International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

A Survey on Various Techniques of Semantic Object Extraction in Videos

Shilpi Arora[1], Aparna A. Junnarkar[2]

[1]Savitribai Phule Pune University, Progressive Education Society's,
Modern College Of Engineering, Shivaji Nagar, Pune – 411005, Maharashtra, India

Abstract: ***Video-based applications are largely being used these days. These applications include video surveillance, criminal detection and sports video analysis etc. In particular, Object –based representation consists of decomposing the video content into a collection of meaningful objects. There are many literatures present on developing technique for efficiently querying videos on their content. The content of the video can be basically defined as the objects and the iteration of the objects. So the Object identification and classification has become challenge these days. Object Extraction strategies can be broadly categorized in three parts: Manual, Semi-automatic and Automatic. In order to make object extraction technique Automatic, Algorithms are applied on extracted features of images for classification. Genetic Algorithm is used to automatically classify the objects. To find the optimal solution from genetic algorithm, it is required to maintain the larger population size which is not cost effective. There is another approach to classify the candidate object which is Fuzzy C-means based GA. In this paper, object extractions strategies are reviewed followed by discussion on Fully Automatic approaches of Object classification.***

Keywords: Semantic Object Extraction, Classification, Image segmentation, Fuzzy C-means (FCM), Genetic Algorithm (GA)**.**

1. Introduction

Image analysis is meant to extract meaningful entities from visual data. Such meaningful entities can be referred to as *semantic video objects*. The meaning of *semantic* may change according to the application. The process of identifying and tracking the collection of image pixels corresponding to meaningful entities is referred to as semantic video object extraction [1]. The main goal of this extraction process is *spatial accuracy*, that is, precise definition of the object boundary and the *temporal coherence*, that is, the property of maintaining the spatial accuracy in time. Image segmentation is used to locate objects and boundaries in image. It's a process to assign a label to every pixel in an image i.e. pixel with the same label share some amount of visual characteristics. Because of this rapid growth of available amount of video data, manual extraction of objects is not an efficient mechanism. Object extraction can be classified into two steps: Feature extraction and classification. When rules are applied on extracted features using algorithm, then it is considered as Automatic approach of classification. When a video is encountered, firstly the keyframes of videos are extracted by using keyframe extraction tools. Then to find possible objects in each frame, segmentation is performed on it. Next is to perform the classification on the candidate objects. This step is also known as Decision making [2]. Image segmentation is a very important field in image analysis object recognition, image coding and medical imaging. Segmentation is very challenging because of the multiplicity of objects in an image and the large variation between them. Image segmentation is the process of division of the image into regions with similar attributes [3].

The paper is organized as follows- Section 2, describes the various Object Extraction Strategies, Section 3, describes the Genetic Algorithm framework, Section 4, explains FCM-GA, hybrid approach for object extraction, Section 5, conclusion.

2. Semantic Object Extraction Strategies

According to the amount of human intervention in the extraction process, we can classify the different approaches to semantic video object extraction in three classes: Manual, Semi-automatic (supervised) and Automatic (unsupervised) [1]:

2.1 Manual extraction

This is the most accurate form of the extraction technique as rules representing the semantic information are applied directly by the user. Though this process is very time consuming and cannot be applied on large amount of visual data but this provides high quality classified image which assures spatial accuracy and temporal coherence.

2.2 Semi-automatic extraction

In this technique semantic information is provided by the user during some stages of the extraction process. Semi-automatic techniques can be further classified as feature-based, contour-based, or region-based.

2.2.1 Feature-based interaction

In this case, user selects some pixels belonging to the object that shows the characteristic such as color/texture properties. The pixels are used as basis for the extraction. There might be a problem of connectivity of objects.

2.2.2Contour-based interaction

Contour can be defined either as control points or as a sketch of the objects. Instead of selecting a set of points belonging to the object, user can mark its contour. This method is slower than feature-based interaction.

2.2.3Region-based interaction

In this case user interacts with the result of preliminary segmentation of the image into regions. The user marks some of these regions as corresponding to a semantic object. These are then automatically merged to obtain the shape of the semantic video object.

2.3 Fully-automatic extraction

In this approach semantic rules are applied in an algorithmic way. These rules are based on the prior knowledge of the object which are derived from specific application or class of application.

2.3.1 Chroma-keying

In this approach, objects are uniformly placed in front of colored background. The pixels are discarded with the known background color. It provides good spatial accuracy and temporal coherence. To avoid shadows, special care should be taken while lighting up the scene.

2.3.2Depth-keying

In this case, depth information is used to segment the image. Two approaches are used for the same: *Active method*, it gives good results when objects are close to the camera. *Passive method*, depth can be computed using information coming from stereo cameras. It is good for both indoor and outdoor scenes.

2.3.3A Priori Information

In this technique, a template of the object which is known a prior can be used to detect the rest of the objects in an image. Face detection, template matching, moving object segmentation are the examples of this method.

3. Genetic Algorithm- Automatic Object Extraction Approach

Genetic Algorithms (GAs) are adaptive methods which is a population-based stochastic search procedure to find exact or approximate solutions to optimization and search problems. Each chromosome in the population is a potential solution to the problem. the aim of the algorithm being to optimize a given objective or fitness function.

In biology, the evolution is defined as the change in a population's inherited trait from generation to generation. Traits are the expression of genes that are produced, copied and passed from ancestor to offsprings. In computer science, Genetic Algorithms generally involve with implementation of evolution steps as follows:

3.1 Initialization

To perform genetic operations, there should be an initial population of individuals. How it is chosen depends on the problem domain and solution strategy; the population can be generated randomly or obtained from a training data.

3.2 Selection

For each successive generation, some part of the existing population is selected for mating (reproduction). There are many different techniques to select the individuals-

- Elitist selection
- Roulette-wheel selection
- Scaling selection
- Tournament selection
- Rank selection
- Generational selection
- Steady-state selection
- Hierarchical selection

3.3 Reproduction

After selecting fittest parents for mating, they are used for reproduction processes; There are two operators for Reproduction:

3.3.1 Crossover (also called recombination)

For each selected pairs of parents, new child individual(s) are reproduced. Crossover is the process that enables gene interchange between parents so that two new individuals are reproduced that are different from parents. For crossover, two chromosomes are randomly chosen from the population. Assuming the length of the chromosome to be l, this process randomly chooses a point between 1 and l-1 and swaps the content of the two chromosomes beyond the crossover point to obtain the offspring. A crossover between a pair of chromosomes is affected only if they satisfy the crossover probability

3.3.2 Mutation

Besides mutation is not affected from parents, it provides a random/rule based gene change on the individuals.

Mutation is the second operator, after crossover, which is used for randomizing the search. Mutation involves altering the content of the chromosomes at a randomly selected position in the chromosome, after determining if the chromosome satisfies the mutation probability.

3.4 Termination

The process of reproducing new generations is repeated until a termination condition. Common terminating conditions can be listed as follows:

- A solution is found that satisfies a defined minimum criteria
- Fixed number of generations reached
- Allocated budget (computation time/money) reached
- The highest ranking solution's fitness is reaching or has reached a plateau such that successive iterations no longer produce better results
- Manual inspection
- Different combinations of the above.

4. FCM based GA– Automatic Object Extraction Approach

FCM based GA is a partially evaluated GA based on fuzzy clustering, which considerably reduces evaluation cost without any loss of its performance by evaluating only one representative for each cluster. The fitness values of other individuals are estimated from the representative fitness values. Following are the computational steps of the FCM based GA:

1) Initial Population:

Each chromosome represents a solution which is a sequence of K cluster centers. Each chromosome contain N- no of genes. For image datasets each gene is an integer representing an intensity value. In Genetic Algorithm, the population size of P is needed. In this proposed method, the FCM is run P times for generating these P chromosomes; each chromosome is of size K.

The Fuzzy C-Means (FCM) algorithm is an iterative optimization that minimizes the cost function. Let X=($x_1 x_2$,., x_n) denotes an image with N pixels to be partitioned into c clusters, where xi represents multispectral (features) data. The cost function is defined as follows:

$$J=\sum_{j=1}^{N}\sum_{i=1}^{c}u_{ij}^{m}\left\|x_j - z_i\right\|^2 \quad (1)$$

Where u_{ij} represents the membership of pixel x_j in the i^{th} cluster, z_i is the i^{th} cluster center, || || is a norm metric, and m is a constant[2].

The membership functions and cluster centers are updated by the following:

$$u_{ij}=\frac{1}{\sum_{k=1}^{c}\left(\frac{\|x_j-z_i\|}{\|x_j-z_k\|}\right)^{\frac{2}{m-1}}} \quad (2)$$

And

$$z_i=\frac{\sum_{j=1}^{N}u_{ij}{}^{m}x_j}{\sum_{j=1}^{N}u_{ij}{}^{m}} \quad (3)$$

2) Fitness Computation:

In fitness computation the pixel dataset is clustered according to the centers encoded in the chromosome under consideration. In the next step values of the cluster centers encoded in the chromosome are adjusted, replacing them by the mean points of the respective clusters. Later on, the clustering metric M is computed as the sum of Euclidean distances of each point from their individual cluster centers.

The fitness function is defined as:

$$f=\frac{1}{M} \quad (4)$$

A low value of intra-cluster spread is a feature of efficient clustering. Hence our purpose is to minimize the clustering metric M i.e. maximize f.

3) Selection:

We use the roulette wheel technique to produce the mating pool of chromosomes. The main idea of the roulette wheel technique is to associate more chance to better chromosomes.

4) Crossover:

Crossover is the next step after the selection of parent chromosomes. In this step, a new offspring is generated as a result of combining two parents.

5) Mutation:

Each iteration of the chromosome changes according to the probability. Mutation is used to perform a search over the entire range of answers.

5. Conclusion

There has been an extensive research in semantic object extraction techniques. This paper briefs about the various object extraction techniques such as manual, automatic and semi-automatic and their applicability and limitations. Genetic Algorithm is one of the Automatic approaches of object extraction which requires large population in order to classify the image and to get the optimal solution. On the other hand FCM based GA is more powerful and is cost effective technique of Object extraction as fitness value is estimated by representative of the cluster and remaining individuals in each cluster get their fitness estimated by their membership value. Video content based application area covers video surveillance, sport event, intrusion detection, video-on-demand systems. Hence it is more important to segment an image and locate the objects and boundary accurately in order to get the deeper understanding of videos.

References

[1] Andrean Cavallaro, Touradj Ebrahimi," Object-based video: extraction tools, evaluation matrics,and applications" Signal processing institute, swiss federal institute of technology(EPFL).

[2] Yilmaz, TurgayOBJECT EXTRACTION FROM IMAGES/VIDEOS USING A GENETIC ALGORITHM BASED APPROACH" January 2008.

[3] Amiya Halder, Soumajit Pramanik, Arindam Kar, " Dynamic Image Segmentation using Fuzzy C-Means based Genetic Algorithm" International Journal of Computer Applications (0975 – 8887), Volume 28– No.6, August 2011

[4] K. Dilpreet Kaur#1, Yadwinder Kaur*2," Intelligent Medical Image Segmentation Using FCM, GA and PSO" (IJCSIT) International Journal of computer Science and Information Technologies, Vol. 5 (5) , 2014, 6089-6093

[5] Omid Jamshidi and Abdol Hamid Pilevar," Automatic Segmentation of Medical Images Using Fuzzy c-Means and the Genetic Algorithm" Journal of Computational Medicine, Volume 2013 (2013), Article ID 972970, January 2013

[6] Yakup Yildirim, Adnan Yazici," Automatic Semantic Content Extraction in Videos Using a Fuzzy Ontology and Rule-Based Model" IEEE TRANSACTIONS ON

KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 1, JANUARY 2013.

[7] P.Pavithra, Dr.N.Uma Maheswari," Automatic Semantic Rule Based Coordination for Content Extraction in Videos Using Fuzzy Ontology" International Journal of Innovative Research in Science, Engineering and Technology, Volume 3, Special Issue 3, March 2014

[8] Changsheng Xu, Jian Cheng, Yi Zhang, Yifan Zhang, Hanqing Lu," Sports Video Analysis: Semantics Extraction, Editorial Content Creation and Adaptation" JOURNAL OF MULTIMEDIA, VOL. 4, NO. 2, APRIL 2009

[9] 1Fagbola, T. Mathew, 2Babatunde R. Seyi. 3Oyeleye, C. Akinwale," Image Clustering using a Hybrid GA-FCM Algorithm" International Journal of Engineering and Technology Volume 3 No. 2, February, 2013

[10] Mr.A.Julius Ceaser1, Mr.S.Rajesh2," RBM: Rule Based Lexical Meaning Extraction in Video Data Using Fuzzy" A.Julius Ceaser et al, International Journal of Computer Science and Mobile Applications, Vol.2 Issue. 1, January- 2014,

Author Profile

Shilpi Arora received the B.E degree in Computer Science and Engineering from IITM, R.G.P.V., Bhopal in 2006. She is now pursuing M.E. degree in Computer Engineering from P.E.S.'s Modern College of Engineering, Savitribai Phule Pune University, Pune and her area of interest is Image Processing