

ASL CHARACTERS RECOGNITION USING PERSON SPECIFIC SUBSPACE METHOD

¹Md. Hasanuzzaman, ¹S. M. Tareeq, ²M. A. Bhuiyan, ³Y. Shirai and ⁴H. Ueno

¹Department of Computer Science and Engineering
University of Dhaka, Dhaka, Bangladesh

²Department of Electronics and Computer Science
Jahangirnagar University, Savar, Dhaka, Bangladesh

³Department of Human and Computer Intelligence, School of Information Science and
Engineering, Ritsumeikan University, Japan

⁴Department of Informatics, National Institute of Informatics, Tokyo, Japan.
E-mail: hzamancsdu@yahoo.com

Abstract: *This paper presents a person-specific subspace method for ASL characters recognition. To implement this person specific subspace method for ASL characters recognition the system first recognizes the faces and then uses skin-color segmentation method for hand poses segmentation. From the segmented hand poses the system recognizes ASL characters using both PCA and person-specific subspace method separately. Experiment result shows that the classification accuracy of person-specific subspace method is better than general PCA method. The experiment result also shows that person-specific subspace method is faster than general PCA method.*

Key Words: *ASL character, Face recognition, PCA, Person-specific subspace method*

1 Introduction

The Gestures associated with speech are referred to as gesticulation. Gestures, which function independently of speech, are referred to as autonomous gestures. Autonomous gestures can be organized into their own communicative language, such as American Sign Language (ASL) [1]. Autonomous gesture can also represent motion commands to use in communication and machine control [2][3]. Researchers are usually concerned with gestures those are directed toward the control of specific object or the communication with a specific person or group of people [4].

In our society there are large numbers of people who cannot speak. On the other hand human are habituated to use gesture with speech. In every country there are some deaf or dumb people who mainly communicate with sign language. Every country has its own

culture though major hand signals meaning are same all over the world. The same gesture is used for different meanings and several gestures are used for the same meanings for different countries [5]. To interact with sign user the machine should recognize and learn sign characters and language. Automatic gesture and sign language recognition has been attracting computer vision researchers for a long time [4]. It offers enhancement of communication capabilities for the speech-impaired and deaf people. Automatic sign recognition has been investigated since around 1995 [6]. Researchers tried a variety of techniques, such as Fuzzy logic [7], neural networks [6][8], Principal Component Analysis (PCA) method [9], and Hidden Markov Models (HMMs) [10].

In this paper we propose person-specific subspace method for ASL characters recognition. Human skin-color differs from person to person and varies on lighting conditions. However, in a specific lighting condition, human skin colors are almost similar for a specific origin. This system utilizes skin color based segmentation approach to separate probable hand pose from the YIQ color spaces [11-12]. We have already developed vision-based gesture recognition system to interact with robot [13-14]. In that paper we used pose specific subspace method for pose classification where each eigenspace represent one hand pose. This system was tested for "TwoHand", "LeftHand", "RightHand", "PointLeft", "PointRight", "One", "Two" and "Three" gestures.

This paper compares PCA and person specific sub-space method for ASL characters classification. In the general PCA method, eigenvectors are calculated from training images that include all the hand poses or ASL characters. But for classification a large number of hand poses of large number of users, needs a large number of training datasets from which eigenvector generation is tedious and may not be feasible using personal computer. Considering these difficulties, we have proposed person-specific subspace method that partitions the comparison area. In PCA method, one PCA is used for all poses of all persons. In person-specific subspace method, all hand poses are grouped based on each person and for each person one PCA is used. Our current system can classify ASL sign characters, but in the future it should understand the ASL sign words or sentences that consider hand poses, its positions relating to other body parts, and direction of motion.

2 Person-Identification

In the last few years, face detection and person identification attracts many researchers due to security concern; therefore, many interesting and useful research demonstrations and commercial applications have been developed. A first step of any face recognition or vision-based person identification system is to locate the face in the image. After locating the probable face, researchers use facial features (eyes, nose, nostrils, eyebrows, mouths, leaps, etc.) detection method to detect face accurately [15]. Face recognition or person identification compares an input face image or image features against a known face database or features databases and report match, if any. Face detection from a single image or an image sequences is a difficult task due to variability in pose, size, orientation, color, expression, occlusion and lighting condition. Face detection researchers summarized the face detection work into four categories: template matching methods, feature invariant approaches, appearance-based methods and knowledge-based approaches [15]. Such approaches typically rely on a static background, so that human face can be detected using image differencing. Many

researches also used skin color as a feature and leading remarkable face tracking as long as the lighting conditions do not varies too much [2, 11, 12, 17].



Fig. 1 Example of face detection scenario

This system detects face from a clutter background using multi-resolutions face template pyramid [13] and recognizes a person using the eigenface method [16]. Figure 1 Shows example scenario of face detection where right part shows the multi-resolution face templates those are scan over the image frame one after another [3][13]. After identifying the person, the system uses person-specific subspace method for ASL character classification from the segmented skin like regions.

3 Skin-Region Segmentation and Normalization

Several color spaces have been utilized to label pixels as skin including RGB, HSV, YCrCb, YIQ, CIE-XYZ, CIE-LUV, etc. However, such skin color models are not effective where the spectrum of the light sources varies significantly. In this study YIQ (Y is luminance of the color and I, Q are chrominance of the color) color representation system is used for skin-like region segmentation because it is typically used in video coding and provides an effective use of chrominance information for modeling the human skin color [12]. To segment probable ASL characters (hand poses), it is assumed that the captured camera images are represented in the RGB color spaces. Each pixel in the images is represented by a triplet $P=F(R, G, B)$.

The RGB images taken by the video camera are converted to YIQ color representation system. Skin color region is determined by applying threshold values $((Y_{Low} < Y < Y_{High}) \&\& (I_{Low} < I < I_{High}) \&\& (Q_{Low} < Q < Q_{High}))$

[2][13]. The thresholds are predetermined from the histogram of YIQ components of the skin region and included Y_High, Y_Low, I_High, I_Low, Q_High and Q_Low as the ranges of threshold values of the Y, I, Q components [13]. **Table 1** shows the attributes of the multiple thresholds of Y, I, Q components for each person.

Table 1 Contents of a person profile

Item	Type	Meaning
Person_ID	String	Name of the User
Y_High	Integer	Maximum luminance (Y) of skin regions
Y_Low	Integer	Minimum luminance (Y) of skin regions
I_High	Integer	Maximum Chrominance (I) of skin regions
I_Low	Integer	Minimum Chrominance (I) of skin regions
Q_High	Integer	Maximum Chrominance (Q) of skin regions
Q_Low	Integer	Minimum Chrominance (Q) of skin regions

Probable hands regions are segmented from the image with the two largest connected regions excluding face poses of skin-colored pixels. The notation of pixel connectivity describes a relation between two or more pixels. In order to consider two pixels to be connected, their pixel values must both be from the same set of values V (for binary images V is 1, for gray images it may be specific gray value). Generally, connectivity can either be based on 4- or 8-connectivity. In the case of 4-connectivity, it does not compare the diagonal pixels but 8-connectivity compares the diagonal positional pixels considering 3×3 matrix, and as a result, 8-connectivity component is more noise free than 4-connectivity component. In this system, 8-pixels neighborhood connectivity is employed.

In order to remove the false regions from the segmented blocks, smaller connected regions are assigned by the values of black-color (R=G=B=0). As a result, after thresholding the segmented image may contain some holes

in the three largest skin-like regions. In order to remove noises and holes, segmented images are filtered by morphological dilation and erosion operations with a 3×3 structuring element. The dilation operation is used to fill the holes and the erosion operations are applied to the dilation results to restore the shape.

After filtering, the segmented skin regions are bounded by rectangular box using height and width information of each segment: $(M_1 \times N_1)$ and $(M_2 \times N_2)$. If the user shirt's color is similar to skin color then segmentation accuracy is very poor. If the user wears short sleeves or T-shirt then it needs to separate hand palm from arm. This system assumes the person wearing full shirt with non-skin color.

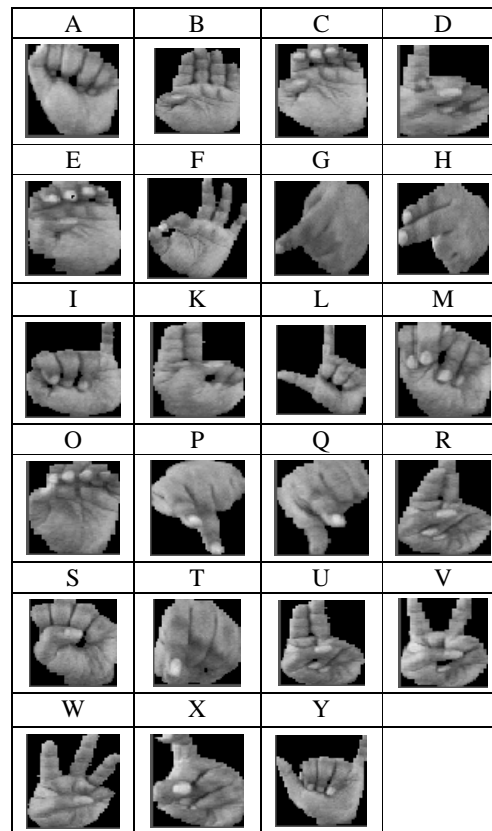


Fig. 2 Examples of training images (23-ASL characters)

Normalization is done to scale the image to match with the size of the training image and convert the scaled image to gray image [17]. Each segment is scaled to be square images with (60×60) and converted it to as gray images (BMP image). Suppose, we have a segment of rectangle $P[(x^l, y^l) - (x^h, y^h)]$ we sample it to rectangle $Q[(0, 0) - (60 \times 60)]$ using following expression,

$$Q(x^q, y^q) = P(x^l + \frac{(x^h - x^l)}{60} x^q, y^l + \frac{(y^h - y^l)}{60} y^q) \quad (1)$$

Using the same segmentation and normalization methods training images and test images are prepared, that is why result of this matching approach is better than others who used different training/template image databases. Beside this, we have included training/template images creation functions in this system so that it can adapt with person and illumination changes. Figure 2 shows the examples of training images for 23-ASL characters.

4 ASL Character Classification

Principal component analysis (PCA) method is a standard pattern recognition approach and many researchers use it for face and hand pose classification [3]. The main idea of the principal component analysis (PCA) method is to find the vectors that best account for the distribution of target images within the entire image space. In the general PCA method, eigenvectors are calculated from training images that include all the poses or classes. But for classification a large number of hand poses for a large number of users, need large number of training datasets from which eigenvectors generation is tedious and may not be feasible for a personal computer. Considering these difficulties we have proposed person-specific subspace method that partitions the comparison area. In person-specific subspace method, ASL characters (hand poses) are grouped based on each person and for each person one PCA is used. In the following subsection we have described the algorithm of person-specific subspace method for ASL characters classification, which is very similar to general PCA based algorithm.

4.1 Person-Specific Subspace Method

In subspace method target image is projected on each subspace separately. Table 2 summarizes the symbols that are used for describing person-specific subspace method for ASL characters classification. The procedure of person-specific subspace method includes following operations:

Prepare noise free version of predefined ASL characters (hand poses) of each person for training images $T_j^{(i)} (N \times N)$, where j is number training images of i^{th} person and $j=1, 2, \dots, M$. Figure 2 shows the example training image of one person for 23 ASL characters.

For each person ASL characters (hand poses), calculate eigenvectors ($u_m^{(i)}$) using Matthew Turk and Alex Pentland technique [16] and chose k -number of eigenvectors ($u_k^{(i)}$) corresponding to the highest eigenvalues to form principal components for that person. These vectors for each person define the subspace of that person.

Calculate corresponding distribution in k -dimensional weight space for the known training images by projecting them onto the subspaces (eigenspaces) of the corresponding person and determined the weight vectors ($\Omega_l^{(i)}$), using equations (2) and (3).

$$\omega_k^{(i)} = (u_k^{(i)})^T (s_l^{(i)} - \Phi_i) \quad (2)$$

$$\Omega_l^{(i)} = [\omega_1^{(i)}, \omega_2^{(i)}, \dots, \omega_k^{(i)}] \quad (3)$$

Where, average image of i^{th} person $\Phi_i = \frac{1}{M} \sum_{n=1}^M T_n$ and $s_l^{(i)} (N \times N)$ is l^{th} known images of i^{th} person.

Each segmented skin-region is treated as individual test image, transformed it into eigenimage components and calculate a set of weight vectors ($\Omega^{(i)}$) by projecting the input image onto each of the subspace as equations (2) and (3).

Determine if the image is a predefined ASL character based on minimum Euclidean distance among weight vectors using equation (4) and (5),

$$\epsilon_j^{(i)} = \|\Omega^{(i)} - \Omega_j^{(i)}\| \quad (4)$$

$$\epsilon = \arg \min \{\epsilon_j^{(i)}\} \quad (5)$$

If ϵ is lower than predefined threshold then

its corresponding ASL character is identified. For exact matching ε should be zero but for practical purposes this method uses a threshold value that is selected through experiment.

Table 2 List of symbols used in subspace method

Symbols	Meanings
$T_j^{(i)}$	Training images for i^{th} person
$u_m^{(i)}$	m^{th} Eigenvectors for i^{th} person
$\Omega_l^{(i)}$	Weight vector for i^{th} person
$\omega_k^{(i)}$	Element of weight vector for i^{th} person

5 Experimental Setup

The goal of the system is to classify the ASL characters using both PCA and person-specific subspace method. In this system a standard video camera is used for data acquisition. Each captured image is digitized into a matrix of 320×240 pixels with 24-bit color. This system is tested for real-time input images as well as static images. First, the system is trained using the known training images of predefined ASL characters (hand poses) of all known persons. All the training images are 60×60 pixels gray images. In the training phase, this system generates eigenvectors and feature vectors for all the known users ASL characters.

5.1 Results of Face Recognition

This method can detect the face from cluttered background if the two eyes are visible. The proposed face detection method in this paper is robust against background, motion and distance. However, it has a large computational cost that is the bottleneck for real-time human-robot interaction. Three factors directly affect the computation costs: step size, dimension of template images and the number of template images. If the step size is 1, the number of comparison for sliding one template on the whole image is $46800(60 \times 60, 320 \times 240)$ where 60×60 is the template image dimension and 320×240 is the input image dimension. In similar cases, if step sizes are 2, 3, 4, 5 then numbers of comparisons are 11700, 5220, 2925, and 1872, respectively.

In our research, we found that the accuracy of frontal face recognition is better than up, down and more left right directed faces [13].

In this person identification system we prefer frontal and a small left or right rotated faces. We have verified this face recognition method for 680 faces of 7 persons, where two are females. Precision is 93% and recall rate is 94 % [3]. The precision (%) is defined by the ratio of the numbers of correct recognition to total numbers of recognition for each person faces. The recall rate (%) is defined by the ratio of the numbers of correct face recognition to total numbers of input faces for each person.

5.2 Results of ASL Characters Recognition

In this study we use two pose classification methods: PCA method and person-specific subspace method. In person-specific subspace method, ASL characters (hand poses) are grouped based on each person and for each person one PCA is used. In the following two subsections we will present the results of ASL character classification using general PCA and person-specific subspace methods.

5.2.1 Comparison of PCA and Person-specific Subspace Methods:

In this section we compare the performance of the general PCA and the person-specific subspace methods for the classification of 23 ASL characters. For this experiment we have trained the system using 4140 ($9 \times 23 \times 20$) training images of 23 ASL characters of 9 persons. For each ASL character of a person 20 images are used as training images. In the case of general PCA method eigenvectors are calculated using 4140 training images. On the other hand, in the case of person-specific subspace method separate eigenvectors are calculated for each person based on 460 (23×20) training images of 23 ASL characters of a person. In case of subspace method, the comparison area is divided by the number of persons. After identification of a person the system uses corresponding person eigenspace as the comparison space for pose classification and thereby reduce the computation cost in the order of number of persons. Nine individuals were asked to act for these predefined ASL character (hand poses) in front of the camera and the sequences of images were saved as individual image frames. Then each image frame is tested using the general PCA method and the person-specific subspace method. The threshold value (for minimal Euclidian

distance) for the pose classifier is selected so that all the poses are classified.

Table 3(a) and 3(b) presents the comparison of PCA and person-specific subspace (PSN_SUB) methods for 23 ASL characters classification for person_1 “Sakurai”. The accuracy is defined by the ratio of the number of correct recognitions to the number of total input for a specific character. Figure 3 shows the graphical representation of the pose classification accuracy of PCA and person-specific subspace methods for person_1 “Sakurai”. From Figure 3, we can see that in the case of person-specific subspace method pose classification accuracy is better than general PCA method, except for the ASL character ‘I’. In the general PCA method, test images of an ASL character of a person are matched with the other ASL character (difference in shape) of another person. For example, in the PCA method 15-test images for ASL character ‘A’ of person “Sakurai” (P_x) are matched with the model images of ASL character ‘M’ of person “Makino” (P_y).

Table 3 (a) Comparison of PCA and person-specific subspace methods for ASL characters classification (person_1 “Sakurai”)

ASL Char	Number of Input	Correct Recognition	
		PCA	PSN_SUB
A	55	40	55
B	30	26	30
C	60	55	60
D	30	30	30
E	60	55	57
F	60	58	60
G	60	60	60
H	45	41	42
I	57	57	55
K	60	60	60
L	60	60	60
M	60	59	60
O	60	56	59
P	60	60	60
Q	60	60	60
R	60	37	53
S	60	60	60
T	60	54	60
U	60	60	60

V	60	56	58
W	60	60	60
X	60	57	60
Y	60	57	60

Table 3(b) Comparison of PCA and person-specific subspace methods for ASL characters classification (person_1 “Sakurai”)

ASL Char	Number of Input	Accuracy (%)	
		PCA	PSN_SUB
A	55	72.72	100
B	30	86.66	100
C	60	91.66	100
D	30	100	100
E	60	91.66	95
F	60	96.66	100
G	60	100	100
H	45	91.11	93.33
I	57	100	96.49
K	60	100	100
L	60	100	100
M	60	98.33	100
O	60	93.33	98.33
P	60	100	100
Q	60	100	100
R	60	61.66	88.33
S	60	100	100
T	60	90	100
U	60	100	100
V	60	93.33	96.66
W	60	100	100
X	60	95	100
Y	60	95	100

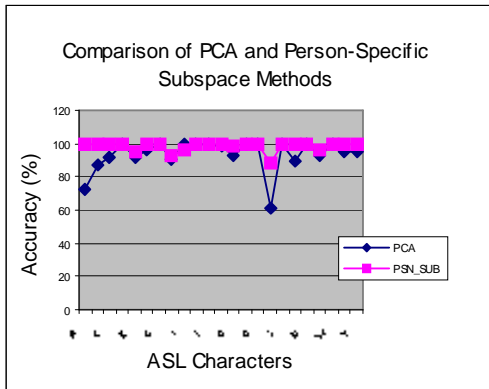


Fig. 3 Comparison of PCA and person-specific Subspace methods for ASL character classification (person_1 "Sakurai")

In Fig. 4, first 20 feature points corresponding to the first 20 images denote the training images of character 'A' of "Sakurai" ($P_{xA_{tr}}$); 15 points for images 21 to 35 denote the test images of character 'A' of "Sakurai" ($P_{xA_{tst}}$); 20 feature points for images 36 to 55 denote the training images of character 'M' of "Makino" ($P_{yM_{tr}}$); 20 feature points for images 56 to 75 denote the training images of character 'M' of "Sakurai" ($P_{xM_{tr}}$) and last 20 feature points for images 76 to 95 denote the training images of character 'A' of "Makino" ($P_{yA_{tr}}$). From the figure we can see that 15 test images of character 'A' of "Sakurai" ($P_{xA_{tst}}$) are nearest to model images of character 'M' of person 'Makino' ($P_{yM_{tr}}$) in terms of vertical distances.

On the other hand, in the case of person-specific subspace method those 15-test images are matched (nearest) with the ASL character 'A' of the same person "Sakurai" as shown in Fig. 5. In the case of ASL character 'I' of "Sakurai", 2-test images are classified as ASL character 'U' in person-specific subspace method, but in PCA method it is correctly classified as 'I'.

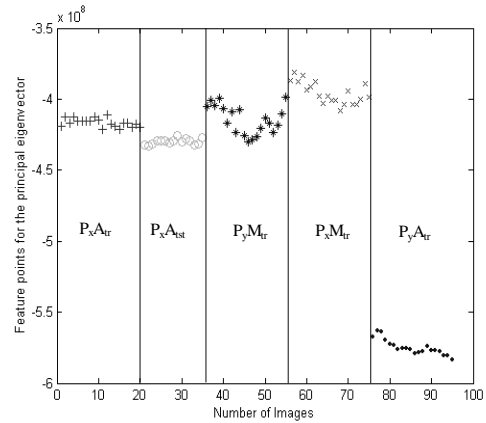


Fig. 4 Sample distribution of feature points in the general PCA method (for char 'A' of 'Sakurai')

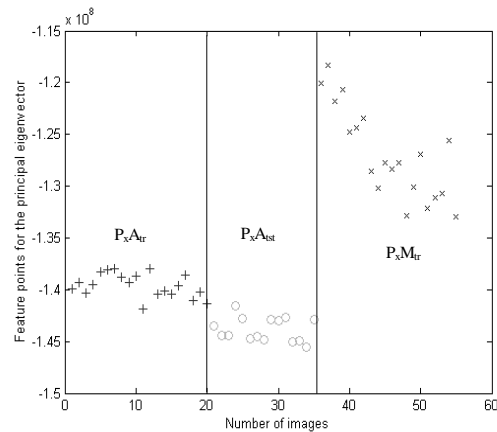


Fig. 5 Sample distribution of feature points in the person-specific subspace method (for char 'A' of 'Sakurai')

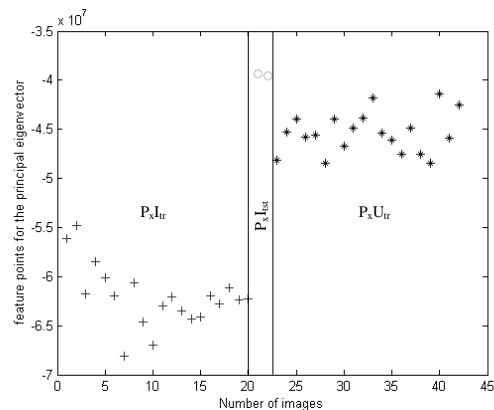


Fig. 6 Sample distribution of feature points in the person-specific subspace method (for Char 'I' of 'Sakurai')

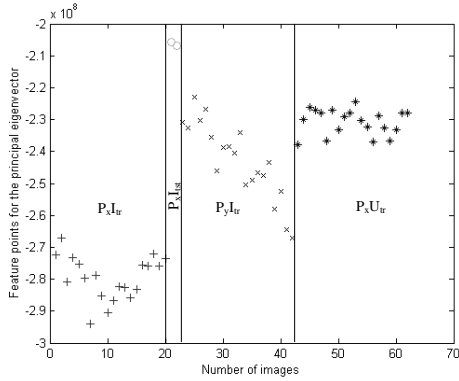


Fig. 7 Sample distribution of feature points in the general PCA method (for Char 'I' of 'Sakurai')

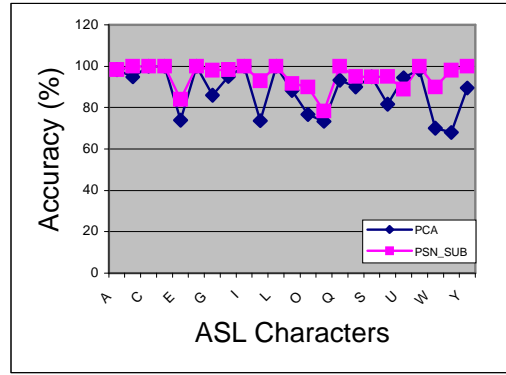


Fig. 8 Comparison of PCA and person-specific Subspace methods for ASL character classification (person_2 "Hasan")

Table 4 Comparison of PCA and person-specific subspace methods for ASL characters classification ("Hasan")

ASL Char	Number of Input	Correct Recognition	
		PCA	PSN_SUB
A	60	59	59
B	58	55	58
C	59	59	59
D	14	14	14
E	50	37	42
F	50	50	50
G	50	43	49
H	60	57	59
I	60	60	60
K	57	42	53
L	60	60	60
M	60	53	55
O	60	46	54
P	60	44	47
Q	60	56	60
R	60	54	57
S	58	55	55
T	60	49	57
U	18	17	16
V	60	59	60
W	40	28	36
X	50	34	49
Y	57	51	57

Table 4 presents the comparison of PCA and person-specific subspace (PSN_SUB) methods for 23-ASL characters classification for the person_2 "Hasan".

Table 5 Comparison of PCA and person-specific subspace methods for ASL characters classification (person_3 "Huda")

ASL Char	Number of Input	Correct Recognition	
		PCA	PSN_SUB
A	60	41	60
B	60	58	60
C	60	60	60
D	60	58	60
E	60	60	60
F	50	25	50
G	60	60	60
H	60	60	60
I	60	60	60
K	60	60	60
L	60	60	60
M	60	53	53
O	60	59	60
P	60	60	60
Q	60	60	60
R	60	59	60
S	60	43	58
T	60	60	60
U	60	56	60
V	60	60	60
W	60	59	60
X	57	57	57
Y	56	56	56

Figure 8 shows the graphical representation of the pose classification accuracy of the PCA and person-specific subspace methods for person_2 "Hasan". In this case, pose classification accuracy in person-specific

subspace method is better than the general PCA method in the major cases except for the character ‘U’.

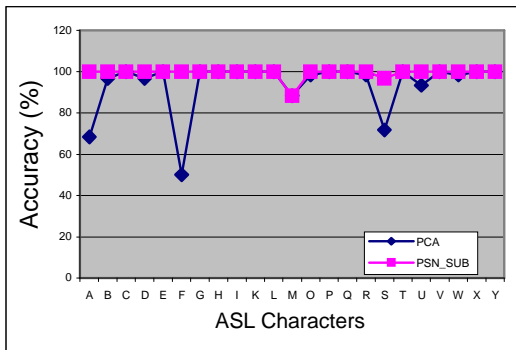


Fig. 9 Comparison of PCA and person-specific Subspace methods for ASL character classification (person_3 ‘Huda’)

Table 5 presents the comparison of PCA and person-specific subspace (PSN_SUB) methods for 23 ASL characters classification for person_3 ‘Huda’. Figure 9 shows the graphical representation of the pose classification accuracy of the PCA and person-specific subspace methods for person_3 ‘Huda’. In his case, for the ASL characters ‘A’, ‘B’, ‘D’, ‘F’, ‘O’, ‘R’, ‘S’, ‘U’, and ‘W’ pose classification accuracy is better in person-specific subspace method than general PCA method, and for other characters classification accuracy is the same in both methods. In the PCA method, 19-test images of character ‘A’ of person_3 ‘Huda’ were wrongly classified as character ‘M’ of person ‘Sakurai’ and ‘Makino’; 25-test images of characters ‘F’ of person ‘Huda’ were wrongly classified as character ‘Y’ of person ‘Hasan’. On the other hand, in the person-specific subspace method those images are correctly classified.

Table 6 presents the comparison of PCA and person-specific subspace (PSN_SUB) methods for 23-ASL characters classification for person_4 ‘Okawa’. Figure 10 shows the graphical representation of the pose classification accuracy of the PCA and person-specific subspace methods for person_4 ‘Okawa’. In his case, for the ASL characters ‘K’, ‘O’, ‘R’, ‘T’, and ‘W’ pose classification accuracy is better in person-specific subspace method than general PCA method, and for other characters classification accuracy is the same in both

methods.

Table 6 Comparison of PCA and person-specific subspace methods for ASL characters classification (person_4 ‘Okawa’)

ASL Char	Number of Input	Correct Recognition	
		PCA	PSN_SUB
A	60	60	60
B	60	60	60
C	60	60	60
D	57	57	57
E	50	50	50
F	60	60	60
G	60	60	60
H	60	60	60
I	60	60	60
K	60	21	60
L	60	60	60
M	60	60	60
O	60	56	60
P	60	60	60
Q	60	60	60
R	60	55	59
S	60	60	60
T	60	43	60
U	60	60	60
V	60	60	60
W	60	27	59
X	60	60	60
Y	60	60	60

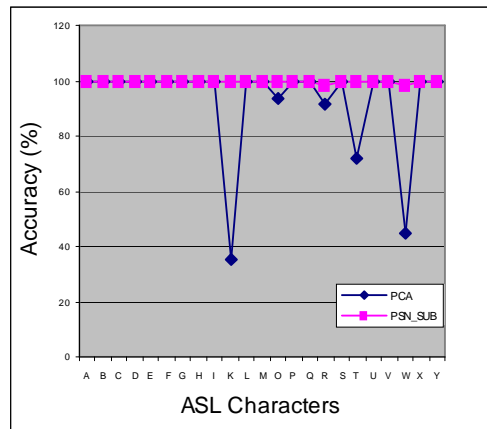


Fig. 10 Comparison of PCA and person-specific Subspace methods for ASL character classification (person_4 ‘Okawa’)

Table 7 presents the comparison of PCA and person-specific subspace (PSN_SUB) methods for 23 ASL characters classification for person_5 ‘Makino’. Figure 11 shows the graphical representation of the pose classification accuracy of the PCA and

person-specific subspace methods for person_5 “Makino”. In his case, for the ASL characters ‘D’, ‘G’, ‘I’, ‘M’, ‘T’, ‘V’ and ‘W’, pose classification accuracy is better in person-specific subspace method than general PCA method; for the character ‘Q’ PCA method is little better than person-specific subspace method; and for other characters pose classification accuracy is the same in both methods.

Table 7 Comparison of PCA and person-specific subspace methods for ASL characters classification (person_5 “Makino”)

ASL Char	Number of Input	Correct Recognition	
		PCA	PSN_SUB
A	50	50	50
B	60	60	60
C	60	60	60
D	20	17	20
E	60	60	60
F	60	60	60
G	60	46	60
H	60	60	60
I	60	55	60
K	60	60	60
L	60	60	60
M	60	56	60
O	60	60	60
P	60	60	60
Q	29	28	27
R	50	50	50
S	60	60	60
T	60	39	50
U	60	60	60
V	60	45	46
W	60	59	60
X	60	60	60
Y	60	60	60

In the case of PCA method, 3-test images of character ‘D’ of person “Makino” are classified as character ‘X’; 14-test images of character ‘G’ of person “Makino” are classified as ‘B’ of “Sakurai”; 5-test images of character ‘I’ of person “Makino” are classified as ‘F’ of “Cho”; 4-test images of character ‘M’ of person “Makino” are classified as ‘T’; 11-test images of character ‘T’ of person “Makino” are classified as ‘O’

of “Hasan”; 15-test images of character ‘V’ of person “Makino” are classified as ‘W’ of the same person.

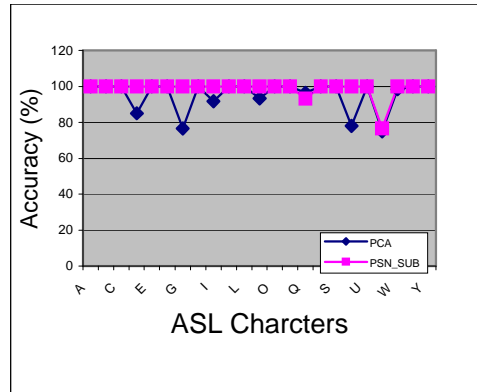


Fig 11 Comparison of PCA and person-specific Subspace methods for ASL character classification (person_5 “Makino”)

Table 8 Comparison of PCA and person-specific subspace methods for ASL characters classification (person_6 “Sato”)

ASL Char	Number of Input	Correct Recognition	
		PCA	PSN_SUB
A	60	60	60
B	60	60	60
C	60	60	60
D	58	58	57
E	60	45	56
F	60	60	60
G	60	59	60
H	60	60	60
I	60	60	60
K	40	40	40
L	60	60	60
M	60	60	60
O	60	60	60
P	60	60	60
Q	60	60	60
R	60	60	60
S	60	60	60
T	60	60	60
U	58	58	58
V	60	60	60
W	60	60	60
X	60	60	60
Y	60	60	60

Table 8 presents the comparison of PCA and person-specific subspace (PSN_SUB)

methods for 23-ASL characters classification for person_6 “Sato”. Figure 12 shows the graphical representation of the pose classification accuracy of the PCA and person-specific subspace methods for person_6 “Sato”. In his case, for the ASL characters ‘E’, and ‘G’, pose classification accuracy is better in person-specific subspace method than general PCA method; and for other characters pose classification accuracy is same in the both methods.

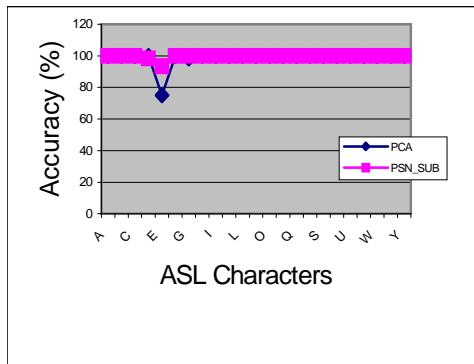


Fig. 12 Comparison of PCA and person-specific Subspace methods for ASL character classification (person_6 “Sato”)

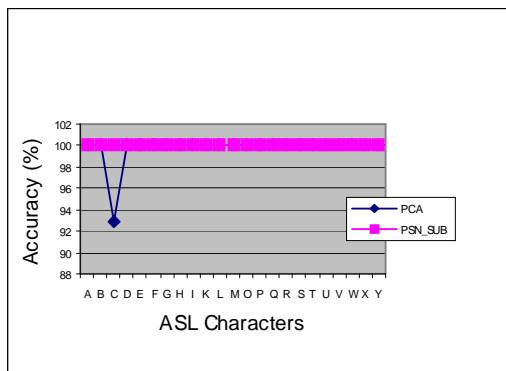


Fig. 13 Comparison of PCA and person-specific Subspace methods for ASL character classification (person_7 “Cho”)

Table 9 presents the comparison of PCA and person-specific subspace (PSN_SUB) methods for 23 ASL characters classification for person_7 “Cho”. Figure 13 shows the graphical representation of the pose classification accuracy of the PCA and person-specific subspace methods for person_7 “Cho”. In his case, for the ASL characters ‘C’, pose classification accuracy is better in person-specific subspace method

than general PCA method; and for other characters pose classification accuracy is same in both methods. In the general PCA method, 4-test images of ASL character ‘C’ of person “Cho” are matched with the ASL character ‘B’ of person “Satomi”, but in the person-specific subspace method those test images are matched with character ‘C’ of the person “Cho”.

Table 9 Comparison of PCA and person-specific subspace methods for ASL characters classification (person_7 “Cho”)

ASL Char	Number of Input	Correct Recognition	
		PCA	PSN_SUB
A	60	60	60
B	60	60	60
C	56	52	56
D	60	60	60
E	58	58	58
F	60	60	60
G	60	60	60
H	60	60	60
I	60	60	60
K	60	60	60
L	60	60	60
M	60	60	60
O	60	60	60
P	60	60	60
Q	60	60	60
R	60	60	60
S	60	60	60
T	60	60	60
U	60	60	60
V	60	60	60
W	60	60	60
X	60	60	60
Y	60	60	60

Table 10 presents the comparison of PCA and person-specific subspace (PSN_SUB) methods for 23 ASL characters classification for person_8 “Satomi”. Figure 14 shows the graphical representation of the pose classification accuracy of the PCA and person-specific subspace methods for the person_8 “Satomi”. In her case, for the ASL characters ‘G’, ‘I’, ‘I’, ‘K’, ‘O’, ‘R’, and ‘T’, pose classification accuracy is better in

person-specific subspace method than general PCA method; and for other characters pose classification accuracy is the same in both methods. In the general PCA method, 26-test images of ASL character ‘G’ of person “Satomi” are matched with the ASL character ‘H’ of person “Satomi”, but in the person-specific subspace method 20 images of those are matched with character ‘G’ of the person “Satomi”.

Table 10 Comparison of PCA and person-specific subspace methods for ASL characters classification (person_8 “Satomi”)

ASL Char	Number of Input	Correct Recognition	
		PCA	PSN_SUB
A	60	60	60
B	60	60	60
C	60	60	60
D	60	60	60
E	60	60	60
F	50	50	50
G	60	34	54
H	60	60	60
I	60	59	60
K	60	53	55
L	60	60	60
M	60	60	60
O	60	57	59
P	60	60	60
Q	60	60	60
R	60	55	57
S	60	60	60
T	60	50	56
U	60	60	60
V	60	60	60
W	60	60	60
X	60	60	60
Y	60	60	60

Table 11 presents the comparison of PCA and person-specific-subspace (PSN_SUB) methods for 23-ASL characters classification for person_9 “Yoshi”. Figure 15 shows the graphical representation of the pose classification accuracy of the PCA and person-specific subspace methods for person_9 “Yoshi”. In his case, for the ASL characters ‘A’, ‘C’, ‘H’, ‘K’, and ‘Y’, pose classification accuracy is better in person-specific subspace method than general

PCA method; and for other characters pose classification accuracy is the same in both methods.

From the experimental results we can conclude that pose classification accuracy of person-specific subspace method is higher than the general PCA method in the same environment. In general PCA method some characters of one person can match with another characters of other persons due to the variation of orientation or imperfect segmentation results. But when person specific subspace method is used there is no way to match with other character of another person. Besides this, the person-specific subspace method is also N times (N is the number of person) faster than the general PCA method.

Table 11 Comparison of PCA and person-specific subspace methods for ASL characters classification (person_9 “Yoshi”)

ASL Char	Number of Input	Correct Recognition	
		PCA	PSN_SUB
A	50	46	50
B	60	60	60
C	60	59	60
D	60	60	60
E	60	53	53
F	60	60	60
G	60	60	60
H	60	57	59
I	60	60	60
K	60	59	60
L	60	60	60
M	60	60	60
O	60	60	60
P	60	60	60
Q	60	60	60
R	60	60	60
S	60	60	60
T	60	60	60
U	60	60	60
V	60	60	60
W	60	60	60

X	60	60	60
Y	60	57	60

This system can classify 23 ASL characters. This system is tested with 3400 ASL character poses of different persons. The accuracy of pose classification depends on the accuracy of segmentation output. Wrong classification mainly happen due to imperfect segmentation and variation of orientation of a particular pose. The success rate for ASL character recognition is about 95% but it is still difficult to distinguish all characters, only using static pose classification method. For example, character I is very similar to J and X is very similar to Z considering static shape, but there are a transitions from I to J and X to Z (dynamic gesture). For correct recognition of such ASL characters, we need to analyze the motion also. ASL sign words (vocabulary) recognition method requires tracking of hand motion as well as the position with respect to other important parts of the body such as the head, chest and shoulders. Vision-based sign language recognition is still difficult and it is the future aim of gesture recognition society.

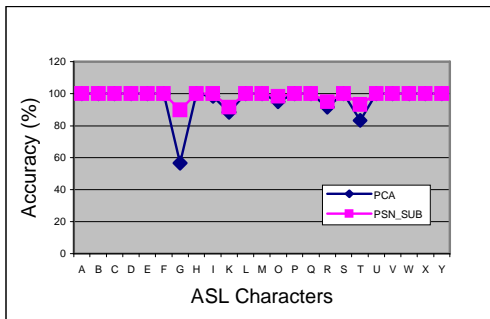


Fig. 14 Comparison of PCA and person-specific Subspace methods for ASL character classification (person_8 "Satomi")

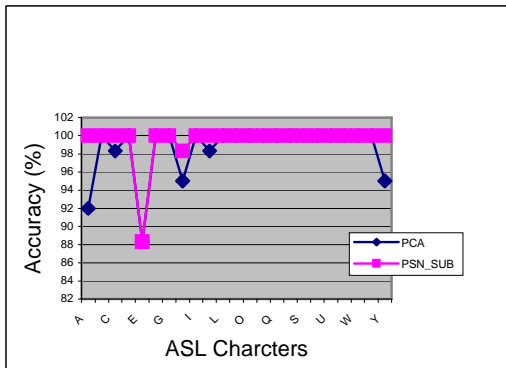


Fig. 15 Comparison of PCA and person-specific Subspace methods for ASL character classification (person_9 "Yoshi")

6 Conclusions

This paper describes person specific subspace method for ASL character recognition and compares it performance with PCA method. In addition, with ASL character recognition, this system is also capable of identifying persons. To utilize person-centric knowledge, this system integrates person identification method with ASL recognition system. From the experimental results we have concluded that, the classification accuracy of person-specific subspace method is better than general PCA method and person-specific subspace method is faster than the general PCA method.

This system uses a skin-color based hand poses segmentation method, which is still suffering from the variation of lighting condition and background color or human sleeve's (shirt's) color. The system assumes that the background contains non-skin color and a user wears non-skin color sleeves. If we use infrared camera then it is possible to overcome this problem using minor modification in segmentation module other modules will remain the same. Since the skin reflects near IR light well, active IR sources placed in proximity to the camera in combination with IR pass filter on the lens makes it easy to locate hand those are within the range of light sources.

Considering the reduction of processing time, so far eigenvectors calculations are performed separately in off-line. The eigenvectors do not change during dynamic learning process. The user has to initiate this calculation function to change the eigenvectors or principal components. In the future, if faster CPUs are available, these components are then possible to be integrated into on-line learning function. We could not claim that our system is more robust against new lighting condition and clutter background. Our hope is to make this sign language recognition system more robust and capable to recognize dynamic gestures for interaction with different intelligent machine.

References

[1] "American Sign Language Browser"
<http://commtechlab.msu.edu/sites/aslweb/browser.htm>, visited on April 2004.

- [2] M. Hasanuzzaman, T. Zhang, V. Ampornaramveth, and H. Ueno, "Gesture-Based Human-Robot Interaction Using a Knowledge-Based Software Platform", *International Journal of Industrial Robot*, Vol. 33, No. 1, 2006, pp. 37-49.
- [3] M. Hasanuzzaman, "Vision and Knowledge-Based Gesture Recognition for Human-Robot Interaction", Ph.D. thesis, NII, Tokyo, Japan, March 2006.
- [4] V. I. Pavlovic, R. Sharma and T. S. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review", *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 19, No. 7, 1997, pp. 677-695.
- [5] R. E. Axtell, *Gestures: The Do's and Taboos of Hosting International Visitors*, John Wiley & Sons, 1990.
- [6] M.B. Waldron, and S. Kim, "Isolated ASL Sign Recognition System for Deaf Persons", *IEEE Transactions on Rehabilitation Engineering*, Vol. 3, No. 3, 1995, pp. 261-271.
- [7] Holden, E.-J., R. Owens, and G. Roy, "Adaptive fuzzy Expert System for Sign Recognition", in *Proceedings of the International Conference on Signal and Image Processing (SIP'2000)*, Las Vegas, USA, 2000, pp. 141-146.
- [8] Vamplew, P. and A. Adams, "Recognition of Sign Language Gestures using Neural Networks", *Australian Journal of Intelligent Information Processing Systems*, Vol. 5, No. 2, 1998, pp. 94-102.
- [9] H. Birk, T. B. Moeslund, and C. B. Madsen, "Real-time Recognition of Hand Alphabet Gesture Using Principal Component Analysis", in *Proceeding of 10th Scandinavian Conference on Image Analysis*, Finland, 1997.
- [10] Vogler, C. and D. Metaxas, "Adapting Hidden Markov Models for ASL Recognition by Using three-Dimensional Computer Vision Methods", in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics SMC97*, IEEE Computer Society: Orlando, Florida, 1997, pp. 156-161.
- [11] M. A. Bhuiyan, V. Ampornaramveth, S. Muto, and H. Ueno, "On Tracking of Eye For Human-Robot Interface", *International Journal of Robotics and Automation*, Vol. 19, No. 1, 2004, pp. 42-54.
- [12] M. A. Bhuiyan, V. Ampornaramveth, S. Muto, H. Ueno, *Face Detection and Facial Feature Localization for Human-machine Interface*, *NII Journal*, Vol. 5, No. 1, 2003, pp. 25-39.
- [13] M. Hasanuzzaman, T. Zhang, V. Ampornaramveth, H. Gotoda, Y. Shirai, H. Ueno, "Knowledge-based Person-centric Human-Robot Interaction by Means of Gestures", *Information Technology Journal*, Vol. 4, No. 4, 2005, pp. 496-507.
- [14] M. Hasanuzzaman, T. Zhang, V. Ampornaramveth, M.A. Bhuiyan, Y. Shirai, H. Ueno, "Face and Gesture Recognition Using Subspace Method for Human-Robot Interaction", in *Proceedings of the PCM'2004 (5th Pacific Rim Conference on Multimedia)*, Tokyo, Japan, LNCS, Vol. 3331, No. 1, Springer-Verlag Berlin Heidelberg, 2004, pp. 369-376.
- [15] M. H. Yang, D. J. Kriegman and N. Ahuja, "Detection Faces in Images: A survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 24, No. 1, 2002, pp. 34-58.0
- [16] M. Turk and A. Pentland, "Eigenface for Recognition", *Journal of Cognitive Neuroscience*, Vol. 3 (1), 1991, pp. 71-86.
- [17] M. Hasanuzzaman, V. Ampornaramveth, T. Zhang, M.A. Bhuiyan, Y. Shirai, H. Ueno, "Real-time Vision-based Gesture Recognition for Human-Robot Interaction", in *Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO)*, China, 2004, pp. 379-384.