Quantitative evaluation of Segmentation algorithms based on level set method for ISL datasets

Ms. M.Krishnaveni
Research Assistant
Department of Computer Science
Avinashilingam Deemed University for Women
Coimbatore ,India
krishnaveni.rd@gmail.com

Dr.V.Radha
Associate Professor
Department of Computer Science
Avinashilingam Deemed University for Women
Coimbatore ,India
radharesearch@yahoo.com

***Abstract*— The enormous potential research efforts have been taken for sophisticated and natural human computer interaction using gestures. This work has got motivated from long ago as 1980's since sign language is the only communicate mode for deaf community people. In signing, the face and a hand of a signer often overlap, thus the system needs to segment these for the purpose of feature extraction. This research work concentrates with the separation of the face and hand by detecting contour of the static object using reference labels and different snake algorithms. Indian sign language dataset is used to evaluate a few level set computer vision algorithms. Specifically different feature sets, segmentation algorithms and color constancy algorithms are evaluated quantitatively. In future, it is possible to evaluate the accuracy of sign on a large scale due to the availability of large annotated databases.**

Keywords –Segmentation; Level set; Sign language; Contour method; objective assessment

I. INTRODUCTION

The face and hand segmentation process for sign language consists of three steps, namely skin-color segmentation, change detection, and face and hand extraction[12]. For skin-color CbCr plane is modeled as a bivariate normal distribution. The change detection method is based on the test and block based motion estimation. Face and hand extraction can be done using segmentation techniques for consistent detection, preserving the shape contour which is therefore useful for feature extraction[10][11]. From the previous research [2] it is therefore concluded that contour methodology identifies both isolated regions and consequential features for sign language from color image. The main aim is to improve segmentation in a more effectivel manner. This therefore requires evaluation of various level set based segmentation methods that overcomes the difficulties of overlapping regions. The goal is achieved using two different datasets namely : (i) general sign language dataset (ii) Indian Sign language dataset. This research work has fairly found out the difference by its change of representation of an image and its curve parts. Here the segmentation methods using level set algorithms are also evaluated with performance metrics. The paper is organized as follows: Section 2 deals with the overview of Indian sign language system Section 3 deals with the need for snake model based segmentation for hand extraction. Section 4 deals the comparison of level set methods and the subjective assessment of the adopted approach Section 5 explores the performance evaluation of these methods. The paper ends with future work and conclusion.

II. OVERVIEW OF RECOGNITION SYSTEM FOR INDIAN SIGN LANGUAGE

One of the biggest challenges in Sign Language Recognition (SLR) is to find a modeling technique that is powerful enough to capture the language, yet it can scale to large vocabulary size[9]. There are approximately 53 categories of signs in Indian Sign Language, and 1750 words of signs can appear in many different forms depending on subject, object, and numeric agreement. In many hand gesture recognition systems the entire gesture is modeled, however, in sign language recognition, where the language is very large, it is not feasible to

model each sign. It is more feasible to break down signs into a limited set of primitive parts, phonemes that can be combined to make up the entire set of signs in Indian Sign Language (ISL). Here a combined innovative scientific theory and vision-based technology is required which can be developed using most advanced techniques in image analysis, computational intelligence.
The Common Indian Sign Language Dictionary has been published by Sri Ramakrishna mission Vidyalaya College of Education, Coimbatore . The dictionary is framed based on five parameters namely hand shapes orientation, Location, Parameters, Movements, Facial Expression. The segmentation methods are tested based on the hand shapes parameters which the most significant datasets that need effective segmentation image processing techniques. The figure 1 depicts the system for ISL recognition system.

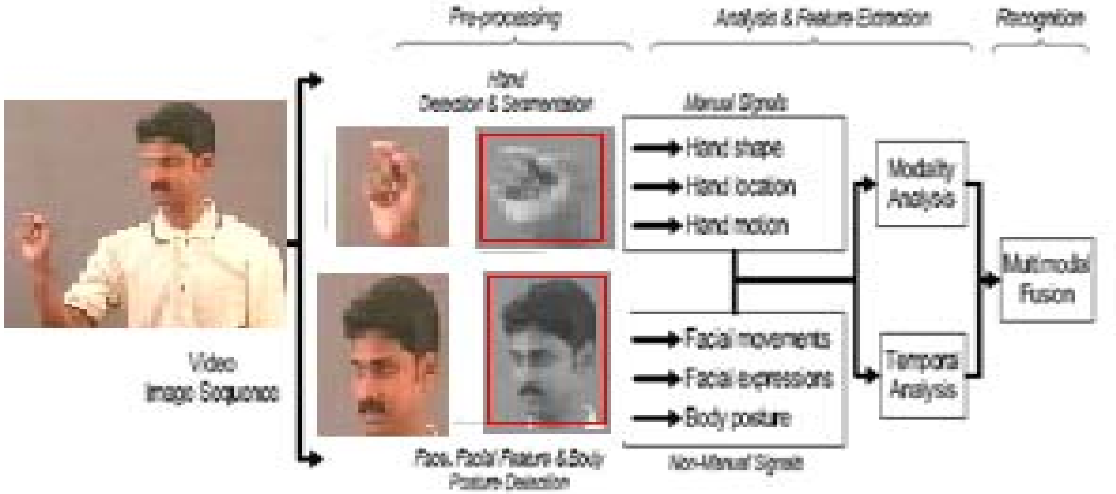


Figure 1: Indian Sign language recognition system

III. THE LEVEL SET METHOD IN IMAGE SEGMENTATION

The level set method is a numerical and theoretical tool for transmitting interfaces. It is basically used to track moving fronts by considering the front as the zero level set of an embedded function, called the level set function. In image processing, it is used for propagating curves in 2D and surfaces in 3D. The application here covers the 2D curves for segmentation purpose. There has been a number of segmentation algorithms proposed that aim at splitting up images into coherent regions using low-level visual cues. Contour cue is encoded using the oriented energy approach. Orientation energy at a pixel is the edge strength at that pixel at a given orientation and scale. At all pixels, the orientation energies are computed at different scales and orientations. Only those pixel orientation energies are retained that are a local maximum over different scales and orientations and others are set to 0. Further, an intervening contour framework is adopted. If the orientation energy along the line joining two pixels *i* and *j* is strong, then the similarity weight between these pixels based on the contour cue, W_{ij}^{IC} , is made low (towards 0). On the contrary, if there is little energy, for example in a constant brightness region, the link between the two pixels is made strong (towards 1). Again W_{ij}^{IC} is forced to be in the range [0, 1] as in the case of texture cue. For each method, the energy criterion is given which is minimized for the derived evolution equation. Usually, the traditional level set method only works when the boundary is clear and complete because of the boundary detector term. This is been attempted using the sign dataset below which is shown in figure 2.

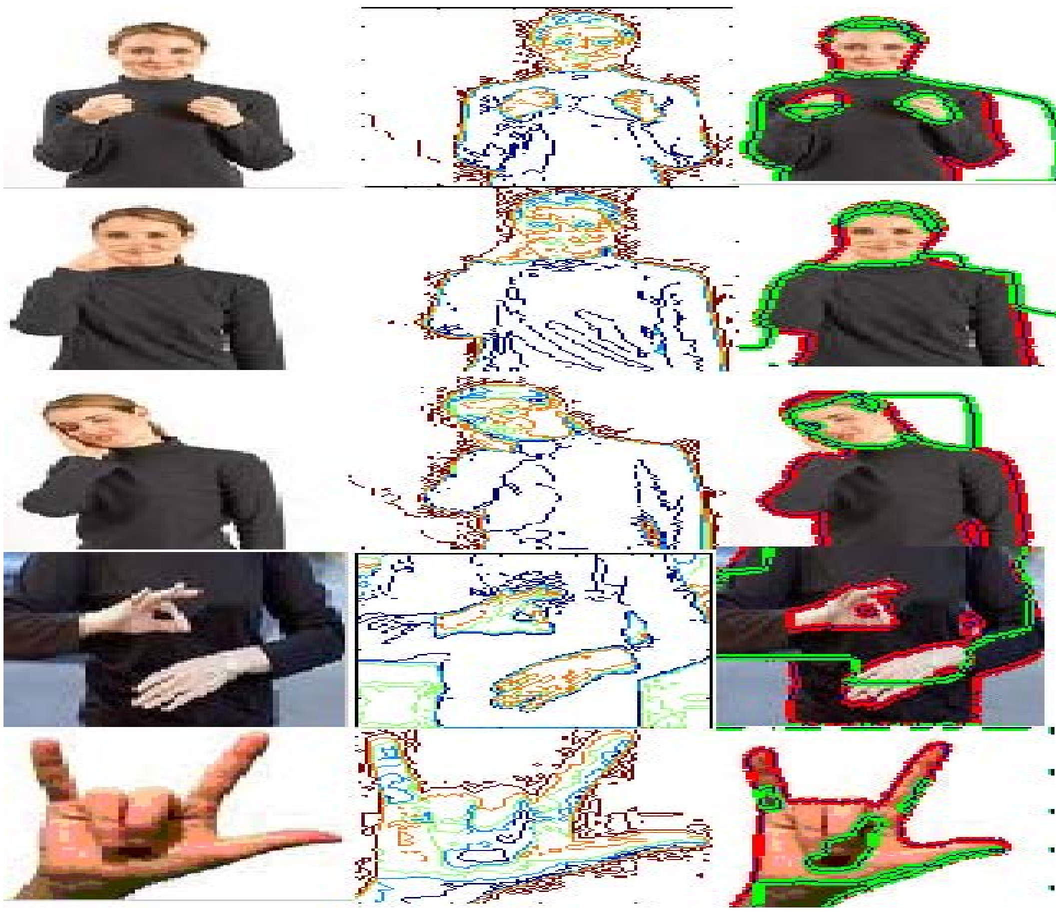


Figure 2: Subjective evaluation of conventional Segmentation methods (a) Original image (b) Local Active method (c) Global active method

IV. COMPARISON OF LEVEL SET METHODS BASED ON SIGN LANGUAGE DATASETS

In this paper, experimental study is done on level set local minima by comparing them with two datasets especially for sign language recognition. To the knowledge, such a comparison has not yet appeared in the literature especially for Indian dataset. There are two ways in which minima seeking algorithms may be compared: First, the value of the energy function at the minima can be compared, and second, the location of the minima can be compared. The first is taken into account that explains the relative power of the algorithms to minimize the energy function.

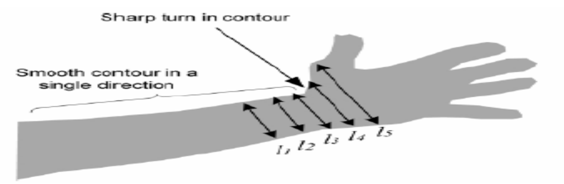


Figure 3 : Contour based Hand Segmentation

A. *Caselles method*

This approach is for object segmentation that allows to connect classical “snakes” based on energy minimization and geometric active contours based on the theory of curve evolution.This algorithm is a contour-based method i.e. the gradient of the image is used to compute the force function[4]. The curve will thus be driven to regions with high gradient. This method does not require any regularization term as it is intrinsic to the method, and Eqn (1) represents the same

$$\frac{\partial \phi}{\partial t}(x) = g(I(x))\nabla \phi(x)(c+k) + \nabla_g (I(x))\nabla \Phi(x) \quad \text{……………..(1)}$$

Where $k = div(\frac{\nabla \phi(x)}{\|\nabla \phi(x)\|})$ corresponds to the curvature of the evolving contour and c is a constant that acts as a balloon force. Energy criterion can be defined as eqn (2)

$$E(T) = \int_0^1 g(I(\Gamma(q))) \|\Gamma^{'}(q)\| \, dq \quad \text{..(2)}$$

$$g(I) = \frac{1}{1 + \|\nabla (G * I)\|^2} \quad \text{……………………………………….(3)}$$

Where I(.) corresponds to the image intensity, Γ is the parametric curve.

B. Chan Vese method

This level set method has some advantages compared to the active contour model. The level set method conquers the difficulties of topological transformations. The level set approach is able to handle complex topological changes automatically. This method takes the advantage of chanvese model in which it detects the contours both with and without gradients. In addition, by using this model and its level set formulation, interior contours are automatically detected, and the initial curve can be anywhere in the image. This frame work is applied to see how the different image properties can be used for segmentation. In order to segment images into more regions, a Multiphase Level Set Method has been developed. Based on the Four-Color Theorem, only four colors are enough to dye all the regions in a partition[5]. Therefore only two level set functions will suffice to represent any partition. The formulation used for this method which will minimize the energy function is defined as follows in eqn (4)

$$\begin{aligned} F(c,\phi) &= \int_{\Omega} (u_0 - c_{11})^2 H(\phi_1) H(\phi_2) dxdy \\ &+ \int_{\Omega} (u_0 - c_{10})^2 H(\phi_1)(1 - H(\phi_2)) dxdy \\ &+ \int_{\Omega} (u_0 - c_{01})^2 (1 - H(\phi_1)) H(\phi_2) dxdy \\ &+ \int_{\Omega} (u_0 - c_{00})^2 (1 - H(\phi_1))(1 - H(\phi_2)) dxdy \\ &+ \mu \int_{\Omega} |\nabla H(\phi_1)| + \mu \int_{\Omega} |\nabla H(\phi_2)|, \end{aligned} \quad \text{………………(4)}$$

This algorithm is a region-based method which tends to separate the image into two homogeneous regions. It is implemented as a signed distance function and is reinitialized at each iteration.

C. Chumming Li method

The energy criterion can be defined here as in eqn (5)

$$\begin{aligned} E(\phi) &= \lambda_1 \int_{\Omega} \int_{\Omega} K_{\sigma}(x - y) |I(y) - f_1(x)|^2 H(\phi(x)) \, dydx \\ &+ \lambda_2 \int_{\Omega} \int_{\Omega} K_{\sigma}(x - y) |I(y) - f_2(x)|^2 (1 - H(\phi(x))) \, dydx \\ &+ \nu \int_{\Omega} \delta(\phi(x)) \|\nabla \phi(x)\| dx + \mu \int_{\Omega} \frac{1}{2} (\|\nabla \phi(x)\| - 1)^2 \, dx \end{aligned} \quad \text{……….(5)}$$

Where I(x) is the image intensity at pixel x, H is the Heaviside function, K_{σ} is a Gaussian kernel defined as :

$$K_{\sigma}(u) = \frac{1}{(2\pi)^{n/2} \sigma^n} e^{-\|u\|^2 / 2\sigma^2} \quad \text{............... (6)}$$

With a scale parameter $\sigma > 0$ .f_1 and f_2 are two functions centered at pixel x and computed at each iteration as :

$$f_1(x) = \frac{K_{\sigma} * (H(\phi(x)) I(x))}{K_{\sigma} * H(\phi(x))} \quad \text{…………….(7)}$$

$$f_2(x) = \frac{K_{\sigma} * (1 - H(\phi(x))) I(x))}{K_{\sigma} (1 - * H(\phi(x)))} \quad \text{………..(8)}$$

The two first integrals correspond to data attached term, which are localized around each point x[7]. The third integral corresponds to the usual regularization term that smoothes the curve during its evolution. The last integral is a regularization term that forces the level-set to keep signed distance properties over the evolution process.

D. Lankton method

The evolution equation used in this method is represented in eqn (9) as

$$E\phi = \int_\Omega \delta(\phi(x)) \int_\Omega B(x,y).F(I(y),\phi(y))dydx + \lambda \int_\Omega \delta(\phi(x)) \|\nabla\phi(x)\| dx \text{ ………………(9)}$$

Where δ is the dirac function, B is a ball of radius r centred at point x and defined as follows :

$$B(x,y) = 1, \|x - y\| \leq r \text{ ……………(10)}$$
$$0, \textit{otherwise}$$

This algorithm is a region-based method and feature term is computed locally[6]. This allows the algorithm to segment non homogeneous objects, in which the method is sensitive to initialization.

E. Bernard method

Here, let Ω be a bounded open subset of R^d and let f: Ω→R be a given d-dimensional image[3]. In the B-Spline level-set formalism, the evolving interface, Γ C R^d is represented as the zero level-set of an implicit function φ(.) expressed as a linear combination of B-Spline basis functions

$$\phi(x) = \sum_{k \varepsilon Z^d} c[k] \beta^n \left(\frac{x}{h} - k\right) \text{ …………(10)}$$

In eqn (10) β^n(.) is the uniform symmetric d-dimensional B-spline of degree n.

The energy criterion here is

$$E(\phi) = \int_\Omega F(I(x),\phi(x))dx \text{ ……………..(11)}$$

F is given as $F(I(x),\phi(x)) = H(\phi(x))(I(x)-\upsilon)^2 + (1-H(\phi(x)))(I(x)-\mu)^2$ …..(12)

Eqn (12) gives the same chan vese model

F. Shi method

This method is a fast algorithm based on the approximation of the level-set based curve evolution[8]. The implicit function is approximated by a piece-wise constant function taking only four values (-3, -1, 1, 3) corresponding respectively to the interior points, the interior point adjacent to the evolving curve, the exterior points adjacent to the evolving curve and the exterior points. The two narrow-bands that enclosed the evolving contours are gathered into two lists that are updated at each iteration from simple rules, making the algorithm particularly fast. The evolution equation used here is the same Chan&Vese one given by eqn (13)

$$F(I(x),\phi(x)) = H(\phi(x))(I(x)-\upsilon)^2 + (1-H(\phi(x)))(I(x)-\mu)^2 \text{ ………………(13)}$$

Here, H is the Heaviside function

Figure 4 : Implementation with simple database and its subjective evaluation

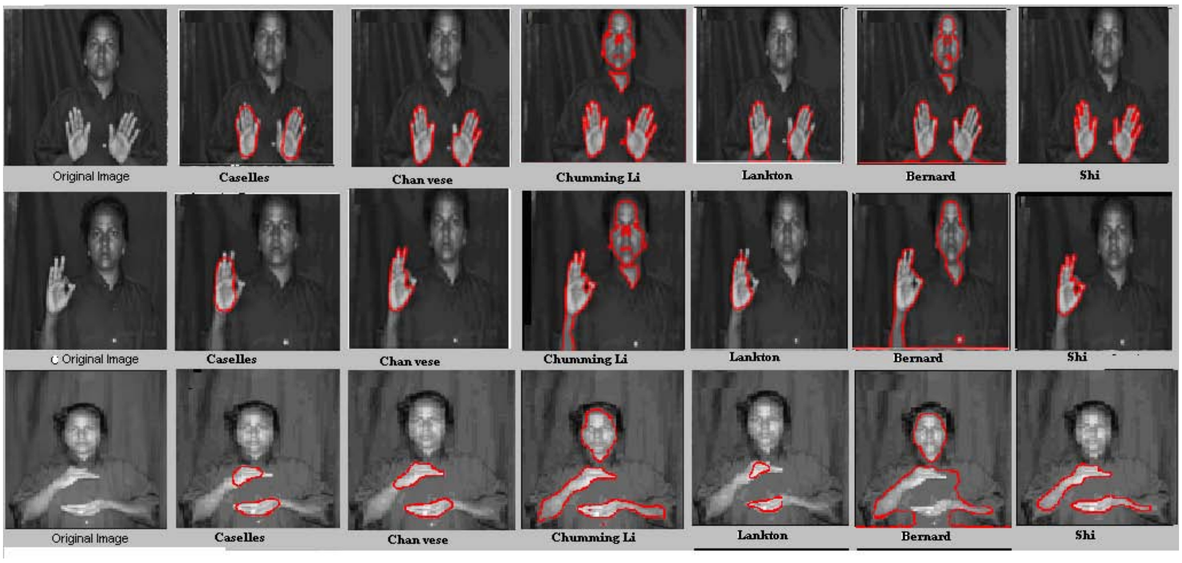


Figure 5 : Implementation with simple database and its subjective evaluation

Figure 4 and 5 explains the subjective assessment of the adopted level set algorithm on two sets of database.

V. PERFORMANCE EVALUATION

It is very much essential in estimating the performance of an algorithm on the specific application. This therefore helps in addressing the characteristic of the algorithm and the application prospectus. The comparative performance of the adopted methods is based on four parameters from the segmentation process. Comparison of two classes of datasets, namely general and Indian database is carried out using creseg tool[1]. Results indicate that the performance is a function of number of regions used for annotation. Segmentation algorithms that produce visually superior is shi which is better than other methods when implemented in the given datasets. Furthermore, shi algorithm is considered to be the best in comparison with detailed experimentation also. The subjective evaluation of the comparison results are shown in figure 6

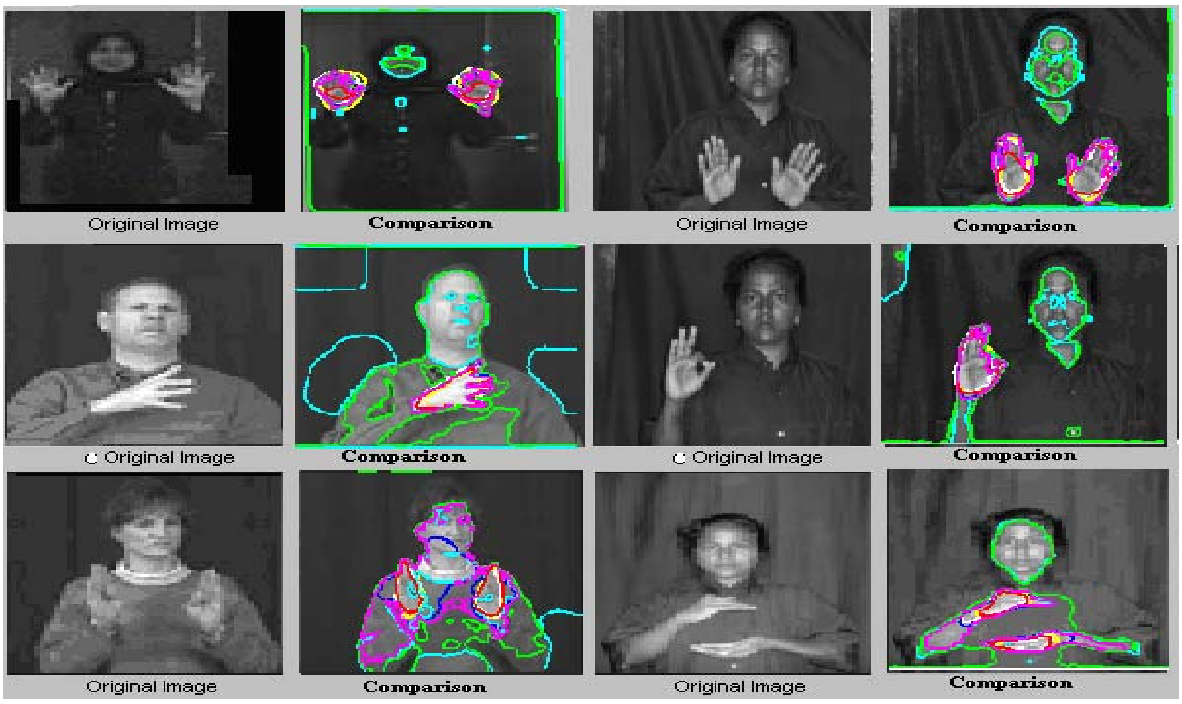


Figure 6: Subjective comparison of level set methods

The evaluation parameters taken for the experiments are Dice, PSNR, HD and MSSD

Dice Criterion : dice = $\frac{2(A \cap B)}{A + B}$,where A and B are the reference mask region and the result mask region of an algorithm.

PSNR :PSNR $= 10\log_{10}(\frac{d}{MSE(A,B)})$, where d is the maximum possible value of the image and MSE (A,B) is the mean square error computed between A and B are given in eqn (14).

MSE(A,B)= $\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\|A(m,n) - B(m,n)\|^2$ ……….(14)

Hausdorff distance:

Hausdorff =max $(D_1(A,B), D_1(B,A)$, where A and B are the reference contour and the result contour of an algorithm and calculated based on eqn (15) $D_1(A,B) = \max_{x \in A}(\min_{y \in B}(\|x - y\|))$ ……………………..(15)

Mean Sum of Square Distance (MSSD):

MSSD $= \frac{1}{N}\sum_{n=1}^{N} D_2^2(A, B(x_n))$, where A and B are the reference contour and the result contour of an algorithm, N is the size of the result contour and $D_2(A, B(x)) = \min_{y \in A}(\|y - x\|)$ …………………….(16)

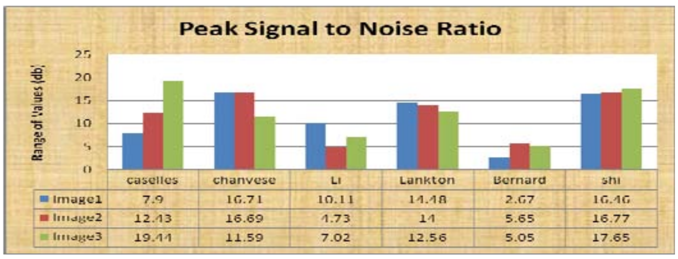


Figure 7: Comparison of PSNR for GSL

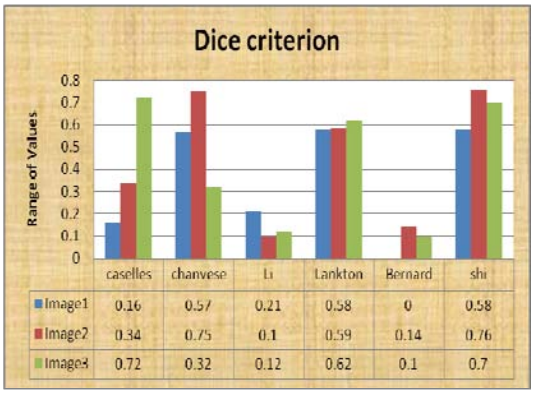


Figure 8: Comparison of DICE for GSL

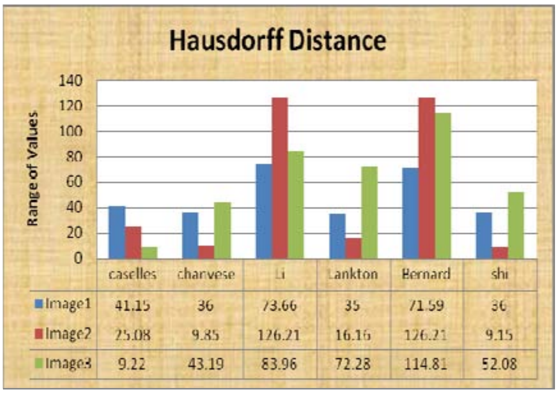


Figure 9: Comparison of HD for GSL

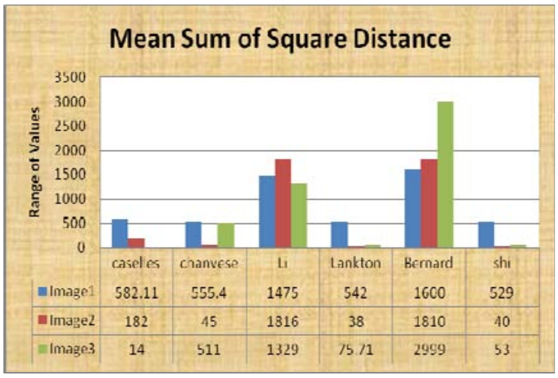


Figure 10: Comparison of MSSD for GSL

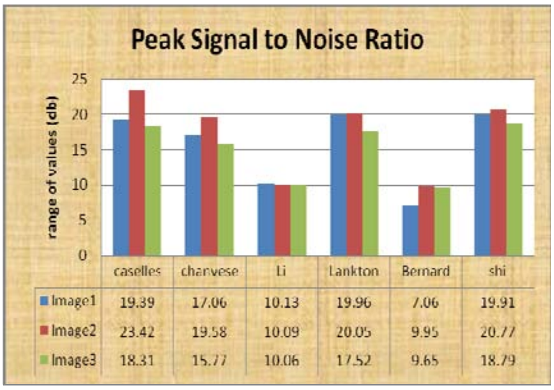


Figure 11: Comparison of PSNR for ISL

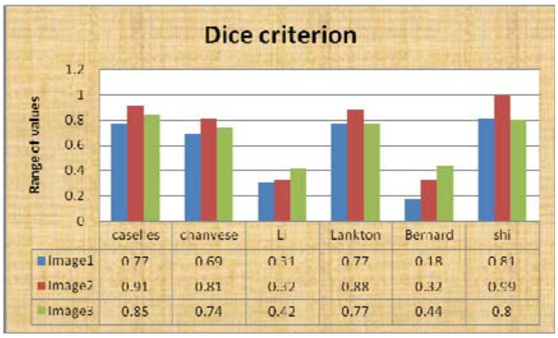


Figure 12: Comparison of DICE for ISL

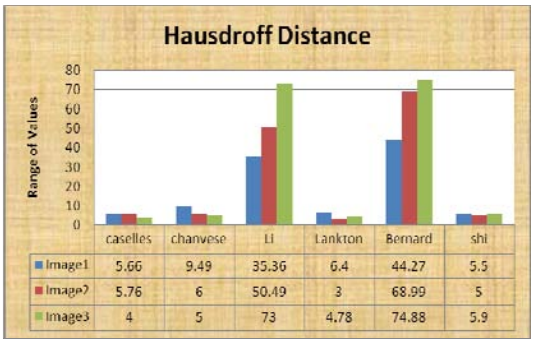


Figure 13: Comparison of HD for ISL

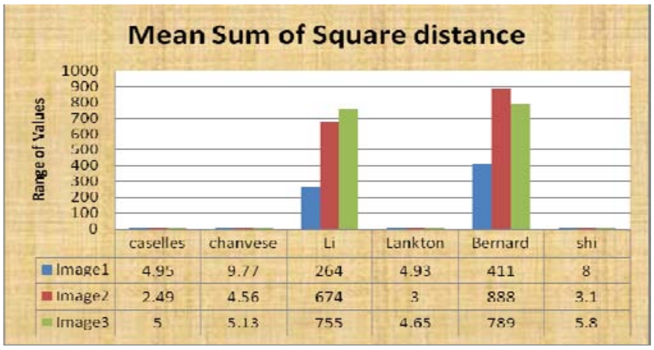


Figure 14: Comparison of MSSD for ISL

The results obtained from the segmentation methods are used for finding appropriate method for SLR system. Figure 7,8, 9,10 are the graphical representation of the segmentation metric results for GSL and figure 11, 12, 13, 14 are the graphical representation of the segmentation metric results for ISL .

VI.CONCLUSION

In this paper, performance of a level set methods is experimented using two various datasets of sign language. A methodology was presented for the objective evaluation of Indian Sign Language image segmentations. The proposed metric is based on examining the spatial accuracy of segmentation results using a pre-existing or manually generated reference mask. These metrics effectively capture deficiencies such as inaccurate boundary localization, over-segmentation and under segmentation, and its output was shown to correlate well with the outcome of subjective evaluation of segmentation masks by a human observer. Hence the research work states that the, shi level set method works well after the comparison of segmentation methods on the sign language representative images. Experimental results with six images show that the level set evolution strategy gives minima whose energy is close to the minimum energy obtained from simulated annealing and random search. Further, progress of the research work can be based on the development of personal algorithm that compete the potentiality of the adopted methods.

ACKNOWLEDGEMENT

The authors thank RamaKrishna Mission Vidyalaya, International Human Resource Development Centre (IHRDC for the disabled) for providing Indian sign language dictionary.

REFERENCES

[1] T. Dietenbeck, M. Alessandrini, D. Friboulet, O. Bernard. CREASEG: a free software for the evaluation of image segmentation algorithms based on level-set. In IEEE International Conference On Image Processing. Hong Kong, China, 2010.

[2] M.Krishnaveni.M, V.Radha, "Feature Extraction based on multiphase level set framework for Sign Language recognition system" ,Proceedings of ACM, Amrita ACM-W Celebration of Women in Computing, Coimbatore, Sept 16th – 17th, 2010.

[3] O. Bernard, D. Friboulet, P. Thevenaz, and M. Unser. Variational B-Spline level-set: "A linear filtering approach for fast deformable model evolution". IEEE Trans. Image Process., volume 18, no.06, pp.1179-1191, 2009.

[4] V. Caselles, R. Kimmel, and G. Sapiro. "Geodesic active contours". Int. J. of Computer Vision, volume 22, pp.61-79, 1997.

[5] T. Chan and L. Vese. "Active contours without edges". IEEE Trans. Image Process., volume 10, no.02, pp.266277, 2001.

[6] S. Lankton and A. Tannenbaum. "Localizing region-based active contours". IEEE Trans.Image Process., volume 17, no 11, pp.2029-2039, 2008.

[7] C. Li, C.-Y. Kao, J. C. Gore, and Z. Ding. "Minimization of region-scalable fitting energy for image segmentation". IEEE Trans. Image Process., volume 17, no.10, pp.1940- 1949, 2008.

[8] Y. Shi and W. C. Karl. "A real-time algorithm for the approximation of level-set based curve evolution". IEEE Trans. Image Process., volume 17, no.05, pp.645- 656, 2008.

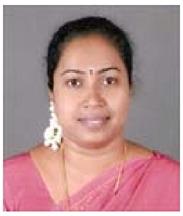
[9] T. Starner and A. Pentland, "Real-time American sign language recognition from video using Hidden Markov Models," In Proc. International Symposium on Computer Vision, pp.265-270, 1995.

[10] J.-S. Kim, W. Jang and Z. Bien, "A Dynamic Gesture Recognition System for the Korean Sign Language", IEEE Trans. on Systems, Man and Cybernetics, Vol. 26, No. 2, pp. 354-359, 1996.

[11] R.-H. Liang, Continuous Gesture Recognition System for Taiwanese Sign Language, Ph.D Thesis, National Taiwan University, Taiwan, R.O.C., 1997.

[12] B. Menser and M. Wien, "Segmentation and Tracking of Facial Regions in Color Image Sequences," SPIE Visual Comm. and Image Processing 2000,vol. 4067, pp. 731-740, June 2000.

AUTHORS PROFILE



Ms. M. Krishnaveni, 4 Years of Research Experience Working as Research Assistant in Naval Research Board project, Area of Specialization: Image Processing, Pattern recognition, Neural Networks. She has 33 publications at national and International level journals and conferences and one major research project under UGC.
Email id:krishnaveni.rd@gmail.com.

Dr. V. Radha, is the Associate Professor of Department of Computer Science, Avinashilingam Deemed University for Women, Coimbatore, TamilNadu, India. She has more than 21 years of teaching experience and 7 years of Research Experience. Her Area of Specialization includes Image Processing, Optimization Techniques, Voice Recognition and Synthesis, Speech and signal processing and RDBMS. She has more than 34 Publications at national and International level journals and conferences. She has one Major Research Project funded by UGC. Email id: radharesearch@yahoo.com.