ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

INDIAN SIGN LANGUAGE RECOGNITION WITH IMPLEMENTATION OF SPEEDED UP ROBUST FEATURES ALGORITHM

Jeevan Musale*, A P Mane
M.E (ETC, Appeared), Department Of Electronics and Telecommunication engineering, C.O.E.Osmanabad, Affiliated to Dr.B.A.M.U, Aurangabad, Maharashtra, India
Assistant Professor, Department Of Electronics and Telecommunication engineering,C.O.E.Osmanabad, Affiliated to Dr.B.A.M.U, Aurangabad, Maharashtra, India

DOI: 10.5281/zenodo.47635

ABSTRACT

In This paper automatic gesture recognition is used for Indian Sign Language (ISL). In Indian sign language for representing alphabet we use both hands. We suggest an a method used in dealing with addresses local-global ambiguity which increases the range of possible interpretations of natural language, and a computer has to find a way to deal with this identification, inter-class variability increases for each hand gesture. Hand parts are firstly separated and then detected by skin color model reference i.e. HSI. In this system the well-built points are removing out for recognition process using Speeded Up Robust Features algorithm of each hand posture. The multi class linear support vector machines (SVM) is used to arrange each hand posture, because of it for the system a identification rate of 93.3% is achieved. The performance of this suggested approach is analyzed with well-known classifiers like SVM and experimental results are differentiated with the conventional and existing algorithms which indicate the better efficiency of the proposed approach.

KEYWORDS: Speeded up robust features algorithm.

.

INTRODUCTION

Sign language is used as a media for among deaf & dumb people to convey their information i.e. specking words or their expressions with each other. A normal person cannot communicate easily without knowing the sign language with Deaf and dumb person unless he/she is familiar with respect to it. Similar case is applicable when a deaf & dumb person she /he want to communicate with a normal person or blind person. In order to avoid the gap of transferring the information among deaf & dumb community and normal community, for that purpose now a days we use Video Relay Service (VRS). The VRS also sometimes known as a video interpreting service (VIS), is a video telecommunication service that allows deaf, hard-of-hearing and speech-impaired to communicate with other and also a normal person. In this system a manual interpreter can translates the hand moments or signs to voice and vice versa which is used for easy communication at both sides.

A lot of research work has been carried out by using image processing and pattern recognition techniques to automate the process of sign language interpretation. These suggestions can be broadly classified into two categories i.e. “Data-Glove based” and “Vision-based” [1]. The low level systems such as skin color, shape, or depth information [2] generally require uniform background, invariable illumination, a single person in the camera view, and/or a single large centered hand in the camera view these are used for Tracking bare hand and recognizing the hand gestures. A lot of researchers initially used morphological operations [3] to detect hand from image frames. Drawback for these methods lies in its huge computational complexity which is further handled with the concept of

integral image. Integral imaging is an auto stereoscopic and multiscopic three-dimensional imaging technique that captures and reproduces a light Field by using a two-dimensional array of micro lenses, sometimes called a fly's-eye lens .The use of integral image for hand detection in viola-Jones [4] method reduces computational complexity and shows satisfactory performance only in a controlled environment. For detecting the hand in a cluttered background format the N.Petersen & D.Stricker he used the color information and histogram distribution model [5]. These models are used for static gesture recognition purpose. Also Some Local orientation histogram technique is [6] used for it. These algorithms work well in a controlled lighting condition, but drawback is that it fails when lighting illumination changes, scaling and rotation. To resist illumination changes, Elastic graphs [2] are applying to get different hand gestures in Triesch's work with local jets of Gabor Filters. Mathias and Turk used Adaboost for wearable computing. It is inconsiderate to camera movement and user variance. Their hand tracking is promising, but segmentation is not authentic. We used Fourier descriptors of binary hand blobs as feature vector to Radial Basis Function (RBF) classifier for pose classification and combined HMM classifiers for gesture classification. It is also defined by Chan &Ranganath [7]. Even though this system reaches the required qualities for performance, it is not tough in opposition to multi variations during hand variation. The SIFT is used To reduce the problem of multi variations like rotation, scaling, translation [8], Haar-like features [9] with Adaboostclassifiers [10], Active learning [11] and appearance based approaches [2] are used. However, time complexity is problem for all these algorithms. We acquire a combined feature selection approach To improve the quality of the hand gesture recognition system [16]. For this paper, we use anew algorithms for real time hand gesture recognition which can identify different hand postures in a robust and faster way are introduced. In a Indian Sign Language (ISL) alphabet signs are used for recognition process. Video and still images datasheet of ISL alphabets is taken with various background and environmental conditions. In this system for single letter recognition Classifier architecture like SVM classifiers is used and accuracy of recognition is discussed.

HAND GESTURE RECOGNITION SYSTEM

1. *Hand Gesture Recognition System*

Hand gesture recognition system consists of the following steps (a) Image Acquisition (b) Pre-processing and hand segmentation, (c) Hand detection and tracking,

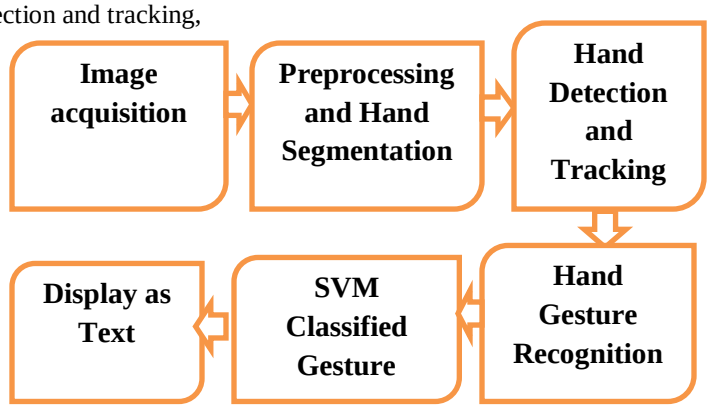

Fig.1block diagram of hand gesture recognition system in ISL.

(d) Hand Gesture recognition and (e) SVM Classified Gesture classification as shown in figure 1. The extracted features are converted into appropriate feature vectors. Multi class linear Support vector machines (SVM) is used for the classification of each hand gestures (Alphabets).

A. The Hand Detection and Segmentation

First, pre-processing i.e. some preprocessing of the concordances is desirable prerequisite and normalization the process of reorganizing data in a database so that it meets two basic requirements are done on the video object

frames. Skin color segmentation is the process of partitioning a digital image into multiple segments performed which is in HSI color space since it reduces the effect of uneven illumination in an image. HSI is an encoded nonlinear RGB signal with simple transformation; explicit separation of Hue, Saturation and Intensity components makes this color space attractive for skin color modeling. RGB color frames I (m, n, p) (where m, n and p are number of rows, number of columns and number of color planes) are converted into HSI images using standard equations.

B. The SURF Feature Extraction for ISL Alphabet Recognition

A combined feature extraction methodology using Speeded Up Robust Features (SURF) [17] and Hu Moment Invariant features [18] is incorporated in which the bounding box of the detected hand in each frame is obtained from the previous section. To recognize the posture of detected hand. Bounding box, BBIm(x, y) is taken as test image. These Features are determined and compared with the present database features. Minimum Euclidean distance between the feature vectors recognizes particular hand posture/letter.

Given an image BBIm(x, y), integral image ii(x, y) is calculated using,

$$ii(x,y) = \sum_{\substack{x1\le x \\ y1\le y}} BBIm(x1,y1) \quad \text{..... (1)}$$

To find out the interest points from the integral image, Fast Hessian Detector [22] is used. Given a point X=(x, y) in image ii(x, y), the Hessian matrix H(X,σ) in X at scale σ is defined as

$$H(X,\sigma) = \begin{bmatrix} L_{xx}(X,\sigma) & L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma) & L_{yy}(X,\sigma) \end{bmatrix} \quad \text{..... (2)}$$

Where $L_{xx}(X, \sigma)$, __ is the convolution of the
Gaussian second order derivative with the image ii in point X, and similarly for $L_{xx}(X, \sigma)$, $L_{yy}(X, \sigma)$.

To localize interest points in the image and over scales, non-maximum suppression in a 3x3x3 neighborhood is applied[19]. The maxima of the determinant of the Hessian matrix are then interpolated in scale and image space with the Browns method [20]. In order to be invariant to rotation, Haar-wavelet responses in x and y direction, within radius 6s around interest point is calculated. For the extraction of the descriptor, the first step consists of constructing a square region centred on the interest point. The region is split up regularly into smaller 4 × 4 square sub-regions. This keeps important spatial information in. For each sub-region, a few simple features at 5×5 regularly spaced sample points are computed. *dx*the Haar wavelet response in horizontal direction and *dy*the Haar wavelet response in vertical direction (filter size 2s). The wavelet responses dx and dy are summed up over each sub-region and form a first set of entries to the feature vector. Absolute values of the responses |*dx*| and |*dy*| provide polarity information. Each sub-region has a four-dimensional descriptor vector v**,** This results in a descriptor vector for all 4×4 sub-regions of length 64.

Figure 2: Extracted SURF feature points from reference images

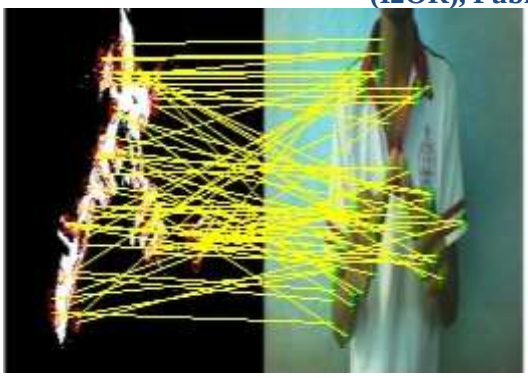

Figure 3: SURF matched key points from detected hand

C. Classification using Support Vector Machines

Then the output is simultaneously the feature vectors of the dataset are given to SVM classifier [23] for training. A classifier that implements the method, which we have just considered, is called the Support Vector Machine. The basic principle of SVM is to find an optimal separating hyper plane (OSH) maximize margin means which can separate different classes in a feature space, that is, the distances between these classes should be the furthest. To perform the classification between two classes, a nonlinear SVM classifier is applied by mapping the input data *(xi,yi)* into a higher dimensional feature space using a non-linear operator φ(x) where x ε R^d. The OSH can be computed as a decision surface:

$$f(x) = sign(\sum_i \alpha_i y_i K(x_i, x) + b), \quad \ldots\ldots(3)$$

Where sign () is the sign function and K (xi, x) =$\varphi(x_i)^T\varphi(x_i)$ is the predefined kernel function. In this approach the radial basis function (RBF) is used and it is defined as:

$$K(x_i, x) = e^{\frac{-||x_i - x||^2}{2\sigma^2}}, \sigma > 0, \quad \ldots\ldots(4)$$

where σ is the Gaussian width. The coefficients α_i and *b* in (4) can be determined by the quadratic problem. This procedure is carried out for the sequence of detected hand from video frames. For each frame classifier recognizes a Single letter as output. The results given by both the classifiers are taken as a combined feature vector for gesture classification.

D. Improving accuracy through Strong SURF features

To further increase the recognition rate and speed of processing, prominent and strong features are derived from the available data set of features using forward selection algorithm available in Matlab. it is possible to make strong orientation of each feature's which improve the performance of classifier on a dataset. Also the orientations and strong points are computed. Features whose feature deviation is minimal and prominent that is strong feature and can be best for recognition rate these are the two forward selection computes the "best features" of the data set, the resultant feature vector is used with SVM for classification. In orientation detection we will take the input of hand movement in any form or any orientation the gesture will be detected through the described section of feature extraction as the SURF algorithm [12].

EXPERIMENTAL RESULTS

All the experiments are carried out using Matlab R2013a, with Intel core i3 processor (CPU 2.27GHz) on a 64bit windows platform thus it is clear that The accuracy and performance of the proposed approaches are further verified using an experimental dataset consisting of single ISL letters and continuous word (finger and gesture spelled)..

A. Dataset

In the dataset of more than 10 videos are taken as test samples. Their will be use Three types of test sample videos (Spelling of Words like SIR, SCHOOL) are considered for experimental purpose videos are captured in home environment without so that there is no any special lighting using a consumer quality web camera. Frame rate considered for processing is 5 and The resolution of the video is 320x240.

B. ISL Alphabet Recognition

Using structure and shape features along with the finger count information for ISL Recognition, the following procedure is carried out to conduct the experiment. The sample database images and videos are fed into the respective classifiers namely KNN, SVM and Multi-class SVM for the training phase. After the training phase, the classifier will be learned and familiar with the signs (gestures). Now testing phase is conducted with the new input image or video which is containing the hand sign or gestures to detect and recognize.

Size of Dataset Considered for training and the testing the classifiers is as follows,

1. No. of Training samples:

Static: 23x2=64

[a,b,c,d,e,f,g,I,,k,l,m,n,o,p,q,r,s,t,u,v,w,x,z]

Dynamic: 3x 4 =12 [h, j, y]

2. No. of Test Samples:

09 videos (length: 46secs)

Resolution: 320x240, 640x480, Color: sRGB, Frame

Rate for processing: 5.

Table 1: Accuracy of Different Person with Processing Time

Person	No of Letter Images	Processing Time (sec)for each letter	False Recognit-ion	True Classifiers Recognition Rate
Person1	26	3.58-4.61	3	23
Person2	26	3.58-4.89	5	21

The accuracy of is as shown in table above where dataset of 2 person with static and Dynamic words are considered and accuracy in terms of number of image correctly classified are given.

C. Word Recognition

To test the proposed model for recognition of continuously spelled words (see section 2.4), the full test set of 10 words are used. The non-letter class is treated identically to the letter classes, with non-letter samples sampled randomly from the training videos. Word recognition accuracy of 96% is achieved. Words like “SIR”, “TABREZ”, “CITY”, “TEA”, “APPLE” have been taken for testing purpose. Figure 3: Frames showing transition from letter “A” to “E” in word APPLE. First 4 frames represents meaningful letter.

Figure 4: Apple Word

CONCLUSION AND FUTURE WORK

It is observed from the experimental results that SURF & HSI based Hand Detection, is durable against multiple like rotation, scale, lighting and view-point and provides good real time performance. The algorithm makes use of classifier SVM. The tradeoff between accuracy and speed of processing is maintained by the methods. Use of derived strong features from available feature of SURF & orientation features further makes the approach highly robust against multiple variations and shows consistent real time performance with improved processing speed. In future, research work will be focused on automatic Indian sign language (ISL) interpretation as text or voice. As ISL uses both hands for signing it involves both local and global hand movements thus the concept of gesture spotting, inter-hand occlusion will be investigated deeply in near future.

REFERENCES

1. G.R.S.Murthy& R.S. Jadon,"A Review of vision based hand gestures recognition", International Journal of Information Technology and Knowledge Management, Vol-2,pp. 405-410, July-December 2009.
2. Y.Fang, K.Wang, J.Cheng&H.Lu, "A Real Time and gesture recognition method", IEEE 2007.
3. Salleh, Jias, Mazalan, Ismail, Yussof, Ahmad, Anuar&Mohamad, "Sign Language to Voice Recognition:Hand Detection Techniques for Vision-Based Approach", In Conference Current Developments in Technology-Asssisted Education, FORMATEX 2006.
4. Paul Viola & Michael Jones, "Robust Real-Time Object Detection", Second International Workshop on Statistical and Computational Theories of Vision- Modelling, Learning, Computing and Sampling, Vancouver, Canada, July 13 2001.
5. Nils Petersen & Didier Stricker, "Fast Hand Detection Using Posture Invariant Constraints", Springer-Verlag Berlin Heidelberg LNAI 5803, pp.106-113, 2009.
6. H.Zhou, D.Lin&T.S.Huang, "Static Hand Gesture Recognition based on Local Orientation Histogram Feature Distribution Model", In IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops CVPRW'04 2004.
7. Chan Wah Ng, SurendraRanganath, "Real-Time Gesture Recognition system and application", Image and Vision Computing 20 (2002) 993-1007.
8. David G.Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision 60(2), 91-110, 2004.
9. Q.Chen, N.Georganas&M.Petriu, "Real Time Vision based hand gesture recognition using haarlike features", Instrumentation and Measurement Technology Conference – IMTC, Warsaw, Poland, May 1-3, 2007.
10. H.Francke, J.Ruiz-del-solar &R.Verschae, "Real-Time Hand Gesture Detection and Recognition Using Boosted classifiers and Active Learning", Springer-Verlag Berlin Heidelberg LNCS 4872, pp.533-547, 2007.
11. T.Dinh, V.Dang&D.Duong, "Hand gesture classification using boosted cascade of classifiers", IEEE 2006.
12. T.Burger, A.Caliper&S.Mancini, "Cued speech hand gestures recognition tool", 9th Conference Speech and Computer St.Petersburg, Russia September 20-22, 2004.
13. A.Caliper, L.Bonnaud, S.Malassiotis, M.Strintzis, "Comparison of 2D and 3D analysis for automated cued speech gesture recognition", 9th Conference Speech and Computer St.Petersburg, Russia September 20-22, 2004.
14. T.Starner&A.Pentland, "Real-Time American Sign Language Recognition from video Using Hidden Markov Models", IEEE 1995.
15. G.Rigoll, A.Kosmala&S.Eickeler, "High Performance Real-time Gesture Recognition Using Hidden Markov Models", Gerhard-Mercator- University, Duisburg-Germany.
16. R.Liang&M.Ouhyoung, "A Real Time Continuous Gesture Recognition System for sign language", Shih-Chien University Taipei-Taiwan.
17. S.Mitra&T.Acharya, "Gesture Recognition: A survey", IEEE Transcations on Systems, Man and Cybernetics, part c: Applications and Reviews, Vol.37, No.3 May 2007.
18. TarekM.Mahmoud, "A New fast skin colour detection technique", World Academy of Science, Engineering and Technology 43 2008.
19. K.S. Ravichandran& B. Ananthi, "Color Skin Segmentation Using K-Means Cluster", International Journal of Computational and Applied Mathematics ISSN 1819-4966 Volume 4 Number 2, pp. 153–157 2009.
20. Chieh-Chih Wang &Ko-Chih Wang, "Hand Posture Recognition Using Adaboost with SIFT for Human Robot Interaction", In Springer-Verlag Berlin-Recent Progress in Robotics, LNCIS 370, pp.317-329, 2008.