ISSN 2278 – 3091

International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE), Vol. 4 , No.4 Pages : 44 - 48 (2015)
Special Issue of ICEEC 2015 - Held on August 24, 2015 in The Dunes, Cochin, India
http://warse.org/IJATCSE/static/pdf/Issue/iceec2015sp09.pdf *(ICEEC 2015)*

Radially Defined Local Binary Patterns for Hand Gesture Recognition

J. V. Megha[1], J. S. Padmaja[2], D.D. Doye[3]
[1]SGGS Institute of Engineering and Technology, Nanded, M.S., India, meghavjon@gmail.com
[2]Visvesvaraya National Institute of Technology, Nagpur, M.S., India, padmaja.jonnalagedda@gmail.com
[3]SGGS Institute of Engineering and Technology, Nanded, M.S., India, dddoye@yahoo.com

Abstract: Hand gestures are extensively used in human non-verbal communication by hearing impaired and speech impaired people. To make communication between the hearing impaired and normal people simple and efficient, it is necessary that this process be automated. Number of techniques have been developed for automatic HGR. The usefulness of a method is analyzed based on various parameters like computational complexity, memory requirement, time required, etc. This paper proposes a new algorithm for static hand segmentation from complex background and for feature extraction for hand gesture recognition, considering the need of faster and simpler algorithms giving high recognition efficiency.

Key words: Feature Extraction, Radial Local Binary Pattern (RLBP), Segmentation, Support Vector Machine

INTRODUCTION

A sign in a Sign Language (SL) consists of three main parts: **Manual features** involving gestures made with the hands (employing hand shape and motion to convey meaning), **non-manual features** such as facial expressions or body postures and **finger spelling,** where words are spelt out gesturally in a local language [1]. To interpret the meaning of a sign, all these parameters are to be analysed simultaneously. Sign language poses a main challenge of being multichannel. Each channel in this system is separately built and analysed and the output of each channel is combined at the final stage to draw conclusion. The research in Sign Language Interpretation (SLI) began with Hand Gesture Recognition (HGR). Hand gestures are extensively used in human non-verbal communication by hearing impaired and speech impaired people. Even normal people sometimes use sign language for communication. Though sign language is spread and used all over the world, it is not universal. Wherever hearing impaired community exists, sign languages develop. To make communication between the hearing impaired and normal people simple and efficient, it is necessary that this process be automated. Number of techniques have been developed for automatic HGR. Some techniques consider geometry of the hand whereas some consider appearance of the hand for extracting features which will help automatically detect the hand gesture. The usefulness of a method is analysed based on various parameters like computational complexity, memory requirement, time required and difficulty in developing the code for the technique. This paper proposes a new algorithm for static hand segmentation from complex background and for extracting features for HGR, considering the need of faster and simpler algorithms giving high recognition efficiency.

LITERATURE REVIEW

Hand Detection and Extraction: Segmentation is used to detect hand from background. Mostly data based segmentation algorithms have been proposed, in which either uniform or plane background is needed or hand in the centre of the image is required. Some methods [2], [3] need to wear specialized data gloves which contain sensors which are used to predict the location and movement of hands during image acquisition. Stergiopoulou and Papamarkos [4] proposed YCbCr segmentation which has a limitation that the background should be plain or uniform. Rokade et.al. [5] proposed a segmentation algorithm which is sensitive to hand position. Peer et. al. [6] proposed RGB segmentation which is sensitive to light conditions. Ghotkar et al. [7] have explored various hand segmentation techniques using various color spaces. Other visual techniques commonly used for hand detection depending on gray level are single threshold value, adaptive thresholding, P-Tile method, edge pixel method and background subtraction [8], [9].

Hand Feature Extraction: Chung et al. [10] proposed a real time HGR based on Haar wavelet representation. Just et al. [11] have proposed to apply Modified Census Transform (MCT) which is similar to Local Binary Patterns (LBP) for feature extraction. They have claimed that MCT is illumination invariant.
Ding et al. [12] presented an improved LBP (local binary pattern) texture descriptor, which is used to classify static hand-gesture images. The descriptor makes full use of correlation and differences of pixel gray value in the local regions. Hrúz et al. [13] have used the Local Binary Patterns for describing the manual component of Sign Language. Wang et al. [14] have defined a kind of mapping score–*Disparity Cost Map* to map the few landmarks of the bare hands, instead of using accurate matching between two views.

ISSN 2278 – 3091
International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE), Vol. 4 , No.4 Pages : 44 - 48 (2015)
Special Issue of ICEEC 2015 - Held on August 24, 2015 in The Dunes, Cochin, India
***http://warse.org/IJATCSE/static/pdf/Issue/iceec2015sp09.pdf** (ICEEC 2015)*

Local Binary Pattern (LBP) is used for feature extraction in this paper. Jasim et al. [15] presented a method for computer vision-based static and dynamic hand gesture recognition. Haar-like feature-based cascaded classifier is used for hand area segmentation. Static hand gestures are recognized using linear Discriminant Analysis (LDA) and Local Binary Pattern (LBP)-based feature extraction methods.

LBP has been reported to perform quite well for texture and facial expression recognition. The authors intend to investigate the use of block LBP for HGR. As LBP considers each and every pixel in recognition, a modification in basic LBP is proposed here with an intention to reduce the computational efforts. Authors have tried to incorporate components of hand that would help differentiate inter-class gestures and group intra-class gestures. The input to feature extraction stage is the output of the hand detection stage. A simple and effective hand detection algorithm is also proposed in this work which helps detect the hand from a complex background.

HAND DETECTION

The experimentation in this work is carried out using two datasets representing hand gestures performed with one hand for alphabets A to Z using American Sign Language (ASL). Dataset 1 adapted from Rokade [16] contains 10 replications of alphabets A to Z and are with plane background whereas dataset 2 adapted from Rokade [17] contains 10 replications of 15 alphabets (A to O) with complex background. Images of dataset 1 having plane background does not need much pre-processing. Simple RGB to Gray conversion is applied to an image and it is then resized to the desired dimension. The images of this dataset before and after pre-processing stage are shown in fig.1.

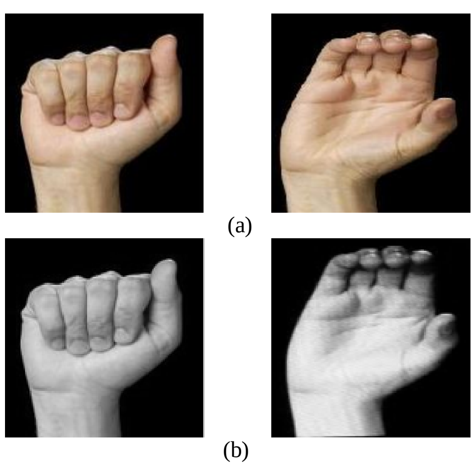


Fig. 1: (a) Original images of alphabet 'A' and 'C', (b) Images after RGB to Gray conversion and resizing.

Images in dataset 2 with complex background needs pre-processing to extract hand from the background. We are proposing a simple and efficient hand detection and extraction algorithm based on the HSV color space. The steps followed in detection and extraction of hand are as given below:

- Acquire Image
- Convert the image into HSV space
- Set threshold in H, S, V plots to allow only skin color to appear
- Create a mask from the chosen histogram thresholds and apply the mask to the image
- Convert image to gray and apply median filter to remove salt and pepper noise
- Analyze 8 neighboring pixels (each at a pre-defined index away from the pixel under consideration) of every pixel to check for any remaining noise
- Set threshold range for allowed pixel values and let 'n' be the number of pixels outside this range
- If n > 5, keep the pixel value as it is else set the pixel value to zero (black)

The output of the pre-processing stage is as shown in fig. 2.

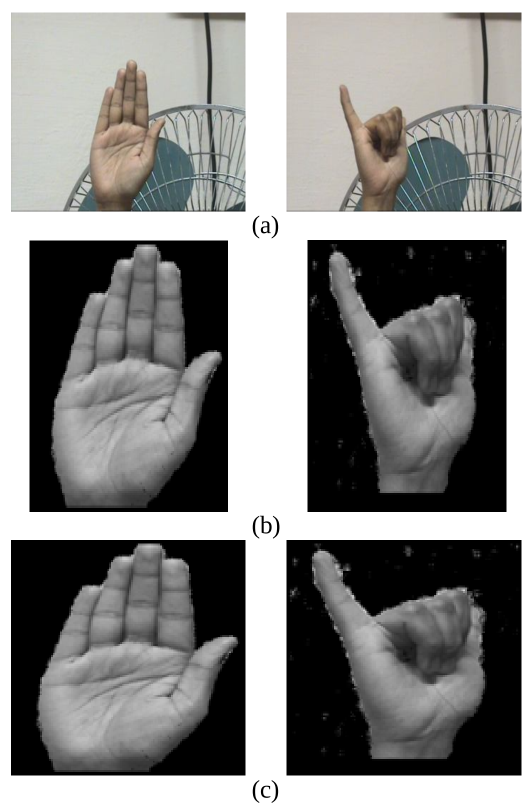


Fig. 2: (a) Original images of alphabets 'B' and 'J' (b) Images after background removal and (c) Resizing

ISSN 2278 – 3091
International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE), Vol. 4 , No.4 Pages : 44 - 48 (2015)
Special Issue of ICEEC 2015 - Held on August 24, 2015 in The Dunes, Cochin, India
***http://warse.org/IJATCSE/static/pdf/Issue/iceec2015sp09.pdf** (ICEEC 2015)*

LBP and RADIAL LBP

The original LBP operator was introduced by Ojala et al. [18] for texture description. The LBP operator labels the pixels f_p (p = 0,1,....7) of an image by thresholding a 3×3 neighbourhood of each pixel with the value of the centre pixel and considering the result as a binary number as shown in fig.3.

165	48	102
210	152	86
189	177	135

Threshold →

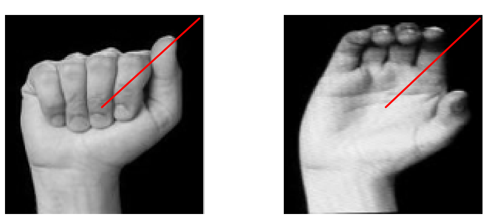
1	0	0
1		0
1	1	0

Binary 10000111

Decimal 135

Fig. 3: Basic LBP Operator

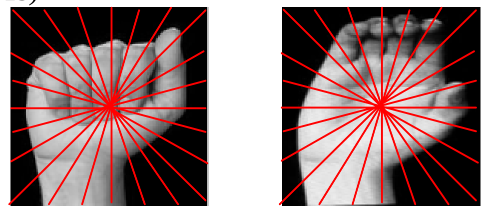
LBPs have proven to be very effective for image representation and have been applied to visual inspection, motion detection and outdoor scene analysis. The LBPs are tolerant against monotonic illumination changes and detect many texture primitives like spot, line end, edge, corner etc. The features in an image are expected to lie along different directions. Hence LBPs can also be used to take into consideration the pixel value variations exhibited in different directions when different gestures are subjected to feature extraction. The approach proposed in this work considers drawing radial lines at fixed angle intervals from centre of an image to the boundary of the image with an assumption that most of these radial lines would encounter different patterns when subject to different gestural images. For example, a radial line drawn at certain angle over an alphabet "B" image would encounter different pixel intensity variations as compared to a radial line drawn at the same angle over an alphabet "J" image. Fig. 4(a) represents one image superimposed with a line drawn at 260^o from the centre of the image to a point on the boundary of the image. Fig. 4(b) represents the same line superimposed on another image representing from the same database. If the LBP values are determined for each pixel on these lines, then it can clearly be observed that they would present substantially different representations for the two images.



(a) **(b)**

Fig. 4: Radial Line drawn over (a) Image of alphabet 'A' and (b)Image of alphabet 'C'

Fig. 5(a) and fig. 5(b) represent the above gesture images with radial lines drawn at some specific incremental angle (β).



(a) **(b)**

Fig. 5: Radial lines drawn with an incremental angle 'β' over (a) Image of alphabet 'A' and (b) Alphabet 'C'

LBP (referred to as Radial LBP / RLBP hereafter) values for each such radial line drawn on the image are computed and represented as a histogram [19]. All such histograms (one for each line) are concatenated to get the feature vector for that gesture image. If an incremental angle (β) of 10^o is considered then 36 radial lines shall be drawn and therefore each image would have a feature vector of size '36 × N' where N represents the number of bins of the histograms. Using LBP values of such radial lines in deriving feature vector for feature extraction and subsequent classification is based on the consideration that pattern exhibited by RLBP values along those radial lines would sufficiently represent the regions containing useful information for hand gesture classification.

EXPERIMENTATION

It is proposed in this paper to evaluate the performance of feature vectors based on Radial LBPs. Five different incremental angle values were considered to understand how recognition efficiency would be affected by the angle at which radial lines are drawn. Two low values (namely 2^o and 5^o) and three high values (namely 10^o, 12^o and 15^o) of β were taken for the purpose of experimentation. Besides, we intend to have as low a value of number of bins (N) as possible to reduce size of the feature vector. However, before fixing the value of N, we varied its value also during the experiments to identify which value caters to better recognition efficiency. Four values namely 8, 12, 16 and 20 were considered for N. The first database used for this experiment consists of 260 images taken from American Sign Language (ASL) database having 10 replications of each of the 26 alphabets. The second database contains 150 images of 15 alphabets with 10 replications of each alphabet. SVM has been used in the work reported in this paper for the purpose of classification. SVM does binary decisions. Multi class classification here is performed using binary classifiers along with a voting scheme. Major steps of the experimentation are enlisted below:

1. For each pair of β and N values from the sets: $S_β$ = {2, 5, 10, 12, 15}; S_N = {8, 12, 16, 20}, perform the following steps.
2. Extract features of all the test and reference images. For this experimentation, feature vector would be of size '(360/β) × N'.

ISSN 2278 – 3091

International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE), Vol. 4 , No.4 Pages : 44 - 48 (2015)
Special Issue of ICEEC 2015 - Held on August 24, 2015 in The Dunes, Cochin, India
http://warse.org/IJATCSE/static/pdf/Issue/iceec2015sp09.pdf *(ICEEC 2015)*

3. Segregate feature vectors of the 260 images so that 10 sets of 26 images each gets grouped. Define one set (not selected earlier) as the set of test images and define the remaining nine sets as reference images.
4. Use SVM for classifying each of the test gesture. As SVM does pairwise comparison, a test feature vector (whose gesture group is not known) is subjected to a series of pairwise comparisons.
5. Repeat steps 3 through 4 till each of the ten sets of 26 images are subjected to classification as the test samples.
6. Repeat the steps 1 through 5 for dataset 2.

RESULTS

The results include representation of overall recognition efficiency achieved (average results over ten cycles are considered as the recognition rate).

For different combination of incremental angle (β) and number of bins (N), the feature vector size can be represented as (360/β) × N (table 1).

Table 1: Feature Vector Size for different combinations of β and N over hand gesture image

	β = 2	β = 5	β = 10	β = 12	β = 15
N = 8	1440	576	288	240	192
N = 12	2160	864	432	360	288
N = 16	2880	1152	576	480	384
N = 20	3600	1440	720	600	480

Overall recognition efficiency obtained in percentage is presented in fig. 6 and table 2 for dataset 1.

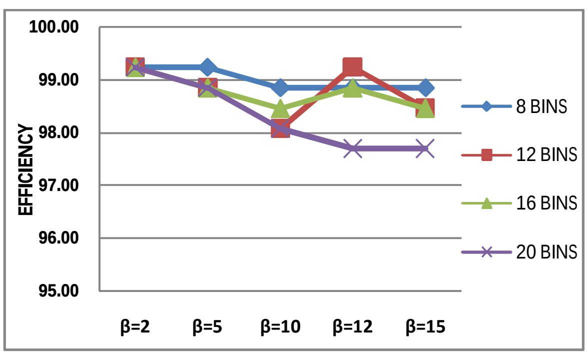


Fig. 6: Overall Recognition Efficiency (%) obtained by using RLBP for dataset 1

Table 2: Overall Recognition Efficiency values (%)

	8 bins	12 bins	16 bins	20 bins
β=2	99.23	99.23	99.23	99.23
β=5	99.23	98.85	98.85	98.85
β=10	98.85	98.08	98.46	98.08
β=12	98.85	99.23	98.85	97.69
β=15	98.85	98.46	98.46	97.69

It can be observed from fig. 6 and table 2 that the maximum recognition efficiency obtained is 99.23 while the minimum is 97.69. The results are quite encouraging as regards the applicability of RLBP to hand gesture feature extraction because RLBP has exhibited very good performance.

The results for dataset 2 are shown in fig. 7 and table 3.

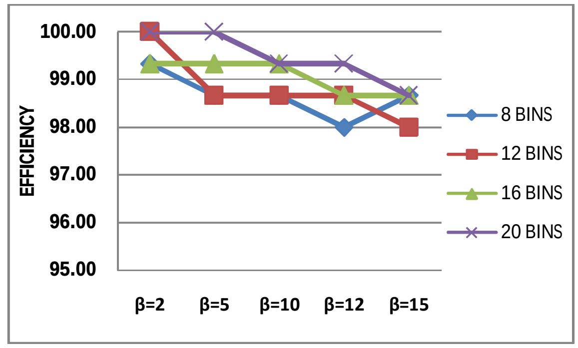


Fig. 7: Overall Recognition Efficiency (%) obtained by using RLBP for dataset 2

Table 3: Overall Recognition Efficiency values (%)

	8 bins	12 bins	16 bins	20 bins
β=2	99.33	100.00	99.33	100.00
β=5	98.67	98.67	99.33	100.00
β=10	98.67	98.67	99.33	99.33
β=12	98.00	98.67	98.67	99.33
β=15	98.67	98.00	98.67	98.67

It can be seen from table 3 that for β= 2 and 5, 100% efficiency is achieved even for 20 bins. In this case the minimum efficiency obtained is 98%.

CONCLUSIONS

Based on the experimentations, it can be concluded that:

The segmentation algorithm applied to the images with complex background is extracting the hand from its background very efficiently.

RLBP applied over the hand gesture images gives higher recognition efficiency even for higher values of β which capture less number of features of an image. Besides, recognition efficiency obtained for other values of β is also very good. The feature vector size for RLBP is comparatively quite small and thereby resulting in lesser time / memory requirement.

RLBP basically considers small areas where features are expected to be present and does not consider the complete hand region which is otherwise considered by LBP technique.

Radial LBP does definitely capture the differences in gestures. Future scope would be to carry out experimentation on real time systems.

ISSN 2278 – 3091

International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE), Vol. 4 , No.4 Pages : 44 - 48 (2015)
Special Issue of ICEEC 2015 - Held on August 24, 2015 in The Dunes, Cochin, India
http://warse.org/IJATCSE/static/pdf/Issue/iceec2015sp09.pdf *(ICEEC 2015)*

REFERENCES

[1] https://en.wikipedia.org/wiki/Sign_language

[2] Cooper H., Holt B. and Bowden R., *Sign Language Recognition,* Visual Analysis of Humans, book by Springer Publication, 2011.

[3] Xu D., "A Neural network Approach for HGR in Virtual Reality Driving Training System of SPG", *18th Int. Conf. on Pattern Recognition,* pp 519-522, 2006.

[4] Stergiopoulou E. and Papamarkos N., "A New Technique for HGR" , *IEEE Int. Conf. on Image Processing, pp 2657-2660, 2006.*

[5] Rokade R., Doye D.D. and Kokare M.B., "Hand Gesture Recognition by Thinning Method", *IEEE International Conf. on DIP,* pp 284-287, March, 2009.

[6] Peer P., Kovac J. and Solina F., "Human Skin Color Clustering for Face Detection", *EUROCON- International Conf. on Computer as a Tool,* Vol. 1, pp 144-148, Aug. 2010.

[7] Ghotkar A. and Kharate G. K., "Hand segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction", *Int. Journal of Human Computer Interaction (IJHCI), Col. 3, Issue 1, pp 15-25,* 2012.

[8] Khan R. Z. and Ibraheem N. A., "Hand Gesture Recognition: A literature Review", *International Journal of Artificial Intelligence and Applications,* Vol. 3, pp 161-174, July, 2012.

[9] Chen F., Fu C. and Huang C., "Hand Gesture Recognition using A Real Time Tracking Method and HMM", *Image and Video Computing,* Vol. 21, pp 745-758.

[10] Chung W. K., Wu X. and Xu Y., "A real time Hand Gesture Recognition based on Haar Wavelet Representation", *Procs. of the 2008 IEEE International Conference on Robotics and Blomimetics, Bangkok, Thailand,* February 21-26, pp 336-341, 2008.

[11] *Just A., Rodriguez Y., Marcel S.,* "Hand Posture Classification and Recognition using the Modified Census Transform", 2006.

[12] Ding Y., Pang H., Wu X. and Lan J., "Recognition of Hand-Gestures Using Improved Local Binary Pattern", IEEE, 2011.

[13] Hrúz M., Trojanová J., and Elezn M., "Local Binary Pattern based features for Sign Language Recognition", *Pattern Recognition And Image Analysis*, Vol. 22, No. 4, 2012.

[14] Wang H., Wang Q. and Chen X., "Hand Posture Recognition from Disparity Cost Map", *ACCV 2012, Part - II, LNCS 7725, Springer - Verlag, Berlin, Heidelberg,* pp 722-733, 2013.

[15] Jasim M., Zhang T. and Hasanuzzaman Md., "A Real-Time Computer Vision-based Static and Dynamic Hand Gesture Recognition System", *Int. Journal of Image Graphics,* 14, 145006, 2014.

[16] Rokade U., "One and Two Hand Gesture Recognition", PhD Thesis, SRTM University, Nanded, M.S., India, 2014.

[17] Rokade R., "Some Aspects of Hand Gesture Recognition", PhD Thesis, SRTM University, Nanded, M.S., India, 2014.

[18] Ojala T., Pietikäinen M. and Harwood D., "Multi Resolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 24 (7), pp 971–987, 2002.

[19] Jonnalagedda M. V. and Doye D. D., “Radially defined Local Binary Patterns for Facial Expression Recognition", *International Journal of Computer Applications*, 119(21):17-22, June 2015.