner shape of hand-signs. This feature makes the system simpler and faster than the previous system [15, 16] and signer independent; (3). We have proposed simple mathematical analysis based classifier instead of traditional classifiers that classifies the hand-signs accurately in real-time.

The rest of the paper is organized as follows. Section 2 describes the proposed system. Section 3 presents the experimental result with discussion. Finally, the paper is concluded in Section 4.

2 Proposed System Description

Fig.1 presents a block diagram of the proposed system (HSSCS). The proposed system uses a USB or CCD camera to capture images with resolution of 640× 480. The system captures image sequence denoted by $I^m_{rgb}(x,y)$ by using the camera, where, m represents the sequence number.

2.1 Hand-sign detection and ROI generation

After capturing the image sequence $I^m_{rgb}(x,y)$, the proposed system uses Haar-like feature-based cascaded classifier (Haar-classifiers) [33, 34] to detect predefined hand-signs from the captured images which was described in details in the previous system [14, 15]. After detecting the predefined hand-sign $I^m_{HS}(x,y) \in I^m_{rgb}(x,y)$, the bounded rectangular box is drawn around the hand-sign $I^m_{HS}(x,y)$ using the Equation 1 which becomes the region of interest (ROI) denoted by $I^m_{ROI}(x,y)$.

$$I^m_{ROI}(x,y) = Rectangle\,(I^m_{HS}(x,y), W, H) \tag{1}$$

Where, $W = (max[r] - min[r]) + 0.5$ and $H = (max[c] - min[c]) + 0.5$; r and c represent row and column of $I^m_{HS}(x,y)$.

As we use the skin-color and motion cues in combination to segment the binary hand-sign from the ROI, this system can be conflicted between the hand-signs and moving face when occlusions occur in the ROI. To solve this problem Xu et al. [40] proposed a system which combines Camshift tracking algorithm, force field and edge detection to estimate hand area during the occlusion but their algorithm is complex

Fig. 1 Block diagram of the proposed system.

and computationally costly. In the proposed system, to overcome this problem, the system detects human face area by using Viola and Jones method [33, 34] from the captured image $I^m_{rgb}(x,y)$ and then the face area is masked with non-skin-color before generating the ROI.

Fig.18(a)(i), Fig.18(b)(i) and Fig.18(c)(i) show the successful results of hand-signs detection in cluttered background, plain background, and cluttered background with illumination variation respectively.

2.2 Fuzzy rule based RGB (FRB-RGB) model for skin-color segmentation

After ROI generation by detecting the hand-signs, skin-color segmentation method starts to segment skin-like area from the ROI using the Algorithm1. This paper presents a fuzzy rule based non-parametric RGB skin-color segmentation method (denoted by 'FRB-RGB') as shown in Fig.2 to segment skin-color in cluttered background with illumination variation. The skin-color segmentation method using fuzzy rule based model is a pixel-based segmentation method [41]. Combination of Haar-classifiers [33, 34] to detect hand-signs from the captured images with the proposed FRB-RGB model, enables the system to segment skin-color robustly only from the ROI instead of whole image. As a result the FRB-RGB model performs faster by reducing the computational cost significantly. The proposed method segments each skin-color pixel according to a set of fuzzy rules from the ROI containing hand-signs using RGB color spaces C_{rgb}, where each color pixel is represented by a color vector $P^k_C = \{P^R_C, P^G_C, P^B_C\} \in [0, 255]$. Given the RGB image $I^m_{ROI}(x,y)$ of size $N \times M = n$ pixels. Hence, $P^{(k,l)}_C = \{P^{(k,1)}_C, P^{(k,2)}_C, P^{(k,3)}_C, ..., P^{(k,n)}_C\} \in I^m_{ROI}(x,y)$.

The histogram will be $H[C_{rgb}] = \{h1, h2, h3, ..., hl\}$ of C_{rgb} as $b \times l$ array, where each hl represents the frequency of the color component $P^{(k,l)}_C$, for $l = 1, 2, ..., n$ with b bins. A fuzzy set is represented by (C_{rgb}, H) where, $H : C_{rgb} \rightarrow [0, 1]$ [42] in which the histogram is normalized by n. For each $P^k_C \in C_{rgb}$, $P^k_C \in (C_{rgb}, H) \Leftrightarrow P^k_C \in C_{rgb}$ and $H[P^k_C] \neq 0$. The outputs of the $H[C_{rgb}]$ and membership functions for the skin-color in the RGB color plane that are modeled using a bell-shaped function [43, 44] are shown in Fig.3.

RGB values are easily affected by illumination and cluttered background. The proposed method can handle the fuzziness due to illumination variation by combining a set of fuzzy rules. In the proposed system, a variation of the FRB-RGB model is considered for identifying the fuzziness in the transition region between the skin-color and the non-skin-color [31, 45]. Shannon entropy [46] is an effective method which was

Fig. 2 Hand-sign detection and segmentation method.

used in image segmentation [47–49]. We have used fuzzy 3-partition entropy criterion [31, 50, 51] for the RGB color plane to measure the fitness of the fuzzy membership functions for skin-color and non-skin-colors. The values of the transition points $\{a_S^k, b_S^k, \gamma_S^k, a_Z^k, b_Z^k, \gamma_Z^k\}$ will be determined by maximizing the values of fuzzy entropy $FE(a_S^k, b_S^k, \gamma_S^k, a_Z^k, b_Z^k, \gamma_Z^k)$ and fuzzy model is generated using Algorithm1.

By experimenting with the training dataset of 100 BdSL hand-signs from 10 different skin-color people, we found that values of $a_S^k = \gamma_Z^k = \underset{(P_C^k)}{\operatorname{argmax}}\left[H[P_C^k]\right]$; $a_Z^k = 0.5\gamma_Z^k$; $\gamma_S^k = 1.5a_S^k$; and $b_S^k = b_Z^k = 0.5$. Fig.14 shows the example training ground truth dataset from a single signer. The output fuzzy model of the Algoritm 1 for the proposed 'FRB-RGB' method for skin-color and non-skin-color area is shown in Fig.4.

Fig.2 shows the output of hand-sign detection and segmentation method where the output of skin-color

Algorithm 1 Skin-color segmentation using FRB-RGB model.

Input: $I_{ROI}^m(x,y)$ of size $M \times N = n$

Output: $I_{skin}^m(x,y)$

1: **Initialization:** $H[P_C^k] = [0]$; $\beta^k = 0$; $\alpha^k = 0$; Membership function for skin-color$=\mu_{skin}^k(P_C^k) = 0$; Membership function for non-skin-color of S partition$=\mu_{non-skin}^{\langle k|S\rangle}(P_C^k) = 0$; Membership function for non-skin-color of Z partition$=\mu_{non-skin}^{\langle k|Z\rangle}(P_C^k) = 0$; $\Psi_{skin}^k(P_C^k) = \Psi_{non-skin}^{\langle k|S\rangle}(P_C^k) = \Psi_{non-skin}^{\langle k|Z\rangle}(P_C^k) = 0$

2: **for** $k = R, G, B$ **do**

3: Compute $H[P_C^k] \in [0, 255]$ from $I_{ROI}^m(x,y)$; $\beta^k = max\left[H[P_C^k]\right]$; $\alpha^k = \underset{(P_C^k)}{\operatorname{argmax}}\left[H[P_C^k]\right]$

4: $\mu_{skin}^k(P_C^k) = \beta^k e^{-\frac{(P_C^k - a^k)^2}{2(\sigma^k)^2}}$ \\$(\sigma^k)^2$ is the variance of each P_C^k

5: **if** $P_C^k \leqslant a_S^k$ **then**

6: $\mu_{non-skin}^{\langle k|S\rangle}(P_C^k) = 0$

7: **else if** $a_S^k \leqslant P_C^k \leqslant b_S^k$ **then**

8: $\mu_{non-skin}^{\langle k|S\rangle}(P_C^k) = 2\left(\frac{P_C^k - a_S^k}{\gamma_S^k - a_S^k}\right)^2$

9: **else if** $b_S^k \leqslant P_C^k \leqslant \gamma_S^k$ **then**

10: $\mu_{non-skin}^{\langle k|S\rangle}(P_C^k) = 1 - 2\left(\frac{P_C^k - a_S^k}{\gamma_S^k - a_S^k}\right)^2$

11: **else if** $\gamma_S^k \leqslant P_C^k$ **then**

12: $\mu_{non-skin}^{\langle k|S\rangle}(P_C^k) = 1$

13: **end if**

14: **if** $P_C^k \leqslant a_Z^k$ **then**

15: $\mu_{non-skin}^{\langle k|Z\rangle}(P_C^k) = 1$

16: **else if** $a_Z^k \leqslant P_C^k \leqslant b_Z^k$ **then**

17: $\mu_{non-skin}^{\langle k|Z\rangle}(P_C^k) = 1 - 2\left(\frac{P_C^k - a_Z^k}{\gamma_S^k - a_Z^k}\right)^2$

18: **else if** $b_Z^k \leqslant P_C^k \leqslant \gamma_Z^k$ **then**

19: $\mu_{non-skin}^{\langle k|Z\rangle}(P_C^k) = 2\left(\frac{P_C^k - a_Z^k}{\gamma_Z^k - a_Z^k}\right)^2$

20: **else if** $\gamma_Z^k \leqslant P_C^k$ **then**

21: $\mu_{non-skin}^{\langle k|Z\rangle}(P_C^k) = 0$

22: **end if**

23: $\Psi_{skin}^k(P_C^k) = \sum_{P_C^k=0}^{255} H[P_C^k]\mu_{skin}^k(P_C^k)$ \\Probability of skin-color

24: $\Psi_{non-skin}^{\langle k|S\rangle}(P_C^k) = \sum_{P_C^k=0}^{255} H[P_C^k]\mu_{skin}^{\langle k|S\rangle}(P_C^k)$ \\Probability of non-skin for S partition

25: $\Psi_{non-skin}^{\langle k|Z\rangle}(P_C^k) = \sum_{P_C^k=0}^{255} H[P_C^k]\mu_{skin}^{\langle k|Z\rangle}(P_C^k)$ \\Probability of non-skin for Z partition

26: $FE(a_S^k, b_S^k, \gamma_S^k, a_Z^k, b_Z^k, \gamma_Z^k) = -\Psi_{skin}^k(P_C^k)\log\left(\Psi_{skin}^k(P_C^k)\right) -\Psi_{non-skin}^{\langle k|S\rangle}(P_C^k)\log\left(\Psi_{non-skin}^{\langle k|S\rangle}(P_C^k)\right) -\Psi_{non-skin}^{\langle k|Z\rangle}(P_C^k)\log\left(\Psi_{non-skin}^{\langle k|Z\rangle}(P_C^k)\right)$

27: Compute $I_{skin}^m(x,y)$ using Eq.2 form the Fuzzy model (Fig.4) based on $FE(a_S^k, b_S^k, \gamma_S^k, a_Z^k, b_Z^k, \gamma_Z^k)$

28: **end for**

Fig. 3 Fuzzy system for Gaussian approximation of $H[C_{rgb}]$ for RGB color space.

Fig. 4 Modeling of the skin-color and non-skin-color area using Algorithm 1 of proposed 'FRB-RGB' method.

segmentation, motion detection, binarization and noise removal are presented sequentially.

After modeling the fuzzy inference system for skin-color and non-skin-color area, the system segments the skin-color region denoted by $I^m_{skin}(x,y)$ from the fuzzy model using Eq.(2). Skin-color detection is a fuzzy problem in which there is an optimal threshold for each method and it is not constant for different images [52]. In our experiments, we use a value $\Psi^k_{skin}(P^k_C) \geqslant T_{skin}$ for skin-color detection, where T_{skin} is a selected threshold value and $\Psi^k_{skin}(P^k_C)$ is the probability of observing skin-color. T_{skin} is normalized to [0,1] for the sake of comparability to the other methods as shown in Fig.17 which is described in details in section 4. Fig.18 shows sample results of skin-color segmentation in three challenging environment using the proposed method comparing with traditional skin-color segmentation method [15].

$$I^m_{skin}(x,y) = \begin{cases} P^i_{skin} & \text{if } ((P^i_{skin}) \in \text{SCA}) \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Where, P^i_{skin} is the ith skin-color pixel and SCA represents the skin-color area of the fuzzy model (Fig.4).

The proposed skin-color segmentation method is also applied using HSV and YCbCr color model using the selected off-line samples of hand-signs, in which the proposed model gives accurate results. If we apply this segmentation methods after detecting hand-signs using Haar-classifiers then the accuracies are increased significantly with faster segmentation rate. The comparison results of the different skin-color segmentation methods with computational times are shown in Table 2 in section 4.

2.3 Probable Binary hand-signs extraction

The system extracts probable binary hand-signs based on the skin-color pixels ($P^i_{skin} \in I^m_{skin}(x,y)$) with specific motions using Eq.(3). Motion of each pixel is used to avoid the occlusion of hand-sign and skin-like backgrounds.The skin-color detector and motion detector work in parallel in the system.

$$I^m_{bin}(x,y) = \begin{cases} 0 & \text{if } ((P^i_{skin} \in I^m_{skin}(x,y) \neq \phi \text{ and} \\ & (T_H \geqslant \overrightarrow{v_i} \geqslant T_L)) \\ 1 & \text{otherwise} \end{cases} \tag{3}$$

Where, $\overrightarrow{v_i}$ is the velocity of ith skin-color pixels P^i_{skin} which is calculated by using Eq.(4). Here, we have used two user defined thresholds denoted by T_H (Higher Threshold) and T_L (Lower Threshold) for motion detection.

$$\overrightarrow{v_i} = \frac{\overrightarrow{\delta}}{\Delta t} \tag{4}$$

Where $\overrightarrow{\delta}$ is the Euclidean distance between the skin-color pixel's position $(r_P, c_Q)_t$ and $(r'_P, c'_Q)_{t+1}$ which is measured by using Eq.(5) and Δt is the time to move that skin-color pixel's position $(r_P, c_Q)_t$ at t time to $(r'_P, c'_Q)_{t+1}$ at $(t = t + 1)$ times. Hence, $\Delta t = 1$.

$$\overrightarrow{\delta} = \sqrt{(r_P - r'_P)^2 + (c_Q - c'_Q)^2} \tag{5}$$

To remove noise, morphological closing (dilation followed by erosion operation) and Gaussian smoothing operations are done on $I^m_{bin}(x, y)$ using Eq.(6) and Eq.(7) [53] and obtains $I^m_{binclosing}(x, y)$ and $I^m_{smooth}(x, y)$ respectively as shown in Fig.2(h).

$$\begin{aligned} I^m_{binclosing}(x, y) =& I^m_{bin}(x, y) \cdot SE \\ =& (I^m_{bin}(x, y) \oplus SE) \ominus SE \end{aligned} \tag{6}$$

$$I^m_{smooth}(x, y) = I^m_{binclosing}(x, y) \cdot G(x, y) \tag{7}$$

Where, SE represents Structuring Element and $G(x, y)$ is Gaussian kernel defined by Eq.(8).

$$G(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{8}$$

Where, x and y are the horizontal and vertical distances from the origin of $I^m_{binclosing}(x, y)$ and σ^2 is the variance of the Gaussian distribution.

Fig.18 shows sample results of skin-color segmentation and binary hand-signs extraction in three challenging environment using the proposed method and comparing with traditional skin-color segmentation method [15].

2.4 Feature Vector Generation

2.4.1 Window-Grid Vector (Ω) Generation

Before generating the Grid-Pattern represented by Window-Grid Vector (WGV), the system makes the binary smoothed image $I^m_{smooth}(x, y)$ rotation and scale invariant. To make the rotation invariant, the system rotates the hand-sign centered on center of gravity (COG) with respect to align the point of wrist (POW) to the vertical line which is drawn through the COG (to ensure $\theta = 0$) as shown in Fig.5. COG denoted by $P_{COG}(r_P, c_Q)$ is calculated using the Eq.(9). To calculate the POW, the system finds the hand-wrist by searching more than 15 consecutive zero pixels (black pixels) at the bottom, right and left of the border row [21]. After finding the hand-wrist, the system calculates the POW denoted by $P_{POW}(r_P, c_Q)$ using the Eq.(10) and draws a line from the POW to COG of the hand-sign. Then the system gets the angle θ in the COG with respect to vertical line and the line joining between COG and POW as shown in Fig.5(a)(i) and Fig.5(b)(i). For two-handed hand-signs, POW is calculated by determining the middle point of the line joining the two POWs (POW-1 for left hand and POW-2 for right hand as shown in Fig.5(b)(i)). If the value of $\theta = 0$ then no rotation was performed; otherwise images are rotated until $\theta \neq 0$. If the system allow $\theta > 45^\circ$ then the recognition accuracy is reduced significantly as shown in Fig.6 which is drawn based on different angles versus recognition accuracies using selected off-line samples of hand-signs by updating the angle θ by 5° from $\theta = 0$ to $\theta = 90^\circ$. After rotation, the system finds horizontal cut point (HCP), $P_{HCP}(r_P, c_Q)$ according to the Fig.5(a)(iii) and Fig.5(b)(iii) based on lowest value of $[r_P] \in [P_{HWi}(r_P, c_Q)]$ using Eq.(11).

$$P_{COG}(r_P, c_Q) = \left\{ \frac{\sum_{P=1}^{n} r_P}{\Lambda}, \frac{\sum_{Q=1}^{n} c_Q}{\Lambda} \right\} \tag{9}$$

Where, Λ is the area of hand-signs which is the total number of black pixel of the smoothed image $I^m_{smooth}(x, y)$; r_P and c_Q represent row and column value of each boundary coordinate of that image.

$$P_{POW}(r_P, c_Q) = \begin{cases} (\frac{r_P + r'_P}{2}, \frac{c_Q + c'_Q}{2}) \in (P_{HW1}, P_{HW2}) \\ (\frac{r_P + r'_P}{2}, \frac{c_Q + c'_Q}{2}) \in (P_{POW1}, P_{POW2}) \end{cases} \tag{10}$$

Where, the hand wrist point for one-handed signs is $(P_{HW1}, P_{HW2}) = (P_{HW1}(r_P, c_Q), P_{HW2}(r'_P, c'_Q))$ and point of wrist for two-handed signs is $(P_{POW1}, P_{POW2}) = (P_{POW1}(r_P, c_Q), P_{POW2}(r'_P, c'_Q))$.

$$P_{HCP}(r_P, c_Q) = P_{HWi}(r_P, c_Q) \textbf{ if } \min[r_P] \in [P_{HWi}] \tag{11}$$

Where, P_{HWi} is the point of hand wrist denoted by $P_{HWi}(r_P, c_Q)$ is the end points of the hand-wrist where $i = 1, 2$ for one-handed signs and $i = 1, 2, 3, 4$ for two-handed signs; r_P and c_Q represent P^{th} row and Q^{th} column position. After getting HCP denoted by $P_{HCP}(r_P, c_Q)$, the system adjusts the ROI by aligning with the horizontal line with respect to HCP and deletes the rows containing value of

$r_P > \min[r_P] \in [P_{HWi}(r_P, c_Q)]$; in which this rotated image is denoted by $I^m_{rotate}(x, y)$. To make it scale invariant, the system applies clipping and normalization operation on the rotated binary image $I^m_{rotate}(x, y)$.

Fig. 5 Example output of rotation (a) for one-handed sign (left side), (b) for two-handed sign (right side); in both side, (a) and (b) each figure represents as (i) center of gravity (COG) and point of wrist (POW) calculation, (ii) after rotation with respect to align $\theta = 0$, (iii) align the ROI based on horizontal cut point, (iv) after ROI adjustment and alignment.

Clipping process is done by using four lines approaching [9] on $I^m_{rotate}(x, y)$. In this process, rows and columns containing no black pixel ('0') are deleted. The clipping image is denoted by $I^m_{clipping}(x, y)$. Then the system normalizes the variable sizes of clipping binary images $I^m_{clipping}(x, y)$ at predefined size of (150×150). The normalized image is denoted by $I^m_{norm}(x, y)$. Example result of the clipping and normalization is shown in Fig.7.

Fig. 6 The graphs of different values of angles versus accuracies for hand-signs rotation.

Fig. 7 Example result of the clipping and normalization.

We plot the computational costs (CC) versus accuracy graphs (as shown in Fig.8(a)) to fine-tune the values of the normalizing size of the clipping binary images using the selected off-line samples of hand-signs. From the observations of the graphs as shown in Fig.8(a), we decide that the normalizing size of clipping binary image is set to 150×150 for achieving the high accuracy with reduced CC.

After clipping and normalization, the system generates the Window-Grid Vector (WGV) denoted by Ω instead of extracting row vector mentioned in [15] from each normalized binary hand-sign $I^m_{norm}(x, y)$. The previous system [15] needs to process every pixels of

Fig. 8 The graphs of computational costs versus accuracy for (a) Different normalizing sizes and (b) Different sizes of the Window-Grid Mask (WGM).

the hand-signs which made the system computationally high costly. The proposed system overcomes these limitations by applying a 5×5 'Window-Grid Mask (WGM)' on each normalized binary hand-sign $I^m_{norm}(x, y)$, where each Window-Grid (ω^η) size of WGM is 30×30 as shown in Fig.9(a) which reduces the CC. Note that we decide that the size of WGM is set to 5×5 for achieving the high accuracy with reduced CC from the observation graph as shown in Fig.8(b). We plot the results of computational costs versus accuracy in a graphs (as shown in Fig.8(b)) to fine-tune the values of the WGM size using the selected off-line samples of hand-signs.

Fig. 9 Window-Grid Vector (WGV) generation process (a) Example 5×5 Window-Grid Mask (WGM), (b) 5×5 WGM is applied on the 150×150 normalized binary hand-sign, and (c) Example values of each Window-Grid (ω^η).

The system extracts the sequences of Window-Grids defined as WGV (Ω) which is represented by the Eq.(12) for each hand-sign $I^m_{norm}(x, y)$. Where, each Window-Grid value (ω^η) is obtained by calculating the percentage of white pixels ('1') presented in each Window-Grid (ω^η) of the WGM for each hand-sign. Fig.9 shows the WGV generation process by applying the WGM.

$$\Omega = (\omega^0, \omega^1, \omega^2, ..., \omega^{M_\omega - 1}) \tag{12}$$

Where, Ω represents the WGV for each hand-sign; ω^η represents each Window-Grid of Ω; $\eta = 0, 1, 2, ..., M_\omega - 1$; where, M_ω= (Size of Input Image)/(Size of Each Window); In this case, Size of Input Image is 150×150 and Size of Each Window is 30×30; Hence, $M_\omega = (150 \times 150)/(30 \times 30) = 25$; and each ω^η is calculated by Eq.(13). Hence, the size of each WGV (Ω) for a single hand-sign is $(5 \times 5) = 25$.

$$\omega^\eta = \frac{NW}{100} \tag{13}$$

Where, NW=Number of white pixels (1) in each ω^η.

In the graph as shown in Fig.12, Case-1 represents the performance using the value of recognition or classification coefficients ($\delta_1 = 1, \delta_2 = 0$, and $\delta_3 = 0$) by using the Eq.(16); that means, the system classifies the hand-sign in this case using only the proposed Grid-Pattern represented by WGV (Ω) and achieves the accuracy of 95.03%. To increase this performance, we use the additional two features such as Area (Λ) and Mean height (λ) in the proposed system according to the Case-8 which achieves the highest success rate of 96.54% accuracy for selected offline sample of hand-signs.

2.4.2 Area (Λ) of a hand-sign

The Area (Λ) of a hand-sign is the total number of black pixels in the normalized binary image $I^m_{norm}(x, y)$.

The area is calculated by scanning each pixel ($r_P c_Q$) according to Fig.10, starting from the left most column of top row ($r_0 c_0$) to right most column of bottom row ($r_P c_Q$) by using unidirectional pixel scanning method [54].

(a) Row by row pixel scanning method. (b) Example representation of a hand sign by r and c position.

Fig. 10 Hand area calculation process by pixel scanning method.

2.4.3 Mean Height (λ) of a hand-sign

The Mean Height of the hand-sign is calculated by using Eq.(14).

$$\lambda = \frac{\Lambda}{N} \tag{14}$$

Where, N = Number of column and Λ is the area of the hand-sign.

Finally, the features are collected as Eq.15 in the feature space for each hand-sign. The size of the total features for each hand-sign is $\xi = [\Omega, \Lambda, \lambda] = [(5 \times 5) + 1 + 1] = 27$ where as the previous system [15] was considered the size of the row vector as the feature as $(150 \times 150) = 22,500$. Hence, the size of the extracted features for each hand-sign in the proposed system using WGV is more than 833 times smaller than the previous system [15] which reduces the computational cost of the proposed system.

$$\xi = [\Omega, \Lambda, \lambda] \tag{15}$$

2.5 Training Method

A training module is generated by storing the extracted features of each hand-sign in each sign class labeled with the training dataset. For each sign class i, the system calculates j=100 features in each feature type of each sign X_i such as $[\Omega_i^0, \Omega_i^1, \Omega_i^2, ..., \Omega_i^{99}]$ for each WGV (Ω_i^j), $[\Lambda_i^0, \Lambda_i^1, \Lambda_i^2, ..., \Lambda_i^{99}]$ for each Area (Λ_i^j) feature and $[\lambda_i^0, \lambda_i^1, \lambda_i^2, \ldots, \lambda_i^{99}]$ for each mean height (λ_i^j) feature; where $i = 0, 1, 2, ..., n$ represents sign classes. The system stores these three types of extracted features $\Omega_i^j, \Lambda_i^j, \lambda_i^j$ in to the three sub-classes of each sign class i and creates the training database as shown in Fig.11. The resulted size of the feature vector (ξ_i^j) for 46 hand-signs in BdSL is $[(25 + 1 + 1) \times 100 \times 46] = 1,24,200$.

Finally, the system stores each extracted feature into the corresponding sign class for each sign, X_i. Two separate feature space are used for BdSL and ChSL. The input hand-signs of 46 alphabet and numeral signs [অ (a), আ (aa), ই (i), উ (u), এ (e), ও (o), ক (ka), খ (kha), গ (ga), ঘ (gha), চ (cha), ছ (chha), জ (ja), ঝ (jha), ট (toa), ঠ (thhoa), ড (doa), ঢ (dhhoa), ত (ta), থ (tha), দ (da), ধ (dha),ণ (na), প (pa), ফ (pha), ব (ba), ভ (bha), ম (ma), য় (wa), র (ra), ল (la), স (sa), হ (ha), ড় (rha), ং (ng), ঃ (hh), ০ (zero), ১ (one), ২ (two), ৩ (three), ৪ (four), ৫ (five), ৬ (six), ৭ (seven), ৮ (eight), ৯ (nine)] in BdSL are to be assigned to the class labels of [0, 1, 2, ..., 45] and the input hand-signs [0, 1, 2, 3, 4, 5, 6, 7, 8, 9] in ChSL are to be assigned to the class labels of [0,1,2, ..., 9] in training databases. The example of training dataset of BdSL and ChSL are presented in Fig.15 and Fig.16 respectively.

Fig. 11 Example structure of training database for i = 0, 1, 2, ..., n sign classes, where each sign class consists of j =0, 1, 2, ..., 99 features for each sub-class of WGV (Ω_i^j), Area (Λ_i^j), and Mean height (λ_i^j).

2.6 Recognition Method

In the Recognition or classification module, the BdSL hand-signs (alphabets and numerals) and ChSL numeral signs are recognized by measuring the maximum similarity between the extracted features $\xi = [\Omega, \Lambda, \lambda]$ and the pre-trained features $\xi_i^j = [\Omega_i^j, \Lambda_i^j, \lambda_i^j]$ of the hand-signs using Eq.(16).

$$Sim_{Max}(\xi, \xi_i^j) = \delta_1 \times Sim_{Max}(\Omega, \Omega_i^j) + \delta_2 \times Sim_{Max}(\Lambda, \Lambda_i^j) + \delta_3 \times Sim_{Max}, (\lambda, \lambda_i^j) \quad (16)$$

Where, δ_1, δ_2, and δ_3 are the recognition coefficient with the condition, $\delta_1 + \delta_2 + \delta_3 = 1$; $Sim_{Max}(\xi, \xi_i^j)$ is the collected maximum similarity between the collected feature of the test hand-sign (ξ) and pre trained hand-sign (ξ_i^j) with the value among 0 to 1; $Sim_{Max}(\Omega, \Omega_i^j)$ is the maximum similarity between test WGV (Ω) and pre-trained WGV (Ω_i^j) of the hand-signs which is calculated using Eq.(17); $Sim_{Max}(\Lambda, \Lambda_i^j)$ is the maximum similarity between test Area (Λ) and pre-trained Area (Λ_i^j) of the hand-signs which is calculated by using Eq.(18); and $Sim_{Max}(\lambda, \lambda_i^j)$ is the maximum similarity between test Mean height (λ) and pre-trained Mean height (λ_i^j) of the hand-signs which is also calculated by using Eq.(19).

$$Sim_{Max}(\Omega, \Omega_i^j) = Max\left[\frac{|ICC(\Omega_i^j, \Omega^{(N)})|}{|\Omega_i^j||\Omega|}\right] \quad (17)$$

$$Sim_{Max}(\Lambda, \Lambda_i^j) = Max\left[1 - \frac{|\Lambda - \Lambda_i^j|}{|\Lambda| + |\Lambda_i^j|}\right] \quad (18)$$

$$Sim_{Max}(\lambda, \lambda_i^j) = Max\left[1 - \frac{|\lambda - \lambda_i^j|}{|\lambda| + |\lambda_i^j|}\right] \quad (19)$$

Where, $|ICC(\Omega_i^j, \Omega^{(N)})|$ represents the Inter Correlation Coefficient between the trained WGV (Ω_i^j) and the test WGV (Ω); $\Omega^{(N)}$ represents a Window-Grid (ω^η) received from test WGV (Ω) by cycle shift by its Window-Grid ω^η on 'N' of elements; $|\Omega_i^j|$ and $|\Omega|$ represent the normalized value of trained WGV (Ω_i^j) and the test WGV (Ω) respectively that are calculated by Eq.(20). $|\Lambda|$ and $|\lambda|$ of the Eq.(18) and Eq.(19) represent the absolute value of Λ and λ respectively.

$$|\Omega| = \sqrt{\sum_{\eta=0}^{M_\omega - 1} |\omega^\eta|^2} \quad (20)$$

When the extracted feature vector ξ is matched with the pre-stored trained feature vectors ξ_i^j of the hand-sign based on measuring the maximum similarity using Eq.(16), the system returns a class label containing the specific hand-signs which represent the recognized hand-sign as text.

In the classification stage, the tuning of the recognition coefficients (δ_1, δ_2,and δ_3) is very important. We have plotted the classification results of accuracies for different values of recognition coefficients (δ_1, δ_2,and δ_3) in a graphs (as shown in Fig.12) to fine-tune the recognition coefficients (δ_1, δ_2,and δ_3) using the selected off-line samples of hand-signs. We decide that the values of recognition coefficients are set as $\delta_1 = 0.75, \delta_2 = 0.15$, and $\delta_3 = 0.1$ (Case-8) from the observation graph of Fig.12 for better performance.

3 Experimental Result and Discussion

The proposed system uses a built-in webcam of ASUS A42F series laptop for image acquisition. The system uses an ASUS A42F series laptop with Intel Corei3 processor of 2.40 GHz and 2GB RAM. The system uses Microsoft® Visual Studio® 2008 and EmguCV (C# and OpenCV wrapper) [55] in 32-bit operating system of Microsoft Windwos7®, as system development platform.

3.1 Training Images Database

Fig.13 presents the 10 different skin-colored signers who perform the hand-signs in system training where four are female and six are male.

3.1.1 Training images database for skin-color segmentation

The proposed FRB-RGB model for skin-color segmentation is trained using 100 BdSL hand-signs from the 10 different skin-color people (Fig.13), where each people performs 10 hand-signs in different environment with illumination variation. To obtain the ground truth training dataset for FRB-RGB model, we manually segment and labeled 100 images which consists of 10 subsets for 10 signers. For simplicity, only one subsets from a single signers are presented here as shown in Fig.14.

3.1.2 Training images database for hand-sign classification

The proposed system is trained for 46 hand-signs (36 alphabets and 10 numerals) in BdSL and 10 hand-signs (0 to 9 numerals) in ChSL. 10 images of each predefined hand-signs are captured from each signer in plain and non-skin-color background with proper

Case-1	$\delta_1 = 1; \delta_2 = 0;\ \delta_3 = 0$
Case-2	$\delta_1 = 0.75; \delta_2 = 0;\ \delta_3 = 0.25$
Case-3	$\delta_1 = 0.75; \delta_2 = 0.25;\ \delta_3 = 0$
Case-4	$\delta_1 = 0.9; \delta_2 = 0.05;\ \delta_3 = 0.05$
Case-5	$\delta_1 = 0.85; \delta_2 = 0.05;\ \delta_3 = 0.1$
Case-6	$\delta_1 = 0.85; \delta_2 = 0.1;\ \delta_3 = 0.05$
Case-7	$\delta_1 = 0.8; \delta_2 = 0.1;\ \delta_3 = 0.1$
Case-8	$\delta_1 = 0.75; \delta_2 = 0.15;\ \delta_3 = 0.1$
Case-9	$\delta_1 = 0.75; \delta_2 = 0.1;\ \delta_3 = 0.15$
Case-10	$\delta_1 = 0.7; \delta_2 = 0.2;\ \delta_3 = 0.1$
Case-11	$\delta_1 = 0.7; \delta_2 = 0.1;\ \delta_3 = 0.2$
Case-12	$\delta_1 = 0.6; \delta_2 = 0.2;\ \delta_3 = 0.2$
Case-13	$\delta_1 = 0.5; \delta_2 = 0.3;\ \delta_3 = 0.2$
Case-14	$\delta_1 = 0.5; \delta_2 = 0.2;\ \delta_3 = 0.3$
Case-15	$\delta_1 = 0; \delta_2 = 0.5;\ \delta_3 = 0.5$
Case-16	$\delta_1 = 0; \delta_2 = 1;\ \delta_3 = 0$
Case-17	$\delta_1 = 0; \delta_2 = 0;\ \delta_3 = 1$

Fig. 12 Recognition coefficient tuning graph for different cases.

Fig. 13 Different skin-colored 10 signers (4 females and 6 males) who took part in system training.

Fig. 14 Example ground truth training dataset for FRB-RGB skin-color segmentation model from a single signer in different illuminations.

lighting for training the system. The system asks 10 different skin-colored signers to perform those signs, where four are female and six are male. This resulted training images are $4600(10 \times 10 \times 46)$ for BdSL and $1000(10 \times 10 \times 10)$ for ChSL. Fig.15 and Fig.16 present the example training datasets of hand-signs in BdSL and ChSL respectively.

3.2 Performance Analysis

For testing the system, six sets of images have been used for 46 hand-signs in BdSL and 10 hand-signs in ChSL. The total number of images have been used by the system for testing are $(4600 \times 6) = 27600$ for BdSL and $(1000 \times 6) = 6000$ for ChSL. For each set of images, 10 new signers who did not take part in training phase, participate to perform hand-signs, where $10 \times 10 = 100$ samples for each hand-sign are prepared for testing. Test data are prepared in the following six different environments:

- Environment-1(E1): plain background with proper lighting;
- Environment-2(E2): plain background with illumination variation;
- Environment-3(E3): cluttered and static background where skin-color static objects are present;
- Environment-4(E4): cluttered background with illumination variation;
- Environment-5(E5): cluttered and dynamic background where other persons are moving behind the signer; and
- Environment-6(E6): cluttered and dynamic background with illumination variation environment where other persons are moving behind the signer.

The system uses accuracy, computational cost (CC) and receiver operating characteristic (ROC) curves for performance evaluation. Accuracy is calculated using Eq.(21) [56]. The computational cost is calculated by measuring the time to capture an image, preprocess,

(a) Example postures of BdSL vowel signs.

(b) Example postures of BdSL consonant signs.

(c) Example postures of BdSL numeral signs.

Fig. 15 Example 46 hand-signs in BdSL (a) Vowel signs (b) Consonant signs and (c) Numeral signs.

Fig. 16 Example 10 numeral hand-signs in ChSL.

extract features and match it with the training dataset in milliseconds per frame (ms/f). ROC curves are the graphs which are drawn by plotting the true positives (TP) versus false positives (FP) values that visualize the relationship between TP and FP.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \qquad (21)$$

Where, the values of TP (True positive), FP (False Positive), FN (False Negative), and TN (True Negative) are generated from separate confusion matrices for six environments (E1, E2, E3, E4, E5 and E6) of both BdSL and ChSL recognition separately. The confusion matrices are not shown here for the save of simplicity and space.

3.3 Result of hand-signs detection

Fig.18(a)(i), Fig.18(b)(i) and Fig.18(c)(i) show the results of hand-signs detection in cluttered background (E3), plain background (E1), and cluttered background with illumination variation environment (E4) successfully respectively. Table 1 presents the hand-signs detection rate in six challenging environments (E1, E2, E3, E4, E5 and E6) using Haar-classifiers for each hand-sign of BdSL and ChSL dataset. Here, correct detection rate is the mean TP rate and incorrect detection rate is the mean FP rate. The test results show good performance under various illumination and backgrounds which proves that the hand-signs detection module of the proposed system is mostly background and illumination invariant.

Table 1 Results of hand-signs detection rate.

environments	detection rate	
	incorrect (%)	correct (%)
E1	0.24	99.76
E2	1.11	98.89
E3	1.45	98.55
E4	1.75	98.25
E5	2.77	97.23
E6	3.97	96.03
mean detection rate:	1.88	98.12

3.4 Result of hand-signs segmentation using FRB-RGB model

The result of the proposed 'FRB-RGB' model for skin-color segmentation and the comparison with some other existing methods using HSV [16]; Cascaded AdaBoost [57]; SPM (skin probability map) based RGB [40]; YCbCr-K-Means [28]; Fuzzy-YCbCr based Anfis [27] are shown in Table 2. The proposed skin-color segmentation method is also applied using HSV and YCbCr color model using the selected 33600 off-line samples ($(4600 \times 6) = 27600$ images for 46 BdSL

hand-signs and $(1000 \times 6) = 6000$ images for 10 ChSL hand-signs) of hand-signs, in which the proposed model gives accurate results for skin-color segmentation with six challenging environments. For each method we compute the average computational time of three runs of 33600 images. The results of average computational time are shown in Table 2. Computational time of the YCbCr-K-Means clustering-based method [28] and Cascaded AdaBoost method [57] is significantly higher because the clustering itself is very expensive. Table 2 shows that the proposed method keeps the computational time low, because only a histogram has to be processed for segmentation. The proposed FRB-RGB skin-color segmentation model can handle the fuzziness due to illumination variation by combining a set of fuzzy rules in the RGB color model instead of using traditional RGB model by considering the fuzziness in the transition region between the skin-color and the non-skin-color. Introducing Haar-classifiers to initialize the ROI by detecting the hand-signs from the captured images before applying the proposed FRB-RGB model, enables the system to segment skin-color robustly only from the ROI instead of whole image and which reduces the computational cost and increases the accuracy significantly.

We have used ROC curves which visualize the relationship between TP and FP as shown in Fig.17 using six different datasets of 46 BdSL hand-signs in six challenging environments based on the value of $\Psi^k_{skin}(P^k_C) \geqslant T_{skin}$. Where, $\Psi^k_{skin}(P^k_C)$ is the skin-color probability and T_{skin} is the skin-color threshold. FPs (False positives) are background pixels but classified as skin and TPs (True positives) are the all pixels classified as skin-color are actual skin-color pixels. The ROC curves are generated by updating the threshold T_{skin} by 0.1. At the point (1,1) in the ROC curves, all pixels are classified as skin. Obviously, the results for the all runs for threshold T_{skin} (11 runs) are located between the two points ($T_{skin} = 0$ and $T_{skin} = 1$). Higher value of thresholds T_{skin} are closer to the point (1,1). The curve which is closer to the left y-axis and top x-axis indicates the best method.

Fig.18 shows the sample results of skin-color segmentation using proposed FRB-RGB model in three challenging environments (E3, E1 and E4 respectively) comparing with traditional methods of skin-color segmentation [15]. Here, Black points indicate the non-skin-color regions in the output of segmented skin-color image. In each figure of Fig.18, (ii) presents output of skin-color segmentation using previous method [15] based on hue and saturation value of HSV model, (iii) presents output of skin-color segmentation using the proposed 'FRB-RGB' method, (iv) presents binary hand-sign extraction using previous method [15] and (v) presents smoothed binary hand-sign $I^m_{smooth}(x, y)$ extraction using Eq.(3), Eq.(6) and Eq.(7) sequentially [our proposed method].

Fig. 17 ROC curves (TP versus FP) for different value of skin-color threshold in six different challenging environments.

3.5 Result of 46 BdSL hand-signs classification

Table 3 presents the summarized results of 46 hand-signs of BdSL alphabets and numerals recognition considering six environments (E1, E2, E3,

Table 2 Hand-signs segmentation accuracy and average computation times[act] for different methods.

skin-color segmentation methods	accuracy (%)		average computation times (milliseconds)	
	R1	R2	R1	R2
HSV [15]	86.03	77.12	148100.25	348200.87
Cascaded AdaBoost [57]	94.05	91.09	700035.69	1420050.93
SPM based RGB [40]	89.97	86.57	69027.35	145070.75
YCbCr- K-Means [28]	94.43	91.47	819040.55	1628035.75
Fuzzy-YCBCR based Anfis [27]	94.48	91.43	140040.45	290027.55
FRB-HSV (implemented by us)	95.97	93.93	81000.95	167050.88
FRB- YCbCr (implemented by us)	94.83	93.17	81070.37	168075.75
FRB-RGB (Proposed)	96.47	94.53	81090.15	168035.95

[act] *accuracy and average computation time are averaged over 3 runs of 33600 sample images using different environments;*
R1=Result-1 which represents the classification accuracies using different skin-color segmentation methods combining with Haar-classifiers;
R2=Result-2 which represents the classification accuracies using different skin-color segmentation methods without combining Haar-classifiers

Fig. 18 Sample results of hand-signs segmentation in (a) cluttered background (E3); (b) plain background (E1); and (c) cluttered background with illumination variation (E4).

E4, E5 and E6). From the test results shown in Table 3, it is evident that BdSL alphabets and numerals are recognized and distinguished properly in cluttered and dynamic background with illumination variation environment. The system achieves mean accuracy of 98.40% in E1 (plain background with

proper lighting), 97.1% in E2 (plain background with illumination variation), 96.60% in E3 (cluttered and static background where skin-color static objects are present), 96.20% in E4 (illumination variation environment with cluttered background), 95.80% in E5 (cluttered and dynamic background where other persons are moving behind the signer), 95.30% in E6 (cluttered and dynamic background with illumination variation environment where other persons are moving behind the signer) and 96.57% for overall system with the computational cost of 8.01 ms/f. The variation of the results for different environments is acceptable. For each environment, the performance of the recognition rate of hand-signs ‘র’ and ‘ল’ are decreased to less than about 92%. Because, the skin color image of hand-signs ‘র’ and ‘ল’ are distinguishable but the binary hand-signs of them are very similar to each other as shown in Fig.19. In several environments, the performances may decrease due to wrong performing of the hand-signs in front of fixed camera. In particular, performing the hand-signs ‘ও’, ‘খ’, ‘জ’ and ‘৯’ is most difficult in front of fixed camera. For each environment, the performance of one-handed numeral signs recognition in BdSL is higher than the performance of two-handed BdSL alphabet signs recognition. However, the test results are satisfactory for all environments, in which Environment-1 (E1) is the best case and Environment-6 (E6) is the worst case. Fig.20 shows the example snapshots of successfully recognition of hand-signs under various illumination and backgrounds.

(a) Distinguishable Skin-Color Images of ‘র’ and ‘ল’. (b) Indistinguishable Binary Hand Postures of ‘র’ and ‘ল’.

Fig. 19 Indistinguishing problem of binary hand-signs ‘র’ with ‘ল’.

3.6 Result of 10 numeral hand-signs of ChSL classification

Table 4 presents the summarized results of 10 hand-signs of ChSL numerals recognition considering six environments (E1, E2, E3, E4, E5 and E6). From the test results shown in Table 4, it is evident that ChSL numerals are recognized and distinguished properly. The system achieves the mean accuracy of 98.40% for E1, 97.1% for E2, 96.60% for E3, 96.20% for E4, 95.80% for E5, 95.30% for E6 and 96.57% for overall system with the computational cost of 8.01 ms/f. Although the results are varied a bit in different environments, these are acceptable. These satisfied

Table 3 Results of 46 BdSL Hand-signs classification.

BdSL	accuracy (%)						
	E1	E2	E3	E4	E5	E6	mean
অ(a)	99	98	98	97	95	93	96.67
আ(aa)	100	100	98	98	97	95	98
ই(i)	98	97	97	97	95	94	96.33
উ(u)	96	97	97	97	96	95	96.33
এ(e)	97	97	96	97	97	95	96.5
ও(o)	96	97	95	96	95	94	95.5
ক(ka)	96	96	96	96	96	94	95.67
খ(kha)	96	95	95	95	95	93	94.83
গ(ga)	100	99	97	97	98	96	97.83
ঘ(gha)	97	96	95	95	94	94	95.17
চ(cha)	95	95	95	94	95	93	94.5
ছ(chha)	95	95	94	94	94	93	94.17
জ(ja)	95	95	95	95	93	93	94.33
ঝ(jha)	94	93	93	93	92	91	92.67
ট(toa)	100	100	98	97	96	95	97.67
ঠ(thhoa)	100	99	98	98	97	95	97.83
ড(doa)	97	96	96	95	95	94	95.5
ঢ(dhhoa)	95	94	94	94	94	92	93.83
ত(ta)	94	94	94	94	94	92	93.67
থ(tha)	96	96	95	95	94	93	94.83
দ(da)	95	95	95	95	94	93	94.5
ধ(dha)	95	95	95	95	94	92	94.33
ণ(na)	94	94	94	94	94	92	93.67
প(pa)	95	95	94	94	95	92	94.17
ফ(pha)	94	94	93	92	92	91	92.67
ব(ba)	95	94	94	94	94	93	94
ভ(bha)	98	97	96	95	95	93	95.67
ম(ma)	96	96	96	95	94	93	95
য়(wa)	96	96	95	95	94	91	94.5
র(ra)	92	91	92	91	89	87	90.33
ল(la)	92	92	92	91	91	89	91.17
স(sa)	97	95	95	94	93	92	94.33
হ(ha)	98	97	96	96	95	93	95.83
ড়(rha)	97	96	96	95	95	93	95.33
ং(ng)	98	97	97	96	95	93	96
ঃ(hh))	97	97	96	96	95	92	95.5
০(zero)	100	100	99	99	98	97	98.83
১(one)	100	100	100	99	98	97	99
২(two)	100	100	98	98	97	96	98.17
৩(three)	100	100	99	98	98	98	98.83
৪(four)	100	99	100	98	97	97	98.5
৫(five)	100	99	99	99	97	97	98.5
৬(six)	99	99	99	98	97	98	98.33
৭(seven)	99	98	99	98	96	95	97.5
৮(eight)	99	99	98	98	97	96	97.83
৯(nine)	98	97	98	96	96	95	96.67
mean :	96.96	96.54	96.11	95.72	95.04	93.67	95.67

results for the proposed system prove that the proposed system is able to recognize any hand-sign of any sign language in any challenging environments.

Table 4 Results of ChSL numeral signs classification.

ChSL	accuracy (%)						
	E1	E2	E3	E4	E5	E6	mean
0	99	97	97	96	95	95	96.5
1	100	98	97	96	96	95	97
2	99	98	96	96	95	95	96.5
3	98	97	97	97	97	96	97
4	97	96	95	95	95	95	95.5
5	98	97	97	97	96	95	96.67
6	97	96	96	95	95	95	95.67
7	98	98	98	97	97	96	97.33
8	99	97	97	97	96	96	97
9	99	97	96	96	96	95	96.5
mean	98.4	97.1	96.6	96.2	95.8	95.3	96.57

3.7 Comparison between different methods for hand-signs classification

From Table 5, we claim that the test results of the proposed system (denoted by 'HSSCS') show better performance than existing reputed BdSL recognition system developed by Rahaman et al. [16] using contour matching algorithm (denoted by 'CM'); by Yasir et al. [20] using scale invariant feature transform (SIFT) based approach and support vector machine (SVM) (denoted by 'SIFT-SVM'); by Yasir et al. [19] using Linear Discriminant Analysis (LDA) and Artificial Neural Network (ANN) (denoted by 'LDA-ANN'); by Rahaman et al. [15] using skin detect and KNN algorithm (denoted by 'Skin-KNN'); by Jasim et al. [14] using Haar like feature based cascaded classifier and KNN Classifier (denoted by 'Haar-KNN'); and by Ayshee et al. [18] using Fuzzy rule-based approach (denoted by 'Fuzzy-rule-based'). We have implemented these systems using the same dataset of our proposed system (HSSCS) for different six environments only for 46 BdSL hand-signs. We have calculated the mean accuracies from the several confusion matrices for six different environments and computational costs (CC) for the several systems.

From the Table 5, the comparison results present that the proposed system (HSSCS) is more than eleven times faster than the mentioned previous systems 'Skin-KNN', 'Haar-KNN', 'Fuzzy Rule-based', 'SIFT-SVM' and 'LDA-ANN' with keeping the higher success rate accept our previous system 'CM'. But the accuracy rate of the proposed system 'HSSCS' is increased significantly from the previous system 'CM' in cluttered background and illumination variation environment with keeping the same computational cost. The recognition accuracies for all methods in Environment-1 (plain background with proper lighting) are satisfactory but the accuracies for other systems in other environments are not satisfactory. Especially, in environments E5 and E6 (cluttered and dynamic background with illumination variation environment), other systems do not perform accurately.

4 Conclusion

This paper presents a computer vision-based hand-signs segmentation and classification system using FRB-RGB model and Grid-Pattern analysis. The system segments human skin-color robustly using proposed Fuzzy rule based RGB (FRB-RGB) model in cluttered background with illumination variation. Probable binary hand-signs are extracted based on segmented skin-color pixels with specific motions. After noise removal followed by rotation, clipping and normalization, the system generates translation, rotation and scale invariant Grid-Pattern represented by WGV by applying a 5×5 Window-Grid Mask (WGM) on the binary image. Then the system forms a feature vector based on the WGV, area and mean height of the binary hand-signs. Finally, the system classifies the hand-signs using the proposed classifier by calculating the maximum similarity between the pre-trained feature vectors and test feature vector. The proposed system uses simple mathematical analysis based classifier instead of traditional classifiers to classify the hand-signs accurately in real-time. The system is trained using 4600 images for 46 hand-signs in BdSL and 1000 images for 10 numeral hand-signs in ChSL. The system is tested using six sets of images where each set is prepared using 4600 images for 46 hand-signs in BdSL and 1000 images for 10 numeral hand-signs in ChSL from 10 signers in six different challenging environments (E1, E2, E3, E4, E5 and E6) achieving satisfactory results. The experimental results prove that the system is capable of recognizing any hand-signs of any sign language in any environment, if it is trained properly. The proposed system is faster

Table 5 Comparison of different methods for BdSL classification.

methods	accuracy (%)						computational cost (ms/f)
	E1	E2	E3	E4	E5	E6	
CM	95.57	89.35	55.34	53.45	45.89	34.76	8.02
HSSCS	96.96	96.54	96.11	95.72	95.04	93.67	8.01
SHIFT-SVM	88.67	87.08	79.35	77.89	67.75	63.55	90.09
LDA-ANN	95.32	90.32	87.48	67.09	56.87	40.09	92.04
Skin-KNN	95.56	93.88	87.07	79.38	73.05	63.85	96.46
Fuzzy-rule-based	92.25	90.83	87.47	85.9	83.46	79.67	100.17
Haar-KNN	93.85	93.9	90.45	89.49	87.48	86.95	94.88

Fig. 20 Example snapshots of successfully recognition of hand-signs in various illumination and cluttered background.

and simpler and achieves higher accuracy when compared with related systems as shown in Table 5. However the system sometimes fails to distinguish two similar binary signs such as 'র' and 'ল' in which the color images of them are distinguishable but the binary images of them are very similar as shown in Fig.19. The system may fail to segment the hand-signs, if any skin-color objects with similar motion of hand-signs are presented in the ROI. These limitations may be overcome by the future development of the system. The system might be applicable for continuous recognition of hand-sign-spelled large vocabulary in any sign language by modeling the syntaxes of the specific language. The proposed system is applicable to assist as an interpreter for communication between sign and non sign people. It can also be used for Human Machine Interaction (HMI) and virtual reality using hand-signs.

Acknowledgements This research is partially supported and funded by the Information and Communication Technology (ICT) Division, Ministry of Posts, Telecommunications and IT, Government of the People's Republic of Bangladesh.

References

1. Choi J, Seo B K, Park J I. Simple efficient load balancing algorithms for peer-to-peer systems. In: Proceedings of the 8th International Conference on Virtual Reality Continuum and its Applications in Industry, VRCAI'09, New York, NY, USA. 2009, 319–321
2. Widjajanto W A, Lund M, Schelhowe H. Wayang authoring: a web-based authoring tool for visual storytelling for chil-

dren. In: Proceedings of the 6th International conference on advances in mobile computing and multimedia. 2008, 464–467

3. Lu F, Tian F, Jiang Y, Cao X, Luo W, Li G, Zhang X, Dai G, Wang H. Shadowstory: creative and collaborative digital storytelling inspired by cultural heritage. In: Proceedings of the the SIGCHI Conference on human factors in computing systems. 2011, 1919–1928
4. Mystakidis S, Lambropoulos N, Fardoun H M, Alghazzawi D M. Playful blended digital storytelling in 3d immersive e-learning environments: a cost effective early literacy motivation method. In: Proceedings of the 2014 workshop on interaction design in educational environments. 2014
5. Rubart J. Face-to-face collaboration points in storytelling using multitouch tabletop systems. In: Proceedings of the 2015 workshop on narrative and hypertext. 2015, 33–35
6. Liang H, Chang J, Kazmi I K, Jian J J, Zhang P. Hand gesture-based interactive puppetry system to assist storytelling for children. Vis Comput, 2017, 33(4): 517–531
7. Kulkarni V S, Lokhande S D. Appearance based recognition of american sign. International Journal on Computer Science and Engineering, 2010, 2(3): 560–565)
8. Ong E, Cooper H, Pugeault N, Bowden R. Sign language recognition using sequential pattern trees. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Providence, RI. 2012, 2200–2207
9. Naoum R, Hussein , Owaied H, Joudeh S. Development of a new arabic sign language recognition using k-nearest neighbor algorithm. Journal of Emerging Trends in Computing and Information Sciences, 2012, 3(8): 1173–1178
10. Bauer B, Hienz V, K. K. Video based continuous sign language recognition using statistical methods. In: Proceedings of 15th International Conference on Pattern Recognition, Barcelona. 2000, 463–466
11. Sako S, Kitamura T. Subunit modeling for japanese sign language recognition based on phonetically depend multistream hidden markov models. Universal Access in Human-Computer Interaction, 2013, 548–555
12. Holden E, Lee G, Owens R. Automatic recognition of colloquial australian sign language. In: Proceedings of the IEEE Workshop on Motion and Video Computing. 2005, 183–188
13. Zhang C, Tian Y, Huenerfauth M. Multi-modality american sign language recognition. In: Proceedings of the IEEE International Conference on Image Processing (ICIP). 2016, 281–285
14. Jasim M, Zhang T, M. H. A real-time computer vision-based static and dynamic hand gesture recognition system. International Journal on Image and Graphics, 2014, 14(01n02): 1–19
15. Rahaman M A, Jasim M, Ali M H, Hasanuzzaman M. Real-time computer vision based bengali sign language recognition. In: Proceedings of the 17th Int'l Conf. on Computer and Information Technology (ICCIT), Dhaka, Bangladesh. 2014, 192–197
16. Rahaman M A, Jasim M, Zhang T, Ali M H, Hasanuzzaman M. Real-time bengali and chinese numeral signs recognition using contour matching. In: Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO), IEEE, Zhuhai, China. 2015, 1215–1220
17. Rahaman M A, Jasim M, Ali M H, Hasanuzzaman M. A real-time appearance based bengali alphabet and numeral signs recognition system. Dhaka University Journal of Applied Science and Engineering, (In Press), 2016
18. Ayshee T F, Raka S A, Hasib Q R, Hossain M, Rahman R M. Fuzzy rule-based hand gesture recognition for bengali characters. In: Proceedings of the IEEE International Advance Computing Conference (IACC). 2014, 484–489
19. Yasir R, Khan R A. Two-handed hand gesture recognition for bangla sign language using lda and ann. In: Proceedings of the 8th International Conference on Software, Knowledge, Information Management and Applications. 2014, 1–5
20. Yasir F, Prasad P W C, Alsadoon A, Elchouemi A. Sift based approach on bangla sign language recognition. In: Proceedings of the IEEE 8th International Workshop on Computational Intelligence and Applications (IWCIA), Hiroshima. 2015, 35–39
21. Jarman A M, Arshad S, Alam N, Islam M J. An automated bengali sign language recognition system based on fingertip finder algorithm. International Journal of Electronics and Informatics, 2015, 4(1): 1–10
22. www.statista.com , May 2016
23. Begum S, Hasanuzzaman M. Computer vision-based bangladeshi sign language recognition system. In: Proceedings of the 12th International Conference on Computer and Information Technology (ICCIT). 2009, 414–419
24. Majumder M S R, Hossain M A, Mamtaz M, Ahmed M, Khan H R, Ali M H, Khaled O, Iqbal M. Bengali Sign Language Dictionary, 2nd. edition. Dhaka, Bangladesh: National Centre for Special Education and Ministry of Social Welfare in Cooperation with The Norwegian Association of The Deaf and Bangladersh National Federation of The Deaf, Bangladesh, 1994
25. Deb K, Mony H P, Chowdhury S. Two-handed sign language recognition for bangla character using normalized cross correlation. Global Journal of Computer Science and Technology, 2012, 12(3)
26. Karmokar B C, Alam K M R, Siddiquee M K. Bangladeshi sign language recognition employing neural network ensemble. International Journal of Computer Applications, 2012, 58(16): 43–46
27. Iraji M S, Tosinia A. Hand gesture segmentation method based on ycbcr color space and k-means clustering. International Journal of Image, Graphics and Signal Processing (IJIGSP), 2012, 4(-): 35–41
28. Qiu-yu Z, Jun-chi L, Mo-yi Z, Hong-xiang D, Lu L. Hand

gesture segmentation method based on ycbcr color space and k-means clustering. International Journal of Signal Processing, Image Processing and Pattern Recognition, 2015, 8(5): 105–116

29. Shaika K B, Ganesana P, Kalista V, Sathisha B S, Jenithab J M M. Comparative study of skin color detection and segmentation in hsv and ycbcr color space. In: Proceedings of the 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015), 57,. 2015, 41–48
30. Kale S, Saradadevi M. Active facial tracking. ASIAN JOURNAL OF COMPUTER SCIENCE & INFORMATION TECHNOLOGY, 2013, 1(3): 60–63
31. Pujol F A, Pujol M, Morenilla A J, Pujol M J. Face detection based on skin color segmentation using fuzzy entropy. Entropy, 2017, 19(1): 1–22
32. Boaventura I A G, Volpe V M, Silva I N D, Gonzaga A. Fuzzy classification of human skin color in color images. In: Proceedings of the 2006 IEEE International Conference on Systems, Man and Cybernetics, Taipei, Taiwan. 2006, 5071–5075
33. Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. In: Proceedings of IEEE CVRP. 2001, 511–518
34. Viola P, Jones M. Robust real-time object detection. International Journal of Computer Vision, 2004, 57(2): 137–154
35. Mufarroha F A, Utaminingrum F. Hand gesture recognition using adaptive network based fuzzy inference system and k-nearest neighbor. International Journal of Technology, 2017, 8(3): 559–567
36. Kumar B P P, Manjunatha M B. A hybrid gesture recognition method for american sign language. Indian Journal of Science and Technology, 2017, 10(1): 1–12
37. Zhang X, Wang J, Wang X, Ma X. Improvement of dynamic hand gesture recognition based on hmm algorithm. In: Proceedings of the International Conference on Information System and Artificial Intelligence (ISAI), Hong Kong, China. 2016, –
38. Ghotkar A, Vidap P, Deo K. Dynamic hand gesture recognition using hidden markov model by microsoft kinect sensor. International Journal of Computer Applications, 2016, 150(4): 1–9
39. Kaur G, Bathla G. Hand gesture recognition based on invariant features and artifical neural network. Indian Journal of Science and Technology, 2016, 9(43): 1–6
40. Xu J, Zhang X. A real-time hand detection system during hand over face occlusion. International Journal of Multimedia and Ubiquitous Engineering, 2015, 10(8): 287–302
41. Bhatia A, Srivastava S, Agarwal A. Face detection using fuzzy logic and skin color segmentation in images. In: Proceedings of the 3rd International Conference on Emerging Trends in Engineering and Technology, Goa, India. 2010, 225–228
42. Zadeh L A. Fuzzy sets. Inf. Control, 1965, 8(3): 338–353
43. Huang W, Yihuali . Bell-shaped probabilistic fuzzy set for uncertainties modeling. Journal of Theoretical and Applied Information Technology, 2012, 42(2): 875–882
44. Koprinkova P. Membership functions shape and its influence on the dynamical behaviour of fuzzy logic controller. Cybernetics and Systems: An International Journal, 2010, 31(2): 161–173
45. Murthy C A, Pal S K. Fuzzy thresholding: Mathematical framework, bound functions and weighted moving average technique. Pattern Recogn. Lett., 1990, 11(3): 197–206
46. Shannon C E. A mathematical theory of communication. The Bell System Technical Journal, 1948, 27: 379–423
47. Abdel-Khalek S, Ishak A B, Omer O A, Obada A S. A two-dimensional image segmentation method based on genetic algorithm and entropy. Optik, 2017, 131(-): 414–422
48. Susan S, Kumar A. Auto-segmentation using mean-shift and entropy analysis. In: Proceedings of the 3rd International Conference on Computing for Sustainable Global Development, New Delhi, India. 2016, 292–296
49. Bhowmik M, Sarkar A, Das R. Shannon entropy based fuzzy distance norm for pixel classification in remote sensing imagery. In: Proceedings of the Third International Conference on Computer, Communication, Control and Information Technology, Hooghly, India. 2015
50. Cheng H D, Chen J R, Li J. Threshold selection based on fuzzy c-partition entropy approach. Optik, 1998, 31(-): 857–870
51. Yin S, Zhao X, W.Wang , Gong M. Efficient multilevel image segmentation through fuzzy entropy maximization and graph cut optimization. Pattern Recognition, 2014, 47: 2894–2907
52. Mahmoodi M R, Sayedi S M. A comprehensive survey on human skin detection. I.J. Image, Graphics and Signal Processing, 2016, 5(1): 1–35
53. Gonzalez R C, Woods R E. Digital Image Processing, 3rd edition. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006
54. Boiroux D, Oke Y, Miwakeichi F, Oku Y. Pixel timing correction in time-lapsed calcium imaging using point scanning microscopy. Journal of Neuroscience Methods, Elsevier, November-2014, 237(-): 60–68
55. sourceforge.net . Emgucv, -
56. Nagarajan S, Subashini T S. Static hand gesture recognition for sign language alphabets using edge oriented histogram and multi class svm. International Journal of Computer Applications, 2013, 82(4): 28–35
57. Lu W, Huang J. Skin detection method based on cascaded adaboost classifier. Journal of Shanghai Jiaotong University (Science), 2012, 17(2): 197–202

Muhammad Aminur Rahaman received his B.Sc. and M.Sc. degree in Computer Science & Engineering from the Department of Computer Science & Engineering, Islamic University, Kushtia, Bangladesh in 2003 and 2004, respectively. He is a candidate for PhD degree under supervision of Prof. Dr. Md. Hasanuzzaman and Prof. Dr. Md. Haider Ali in the Department of Computer Science & Engineering, University of Dhaka, Bangladesh. He is a founder Director of Worldgaon (Pvt.) Limited which is a one of the famous software development company of Bangladesh. His current research interests including Computer vision, Sign language recognition and Human-Computer Interaction. He is a member of IEEE.

Mahmood Jasim received his B.Sc. and M.Sc. degree in Computer Science & Engineering from the Department of Computer Science & Engineering, University of Dhaka, Bangladesh in 2011 and 2013, respectively. He joined the Department of Computer science and engineering as a lecturer in 2014. Currently he is pursuing his PhD degree in the University of Massachusetts Amherst. His current research interests include Human-Computer Interaction, Image Processing, Computer Vision and Artificial Intelligence.

Dr. Md. Haider Ali received Ph.D. degree from the dept. of Electronics & Information Engineering, Toyohashi University of Technology, Japan in 2001. Prof. Ali has completed his Bachelor and Master degree from the dept. of Applied Physics & Electronics (presently EEE) University of Dhaka 1984 and 1985 respectively. He is a professor of the dept. of Computer Science & Engineering, University of Dhaka since June 2007. His current research interests include Human face recognition and expression detection, Post surgical expression simulation, Soft-tissue deformation modeling, Polygonal mesh simplification, Narrow band video transmission/video conferencing, Etc.

Dr. Tao Zhang received his B.S., M.S. and Ph.D. degrees in Electrical Engineering from Tsinghua University of China, in 1993, 1995 and 1999, respectively. In 2002, he received his second Ph.D. degree in Electrical Engineering in the Department of Advanced System Control Engineering, Graduate School of Science and Engineering, Saga University, Japan. Since 2012, he has been serving as a Professor in the Dept. of Automation, Tsinghua University, Beijing, China. His current research interests include Nonlinear System Control Theory and Application, Intelligent Control of Robot, Micro-satellite Engineering, Fault Diagnosis and Reliability Analysis and Signal Processing. He is a member of IEEE, IEIEC and IFAC.

Dr. Md. Hasanuzzaman received Ph.D. degree from the Department of Informatics, National Institute of Informatics (NII), The Graduate University for Advanced Studies, Tokyo, Japan. He graduated (with Honors) in 1993 from the Department of Applied Physics & Electronics, University of Dhaka, Bangladesh. He completed Masters of Science (M.Sc.) in Computer Science in 1994 from the University of Dhaka. He joined as a Lecturer in the Department of Computer Science & Engineering, University of Dhaka, Bangladesh in 2000. Since March 2013, he has been serving as a professor in the Department of Computer Science & Engineering, University of Dhaka, Bangladesh. His current research interests include Human-Computer Interaction, Image Processing, Computer Vision and Artificial Intelligence.