Synergy between Object Recognition and Image Segmentation

Y.Ramadevi, B.Kalyani, T.Sridevi
Department of CSE , Chaitanya Bharathi Institute of Technology
Gandipet, Hyderabad, India.

ABSTRACT - Image segmentation is to partition an image into meaningful regions with respect to a particular application. Object recognition is the task of finding a given object in an image or video sequence. This paper discusses the interaction between image segmentation and object recognition in the framework of the Expectation-Maximization (EM) algorithm. Threshold is image processing technique for converting grayscale or color image to a binary image based upon a threshold value. The OTSU method is one of the applied methods of image segmentation in selecting threshold automatically for its simple calculation and good adaptation. Genetic Algorithms (GAs) tries to find structure in data that might seem random, or to make a seemingly unsolvable problem more or less 'solvable'. In this paper a synergy between Image segmentation and object recognition using EM algorithm, OSTU and GA.

KEYWORDS- EM algorithm, OSTU, Genetic Algorithm, Image Segmentation, Object Recognition.

1. INTRODUCTION

Image segmentation is the foundation of object recognition and computer vision. In general, image noise should be eliminated through image preprocessing. And there is some specifically-given work (such as region extraction and image marking) to do after the main operation of image segmentation for the sake of getting better visual effect. Two major computer vision problems, image segmentation and object recognition, have been traditionally dealt with using a strict, bottom-up ordering.

Image segmentation is to partition an image into *meaningful* regions with respect to a particular application. The segmentation is based on measurements taken from the image and might be *grey level*, *colour*, *texture*, *depth* or *motion*. The result of image segmentation is a set of segments that collectively cover the entire image.

Object recognition is the task of finding a given object in an image or video sequence. For any object in an image, there are many 'features' which are interesting points on the object that can be extracted to provide a "feature" description of the object. This description extracted from a training image can then be used to identify the object when attempting to locate the object in a test image containing many other objects.

A synergy between Segmentation and Object recognition is done using EM algorithm, OSTU and Genetic Algorithm. Experimentation was done on color and gray scale image using MATLAB 7.9. In this paper, Section 2 discusses about segmentation, and comparison of various segmentation techniques. Section 3 discusses about image segmentation and object recognition using various techniques. The implementation and results are shown in section 4. The paper is concluded in section 5.

2. SEGMENTATION

Edge-based segmentation partitions an image based on abrupt changes in intensity near the edges whereas region-based segmentation partitions an image into regions that are similar according to a set of predefined criteria. Region-based segmentation looks for uniformity within a sub-region, based on a desired property, e.g. intensity, color, and texture as shown figure 1. The difference between Region-based segmentation and edge-based segmentation is shown in Table 1.

Original Image | Region Based Segmentation

Figure 1. Region based Segmentation

Table 1. Differences Between Region-Based Segmentation And Edge-Based Segmentation

Region-based segmentation	Edge based segmentation
Closed boundaries	Boundaries formed not necessarily closed
Multi-spectral images improve segmentation	No significant improvement for multi-spectral images
Computation based on similarity	Computation based on difference

2.1 Active Contours

Active contours are popular technique for image segmentation. An advantage of active contours as image segmentation methods is that they partition an image into sub-regions with continuous boundaries. There are two kinds of active contour models according to the force evolving the contours: edge- and region-based. Edge-based active contours use an edge detector, usually based on the image gradient, to find the boundaries of sub-regions and to attract the contours to the detected boundaries. Region-based active contours use the statistical information of image intensity within each subset instead of searching geometrical boundaries as shown in Figure 2.

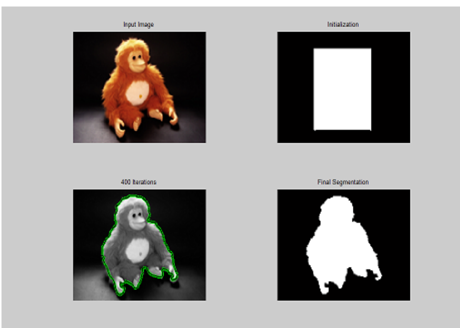


Figure 2. Active contours

3. OBJECT RECOGNITION & IMAGE SEGMENTATION

In Image Segmentation the generative models are used to decide which part of the image a model should occupy. Active Appearance Model (AAMs) is used as generative models and addresses the problem of jointly detecting and segmenting objects in images. Regarding recognition, each object hypothesis is validated based on the image area assigned to the object, as well as the estimated model parameters, which indicate the familiarity of the object appearance. On one hand, knowing the area occupied by an object is needed for the estimation of the model parameters and, on the other hand, the model synthesis is used to assign observations to the model. Since neither is known in advance, we cannot address each problem separately. We view this problem as an instance of the broader problem of parameter estimation with missing data: In our case, the missing data are the assignments of observations to models. A well-known tool for addressing such problems is the EM algorithm.

An **Expectation-Maximization** (**EM**) **algorithm** is used for finding maximum likelihood estimates of parameters in probabilistic models, where the model depends on unobserved latent variables. In order to find maximum likelihood estimate we have to find probability density function and loglikelihood.

3.1 The Expectation-Maximization Algorithm

The EM algorithm is an efficient iterative procedure to compute the Maximum Likelihood (ML) estimate in the presence of missing or hidden data. Each iteration of the EM algorithm consists of two processes: The E-step, and the M-step. In the expectation, or E-step, the missing data are estimated given the observed data and current estimate of the model parameters. This is achieved using the conditional expectation, explaining the choice of terminology. In the M-step, the likelihood function is maximized under the assumption that the missing data are known. The estimates of the missing data from the E-step are used in lieu of the actual missing data.

The EM algorithm seeks to find the MLE by iteratively applying the following two steps:

Expectation step: Calculate the expected value of the log likelihood function, with respect to the conditional distribution of z given x under the current estimate of the parameters $\theta^{(t)}$.

Maximization step: Find the parameter which maximizes this quantity.

3.2 Image segmentation using OTSU

OTSU algorithm is based on the point of image segmentation in the global threshold selection. This method of its calculation is simple, stable and effective, has been widely used still image processing toolbox MATLAB gray image as the threshold value automatically select a single standard algorithm. The OTSU method is one of the applied methods of image segmentation in selecting threshold automatically for its simple calculation and good adaptation.

In image processing, OTSU's thresholding method is used for automatic binarization level decision, based on the shape of the histogram. The algorithm assumes that the image is composed of two basic classes: Foreground and Background. It then computes an optimal threshold value that minimizes the weighted within class variances of these two classes. It is mathematically proven that minimizing the within class variance is same as maximizing the between class variance.

The thresholding techniques are categorized into six groups as:

1. Histogram shape-based methods, where histogram of image is viewed as a mixture of two Gaussian distributions associated to object and background classes, such as convex hull thresholding and peak and valley thresholding.
2. Clustering-based methods, where gray-level pixels are clustered in two classes as either background and foreground objects or alternately modeled as mixture of two Gaussians, such as iterative thresholding, clustering thresholding, minimum error thresholding and fuzzy clustering thresholding.
3. Entropy-based methods use difference in entropy between foreground and background regions, such as, entropy thresholding.

4. Object attribute-based methods; find measure of similarity (fuzzy shape similarity, edge coincidence, etc) between gray-level and binarized images, such as edge field matching thresholding and topological stable-state thresholding.
5. Spatial methods, use higher-order probability distribution and/or correlation between pixels, such as higher order entropy thresholding.
6. Local methods, calculate threshold value at each pixel based on local image characteristics, such as local contrast method and surface-fitting threshold.

The major problem with thresholding is that we consider only the intensity, not any relationships between the pixels. There is no guarantee that the pixels identified by the thresholding process are contiguous.

3.2.1 Mathematical Formulation

q1 and q2 represent the estimate of class probabilities defined as:

$$q_1(t)=\sum_{i=1}^{t}P(i) \quad \text{and} \quad q_2(t)=\sum_{i=t+1}^{I}P(i)$$

and sigmas are the individual class variances defined as:

$$\sigma_1^2(t)=\sum_{i=1}^{t}\left[i-\mu_1(t)\right]^2\frac{P(i)}{q_1(t)} \quad \text{and}$$

$$\sigma_2^2(t)=\sum_{i=t+1}^{t}\left[i-\mu_2(t)\right]^2\frac{P(i)}{q_2(t)}$$

and the class means: $\mu_1(t)=\sum_{i=1}^{t}\frac{iP(i)}{q_1(t)}$ and

$$\mu_2(t)=\sum_{i=t+1}^{I}\frac{iP(i)}{q_2(t)}$$

Here, **P** represents the image histogram. The problem of minimizing within class variance can be expressed as a maximization problem of the between class variance. It can be written as a difference of total variance and within class variance:

$$\sigma_b^2(t)=\sigma^2-\sigma_w^2(t)=q_1(t)\left[1-q_1(t)\right]\left[\mu_1(t)-\mu_2(t)\right]^2$$

Finally, this expression can safely be maximized and the solution is **t** that is maximizing $\sigma_b^2(t)$.

3.2.2 OTSU ALGORITHM

The steps of the OTSU algorithm: For each potential threshold **T**,

1. Separate the pixels into two clusters according to the threshold.
2. Find the mean of each cluster.
3. Square the difference between the means.
4. Multiply by the number of pixels in one cluster times the number in the other.

If we incorporate a little math into that simple step-wise algorithm, such an explanation evolves:

1. Compute histogram and probabilities of each intensity level.
2. Set up initial $q_i(0)$ and $\mu_i(0)$.
3. Step through all possible thresholds maximum intensity.
4. Update q_i and μ_i.
5. Compute $\sigma_b^2(t)$.
6. Desired threshold corresponds to the maximum.

3.3 Image segmentation using Genetic Algorithm

Genetic Algorithms (GAs) can be seen as a software tool that tries to find structure in data that might seem random, or to make a seemingly unsolvable problem more or less 'solvable'. GAs can be applied to domains about which there is insufficient knowledge or the size and/or complexity is too high for analytic solution.

Basically, a genetic algorithm consists of three major operations: selection, crossover, and mutation. The selection evaluates each individual and keeps only the fittest ones in the population. In addition to those fittest individuals, some less fit ones could be selected according to a small probability. The others are removed from the current population. The crossover recombines two individuals to have new ones which might be better. The mutation operator induces changes in a small number of chromosomes units. Its purpose is to maintain the population diversified enough during the optimization process

The existing GA's are founded upon the following main principles: Reproduction

1. Fitness
2. Crossover
3. Mutation

The algorithms code crossover as either a swap of one bit, which is more like a single-point mutation than a 'real crossover', or of several bits (used with genetic programming) and a distinction is made between two parents (bit strings, but called chromosomes in GA terminology) who are the identical, two different parents and single parent. Anyhow, the process has the following procedure:

1. Select two bit strings (chromosomes), or in case of the genetic programming: select a branch of each parent.
2. Cut the chromosome (or branch) at a particular location.
3. Swap the bits/branches of the two parents.

4. IMPLEMENTATION AND RESULTS

A sample grey scale image is considered for segmentation using EM Algorithm and the object is recognized as shown in figure 4.

Input Image and Segmentation of Image using EM Algorithm.

Black and White Image of the given input image.

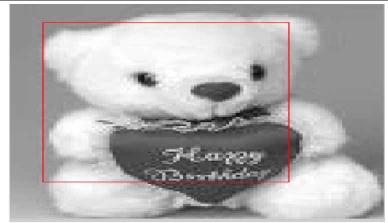


Image Recognition using EM algorithm

Figure 4. Segmented image and object Recognition using Expectation Maximization

RESULTS OF IMAGE SEGMENTATION AND RECOGNITION

The segmentation and object recognition using OSTU algorithm is shown in figure 5. Figure 6 depicts the image segmentation and object recognition using Genetic Algorithm.

The EM algorithm produced stable segmentation effect on different types of images. OTSU showed good and stable segmentation effect. Genetic algorithm exhibited normal segmentation effect on all types images.

a. Original Image

b. OTSU Segmentation

c. OTSU Segmentation

d. OTSU Segmentation

e. OTSU Segmentation

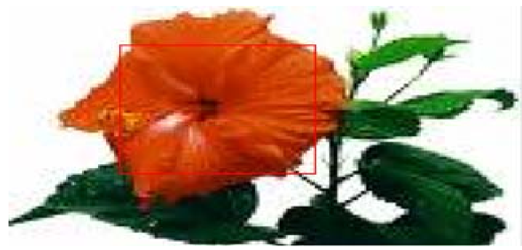
f. Recognition of the image

Figure 5. Image Segmentation and Recognition using OTSU Algorithm

a. Original Image

b. Black and White image

c. Segmentation using Genetic Algorithm



d. Recognition of the image

Figure 6. Image Segmentation and Recognition using Genetic Algorithm

5. CONCLUSIONS

The interaction between image segmentation and object recognition in the framework of the Expectation-Maximization (EM), OSTU and Genetic Algorithms main Categories and Properties of image segmentation are studied. Approaches of Object Recognition are observed. Comparison and Evaluation of different image segmentation algorithms is done. Segmentation and Recognition of an image is done by using Expectation – Maximization Algorithm in MATLAB 7.9. Segmentation and recognition of an image is done using OTSU Algorithm and Genetic Algorithm. The algorithms can be hybridized to get better results.

6. REFERENCES

1. Synergy between Object Recognition and image segmentation using Expectation and Maximization Algorithm. Iasonas Kokkinos, Member, IEEE, and Petros Maragos, Fellow, IEEE., 2009.
2. Wen-Xiong Kang, Qing-Qiang Yang, Run-Peng Liang, “The Comparative Research on Image Segmentation Algorithms,” 2009 First International Workshop on Education Technology and Computer Science.
3. V. Ferrari, T. Tuytelaars, and L.V. Gool, “Simultaneous Object Recognition and Segmentation by Image Exploration,” Proc. Eighth European Conf. Computer Vision, 2004.
4. B. Leibe, A. Leonardis, and B. Schiele, “Combined Object Categorization and Segmentation with an Implicit Shape Model,” Proc. ECCV Workshop Statistical Learning in Computer Vision, 2004.
5. A. Levin and Y. Weiss, “Learning to Combine Bottom-Up and Top-Down Segmentation,” Proc. Ninth European Conf. Computer Vision, 2006.
6. E. Borenstein and S. Ullman, “Class-Specific, Top-Down Segmentation,” Proc. Seventh European Conf. Computer Vision, 2002.
7. J. Winn and N. Jojic, “LOCUS: Learning Object Classes with Unsupervised Segmentation,” Proc. 10th Int’l Conf. Computer Vision, 2005.

Y Rama Devi received BE from Osmania University in 1991 and M.Tech (CSE) degree from JNTU in 1997. She received her Ph.D. degree from Central University, Hyderabad in 2009. She is Professor, Chaitanya Bharathi Institute of Technology, Hyderabad. Her research interests include Image Processing, Soft Computing, Data Mining, and Bio-Informatics. She is a member for IEEE, ISTE, IETE, and IE. She has published more than 25 research articles in various National, Intl’Conferences, Proceedings and Journals.

B.Kalyani received B.Sc computers from SKU, Anantapur in 2003, M.Sc Mathematics from SKU, Anantapur in 2006 and M.Tech(C.S.E) from Osmania University. Her area of interest is image processing.

T.Sridevi received B. E. from Osmania University in 1992 and M.Tech (CSE) from JNTU in 2002. She is pursuing her Ph. D. from Osmania University in CS under the guidance of Dr V. Vijaya Kumar. She has 14 years of teaching/industry experience. She joined as Assistant Professor in Chaitanya Bharathi Institute of Technology, Hyderabad, India in 2002. Presently she is an Associate Professor, Chaitanya Bharathi Institute of Technology, Hyderabad. Her research areas include Water Marking, Image Processing and Soft Computing. She is a life member of IETE. She has published more than 5 research publications in various National, Intl' conferences, proceedings and Journals.