International Journal of Electronics Communication and Computer Engineering
Volume 4, Issue 3, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

Indian Sign Language Gesture Segmentation Using Active Contour

Padmavathi.S
Asst. Professor, Deptt. of IT
Amrita Vishwa Vidyapeetha
Coimbatore, India

Shangeetha.R.K
Deptt. of IT
Amrita Vishwa Vidyapeetha
Coimbatore, India
Email: shangy1906@gmail.com

Karthik.K.S
Deptt. of IT
Amrita Vishwa Vidyapeetha
Coimbatore, India

***Abstract* – Deaf and dumb people communicate using sign language which everyone can't understand. The sign language can be conveyed to others using computers. This paper proposes a segmentation method using active contour algorithm that identifies the hands, which can be converted to medium so that everyone can understand. The gestures for Indian language are segmented and the contour is formed based on local information. The contour is the outline or boundary of an image. The local contour can segment objects with heterogeneous features. First the contour based segmentation is compared with edge based segmentation. Then the study of the energy in response to changes in localization is observed. And then the results of overlapping images that are possible with contour based segmentation are illustrates. This proposed method gives the exact contour of every gesture.**

***Keywords* – Indian Sign Language Segmentation, Active Contour, Multi Region Segmentation.**

I. Introduction

The image segmentation is one of the most complex tasks in image processing. Moreover, the gesture recognition is one of the most interesting and important issue in the field of pattern recognition. Communication is the only means through which human can express their thoughts and ideas. However, the deaf and dumb people find it difficult to interact with the others. They generally use the sign language to communicate within themselves for which they form the gestures in hand that represent the alphabets. Different types of sign language are being used around the world. The sign language in America uses only one hand to represent the alphabet whereas the Indian sign language is completely different. It uses both the hands to represent the alphabets. The Indian sign language resembles the English alphabets. Though the Indian sign language is one of the first known sign language and is considered important in the history of sign languages, there is not much software that converts it to normal message. However, there are many for American Sign Language recognition. There is also gesture recognition techniques that require its users to wear gloves for easier segmentation, which is not required in the method proposed in this paper. The social interest to develop an interfacing for the deaf and people is very low due to its complexity. The signing may differ for every person; and the disturbing elements like illumination problem can be very difficult to ignore. This paper suggests a technique where the images captured through webcam are segmented using active contour method and using image-processing techniques.

The input image is segmented for processing. Segmentation is a process in which a digital image is portioned into many sets of pixels (multiple segments). The main advantage of the contour based segmentation, proposed in this paper is that it does not require an additional hardware. The Chan-Vese method for segmentation is one among the most powerful and flexible. It can segment different kinds of images. It is based on Mumford-shah function for segmentation that is mostly used in medical field. It is an energy minimization method that will be reformulated to level set formulation to solve the problem in an easier way. This paper also tells about the significance of parameters corresponding to the local statistical methods like the degree of localization. Because it is very important to choose the correct one that gives a successful output. Since Indian language uses both the hands, the contour based approach will be easier for segmentation. The contour is highly dependent on the initial curve placement and is sensitive to noise in the image. The features identified from this segmentation will help to find the gestures used in Indian Sign Language. Some of these points are discussed in this paper, which is organized in the following sections. Section 2 talks about the various techniques available, their advantages and disadvantages. Section 3 deals with the method for segmenting the hand. The results obtained for inputs are discussed in section 4. In section 5, the concluding remarks are future works are discussed.

II. Literature Survey

This section is a survey about the various techniques available for segmenting the hand. The segmentation can be easier with the usage of instrumented gloves that have sensors attached with it as explained in [1], but the complex hardware structure is a drawback. The contour finds the outline of a digital image that helps in detecting various objects in an image. The edge also finds the boundary in an image, but the difference between them is that, in edges the outline is drawn based on the change in intensity between every pixel whereas the contour finds the outline of a salient region in an image. The output of an edge will result in confusion because it finds every small edge in the particular input image and not any individual object. Remove the unwanted information. Fig 1 shows the output of edge detection and contour detection. The output of contour finds only the animal, which can be found as such in the output of edge detection. The contour based method segments the image with localized energy even when the global energy

International Journal of Electronics Communication and Computer Engineering
Volume 4, Issue 3, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

corresponding to it fails. The edge-based methods usually depend on the gradient information to stop the evolution of the curve.

Fig.1. The output for edge and contour

There are various methods for contour detection. The Kass et al [2] proposed a contour detection to be efficient in which the model starts with the curve around the object and it minimizes accordingly to fit the outline of it.

Nevertheless, the drawbacks are they are highly sensitive to the initial conditions; they do not form a proper contour for complex images that have overlapping objects.

The Hui Li [3] method for contour detection is based on discrimination for color detection. The thought to include the local statistics begins with Brox and Cremers [4]. All the work proposed in this paper corresponds to the localized energy based on piecewise constant model of Chan and Vese [5]. The Osher and Sethian [6] proposed a level set method that was successful. The traditional active contour uses the spline curve model to get the boundary of an object. However, the level set formulation used in this paper uses deformable curves to find the boundary. The zero level set formulation of a smooth function is used to represent the framework. There are also other methods for contour detection [7][8].

III. Proposed Methodology

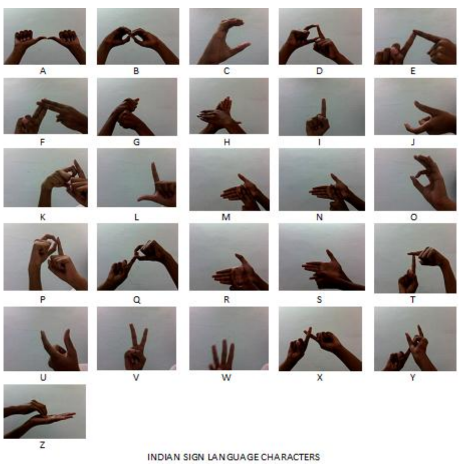


Fig 2: Indian Sign language – Characters

The Indian sign language as shown in Fig 2 uses both the hands that resemble the English alphabets themselves that makes it differ from the other sign languages. The input will be the gestures captured from the webcam directly. Moreover, the output will give an image that segments the hand region from the image. The Chan and Vese model for image segmentation is one among the most successful minimization problems that uses level set formulation and utilizing the image statistics inside and outside the curve to form the contour. Initially a curve is drawn around the object considereing the fact that the required object is in the middle of the frame.

A. Local region-based framework

The basic idea of the proposed method is to deform the contour that minimizes the given energy function to result in a desired segmentation. The local region-based framework is described in this section for the proposed method. The foreground and background are described in a local small regions because this method is not based on global region models.

As we analyse the local region it will lead to a family of local energies at every point, so we optimise the local energy at each point seperately which will minimize or maximize the computed energy in that particular local region.

Let Ω be a bounded openset on R, and the given image be represented as I: $\Omega \rightarrow R^2$. And let C be the piesewise parametrised curve represented in the zero level set. Let the region inside C be denoted as ω and the reigion outside it be $\Omega \backslash \omega$. Let c_1 be the average of the pixels' intensity inside C, and c_2 be the average of the intensity outside C.

B. The Mumford-Shah

The Chan and Vese Model uses a case in the Mumford-Shah method to evolve the cruve. It finds the pair of (u,C), where u is the piecewise smooth approximatin of the image I, and C will be the closed curve.

$$\text{Ems} = \int_{\Omega} |I(x,y) - u(x,y)|^2 + \mu \int_{\Omega/C} \text{Length(C)} + \nu.\text{Area(C)} \tag{1}$$

where (x,y) is the spatial co-ordinate of the pixels in the image and μ, is a positive integer value to smoothen the contour [9]. Minimization of the energy function is the only way to solve this problem.

C. The Chan And Vese Algorithm

The Chan and Vese is an alternative of the Mumford method. The objective of Chan and Vese algorithm is to minimize the energy funtional represented in $F(c_1,c_2,C)$ that is defined as:

$$F(c1, c2, C) = \mu.\text{Length}(C) + \nu.\text{Area}(\text{inside}(C)) + \lambda 1 \int |I(x,y) - c_1|^2 dxdy + \lambda 2 \int |I(x,y) - c_1|^2 dxdy \tag{2}$$

where I(x,y) will be the spatial co-ordinate of the pixel. And $\mu \geq 0$, $\nu > 0$, $\lambda_1, \lambda_2 > 0$ are all fixed parameters that will set by the user according to the classs of the image, which also minimizes the Mumford shah method. Here μ is the penalty on the total length of the curve. It decides on the smoothening term of the curve. To get a smoother curve, the value has to increased. The ν is the penalty on the total area in the foreground of an image. The term (I(x,y)-c_1) is propotional to the variance of the gray scale imag in the forground of an image and it measeures the uniformity in the terms of pixel intensity. The term (I(x,y)-c_2) does the same for the background of an image. The sum of these two terms will result in an uniform foreground and background region. For eg, when λ_1 is given an higher value than λ_2 the final segmentation will be more unifrom in the foreground region than the background [10].

D. Level set formulation

This problem can be solved easier by the level set formulation where the instead of solving it in terms of C. It is represented by the Lipshitz function(where C is independent of (x,y) in the zero level set formulation, $\Phi: \Omega \rightarrow R$

$$C = \partial\omega = \{(x,y)\ \varepsilon\ \Omega: \Phi(x,y) = 0\}$$
$$\text{Inside}(C) = \omega = \{(x,y)\ \varepsilon\ \Omega: \Phi(x,y) > 0\}$$
$$\text{Outside}(C) = \Omega\backslash\omega = \{(x,y)\ \varepsilon\ \Omega: \Phi\phi(x,y) < 0\} \quad (3)$$

given a contour C, Φ(x,y) is the signed distance funtion, where the value outside the curve will be zero.

To find the zero level set

1. *Length(C) for the zero level set* Φ(x,y)=0 will be

$$\text{Length}(C) = \int_{\Omega} \left|\nabla H(\Phi(x,y))\right| dxdy = \int_{\Omega} \delta_0(\Phi(x,y)) |\nabla\Phi(x,y)| dxdy \quad (4)$$

where H(x) is the Heavside funtion defined as the

$$H(x) = \begin{cases} 1, if\ x \geq 0 \\ 0, if\ x < 0 \end{cases} \quad (5)$$

and Φ(x,y) is the signed distance funtion[11].

2. Area(C) will become

$$\text{Area}(\text{inside}\ (C)) = \int_{\Omega} H(\Phi(x,y)) \quad (6)$$

Φ(x,y) is the signed distance funtion, the value inside the contour will be less than 0 [11].

3. The (I(x,y)-c_1) will bwcome

$$\int_{inside\ \ c} |u_0(x,y) - c_1|^2 dxdy = \int_{(x,y):\Phi(x,y)>0} |u_0(x,y) - c_1|^2 dxdy = \int_{\Omega} |u_0(x,y) - c_1|^2 H(\Phi(x,y)) dx\, dy \quad (7)$$

where c_1 is inside the contour, (x,y) are spatial co-ordinayes, Φ(x,y) is the signed distance funtion, the value inside the contour will be less than 0 [11].

4. Similarly (I(x,y)-c_2) will be

$$\int_{outside\ \ c} |u_0(x,y) - c_2|^2 dxdy = \int_{(x,y):\Phi(x,y)<0} |u_0(x,y) - c_2|^2 dxdy = \int_{\Omega} \frac{|u_0(x,y) - c_2|^2}{H(1 - \Phi(x,y))} dxdy \quad (8)$$

where c_2 is inside the contour, (x,y) are spatial co-ordinayes, Φ(x,y) is the signed distance funtion, the valueoutside the contour will be greater than 0 [11].

5. The avaerage intensities will be calculated using

$$c_1 = \frac{\int_{\Omega} |u_0(x,y) - c_1|^2 H(\Phi(x,y)) dx\, dy}{\int_{\Omega} H(\Phi(x,y)) dx\, dy}$$
$$c_2 = \frac{\int_{\Omega} |u_0(x,y) - c_2|^2 H(\Phi(x,y)) dx\, dy}{\int_{\Omega} H(1 - \Phi(x,y)) dx\, dy} \quad (9)$$

The funtion F(c1,c2,C) will be redefined are [11]

$$F(c_1, c_2, \Phi) = \mu \int_{\Omega} \delta_0(\Phi(x,y)) |\nabla\Phi(x,y)| dxdy + \nu \int_{\Omega} H(\Phi(x,y)) dxdy + \lambda_1 \int_{\Omega} |u_0(x,y) - c_1|^2 H(\Phi(x,y)) dx\, dy + \lambda_2 \int_{\Omega} |u_0(x,y) - c_2|^2 H(1 - \Phi(x,y)) dxdy \quad (10)$$

The curvature of the evolving curve has to identifiedusing the kappa method becaude it can find it for the spatial co-ordinates

$$\kappa(\Phi) = \frac{\Phi_{xx}\,\Phi_y^2 - 2\Phi_{xy}\,\Phi_x\,\Phi_y + \Phi_{yy}\,\Phi_x^2}{(\Phi_x^2 + \Phi_y^2)^{3/2}} \quad (11)$$

Where the above are first and second derivatives. The contour will be identified until themaximum iteration is achieved or after the exact boundary of the desired object has be found out [11] . The function has to be normalized Instead of computing the level set function on the whole image the narrow band near the contour alone can be considered to reduce computational complexity.

IV. Expermental Results

To identify the advantages of the proposed methodology the experiments were conducted with different parameters. The results for the desired outputs are given in the table I. The input will be the gesture through the webcam. An initial curve is drawn around the object as shown in Fig 3. The parameters were changed for every image and the outputs were analyzed. The desired output was obtained only when particular parameters were changed. As seen in the previous sections, the parameters in that derivation can vary within a certain range, which can be set according to the required output. The smoothening term can be increased to get a more smother contour. The normalization factor can be used for two extreme cases. When the normalization factor for the function is increased, the output is obtained at its earliest but when it is decreased the contour takes time to complete. The normalization factor can have a positive value below 0.9. In this paper the normalization factor is set to 0.45 so that even the minute change in the gesture is segmented properly. The only condition is that the object for which the contour has to be obtained should be inside the initial contour. This method can also be obtained for overlapping images, as shown in table I.

International Journal of Electronics Communication and Computer Engineering
Volume 4, Issue 3, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

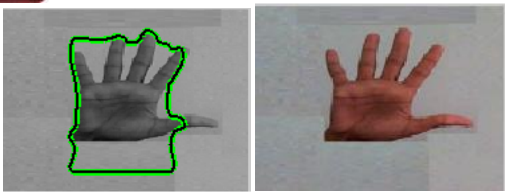
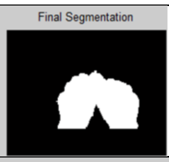
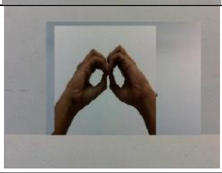
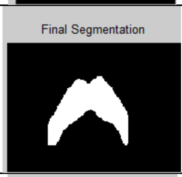
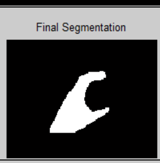
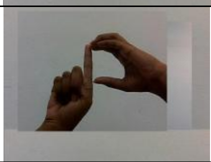
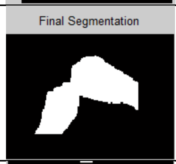
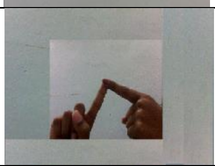
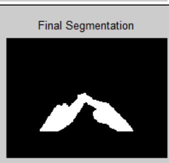
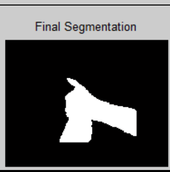
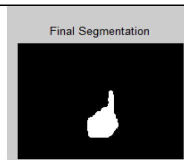
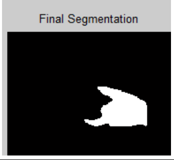
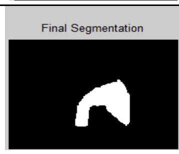
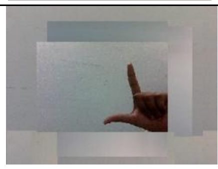
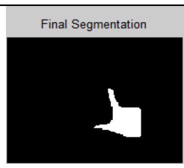
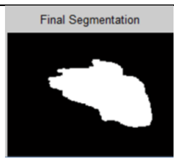
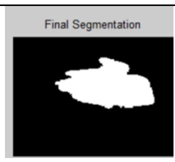
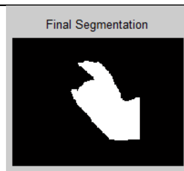
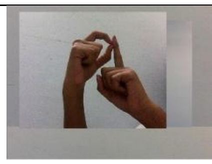
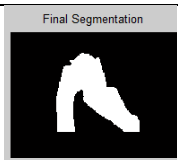
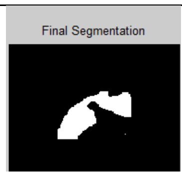
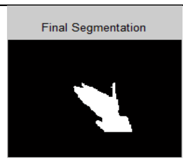
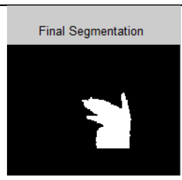


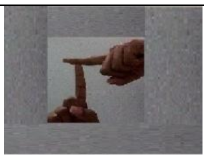
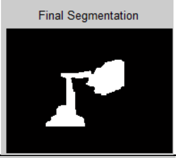
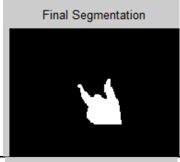
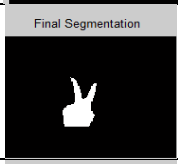
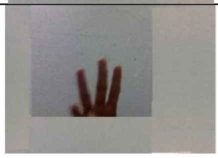
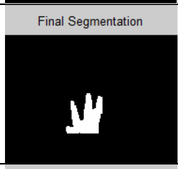
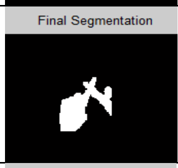
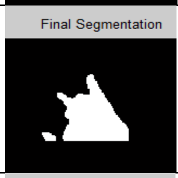
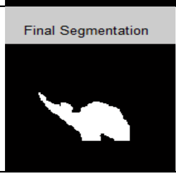
Fig.3. Initial contour

The λ_1, λ_2 decide the uniformity in the foreground and background respectively. In the proposed method they are always equal. The value of the smoothening term can vary from 0.1 to 0.9 according to the desired output. In the proposed method the value of the smoothening term is fixed to 0.2 to get the contour of the gestures. The energy is minimized for all the iterations and the desired contour is obtained.

Table I: Contour segmentation of the alphabets

ALPHABET	INPUT	OUTPUT
A		
B		
C		
D		
E		
F		
G		
H		
I		
J		
K		
L		
M		
N		
O		
P		
Q		
R		
S		

International Journal of Electronics Communication and Computer Engineering
Volume 4, Issue 3, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

T		
U		
V		
W		
X		
Y		
Z		

V. Conclusion

The Chan and Vese model identifies the contour for the desired object by energy minimization. The input image that shows the gesture is taken through the webcam, an initial curve is drawn around the desired object, considering the fact that it is in the middle of the frame, and the energy minimization function is used to deform the curve around the boundary of the object. The level set formulation is used for easier computation. The final output is obtained when the maximum numbers of iterations are achieved or when the contour is formed for the desired object. This method is one among the most flexible and powerful one for contour detection. However, there are certain limitations. They work well for homogenous intensity because it is based on the fact that a constant intensity is maintained. It is always based on the initial contour. The computational cost is very high and the time taken to complete the contour is also high.

Future Work

The features will be identified to recognize every alphabet. The neural network will be used to train and to identify the output. The limitations of this algorithm will be reduced to get a better output. This method will also be implemented for video.

Acknowledgement

We sincerely acknowledge our co team mates V.Valliammai, J.Ajith, M. Niranjan and M.S. Sai Preethy who helped us to go through till this level of the project. We sincerely acknowledge our seniors AshokGautham.M, Arunachalam R.M, Nagarjunan and Ram Rajesh for helping us to get knowledge about the project.

References

[1] Isaac García Incertis,isagar@cartif.es, Jaime Gómez García-Bermejo, University of Valladolid,jaigom@eis.uva.es, Eduardo Zalama Casanova, Hand Gesture Recognition for Deaf People Interfacing

[2] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, Int. J. Comput. Vision 1 (4) (1987) 321–331.

[3] C.M. Li, C.Y. Xu, C.F. Gui, M.D. Fox, Level set formulation without Re- initialization: a new variational formulation, in: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) 2005, vol. 1, 2005, pp. 430–436

[4] T. Brox, D. Cremers, On the statistical interpretation of the piecewise smooth Mumford–Shah functional, Proc. Scale Space Var. Met. Comput. Vision 4485 (2007) 203–213.

[5] T.F. Chan, L.A. Vese, Active contours without edges, IEEE Trans. Image Process. 10 (2) (2001) 266–277.

[6] S. Osher, J.A. Sethian, Fronts propagating with curvature-dependent speed: algorithms based on Hamilton–Jacobi formulations, J. Comput. Phys. 79 (1) (1988) 12–49.

[7] V. Caselles, F. Catte, T. Coll, F. Dibos, A geometric model for active contours in image processing, Numer. Math. 66 (1) (1993) 1–31.

[8] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, Int. J. Comput. Vision 22 (1) (1997) 61–79. [8] R. Malladi, J.A. Sethian, B.C. Vemuri, Shape modeling with front propagation: a level set approach, IEEE Trans. Pattern Anal. Mach. Intell. 17 (2) (1995) 158–175.

[9] D. Mumford, J. Shah, Optimal approximation by piecewise smooth functions and associated variational problems, Commun. Pure Appl. Math. 42 (1989) 577–685.

[10] Rami Cohen, rc@tx.technion.ac.il ,The Chan-Vese Algorithm.

[11] Shawn Lankton, *Student Member, IEEE*, and Allen Tannenbaum, *Member, IEEE*, Localizing Region-Based Active Contours.

Author's Profile

Padmavathy.S

Born on 1st of November. Currently working as an assistant professor in Amrita School of engineering. Graduated BE(EEE) in GCT(Coimbatore,India) in the year 1997. Completed ME(CSE) in same college in the year of 2000. Published papers on around 20 journals. Areas of interests are Digital Image Processing, Data Structures and C.

Shangeetha.R.K
Born on 19th of June 1992. Currently doing B.Tech(CSE) in Amrita Vishwa Vidyapeetham, Coimbatore, India. Published a paper in an International Conference. Areas of Interests are Digital Image Processing, C and DBMS. (*)Has equal rights as the first author in the paper.

Karthik.K.S
Currently doing B.Tech. (CSE) in Amrita Vishwa Vidyapeetham, Coimbatore, India. Areas of Interests are Digital Image Processing, Networks and Wireless.