ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

Tamil Sign Language Recognition using Active Shape Modeling

S. Sudha[1], S. Jothilakshmi[2], R. Soundaramani[3]

Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar, Chidambaram[1,2,3]

Abstract: Communication is the process of exchanging information, views and expressions between two or more persons, in a verbal and non-verbal manner. Hand gestures are the non verbal method of communication used along with verbal communication. A more organized form of hand gesture communication is known as sign language. Physically disabled persons like the deaf and the dumb use this language to communicate among themselves and with others. The main aim of this paper is to design a system that can modify the sign language accurately so that the less fortunate people may communicate with the outside world without an interpreter. By keeping in mind the fact that in normal cases every human being has the same hand shape with four fingers and one thumb, this paper aims at designing a real time system for the recognition of some meaningful shapes made using hands. In this work Tamil signs are synthesized by combining the texture and the shape extracted from the already trained model. K-nearest neighbor (KNN) classifier, Sequential forward and backward features are used for the synthesis.Experimental results demonstrate the effectiveness of the proposed work (recognizing efficiency 91%).

Keywords: Active shape model, Landmark points, Sign Language, Tamil Language.

I. INTRODUCTION

Communication is one of the basic abilities of human beings, who are capable of using oral and body language to express thoughts and, likewise, understand what others express. The ability to communicate allows us to share experiences and interact with each other. It helps to establish social relationships.Speech could provide a more efficient and user-friendly form of human computer interaction (HCI), where input commands are given in the form of natural speech and output is returned in the form of realistic synthetic audio-visual speech. The increasing use of the Internet has provided many new applications for multimodal HCI. In particular, on-line shopping and banking, information services and instant messengers provide alternatives to usual face-to-face dialogue. Such services are usually text-based, while those that do employ virtual charactersresort to unrealistic, cartoon-like models. Since the majority of orally-disabled people belong tothe deaf-mute category, sign language becomes the most popular tool for communication among them;there are 10 to 15 billion deaf people in India of whom 4.5% are in Tamilnadu. Tamil Sign Language (TSL) is the native language for many Tamil deaf people. The spoken language uses vocal organs to express information to others. But the deaf people use hand-movements, palm-orientation and fingers (manually), or perform facial expressions and body changes that are reflexive. Investigations indicate that the hearing impaired with lower literacy have limited reading ability, and information expressed in Sign Language (SL) is easier for them to access, since Sign Language is a natural language for the deaf.Sign Language synthesis aims to translate text to Sign Language animation demonstrated by a virtual human. The sign in a Sign Language is equivalent to a word in written language. Similarly a word in a written language is equivalent to a sentence in a sign language. This is because one requires more than one sign rather a series of signs to define a word. Sign Language as a series of gestures is one of the most natural means of exchanging information for most people. It varies from place to place. Different countries have unique sign languages, with grammar and vocabularies that are distinct from each other's. The sign language in a particular country is also typically independent of the local spoken and written language of that country. TheSign Language often has a grammar and vocabulary distinct from the spoken one. A single sign can represent different meanings in different language communities[2]. The rest of the paper is organized as follows. Section 2 deals with related work, Section 3 describes the proposed work, Section 4 describes the Methodology used, Section 5deals with Training and Classification, Section 6 and Section 7 describes the Experiments and the Results and Conclusion respectively.

II. RELATED WORK

Recognition is one of the key areas in computer-based Sign-Language communication. During the 1970's and 1980's, many researchers approached the recognition problem in a bottom-up fashion; theemphasis being on the design of filters for the detection of local structures such as edges, ridges, corners and T-junctions. The structure of an image can be described as a collection of such syntactical elements and their (spatial) relations, and such descriptions can be used as input for generic recognition schemes. Unfortunately, these recognitions are often not very meaningful. On the other hand, top-down strategies (also referred to as model-based or active approaches) for recognition were used successfully in highly constrained environments, like industrial inspection tasks. Often these methods are based on template-matching. Templates

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

incorporate knowledge about both the shape of the object to be segmented and its gray-level appearance in the image, and are matched for instance by correlation or with generalized Hough-transform techniques. But template matching, or related techniques, are likely to fail if the object and/or background exhibit a large variability in shape or gray-level appearance, as is often the case in real-life images and medical data.Two major techniques have emerged for synthesizing the movements of the features of the face, namely model-based and image-based synthesis. Model-based systems generally use computer graphics rules to deform the geometryof 3D mesh models by manipulating a set of parameters. These parameters can represent a specific variation in the features of the face, e.g. lip reading, or eyebrow rising knownas terminal analogue synthesis. For example Cohen and Massaro [34] and Parke [35], experimented with anatomical models, which in turn could act on the vertices of the mesh through a process, known as physically-based synthesis, e.g. Waters [36]. Many publications use active shape models or active appearancemodels for tracking facial point features and then inferring facial cues from them in the context of video-based automatic sign language analysis. [33,31] focus on recognizing a predefined set of facial expressions. [32,30] provide facial features for use in a Sign Language recognition framework, i.e. they integrate low-level facial features into their system to improve recognition levels.Active contours or snakes [4,5] and wave propagation methods such as level sets [6], have been evolved as new paradigms for recognition. It was their ability to reform freely instead of rigidly that spurred this enthusiasm well in the study of ridges/edges and contours. Nevertheless, such methods have two inherent limitations which make them unsuited for many medical recognition tasks. First, little priori knowledge about the shape to be segmented can be incorporated, except for adjusting certain parameters. Second, the image structure at object boundaries is prescribed by letting the snakes attract to edges or ridges in the image, or by termination of conditions for propagating waves. In practice, object boundaries do not necessarily coincide with edges or ridges. To overcome these limitations, researchers have tried out hand-crafted parametric models. An illustrative example is the work of Yuilleet al. [7] where a deformable model of an eye is constructed from circles and parabolic patches and a heuristic cost-function is proposed for the gray-level appearance of the image inside and on the border of these patches. There are two problems with parametric models. First of all they are dedicated, that is, limited to a single application. Second, there is no proof that the shape model and cost function proposed by the designer of the model are the optimal choice for the given application.Consequently, there is a need for generic recognition schemes that can be trained with examples to acquire a model of the shape of the object to be segmented (with its variability) and the gray-level appearance of the object in the image (with its variability). Such methods are prototype-based which makes it easy to adapt them to new applications by replacing the prototypes; they use the statistical techniques to extract the major variations from prototypes in a principled manner.

III. PROPOSED WORK

The shape model in ASMs is given by the principal components of vectors of landmark points. The gray-level appearance model is limited to the border of the object and consists of the normalized first derivative of profiles centered at each landmark that run perpendicular to the object contour. Shapes and objects have been modeled by landmarks, finite-element methods, Fourier descriptors and by expansion in spherical harmonics (especially for surfaces in three dimensions [11], [12]). Jain et al. [13] have presented a Bayesian framework in which templates are deformed and more probable deformations are more likely to occur. They use a coarse-to-fine search algorithm. Ronfard [14] has used statistics of object and background appearance in the energy function of a snake. Brejl and Sonka [15] have described a scheme similar to ASMs but with a nonlinear shape and appearance model that is optimized with energy function after an exhaustive search to find a suitable initialization. Pizeret al. [16] describe an object model that consists of linked primitives which can be fitted to images using methods similar to ASMs. Cootes and Taylor have explored active appearance models (AAMs) [2], [17], [18] as an alternative to ASMs. In AAMs, a combined principal-component- analysis of the landmarks and pixel values inside the object is made which allows one to generate plausible instances of both geometry and texture. The iterative steps in the optimization of the recognition are steered by the difference between the true pixel values and the modeled pixel values within the object. Sclaroff and co-workers [19, 20] have proposed a comparable method in which the object is modeled as a finite-element model. ASMs have been used for several recognition-tasks in medical images [21, 27].

This paper consists of a new type of appearance model for the gray-level variations around the border of the object. Instead of using the normalized first derivative profile, a general set of local image structure descriptors, viz. the moments of local histograms extracted from filtered versions of the images using a filter bank of Gaussian derivatives is considered. Subsequently a statistical analysis is performed to learn which descriptor is the most informative at each resolution, and at each landmark. This analysis amounts to feature selection with a k-nearest neighbor (KNN) classifier and sequential forward and backward feature selection. The KNN classifier with the selected set of features is used to compute the displacements of landmarks during optimization, instead of the Mahalanobis distance for the normalized first derivative profile.

IV. METHODOLOGY USED

A sign image can be recognized by combining the texture and the shape which is extracted from the already trained model. In this work an active shape model is trained with a set of training images. For sign to be synthesized a separate model is trained and during the synthesis the required parameters are extracted from the model and the corresponding sign to be recognized is generated. The block diagram of the proposed sign recognition system is shown in Fig.1.

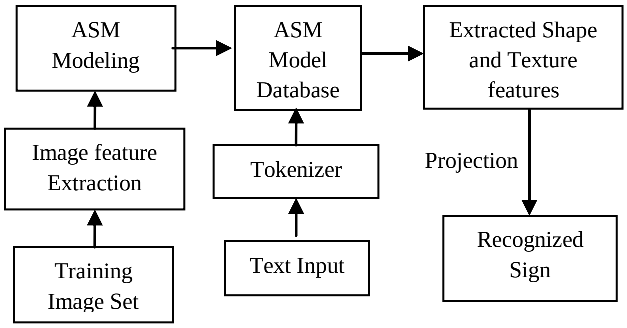


Fig.1. Block Diagram of the Proposed Sign Language Recognition System

An appearance model can represent both the shape and texture-variability seen in a training set. The training set consists of labeled images, where key landmark points are marked on each sample object. For instance, to build a model of the central brain structures in 2D MR images of the brain one needs a number of images marked with points at key positions to outline the main features. Similarly a face model requires labeled face images. Given such a set,one can generate a statistical model of shape and texture variation. The shape of an object can be represented as a vector x and the texture (or grey-levels) represented as a vector g. The appearance model has parameter c, controlling the shape and texture according to

$$x = \bar{x} + Q_s c$$
$$g = \bar{g} + Q_g c$$

Where $\bar{x}$ is the mean shape, $\bar{g}$ the mean texture and $Q_s, Q_g$are matrices describing the modes of variation derived from the training set.A sign image can be synthesized for a given c by generating a texture image from the vector gand warping it using the control points described byx.

A. Active Shape Model Matching

The Active Shape Model algorithm is a fast and robust method of matching a set of points, on a shape model with new images. The shape parameters, b for the model, along with parameters defining the global pose (the position, orientation and scale) define the position of the model points in an image, X. Each step is an iterative approach to improve the fit of the points , to an image Xinvolves first examining the region of the image around each current model point X_i, Y_i to find the best nearby match(X_i', Y_i') and then updating the parameters (t_x, t_y, s, θ, b) to best fit the model to the new found points ofX'. This is repeated until convergence.

In this work, a gray-level appearance model is described that is an alternative to the construction of normalized first-derivative profiles and the Mahalanobis distance-cost-function of the original ASM. The aim is to move the landmark points to better locations during optimization, along a profile perpendicular to the object contour. The best location is the one for which everything on one side of the profile is outside the object, and everything on the other side is inside it. Therefore, the probability that a location is inside/outside the object is estimated, for the area around each landmark separately. This classification is based onoptimal local image features obtained by feature selection and a nonlinear KNN-classifier, instead of using the fixed choice of the normalized first-derivative profiles and the Mahalanobis distance.

B. Image Features

A Taylor expansion approximates a function f around a point of interest x_0 by a polynomial of (some) order N. The coefficients in front of each term are given by the derivatives f^n at x_0

$$f(x) = \sum_{n=0}^{N} \frac{1}{n!} f^{(n)}(x_0)(x - x_0)^n$$

Derivatives of images are computed by convolution with derivatives of Gaussians at a particular scale. This motivates the use of a filter bank of multistate Gaussian derivatives to describe local image structure. Given a set of filtered images, features are extracted for each location by taking the first few moments of the local distribution of image intensities (the histogram) around each location.The most suitable choice for a window function to compute this histogram is a Gaussian, since every other choice induces spurious resolution. The size of this window function is characterized by a second scale parameter. The construction of local histograms, extracted from a Gaussian aperture function, is called a locally order lessimage.The useof histograms of responses of an image and subjecting them to a bank of filters is a standard technique in texture-analysis.There are a very few parameters to vary, like the order of the Taylor- expansion (i.e., the number of filters in the filter bank), the number of scales to consider, the number of scales to use for the local window, and the number of moments to extract from the local histograms. Which combinations are optimal for a given application and a given location remains an open question. The strategy is to compute an extensive set of features, and use feature-selection techniques in the subsequent classification stage to determine the optimal features. However, it will be better tohave$\propto > \sigma$, otherwise the histogram will be computed over a homogeneous region and will, therefore, be uninteresting. In this work, two moments are used i.e., $(m = 1,2)$, all derivatives up to second-order $(L, L_x, L_y, L_{xx}, L_{yy})$.Five inner scales $(\sigma = 0.5, 1, 2, 4, 8)$), pixels and a fixed relation between the inner scale $\sigma$and the histogram extent $\propto$of$\propto = 2\sigma$areused.

For the first moment this yields an effective scale of 1.12, 2.23, 4.47, 8.94 and 17.89 pixels respectively. The total number of feature images is 2* 6* 5= 60. Obviously the method can be extended by using more scales and higher-

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

order derivatives, higher-order moments, or by releasing the fixed relation between σandα.

V. TRAINING AND CLASSIFICATION

The next step is to specify how to construct training set from the training images, which classifier to use, and how to perform feature-selection. Consider again the optimization procedure. At each iteration, each landmark is positioned at $2k + 1$ locations along a profile perpendicular to the current object location. Obviously the image structure is different for each landmark, but the positions that are evaluated are also different for each resolution. Therefore, one has to select a distinct optimal set of features for each landmark andfor each resolution, amounting to nL_{max} feature sets. In the original ASMs the same strategy was followed: nL_{max} mean profiles and the covariance matrices S_g as they appear in (7) are computed: for each landmark, at each resolution. From each training image and for each landmark a square grid of N_{grid}X N_{grid} points is defined with an odd integer, and the landmark point, at the center of the grid. The spacing is $2^{(i-1)}$ pixels for the i^{th} resolution level.$N_{grid}$is fixed to 5, which means that for each landmark and for each resolution level, a feature vector with 60 elements is sampled at 25 points. The output of each feature vector is either inside (1) or outside (0).The landmark points themselves are considered to be inside the objects (this is an arbitrary choice). The set of training images is divided in two equal parts. This leads to two sets of samples,trainingand validation set. A K-NN classifier [38] with weighted voting is used. $k = 5$ was used and the weight of each vote is $\exp(-d^2)$, where d is the Euclidean distance to each neighbor in the feature space. Sequential feature forward selection is used to find a feature set of at the most f_{max}features. This set is subsequently trimmed by sequential feature backward selection, that is, features are removed such removal improves performance. This procedure of forward selection followed by backward selection is as or almost as effective as optimal “floating” feature selection schemes.

The resulting set is the “optimal” set of features that will be used during segmentation. After feature selection, the samples from the training and the validation set are merged and a list of selected features for each landmark and each resolution is stored. When the model is fitted to an input image, the scheme starts by computing the 60 feature images. Instead of sampling the normalized derivative profiles, the optimal feature set at each position along the profile is fed into a KNN classifier to determine the probability that this pixel is inside the object.

The objective function f(g)to be minimized is the sum of absolutedifferences between the expected probability (0 or 1) for points outside or inside the object, respectively and the predicted probability, for each point $g_i$along the profile g

$$f(g) = \sum_{i=-k}^{-1} g_i + \sum_{i=0}^{k} (1 - g_i) \text{-------------- (A)}$$

where the index along the profile , that is oriented from the outside to the inside of the object, runs from -k, to +k . This metric replaces the Mahalanobis distance from (7).

A. Training the shape model.

1) Construct shape model. Train the gray-level appearance model.
2) Compute the 60 feature images for each training image.
3) For each landmark, at each resolution, construct a set of training samples with 60 features as input and output zero or one depending on whether the sample is in or outside the object. Samples are taken from a grid around the landmark for each training image.For each training set, a KNN classifier is constructed with selected optimal features.

VI. EXPERIMENTS AND RESULTS

The recognition of sign languages in image sequences is an important and challenging problem that enables a host of human-computer interaction applications. In thiswork a new method for recognizing sign language has been proposed. Most studies on continuous sign language recognition have been done with frames obtained by processing the videos with regular/equal intervals. If a system is strong enough for processing the static gestures, then it would be the best system to process the frames obtained while processing continuous gestures.

A. Preprocessing Steps

The images are resized to 640*480 pixels to reduce memory space and execution time. For giving the edge input while training the model a binary mask of the input training images are needed. In binary mask creation,the image is read first and the dimensions are captured. The output imagesare initialised and Skin Pixels on the image are detected and the image converted from RGB to Gray scale and then into Binary scale. The results of the preprocessing stage in the recognition task implemented in the study are shown infig2.

B. Configuration of Landmarks

The first stage takes the sample image and marks the points by loading the picture on a GUI screen. Intermediate points are added to better describe the shape of the contour region in the sign image. Configurations x and x' are considered to have the same shape if they can be merged by an appropriate transformation T.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

Fig.2. Preprocessing

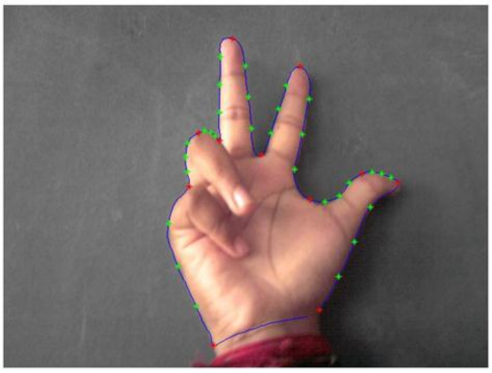


Fig.3. Image with the landmarks position

C.Shape-freetexture

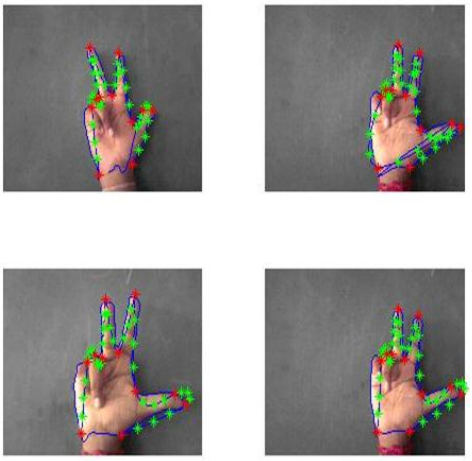


Fig.4. Results obtained for the shape-free texture**.**

In the second step,other training data are marked by contour points. The location of the contour points of the image is the same in length as the coordinates.

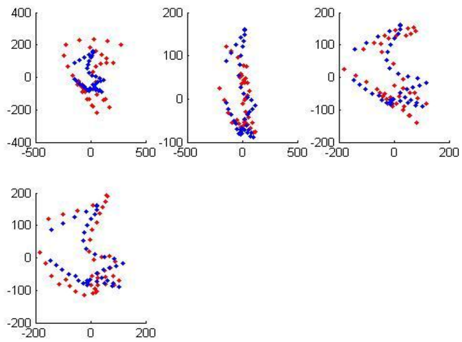
D. Testing Image

The next step marks for the area contours as shown in fig.5 and it is displayed with the help of the testing image and matched to other training sets. After approximating model points that may be related to the training sets, they are displayed as final output. In the context of sign image analysis, the hand is the object and the ASM can be visualized as a hand graph that is iteratively matched to a new sign image.

Fig.5. Mean Contour Region obtained

E. Plot the points for images

The gray world algorithm makes use of the PCA, which transforms the hand region into a data space of Eigenvectors that results from different reference hand poses. This data space is called Pose Eigen Space (PES). The reference poses are generated by means of a rotating virtual hand and are distributed equally. Before projection into PES, the monochrome images used are normalized first by subtracting the average intensity from individual pixels and then dividing the result by the standard deviation. Variations resulting from different illuminations are averaged.



F. Identification of hand outline with an active shape model

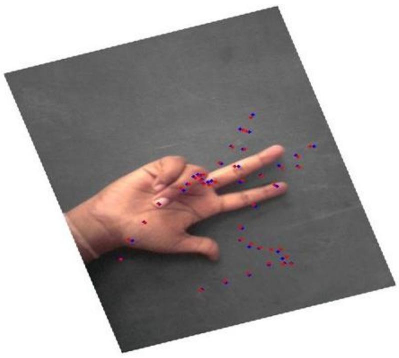


Fig.7. Results for hand outline with active shape

The extraction of hand outlines is based on an Active Shape Model (ASM), an iterative algorithm for matching a statistical model of object shape to a new image as shown in Fig.7. Though related to active appearance models,

DOI 10.17148/IJARCCE.2015.4934

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

ASMs do not incorporate any texture information. The statistical model represents the shape and possible deformations of the hand outline. For ASM initialization the hand borders must be segmented from the image as accurately as possible.

G. Hand Area Contour

The fitting process is done in two steps. First,the initial value is acquired using the algorithm, and iterated using the original ASM to find the coefficients of shape which are close to the final result. The contour of the hand the nearest points of every feature point on the contour are found. The result of the fitting process is shown in Fig. 8.

Fig.8. Resulting sign area contour in Test image

H. Recognizing the Sign Using Disparity Matching

Disparity refers to the distance between two corresponding points on the left and right of the images in a stereo pair. The sign image which is at the least distance from the resulting contour image is considered as the sign present in the test image.

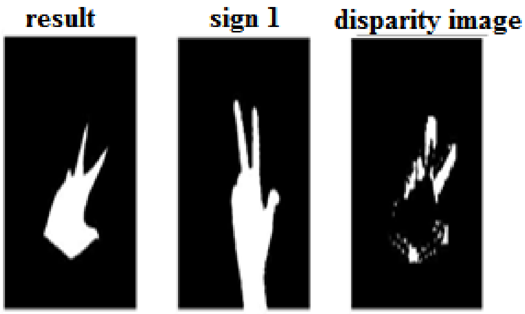


Fig.9. Similarity between the Sign-1 and the Test Image

The improvement in performance is obtained for simulated data in which the textural appearance of the image inside and outside the object is different. This indicates that the proposed method may be especially useful to segment textured objects from textured backgrounds.

VII. CONCLUSION

In the firstphase the ASM modeling and Sign recognition tasks of the sign-synthesis process have been completed. The ASM model is trained with a set of optimal features extracted from a set of training images. A set of test images are used to measure the performance of the trained ASM model by finding the sign contour region in them and then matching them with the binary masks of the sign images to recognize the sign. The correctness of the recognition process greatly depends on the contour resulting from the ASM model. On reviewing the experimental results, the performance of the ASM model trained with the optimal image features are found satisfactory.

The computational complexity of the improved ASM method is roughly 20-fold that of the original scheme. The computational burden can be reduced if the feature images are computed only at those points significant where the value changes. Using a faster classifier mechanism will also reduce computation time. Nevertheless, the algorithm still requires only a few seconds on standard PC hardware. So it is concluded that active shape models provide a fast, effective, automatic, model-based method to meet segmentation-problems in medical imaging. The new ASM method proposed in this work significantly improves on the original method through the use of an adaptive gray-level appearance model based on local image features.Experimental results demonstrate the effectiveness of the proposed work showing a recognizing-efficiency of 91%

REFERENCES

[1] T. F. Cootes, C. J. Taylor, D. Cooper, and J. Graham, "Active shape models—their training and application," Comput. Vis. Image Understanding, vol. 61, no.1, pp. 38–59,95.
[2] T. F. Cootes and C. J. Taylor, "Statistical model of appearance for computer vision," Wolfson Image Anal. Unit, Univ. Manchester, Manchester, U.K., Tech. Rep., 1999.
[3] "Statistical models of appearance for medical image analysis and computer vision," Proc. SPIE, vol. 4322, pp. 236–248, 2001.
[4] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," Int. J. Comput. Vis., vol. 1, no. 4, pp. 321–331, 1988.
[5] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: A survey," Med. Image Anal., vol. 1, no. 2, pp. 91–108, 1996.
[6] J. A. Sethian, Level Set Methods and Fast Marching Methods, 2 ed. Cambridge, U.K.: Cambridge Univ. Press, 1999.
[7] A. L. Yuille, P. W. Hallinan, and D. S. Cohen, "Feature extraction from faces using deformable templates," Int. J. Comput. Vis., vol. 8, no. 2, pp. 99–112, 1992.
[8] A. Blake and M Isard, Active Contours. Berlin, Germany: Springer-Verlag, 1998.
[9] A. K. Jain, Y. Zhong, and M. P. Dubuisson-Jolly, "Deformable template models: A review," Signal Processing, vol. 71, no. 2, pp. 109–129, 1998.
[10] T. F. Cootes and C. J. Taylor, "A mixture model for representing shape variation," Image Vis. Computing, vol. 17, no. 8, pp. 567–574, 1999.
[11] A. Kelemen, G. Székely, and G. Gerig, "Elastic model-based segmentation of 3-D neuroradiological data sets, " IEEE Trans. Med. Imag., vol.18, pp. 828–839, Oct. 1999.
[12] B. M. ter Haar Romeny, B. Titulaer, S. Kalitzin, G. Scheffer, F. Broekmans, E. te Velde, and J. J. Staal, "Computer assisted human follicle analysis for fertility prospects with 3D ultrasound," in Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 1999, vol. 1613, IPMI '99, pp. 56–69.
[13] A. K. Jain, Y. Zhong, and S. Lakshaman, "Object matching using deformable templates," IEEE Trans. Pattern Anal. Machine Intell., vol. 18, pp. 267–278, Mar. 1996.
[14] R. Ronfard, "Region-based strategies for active contour models," Int. J.Comput. Vis., vol. 13, no. 2, pp. 229–251, 1994.
[15] M. Brejl and M. Sonka, "Object localization and border detection criteria design in edge-based image segmentation: Automated

DOI 10.17148/IJARCCE.2015.4934

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

learning from examples," IEEE Trans. Med. Imag., vol. 19, pp. 973–985, Oct. 2000.
[16] S. M. Pizer, D. S. Fritsch, P. A. Yushkevich, V. E. Johnson, and E. L. Chaney, "Segmentation, registration, and measurement of shape variation via image object shape," IEEE Trans. Med. Imag., vol. 18, pp.851–865, Oct. 1999.
[17] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," in Proc. Eur. Conf. Computer Vision, vol. 2, H. Burkhardt and B. Neumann, Eds., 1998, pp. 484–498.
[18] "Active appearance models," IEEE Trans. Pattern Anal. Machine Intell., vol. 23, pp. 681–685, June 2001.
[19] S. Sclaroff and A.P. Pentland, "Modal matching for correspondence and recognition," IEEE Trans. Pattern Anal. Machine Intell., vol. 17, pp. 545–561, July 1995.
[20] S. Sclaroff and J. Isodor, "Active blobs," in Proc. IEEE ICCV, 1998, pp. 1146–1153.
[21] T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam, "The use of active shape models for locating structures in medical images," Image Vis. Computing, vol. 12, no. 6, pp. 355–366, 1994.
[22] P. P. Smyth, C. J. Taylor, and J. E. Adams, "Automatic measurement of vertebral shape using active shape models," in Proc. Br. Machine Vision Conf., 1996, pp. 705–714.
[23] S. Solloway, C. J. Taylor, C. E. Hutchinson, and J. C.Waterton, "The use of active shape models for making thickness measurements from MR images," in Proc. 4th Eur. Conf. Computer Vision, 1996, pp. 400–412.
[24] N. Duta and M. Sonka, "Segmentation and interpretation of MR brain images: An improved active shape model," IEEE Trans. Med. Imag., vol. 17, pp. 1049–1067, June 1998.
[25]G. Behiels, D. Vandermeulen, F. Maes, P. Suetens, and P. Dewaele, "Active shape model-based segmentation of digital X-ray images," in Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 1999, MICCAI '99, pp. 128–137.
[26] F. Vogelsang, M. Kohnen, J. Mahlke, F. Weiler, M. W. Kilbinger, B. Wein, and R.W. Günther, "Model based analysis of chest radiographs," Proc. SPIE, vol. 3979, pp. 1040–1052, 2000.
[27] G. Hamarneh and T. Gustavsson, "Combining snakes and active shape models for segmenting the human left ventricle in echocardiographic images," in Computers in Cardiology. Piscataway, NJ: IEEE, 2000, vol. 27, pp. 115–118.
[28] I. Dryden and K.V. Mardia, The Statistical Analysis of Shape. London, U.K.: Wiley, 1998.
[29] C. Goodall, "Procrustes methods in the statistical analysis of shapes," J.Roy. Statist. Soc. B, vol. 53, no. 2, pp. 285–339, 1991.
[30] U. Von Agris, J. Zieren, U. Canzler, B. Bauer, and K.-F. Kraiss. Recent developments in visual sign language recognition. Universal Access in the Information Society, 6(4):323-362, 2008.
[31] I. Rodomagoulakis, S. Theodorakis, V. Pitsikalis, and P. Maragos. Experiments on global and local active appearance models for analysis of sign language facial expressions. In 9th International Gesture Workshop on Gestures in Embodied Communication and Human-Computer Interaction, pages 96-99, 2011.
[32] J. Piater, T. Hoyoux, and W. Du. Video analysis for continuous sign language recognition. In Proceedings of 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies, pages 22-23, 2010.
[33] I. Ari, A. Uyar, and L. Akarun. Facial feature tracking and expression recognition for sign language. In Computer and Information Sciences, 2008. ISCIS'08. 23rd International Symposium on, pages 1-6, 2008.
[34] M. Cohen and D. Massaro. Modeling coarticualtion in synthetic visual speech. In N. Thalmann and T. D, editors, Models and Techniques in Computer Animation, pages 141–155. Springer-Verlag, 1994.
[35] F. Parke. Parametric models for facial animation. Computer Graphics and Applications, 2(9):61–68, 1982.
[36] K. Waters. A muscle model for animating three-dimensional facial expressions. pages 17–24, 1987.