International Journal of Electronics, Electrical and Computational System
IJEECS
ISSN 2348-117X
Volume 6, Issue 7
July 2017

ISL Gesture Matching Using Distance Ratio and Rotation Angle.

Sandeep B. Patil[1], Rajesh H. Talwekar[2]
[1]PhD Scholar (Electronics & Telecommunication),
Faculty of Engineering and Technology of Shri Shankaracharya Technical Campus,
Chhattisgarh Swami Vivekanand Technical University, Bhilai, INDIA
[2]Professor (Electronics & Telecommunication)
Government Engineering College, Raipur, Chhattisgarh, INDIA.

Abstract:

Objective: *The matching of two images having different angle and resolution is a fundamental problem in computer vision. Hand gestures used for sign language may have different angle of rotation. Indian Sign Language gestures were used in this paper as they are not standard and may differ from region to region.*

Methods: *This paper addresses the use of SIFT algorithm for the matching of two ISL images using various distance ratio and angle of rotation. The scale invariant feature transform has been used because it extracts the features which are invariant to translation, rotation and scaling. The experimental result shows that the number of matching point increases by incrementing distance ratio for any particular value of angle of rotation.*

Findings: *It has been observed from literature review that no standard database is available for Indian sign Language (ISL) and ISL is varying from region to region. We prepared standard database for ISL and performs reliable matching between two ISL by varying validity ratio and angle of rotation. The accuracies of key point matching is increasing by varying validity ratio.*

Novelty: As *SIFT algorithm has been developed by LOWE and used for image matching. The experiments performed by LOWE and other researchers based on image matching using different scales with fixed validity ratio. This paper describes the image matching by varying validity ratio and rotation angle by 45^0 and 90^0.The accuracies has been improved for both the cases.*

Key points: *Indian sign language (ISL), Gesture, scale invariant feature transform, distance ratio, angle of rotation.*

Introduction:

The problem of performing the reliable matching between two images has been an emerging topic of research in computer vision and biometrics. In recent years, there have been numerous research contributions in the field of Indian Sign Language (ISL) which helps in identification of persons based on their traits or characteristics [1],[5],[7]. Arulkarthick et al. (2012) presented a real-time system for hand gestures recognition. Initially hands postures were detected; hand gestures are then recognized based on features of hand postures that are extracted using the Haar transform. K-means clustering algorithm was also used to reduce the computational complexity. The suggested method is limited for small database only [6]. Alon et al. (2009) introduced a framework that performed spatial segmentation, temporal segmentation, and recognition. The system recognized a gesture even when the hand location is highly ambiguous and gesture's begining and ending information are unavailable. The performance of the approach was evaluated on two challenging applications: recognition of hand-signed digits gestured by users wearing short-sleeved shirts, and retrieval of occurrences of signs of interest from a video database containing continuous, unsegmented signing of American Sign Language (ASL). This system is limited to five signs of ASL. The algorithm has to undergo through three major components that is feature extraction, spatio- temperol matching and intermodal competition that requires large processing time[7]. Nguyen et al. (2012) utilized the facial expressions where

Academic Science

International Journal of Electronics, Electrical and Computational System
IJEECS
ISSN 2348-117X
Volume 6, Issue 7
July 2017

features are tracked from the signers face using probabilistic principal component analysis (PPCA). A test was conducted to recognize six isolated facial expressions representing grammatical markers in American Sign Language (ASL). The recognition accuracy reported for ASL facial expressions was 91.76% in person dependent tests and 87.71% in person independent tests. The proposed system is not fully automatic and subject to tilt/rotate their head while thrusting the head forward, contributing to the lower accuracy for these two classes [8]. Paulraj et al. (2011) proposed a sign language recognition system that helps the hearing impaired to communicate more fluently with the normal people. Sign language recognition system employs skin color segmentation and neural network method for training. A segmentation process is carried out to separate the right and left hand regions from the image frame and a simple vertical interleaving method has been used in the preprocessing stage to reduce the size of the image. The system requires about 3218 average mean epochs to train the network model that exceeds the training time. Moreover there was confusion in first, fourth, eighth and ninth sign that greatly reduced the accuracy [9].

This area is not very easy because no standard databases are available for the Indian Sign Language (ISL). The significance of the problem can be easily illustrated by using natural gestures applied together with verbal and nonverbal communication. The use of hand gestures in support of verbal communication is very useful for the people having no visual contact.

The basic principle applied on gesture recognition is direct correlation of the images consuming ample amount of time, as the size of images are too large and the number of comparisons are many. To overcome the problem of time consumption, a new approach of Feature extraction followed by pattern recognition is used. In this paper scale invariant feature transform is used to extract the features from the image.

To perform the reliable matching between two images, scale invariant feature transform has widely been used [2]-[4], the SIFT features give the highest matching accuracy for an affine transformation of various degrees. SIFT descriptors perform better than other local descriptors for various scale change and image rotation. In this paper, we emphasis is on image rotation by various angles at various distance ratios and shows the reliable matching of two different images based on this rotation.

2 ISL Gestures

Indian Sign Language (ISL) using hand gesture shown in figure 1. The static image shown in figure 1 is the hand gesture images of Indian sign language from gesture A to Z. No standard data base is available for Indian sign language, this force us to prepare the data base based on the image shown in figure 1. Based on this image the hand gesture databases from 10 different persons were collected. The database consist of total 260 images out of which 78 images were used for testing purpose and remaining 182 images were used for training. The database consists of 26 images for each class. Figure 2 shows the database of one of the class. Each image was taken using the digital cameras having a resolution of 284 x 215 pixels.

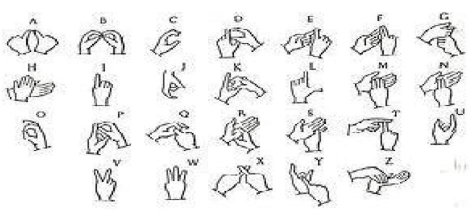


Figure 1. Indian sign language

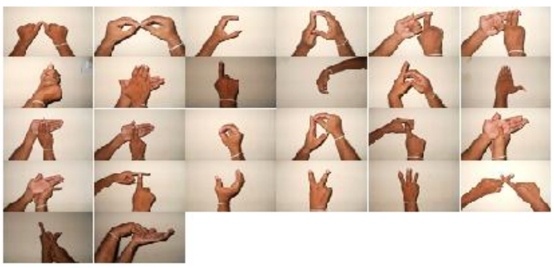


Figure 2. 26 ISL gesture from class 1.

3. Methodology

To perform the reliable matching between two images having two sets of point in space, point pattern matching algorithm is always used. The extracted features by point pattern matching algorithm are pixel

subset of the original image which provides a novel approach to achieve a matching of highest quality in an efficient manner[2]-[4]. The flow chart of proposed algorithm is shown in figure 3 and follows the following steps.

- *Consider two images for matching with different angles.*
- *Manually Initialize the distRatio = 0.65, 0.75 and 0.85*
- *Run the SIFT match algorithm*
- *Extract SIFT key points of two images for various values of distance ratio.*
- *Shows match point for two images.*

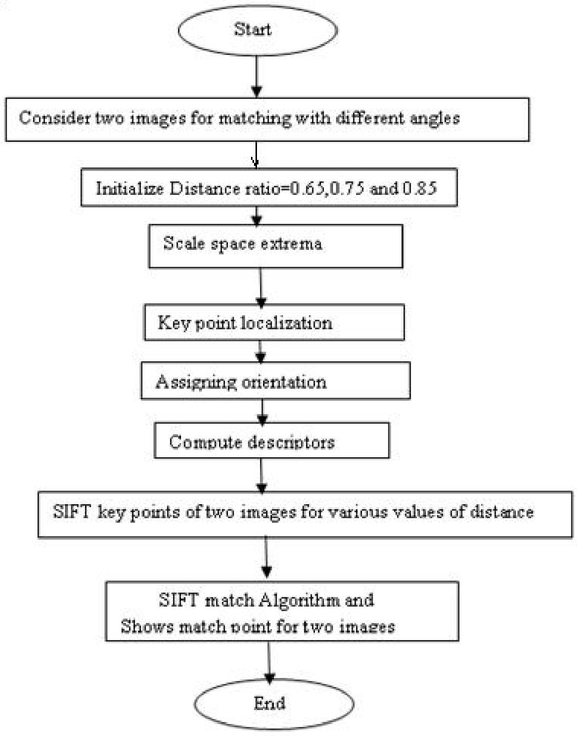


Figure 3. Flow chart for point pattern matching algorithm

The discussion of the data flows in each phase of SIFT algorithm is as follows. The input to a SIFT algorithm is a set of N^2 pixels of an NxN image. Only a small fraction of these pixels typically turn out to be extrema. Let $0 < \alpha < 1$, be this fraction. So αN^2 extrema will move on to the next key point detection phase. Only the small fraction of these extrema will qualify as a key point. Let $0 < \beta <1$, be this fraction. So nominally there is αβN2 key point at this stage. Orientation assignment will re-examine all the N^2 points in the image to check if any point of significant magnitude have been missed. If so they are added to the set of key point. Let the fraction γ of the image pixel qualify to be these added key point. The compute descriptor phase converts these points into vector which are then turned into features. The number of feature descriptor output by SIFT algorithm is nominally $(\propto \beta + \gamma)N^2$ for an N x N image. This process is repeated for original image and rotated image (45^0 and 90^0). The SIFT keypoints are extracted for various values of distance ratio. Once the SIFT keypoints are obtained, SIFT match algorithm (MK-Rod) is applied to show matched keypoints between two images by joining lines. The above process is briefly described below.

3.1 Scale Space Extrema Detection:

This is the first stage where key points are detected. To achieve this, the input image is convolved with Gaussian filters at different scales, the difference of successive Gaussian blurred images are further producing a difference of Gaussian filters (DoG). Thus once the DoG have been obtained, the keypoints are identified as local minima/maxima of the DoG images across various scales. The candidate keypoint selection is done by comparing each pixel in the DoG images to its eight neighbour at the same scale and nine corresponding

neighbouring pixel in each of the neighbouring scale. The noted pixel value is maximum or minimum among all compared pixel, the selected pixel is called candidate keypoint [3],[4].

3.2 Keypoint Localization:

As discussed above, the first stage detect too many keypoints and very few of them are stable. This stage is used to perform a detailed fit to the surrounding data for accurate localization over scale and ratio of principal of curvature. This extracted information is used to reject the keypoints which have low contrast or poorely localized along an edge.

The interpolation of surrounding data is done by using the quadratic Taylor's expansion of difference of Gaussian scale function [1]-[2]. In order to eliminate the low contrast keypoints, the Taylor's expansion is computed for an offset value less than 0.03.

3.3 Orientation Assignment:

In this stage, the stable keypoints is assigned one or more orientations based on the local image gradient directions. These orientations of the keypoints descriptors achieve invariance to image rotation [1],[2],[3],[4],[5].

3.4 Key Point Descriptor:

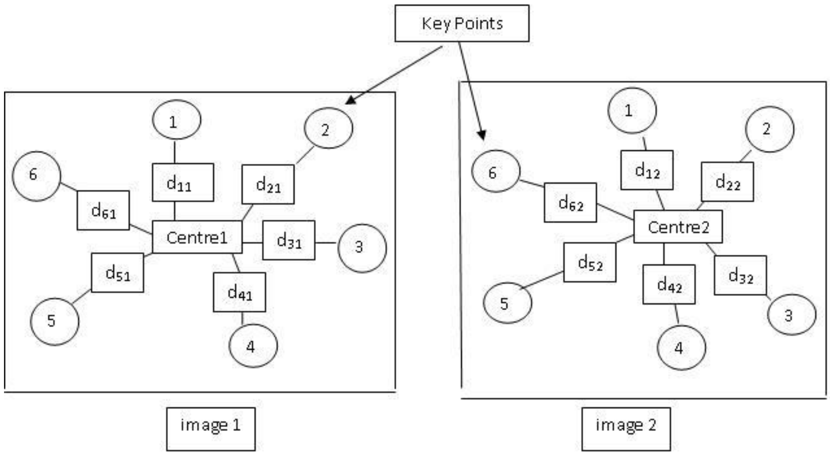
In the previous stages, the keypoint location were found at particular scales and assigned orientations. This final stage computes descriptor vector for these keypoints in such a way that the descriptor are highly distinctive and partially invariant to remaining variations [1],[2],[3],[4],[5].

3.5 SIFT Match algorithm (MK-RoD Algorithm)

This algorithm reads two images, finds their SIFT features, and displays lines connecting the matched key points. A distance ratio is the ratio of vector angles from the nearest to second nearest neighbor. A match is accepted only if its distance is less than distance Ratio times the distance to the second closest match. It returns the number of matches displayed. For efficiency, it is cheaper to compute dot products between unit vectors rather than Euclidean distances. The process of matching using MK-RoD algorithm[8]-[10] is shown in figure 4.

For matching purpose we use MK-Rod algorithm as follows

- *For each descriptor in the first image, select its match to second image.*
- *Precompute matrix transpose*
- *Computes vector of dot products*
- *Take inverse cosine and sort results*
- *Check if nearest neighbor has angle less than distRatio times 2nd.*
- *Show a figure with lines joining the accepted matches.*



(a) (b)

Figure 4. (a) Original image with keypoints (b) Rotated image with keypoints

International Journal of Electronics, Electrical and Computational System
IJEECS
ISSN 2348-117X
Volume 6, Issue 7
July 2017

Step 1: Let image 1 is the original image and image 2 is the rotated (45^0 and 90^0) image as shown in Fig. 4. Assuming further, let d11 to d61 are the distance factors of key point from the center of image 1; and d12 to d62 are the distance factors of key point from the centre of image 2.

Step 2: If the numbers of matched key points are greater than 2 then calculate the distance of the matched key point to the centre of the key points using the formula given below.

$$dT_1 = \sum_{i=1}^{M} d_{i1} \qquad (1)$$

$$dT_2 = \sum_{i=1}^{M} d_{i2} \qquad (2)$$

where dT_1 a dT_2 are the summation of all distance for image 1 and image 2 respectively; M denotes the number of matched points.

Step 3: The distance array ratio (DAR) is defined as:

DAR = (distance between the matched key point and centre of the key point/total distance)

The DAR for original image and rotated image is calculated as:

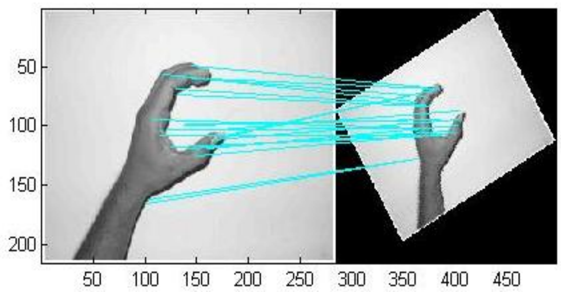
$$D\quad 1 = \frac{d_1}{d_{T1}}\frac{d_2}{d_{T1}}\frac{d_3}{d_{T1}}\frac{d_4}{d_{T1}}\ldots\ldots\frac{d_{M1}}{d_{T1}} \qquad (3)$$

$$D\quad 2 = \frac{d_2}{d_{T2}}\frac{d_3}{d_{T2}}\frac{d_4}{d_{T2}}\ldots\ldots\frac{d_{M2}}{d_{T2}} \qquad (4)$$

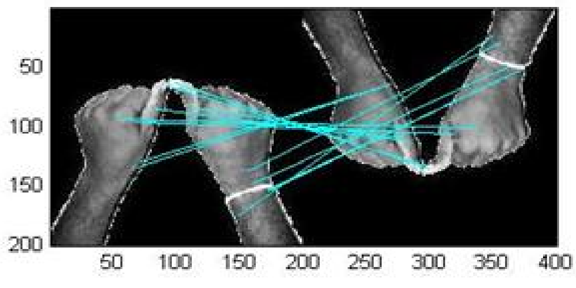
Step 4: The distance masking is necessary in order to determine the similar pattern of matched key point and the centre of the matched key point. The distance masking can be calculated by taking the absolute value of DAR below the threshold value.

Distance mask = $\text{abs}[D\quad 1 - D\quad 2] < Thr\ ho\ v$ (5)

The threshold value we choose here is 0.05, and it is incremented by 0.03 depending on the distance ratio.Figure 5 (a) shows the example of matching between original and rotated (45^0) images. As we increase the distance ratio gradually from 6.5 to 7.5 the keypoint matching will be increase. If we further increase the distance ratio from 7.5 to 8.5 will result in improving the matching performance. Figure 5 (b) shows the matching between original and rotated (90^0) image.



(a)



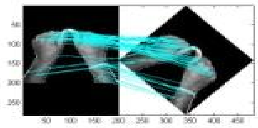
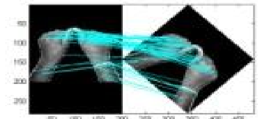
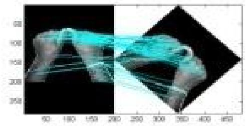
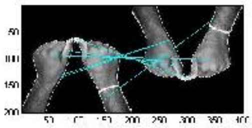
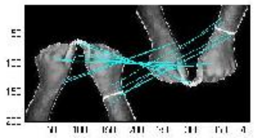
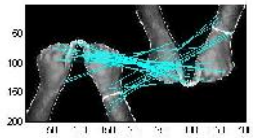
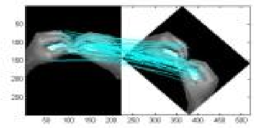
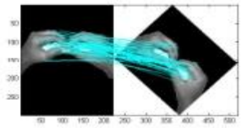
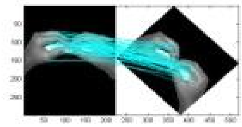
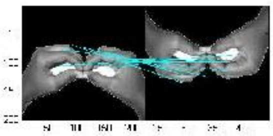
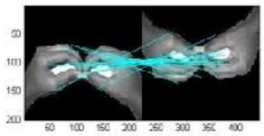
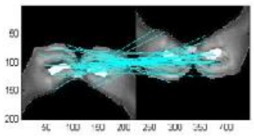
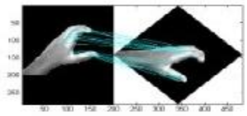
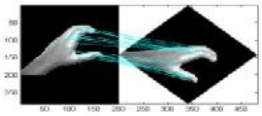
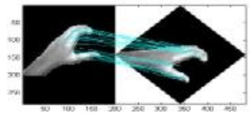
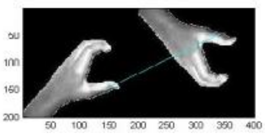
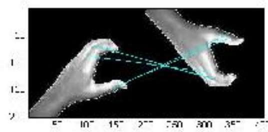
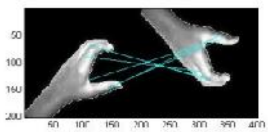
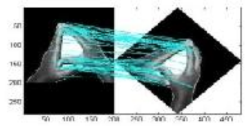
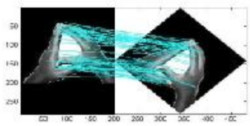
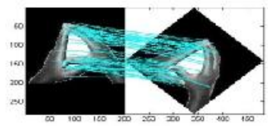
(b)

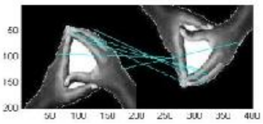
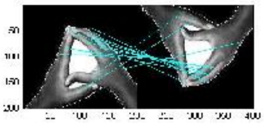
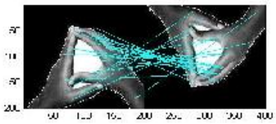
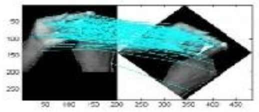
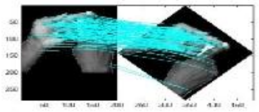
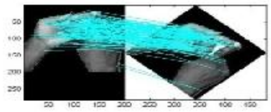
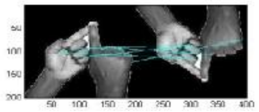
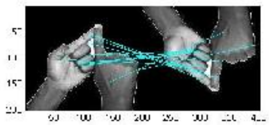
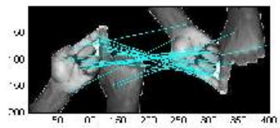
Figure 5. (a) Keypoint matching with 45^0 image rotation (b) Keypoint matching with 90^0 image rotation

4. Result and Discussion

Table 1 shows the result of matching for various distance ratio and rotation angle. We have considered five Indian sign language gestures for reliable matching purpose. For Gesture A if we rotate it with 45 degree angle keeping distance ratio is 6.5, it is found that 64 keypoints are correctly matched between two images. The same numbers of key points are matched with distance ratio 7.5. If we further increase the distance ratio to 8.5 the number of keypoints matching is 69. Similarly if we rotate the gesture A by 90 degree with fixed distance ratio of 6.5, it has been observed that 7 key points are correctly matched between two images. Further increase in distance ration will result in improving the matching performance.

Table 1: Result of image matching with various distance ratio and angle of rotations.

ISL Gesture	*Angle of rotation*	*Distance ratio=6.5*	*Distance ratio=7.5*	*Distance ratio=8.5*
A 	45^0	 Matched key point(64)	 Matched key point (64)	 Matched key point (69)
	90^0	 Matched key point (7)	 Matched key point (14)	 Matched key point (33)
B 	45^0	 Matched key point (59)	 Matched key point (66)	 Matched key point (74)
	90^0	 Matched key point (12)	 Matched key point (24)	 Matched key point (42)
C 	45^0	 Matched key point (16)	 Matched key point (18)	 Matched key point (21)
	90^0	 Matched key point (1)	 Matched key point (5)	 Matched key point (6)
D 	45^0	 Matched key point (53)	 Matched key point (57)	 Matched key point (64)

	90^0	 Matched key point (8)	 Matched key point (15)	 Matched key point (26)
E 	45^0	 Matched key point (112)	 Matched key point (119)	 Matched key point (127)
	90^0	 Matched key point (6)	 Matched key point (13)	 Matched key point (24)

Conclusion:

Image matching is the fundamental problem in object recognition. This paper addresses the matching of two Indian sign language gestures having various distance ratio and angle. Initially for a fixed value of distance ratio and angle, the system extracts the matching keypoints between two images. Further, manually incrementing the distance ratio by keeping angle fixed, the keypoint matching between two images has been increased. The SIFT algorithm is used to extract features because it gives stable features over translation, rotation and scaling. The feature matching is performed using MK-RoD algorithm.

References:

[1] G.R. Sinha ans Sandeep B. Patil, “Biometric: Concept and Applications” Wiley India Pvt. Ltd. ISBN-13: 9788126538652, ISBN-10:8126538651, 2013.

[2] Sandeep B. Patil and G.R. Sinha, 2015, “Intensity Based distinctive feature extraction and matching using Scale invariant feature transform for Indian Sign Language” 17th International Conference on Mathematical Methods, Computational Techniques and Intelligent Systems (MAMECTIS15) held in Tenerife, Canary Islands, Spain, January 10-12, 2015. Pp-245-255. The ISI Journals (with Impact Factor from Thomson Reuters), ISBN: 978-1-61804-281-1

[3] David G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, 60, 2 (2004), pp. 91-110,3-11-2011. http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf.

[4] David G. Lowe, "Object recognition from local scale-invariant features," International Conference on Computer Vision, Corfu, Greece (September 1999), pp. 1150-1157 r : http://www.cs.ubc.ca/~lowe/papers/iccv99.pdf

[5] Lindeberg, Tony, “ Feature Detection with Automatic Scale Selection” International Journal of Computer Vision, vol 30, number 2, pp 77--116, 1998

[6] Arulkarthick V.J, Sangeetha D. and Umamaheswari S, 2012, Sign Language Recognition using K-Means Clustered Haar-Like Features and a Stochastic Context Free Grammar, European Journal of Scientific Research, 78(1):74-84.

[7] Alon J. and Athitsos V, 2009, A Unified Framework for Gesture Recognition and Spatiotemporal Gesture Segmentation, IEEE Transactions on Pattern analysis and Medical Intelligence, 31(9):1685-1699.

[8] Nguyen T.D and Ranganath S, 2012, Facial expressions in American sign language: Tracking and recognition, Pattern Recognition: 1877–1891.

[9] Paulraj M P, Palaniappan R, Yaacob S, Zanar A, 2011, A Phoneme Based Sign Language Recognition System using 2D Moment Invariant Interleaving feature and Neural Network, IEEE Student Conference on Research and Development, 2(11):111-116.

[10] Ulrike Zeshan, Madan M. Vasishta and Meher Sethna, "Implementation of Indian Sign Language in Educational settings" *Asia Pacific Disability Rehabilitation Journal*,Vol.16, No.1, pp.16-40, 2005

[11] Bhosekar A, Kadam K, Ganu R. & Joshi S.D, 2012, American Sign Language Interpreter, IEEE Fourth International Conference on Technology for Education, 6(12): 157-159

[12] Bhuyan M.K, Kar M.K & Debanga R.N, 2011, Hand Pose Identification from Monocular Image for Sign Language Recognition, IEEE International Conference on Signal and Image Processing Applications (ICSIPA2011), 3(12):378-383

[13] Bharti P, Kumar D and Singh S, 2012, Sign Language to Number by Neural Network, International Journal of Computer Applications, 40(10): 38-45.

[14] Balakrishnan G, Rajam P.S, 2011, Real Time Indian Sign Language Recognition System to aid Deaf-dumb People, IEEE International Conference on Computing Communication and Networking Technologies (ICCCNT) :737-742**.**

[15] Chiu Y.H, Cheng C.J, Su H.W, Wu C.H, and 2007, Joint Optimization of Word Alignment and Epenthesis Generation for Chinese to Taiwanese Sign Synthesis, IEEE Transaction on Pattern Analysis and Machine Intelligence, 29(1):28-39.

[16] Chakraborty P, Mondal S, Nandy A, Prasad J.S , 2010, Recognizing & Interpreting Indian Sign Language Gesture for Human Robot Interaction, Int'l Conf. on Computer & Communication Technology |ICCCT'10|,2(10), 52(11): 712-717.

[17] Chen X and Zhou Y, 2010, Adaptive Sign language recognition with exemplar extraction and MAP/IVFS, IEEE signal processing letter, 17(3):297-300

[18] Chung H. W, Chiu Y. H and Chi-Shiang Guo, 2004, Text Generation From Taiwanese Sign Language Using a PST-Based Language Model for Augmentative Communication, IEEE Transcation On Neural System And Rehabilitation Engineering, 12(4):441-454.

[19] Debevc M, Kosec P & Rotovnik M, 2009, Accessible Multimodal web pages with Sign Language Translations for Deaf and Hard of Hearing Users, 20th International Workshop on Database and Expert Systems Application 3(9): 279-283.

[20] Dasgupta T, Basu A, Mitra P, 2009, English to Indian sign language machine translation system: A structure transfer framework, international conference on artificial intelligence (IICAI):118-124.

[21] Ghotkar A.S, Hadap M, Khatal R, Khupase S, 2012, Hand Gesture Recognition for Indian Sign Language, International Conference on Computer Communication and Informatics (*ICCCI* -2012), Coimbatore, INDIA, 2(4):9-12.

[22] Gaolin F ,Wen G and Zhao D, 2007, Large-Vocabulary Continuous Sign Language Recognition Based on Transition-Movement Models, IEEE Transaction On System, man, And Cybernetics, 37(1):1-9.

[23] Habili N and Moini A, 2004, Segmentation of the Face and Hands in Sign Language Video Sequences Using Color and Motion Cues, IEEE Transaction On circuit and System for Video Technology, 14(8):1086-1097.

[24] Huang Q, Kong D, Li J, Sun Y, Wang L, Zhang X, 2012, Synthesis of Chinese Sign Language Prosody Based on Head, International Conference on Computer Science and Service System, 1(8):1860-1864.

[25] Ittisam P, Toadithep N, 2010, 3D Animation Editor and Display Sign Language System Case Study: Thai Sign Language, 2(3):633-637.

[26] Patil S.B and Sinha G.R, 2012, Real Time Handwritten Marathi Numerals Recognition Using Neural Network, International Journal of Information Technology and Computer Science, 4(12): 76-81.

[27] Patil S.B, Sinha G.R and Thakur K, 2012, Isolated Handwritten Devnagri Character Recognition Using Fourier Descriptor and HMM in International Journal of Pure and Applied Sciences and Technology (IJPAST), 8(1): 69-74.

[28] Patil S.B. and Sinha G.R, 2012, Off-line mixed Devnagri numeral recognition using Artificial Neural Network, Advance in computational Research, Bioinfo journal, 4(1): 38-41.

[29] Patil S.B, Sinha G.R and Patil V.S, 2011, Isolated Handwritten Devnagri Numerals Recogntion Using HMM, IEEE Second International conference on Emerging Applications of Information technology, Kolkota, ISBN: 978-1-4244-9683-9:185-189.