Point Pattern Matching Algorithm for Recognition of 36 ASL Gestures

Deval G. Patel

Abstract*— Hand gesture recognition is a way to create a useful, highly adaptive interface between machines and their users. The recognition of gestures is difficult because gestures exhibit human variability. Sign languages are used for communication and interface . There are various types of systems and methods available for sign languages recognition. Our approach is robust and efficient for static hand gesture recognition. The main objective of this paper is to propose a system which is able to recognize 36 static hand gestures of American Sign Language (ASL) for letter A- Z and digits 0-9 successfully and also it is able to perform the classification on static images correctly in real time. We proposed a novel method of pattern recognition to recognize symbols of the ASL based on the features extracted by SIFT algorithm and its performance is compared it with widely used methods such as PCA and Template Matching.

Index Terms*— ASL, Hand Gesture Recognition System, PCA, Point Pattern Matching Algorithm.

I. INTRODUCTION

Gestures are expressive, meaningful body motions. Interpretation of human gestures such as hand movements or facial expressions, using mathematical algorithms is done using gesture recognition. Gesture recognition is also important for developing alternative human-computer interaction modalities [1]. It enables human to interface with machine in a more natural way.

Some commercial products using gesture recognition are already available, but there is much potential for the technology to improve. Our claim is that machines can also be programmed to interpret gestures made in any frame of reference by automatically setting the coordinate system as performed by human brains. This would enable the machine to better understand and interpret gestures, somewhat like the human brain does [2].

Sign language gestures are considered as the most structured sets of gesture. In sign language, each gesture has an assigned meaning (or meanings). Our focus is on American Sign Language (ASL). ASL is the language of choice for most deaf people. The main factor behind the invention of ASL is to allow deaf people communicate with normal people and also with each other. ASL consists of approximately 6000 gestures of common words with finger spelling which are use to communicate proper nouns. Finger spelling can be performs by one hand and 36 gestures to communicate the 26 letters of the alphabets and 10 digits.

Manuscript received on June, 2013.

Deval G. Patel, Computer Engineering, B.V.M. Engineering College, India.

There are various types of systems and methods available for sign languages recognition. Our approach is robust and efficient for static gesture recognition and translation of ASL. We used pattern recognition system to recognize hand gesture for ASL, because it is fast and simple algorithms.

Pattern recognition is the assignment of some output value to a given input value, according to the key-points obtained y SIFT algorithm. Proposed Point pattern matching algorithm finds the matching points between test and target images. By using this approach we can recognize gestures with open as well as closed fingers more effectively.

II. RELATED WORK

There are several methods developed for Sign recognition for various languages. Specifically, [4] used finger tip detection. In [4] edge detection algorithm (Canny edge operator) and boundary tracing are used. Hand gestures are recognized automatically using the data such as the shape and the kinematics of the compressed arm trajectories [4]. [3] used features like palm region, centre and centre of gravity as an input to Self-Growing and Self-Organized Neural Gas algorithm for gesture recognition. The hand is detected using attributes like its motion and the skin colour [5]. Hand shape estimation under complex backgrounds is done by adding the models having only the position and velocity of the hand [6]. The image of the hand gesture is captured and converted into feature vectors [7]. The hand gesture input is taken with the help of a data glove and artificial neural networks are used to recognize the gesture [8]. Hand gestures are represented in terms of hierarchies of multi scale colour images [9]. In some systems more than one feature extraction methods and neural networks are implemented to recognize the gestures made by hand [10].

III. PROPOSED SYSTEM

Hand gesture recognition system recognizes 36 static ASL hand gestures in real time. It takes hand gesture images and performs action based on the category of pattern. Key-points of the image (i.e. the points lying on the high contrast region) are calculated and matched by point pattern matching algorithm. After analyzing the raw data based on a certain purpose and method, we can do actions such as classification [11].

Here, Hand gesture recognition system consisting of point pattern matching and SIFT is used to recognize hand gesture for ASL signs, because it is fast and simple approach. Point pattern matching is the assignment of some sort of output value to a given input value, according to some specific algorithm. An example of recognition is classification, which attempts to assign each input value to one of a given set of classes. It looks for exact matches of the input image features with pre-existing class features(patterns). Fig. 1 shows how the Hand Gesture Recognition system works.

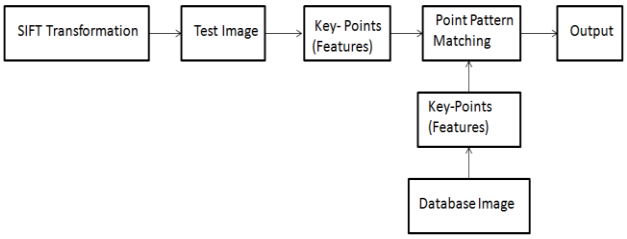


Figure 1. Hand Gesture Recognition System with Point Pattern Matching .

As shown in Fig. 1, first we have to take an image from webcam or from database. These images will go through a SIFT Transformation. This transformation will find key-points (Features) of that particular image into a feature vector, which will be then compared with other feature vectors of a Database images of ASL hand gestures. For comparison we use Point Pattern Matching Algorithm. Hand Gesture Recognition system is developed by using SIFT algorithm with point pattern matching algorithm to perform hand gesture recognition of 36 ASL signs.

A. Point Pattern Matching Algorithm

Hand Gesture recognition systems often require matching two sets of points in space. This is because the analysed images are raster graphics or the extracted features are pixels subset of the original image. For this point pattern matching algorithm is the best solution. Point pattern matching algorithm provides a navel approach to achieve a matching of adequate quality in a efficient and robust manner. For hand gesture recognition of ASL sign, we used point pattern matching with SHIFT match algorithm. The flowchart of proposed algorithm is as shown in following Fig. 2.

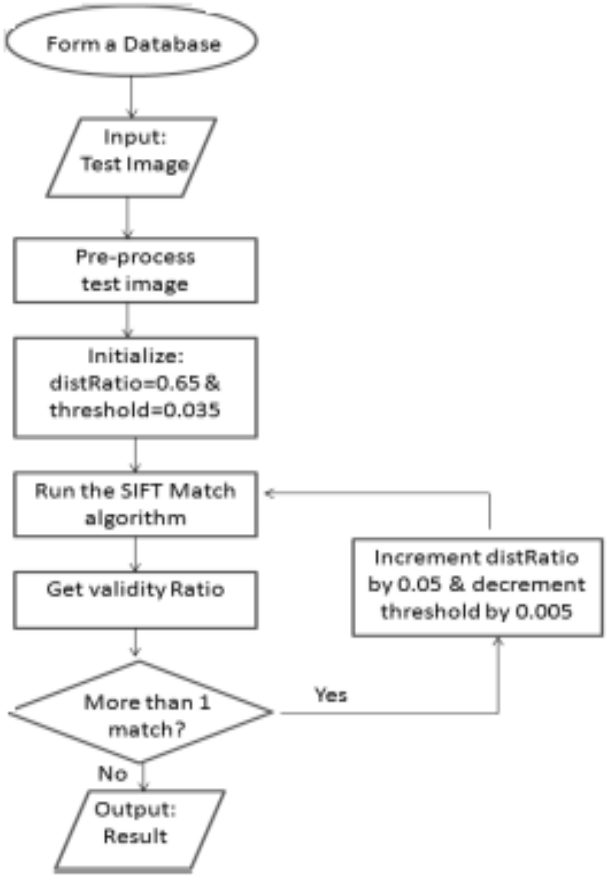


Figure 2. Flow Chart of Point Pattern Matching Algorithm

The working of point pattern matching algorithm is as follows:

1. Take a test image
2. Pre process the test image.
3. Initialize the distRatio = 0.65 and threshold= 0.035
4. Run the SHIFT match algorithm
5. Key point matching starts its execution by running the threshold. It gets the key point matched between test and all 36 trained images. We get the validity ratio.
6. Check that we got more than one result or not.
7. If we get more than1 result then increment the SHIFT distRatio by 0.05 and threshold by 0.005 and repeat the steps from 4 to 7.
8. If we get only one result then display the result.

IV. IMPLEMENTATION

A. Implementation of Point Pattern Matching Algorithm

During the test implementation of Hand Gesture Recognition System, the point pattern matching algorithm is executed. First test image is taken from the database or from the webcam. Then point pattern matching algorithm start its execution to find the matching key-points between test and train database images. After executing this algorithm it recognizes ASL input (query) images by comparing test image with all the database images and outputs the equivalent ASCII representation of it. This algorithm is implemented in two parts.

i) SIFT Algorithm

For any object in an image, points of interest of that object can be extracted to provide a "feature description" of the object. This description, extracted from a training image, can then be used to locate and identify the object in a test image containing many other objects. For accurate and reliable recognition, the features extracted from the training image must be detectable even under changes in image scale, noise and illumination. Also, the relative positions between them in the original image shouldn't change from one image to another. SIFT detects and uses a much larger number of features from the images, which reduces the contribution of the errors caused by these local variations in the average error of all feature matching errors.

The SIFT algorithm consists of following steps:

1. Constructing a scale space .
2. LoG Approximation
3. Finding key-points
4. Get rid of bad key points (Edges and low contrast regions)
5. Assigning an orientation to the key-points.
6. Generate SIFT features .

During the test implementation, the point pattern matching algorithm uses the SIFT algorithm to find the key-points of the images. These key-points are the Scale invariant features located near the high contrast regions of the image that can be used to distinguish them. In SIFT algorithm, image1 and image2 are taken as a two images to match. For our case the first image is one of the database images and image2 is the input (query) image. distRatio is the parameter of SIFT algorithm. In the original implementation, this parameter is set as a constant. For our algorithm's recursivity we made it a variable parameter and threshold is the threshold value for the MK-RoD algorithm.

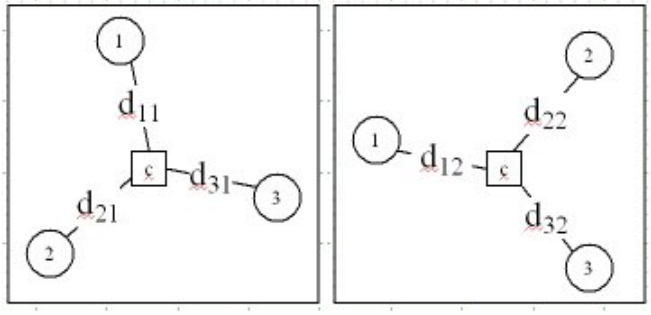
Here, for finding the shift key-points of an image the function sift is called which

finds the key-points with the combination of image description and location of given image. These terms are:

- Image(im): the image in double array format
- Descriptors (des): a K-by-128 matrix, where each row gives an invariant descriptor for one of the K key-points. The descriptor is a vector of 128 values normalized to unit length.
- Locs (loc): K-by-4 matrix, in which each row has the 4 values for a key-point location (row, column, scale, orientation). The orientation is in the range [-PI, PI] radians.

ii) MK-RoD Algorithm

For finding the validity ratio MK-RoD algorithm is used. For example, Fig. 3 shows the two images for finding the validity ratio.



(a) (b)

Figure 3 . Representation of (a) Trained Database Image (b) Test Input Image with Key points

C - Denotes the center points
D - Denotes the distance mask
T - Denotes the No. of test image to match
M - Denotes the No. of Matched Points 1, 2, 3 are the key points.

The procedure to find the Validity ratio of One Database Image versus Test Input Image is as follows:

(1) $d_{T1} = \sum_{i=1}^{M} d_{i1}$

(2) $d_{T2} = \sum_{i=1}^{M} d_{i2}$

(3) $Ratios\ 1 = \left[\frac{d_{11}}{d_{T1}} \frac{d_{21}}{d_{T1}} \frac{d_{31}}{d_{T1}}\right]$

(4) $Ratios\ 2 = \left[\frac{d_{12}}{d_{T2}} \frac{d_{22}}{d_{T2}} \frac{d_{32}}{d_{T2}}\right]$

(5) $Distance\ Mask = abs[Ratios\ 1 - Ratios\ 2] < (Threshold\ Value)$

(6) $Valid\ Points = sum(Distance\ Mask)$

(7) $Validity\ Ratio = \frac{No.of\ Valid\ Points}{No.of\ Matc\ hes\ Points}$

Once we got the validity ratio, mask the distances by taking the absolute which are below the algorithm's threshold. This operation is done in order to determine the similar pattern of the matched keypoints from the center of the matched keypoints. The absolute of the difference of the points which are below the given threshold are treated as valid matched keypoint.

B. Implementation of Other Methods

i) Template Matching

Template matching determines whether an obtained gesture image can be classified as a member of a set of stored gestures. Hand gesture recognition using template matching is done in two steps. The first is to create the templates by collecting data values for each gesture in the image set. Generally, each gesture is performed a number of times with a slight natural variations and the average of the data for each image is taken and stored as the template. The second part is to compare the obtained image with the given set of templates to find the gesture template most closely matching the current image.

An example of template matching comparisons is the use of distance measurements between the current gesture image and each of the image in the gesture image set recognizing the posture with the lowest distance measurement. The distance measurement must be below some threshold value to avoid false positive recognition. distance measurements used for hand posture template matching can be the sum of the absolute differences or sum of squares.

The mathematical formula for the sum of absolute differences is given below.

$$Diff(x_s, y_s, x_t, y_t) = |I_s(x_s, y_s) - I_t(x_t, y_t)|$$

$$SAD(x,y) = \sum_{i=0}^{T_{rows}} \sum_{j=0}^{T_{cols}} Diff(x+i, y+i, i, j)$$

where (x, y) pixel in template is represented as $T(x_t, y_t)$ and $S(x, y)$represents the search image.

ii) Principal Component Analysis (PCA)

Principal Component Analysis(PCA) computes and study the Eigenvectors of the different pictures and then express each image with its principal components (Eigenvectors).In this method a data set is created choosing a considerable no of images of good resolution for better recognition with smallest database. Then mean is subtracted from each of the data dimension. The next step is to calculate the covariance matrix of database of the principal eigenvectors Covariance matrix C of matrix A is calculated as

$$C = A * A'$$

Then, the eigenvectors and the Eigen values of C are the principal components of data set. Next, good components are chosen to feature vector. The principal (most important) eigenvectors with which data is expressed with the lowest information loss are chosen. finally, new data set is created(Eigen set). To compare the different pictures, each image of the data set should be expressed with these principal eigenvectors. Then compare by calculating the Euclidian distance between the coefficients that are before each eigenvector.

V. EXPERIMENTS AND RESULTS

For this proposed method implementation, Messy image database is used for train and test images. Images used are in '.jpg' file format. Fig. 4 shows the image database of 36 hand images. For experimental implementation of proposed method we take 15 ASL signs as an input query images. The input query images consists of images of gestures A, B, C, H, I, L, O, P, X, Y, 1, 3, 5, 6 and 9. For implementing this method, we used MATLAB programming. By using image processing toolbox we perform the SIFT transformation, feature vector processes on images. These feature vectors are matched by point pattern matching algorithm to find the image similar to the test image. Then proposed method is applied on input query images. The recognized result is then displayed as output.

The first experiment considers the complete data set of 36 ASL images 26 alphabets and 10 numbers. testing of one image of each type was performed which resulted into 77.7% recognition and 8.33% false rejection rate. The database used for the experiments is the Messy Database for Hand Gestures.

Figure 4. Image database of Trained Sets for ASL Signs

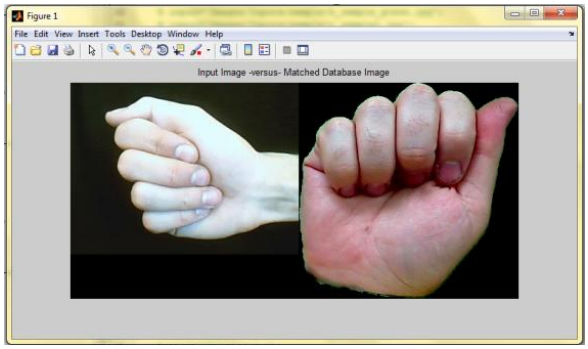
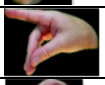


Figure 5. Point Pattern Matching algorithm output of test image A.

The plot of the result obtained by testing images of three sets with different size for the alphabet gestures A,B,C,H,I,L,O,P,X and Y and Numeric gestures 1,3,5,6and 9 is given in Fig 6. It shows that some gestures have higher recognition rates. The recognition rates are irrespective of the test data set size.

Table I : Results of HGR System with Point Pattern Matching.

Sr. No.	Image	Sign	Recognition Rate	False Recognition Rate
1		A	93.3%	6.6%
2		B	73.3%	6.6%
3		C	80%	0%
4		H	73.3%	0%
5		I	86.6%	6.6%
6		L	86.6%	0%
7		O	73.3%	0%
8		P	66.6%	0%
9		X	73.3%	0%
10		Y	86.6%	0%
11		1	80%	6.6%
12		3	80%	13.3%
13		5	100%	0%
14		6	66.66%	13.3%
15		9	60%	6.6%

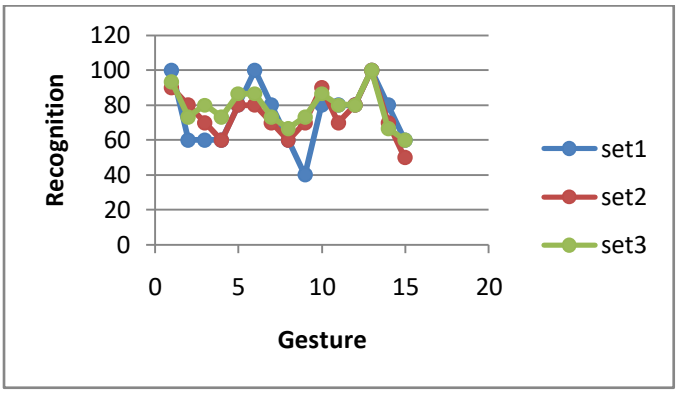


Figure 6. plot of gestures v/s their recognition rates with different sized test data set.

Template matching algorithm, Principal Component as well as point pattern matching algorithm performs HGR in real time for this dataset. Experiments were performed on a test set for all three methods. The graph comparing their recognition rates is given in Fig 7.

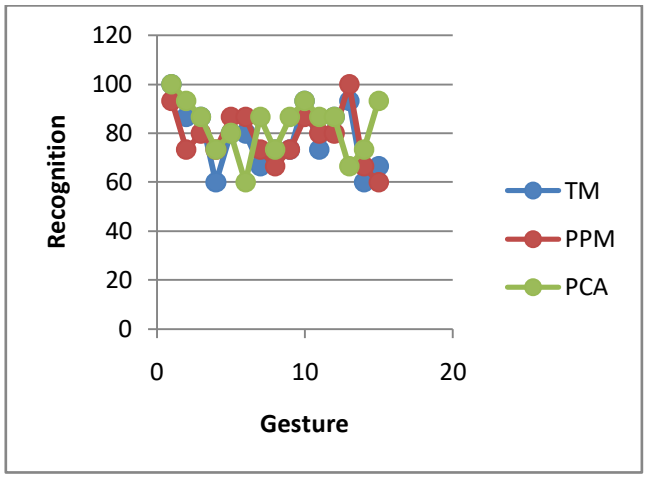


Figure 7. plot of recognition rates obtained for different gestures v/s recognition method used.

VI. CONCLUSION

Hand Gesture Recognition System recognizes all the static symbols of American Sign Language successfully by using point pattern matching algorithm with SIFT. This algorithm can recognize gestures having open as well as closed fingers more effectively. So it provides the simple and novel way to recognize the ASL signs. Hand gesture recognition system using point pattern matching algorithm and SIFT for feature extraction works in real time. Recognition rates of alphabet A (93.3%)and number 5(100%) are high compared to alphabet P and Numbers 6 and 9. It has higher recognition rate compared to template matching. For the Messy database PCA gives higher recognition rate than point pattern matching algorithm. It can overcome effects of the environmental factors like scaling, orientation, illumination& viewpoint that affect the recognition rate which is not the case with PCA or Template Matching. It is also observed that some gestures that are similar to other gestures like gesture C and O or P and H of ASL can be recognized efficiently if high quality images are used.

REFERENCES

1. Daniel ThalmNN, Gesture Recognition Motion Capture, Motion Retargeting, and Action Recognition, EPFL – VRlab, pp. 1-22.
2. Aditya Ramamoorthy et al. “Recognition of dynamic hand gestures “,Department of Electrical Engineering IIT New Delhi-110016 India. October 2002, pp. 1-13.
3. Rafiqul Zaman Khan, Noor Adnan Ibraheem, Hand Gesture Recognition: a literature review,IJAIA 2012.
4. Ravikiran.J, Kavi Mahesh, Suhas Manishi, Dheeraj.R, Sudheendar.S, Nithin.V.Pujari, “On finger detection for sign language recognition”, Hong Kong, March 18, 2009.
5. Sylvie Gibet and Pierre Francois Marteau, “On Approximation of curvature and velocity for gesture segmentation and synthesis”,Universite de Brertagne France, 2009.
6. David Rybach, Prof Dr. J. Brochers, Prof Dr.H. Ney, Appearance based features for automatic continuous sign language recognition”.
7. Yasushi Hamada, Nobutaka Shimada and Yoshiaki Shirai, “on Hand shape estimation for complex background images”, Osaka University, Japan, 2004.
8. Manglai Zhou, “3D model based hand gesture recognition and tracking”, Pami lab, university of Waterloo, December 3,2009.
9. Lars Bretzner, Ivan Laptev, Toney Lindberg, On Hand gesture recognition using multiscale colour features hierarchiel models and partial filtering, CVAP Laboratory, Department of numerical analysis and Computer Science, Sweden,2002.
10. Vaishali S Kulkarani, ME Digital Systems, S. D. Lokhande, “Appearance based segmentation of sign language using gesture segmentation”, Sinhgad College of Engineering, 2010.
11. Wei-Lun Chao, “Introduction to Pattern Recognition”, National Taiwan University, Taiwan, October, 2009, pp. 1-31.

AUTHOR PROFILE

Deval G. Patel has completed her Diploma in Engineering (IT) from BBIT V.V.Nagar in 2007 and her Bachelor of Engineering (IT) from GCET V.V.Nagar, S.P. University in 2010. Currently she is pursuing her M.E.(Computer Engineering) from B.V.M. Engineering College, V.V.Nagar, Gujarat Technological University.