Genetic Algorithm for Inter-frame Region Object Temporal Correlation in Binary Partition Tree

Arief Setyanto[#*1], John Charles Woods[#2], Mohammed Ghanbari[#3]

[#]*Computer Science and Electronics Engineering, University Of Essex*
Wivenhoe Park, Colchester United Kingdom CO4 3SQ

[2]woodjt@essex.ac.uk

[3]ghan@essex.ac.uk

[*]*Informatics Engineering, STMIK AMIKOM Yogyakarta, Indonesia*

[1]asetya@essex.ac.uk , arief_s@amikom.ac.id

***Abstract* - Video contains rich information in the spatial and temporal domains. A single video shoot consists of a number of frames which are generally composed of similar objects which may have changed in position and size. The ability to recognize regions in the current frame is an important task in video analysis, but their subsequent recognition in the next frame receives little attention. This research utilizes a region based binary partition tree (BPT) as the content representation; object searching is conducted inside the binary tree and not in the original pixel domain. This work was conceived from the postulate: if an object exists inside a binary tree partition for a given video frame, can a corresponding branch be found in the BPT of the next frame? This is a difficult problem with a one to many and a many to one mapping and requires an innovative solution in the form of a Genetic algorithms (GA). GA's are ideally suited to this problem as the start and end conditions are known. This research proposes and achieves temporal correlation in region based BPTs using a genetic algorithm.**

***Keyword* –Image Segmentation, Region Based Image analysis, binary partition tree, Genetic Algorithm, inter frame video analysis, region based video analysis, temporal correlations**

I. Introduction

Video sequences are rich in information, which is present in both the temporal and spatial domains. Video contains distinct objects with temporal activity along with accompanying audio content. Video can be divided into a series of scenes which contain several consecutive frames which have similar object content. These frames and associated objects can be analysed in the temporal domain. A series of frames in a single scene are composed of some similar objects which may change in position and size. Basically a frame is an image that consists of an *n x m* picture of elements (pixel) each of which has an intensity value. Processing data in the pixel domain is ill posed and can lead to high computational cost. Hence we perform our analysis in the region domain.

Region based image analysis was originally proposed to deal with the drawbacks of pixel analysis. In doing so images are segmented into small homogenous regions based on features such as colour, texture and spatial frequency. Once pre-processing has been performed the computation cost is considerably reduced since the number of homogenous regions is much less than the number of pixels. However in most cases the result of the application of a pre-segmentation algorithm is an over-segmentation with a huge number of meaningless small regions. In order to have semantically meaningfull regions merging algorithms are applied to adjacent regions. Merging algorithms inspect every adjacent region, if they have similar features they will be merged. The history of the merging process is recorded in a binary tree, as was proposed by Selembier (2000)[1].

Genetic algorithms (GA) are search procedures inspired by biology and the workings of natural selection which was proposed by John Holland in 1960[2]. Many implementations of this algorithm can be found in the image processing field. GA's are particularly useful when an exhaustive search for the solution is expensive in terms of computational cost. In our implementation the start and desired end conditions are known therefore making this problem particularly suited to a GA solution.

The purpose of this research is to temporaly correlate objects between succesive frames in a video.Temporal correlation is important in video analisys for purposes such as: keyframe identification, object tracking, activity recognition and cognition. A series of frames in a single scene generally contain the same objects. Therefore an object in the previous frame is likely to appear in the current frame and in future frames. The problem we propose is finding them in the corresponding binary trees. The object may move slightly or be subject to rotation, scaling or warping.

The research question this work identifies is as follows: If given a single target branch which contains a number of region nodes of a BPT representing a salient and/or semantic object, where are the corresponding branches on the BPT of the next video frame? This is not an easy problem to solve as the branch in one frame may map to many branches in the next frame or vice versa. This problem is well suited to a GA solution.

The organisation of this paper is as follows. Firstly, several previous works in the BPT and the GA are presented. Secondly, region based analysis, image segmentation, region merging and binary the partition tree are discused in Section 3. The description of our algorithm is presented in Section 4. The next Section talks about the results and the conclusions and future work is discusssed in Section 6.

II. Related Work

A collection of work can be found which employs region based Binary Trees (BPT). Selembier et al (2000) proposed the use of a pre segmentation, merging and a binary tree representation [1]. An implementation of colour based segmentation, region merging along with a user interface was developed by Lu Huihai et al in 2007[3].

A region which has marked differences from the previous frame indicates motion within the scene i.e. objects are moving or a pan is taking place. By frame differencing Dorea et al proposed an algorithm to recognize an object in a video sequence [4]. The BPT was extended into a Multi dimensional Binary Partition Tree (MBPT) by Ghanbari et al (2010) to convey not only color but also texture edge and motion in order to produce more robust segmentation results[5].

Genetic Algorithm (GA) has been used by many researchers in image and video processing. A GA is used for image retrieval using local similarity in [6]. Many image enhancement and image segmentation work takes advantage of the GA e.g. [7]. The Genetic algorithm is successfully implemented on shape based object recognition based on Fourier descriptors in [8]. An interactive genetic algorithm (IGA) is used on low level image properties; colour, texture and edge description as a basis of machine classification with a user in the loop in [9]. The GA is employed for emotion based video scene retrieval using an interactive genetic algorithm in 2007 [10].

The main contribution of this work is to propose a novel application of the GA in a region based binary partition tree to identify temporal correlation between different frames in a video. Although there are a number of previous works examining the BPT, generally they work on single frames only (with the exception of Dorea et al who use motion vector to make a trajectory tree based on the BPT [4]).This work uses a region based model in the BPT and a genetic algorithm to successfully identify correlation beetween region branchs in succesive frames; the first time this has ever been proposed.

III. Binary Partition Tree

Binary partition tree generation consists of a number of steps: pre segmentation, region labelling, a region adjacency graph (RAG), tree representation and region merging (tree simplification). Pre segmentation is a process by which the image is divided into many small regions based on colour homogeneity. Several thousand regions will be the result even for a CIF image. In most cases the resultant regions are salient but semantically meaningless. Each region has several neighbouring regions that are entered into a region adjacency graph. These small regions will then become the input for the region merging process that aims to join regions based on their colour similarity.

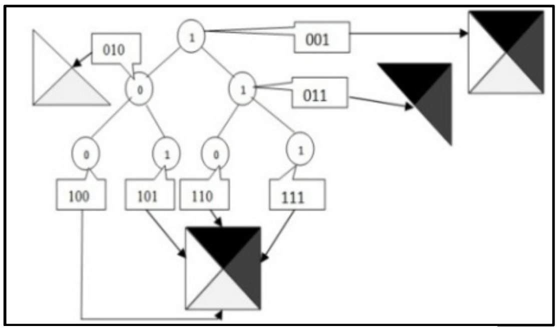

Figure 1 Chromosome Representation

The region merging order is a function of the colour distance between regions. The BPT is constructed to record the history of the region merging process; from the initial partition (as the result of pre-segmentation step) until the root of the BPT is achieved [1]. Let us say that an image is represented by a number (n) of regions Ri in the BPT where i denote the index of the regions. The region model $M(Ri)$ *is* the colour average of the region. The region area denoted by $A(Ri)$ is a number of pixels in region Ri. Each step of the merging algorithm looks for the most similar colour average, the merging order for a pair of adjacent regions Ri and Rj is defined by the following expression:

$$O(R_i, R_j) = A(R_i)\left\|M(R_i) - M(R_i \cup R_j)\right\|_2 + A(R_j)\left\|M(R_i) - M(R_i \cup R_j)\right\|_2 \tag{1}$$

Where $\| \quad \|_2$ denotes L2-norm.

A pair of adjacent regions which have the smallest colour difference will be merged first. The process of merging is recorded in the binary partition tree. Once regions are merged, a new single region is created which has a new area which is the sum of all its descendants. The new colour is the average of all descendants along with a recomputed centroid etc.. Therefore a new region will be the parent of the two direct descendants. The descendent with the smallest area is drawn the left and the larger to the right below the parent. The merging process continues until the root of the binary tree is reached where a single region represents the average of the entire image.

IV. Proposed Genetic Algorithm Region Temporal Correlation

There are three main steps in this proposed approach. The first task is to divide the video into a series of frames. The second stage was described in section 3 and the final step, the genetic algorithm uses the BPT from the current and the next frame as the start condition and target respectively.

In order to make the genetic algorithm run, some data structures and rules must be applied to the chromosome coding, fitness function, crossover and mutation. The problem space is represented in a binary partition tree therefore a binary chromosome representation is selected. The first 1 digit represents the root of the binary tree. The following digits represent the left leaf where the value is 0 and the right leaf where the value is equal to 1. The number gene in a

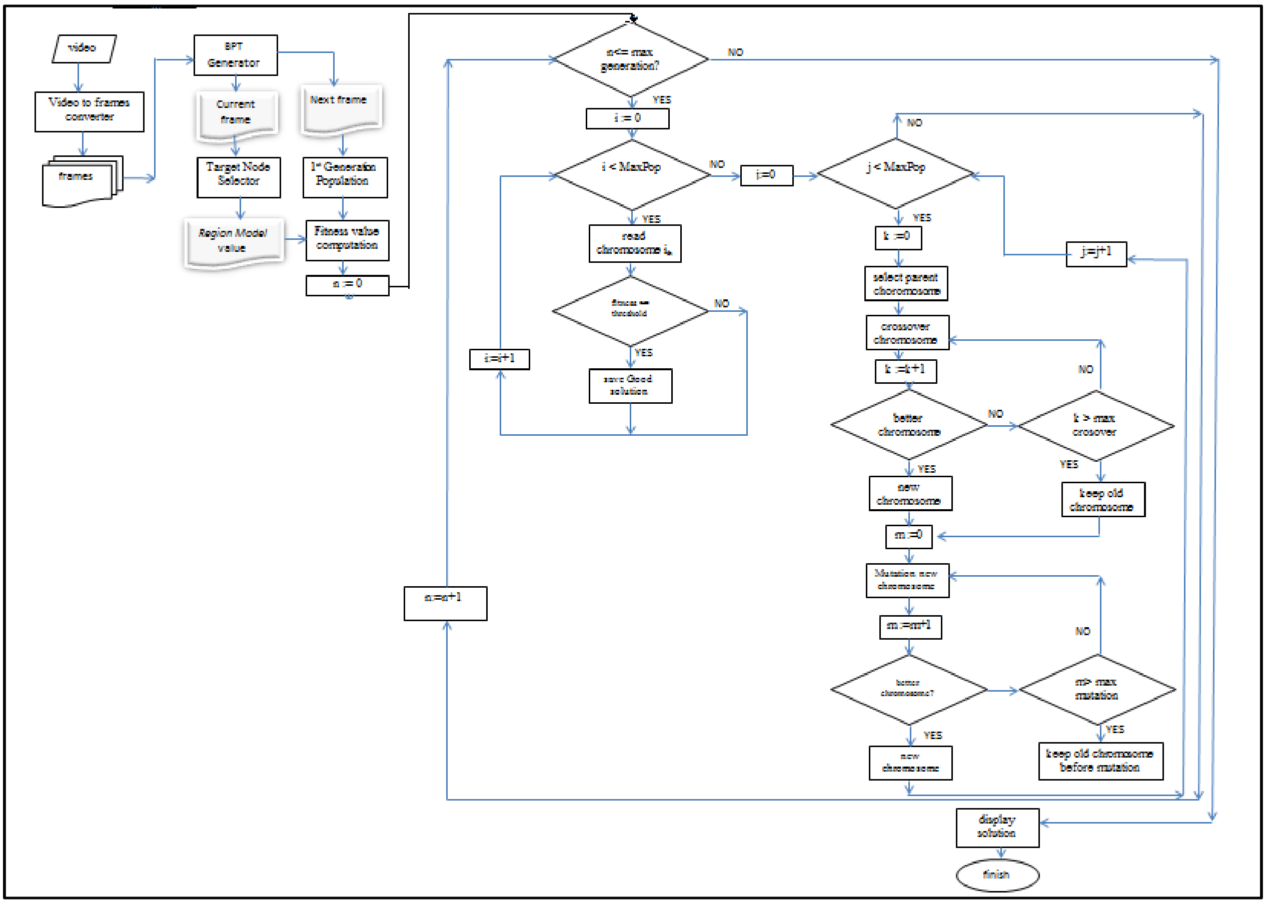

Figure 2 Proposed Flow Chart

chromosome is based on the level inside the binary tree however some genes are valued 0 for the most significant bit. For example in figure 1 because the binary tree consists of three levels the chromosome is coded as 3 binary digits. For a real image the number of levels can tree vary from tens to hundreds of levels.

The fitness function plays a central role in genetic algorithms. A fitness function provides a measure of similarity between a target and the candidate solution. Each chromosome is considered as a solution candidate and compared with the target. Based on figure 1, each chromosome represents one node in the BPT and a region area in the original image. In this research, provided that each region target and solution candidate has some statistical data (e.g. pixel intensity), it is compared according to the fitness function φ *is* defined with the following equation

$$\varphi\big(F_t(Ri), F_{t+1}(Rj)\big) = \begin{cases} 1000, & a = 0 \\ \frac{999}{|a|}, & otherwise \end{cases} \qquad (2)$$

$$where\ a = \ F_t(M(R_i)) - F_{t+1}\big(\mathrm{M}(\mathrm{R_j})\big)$$

F denotes frame, t denotes index of $the\ t_{th}$ frame in a video sequence. R_i denotes region i_{th} of frame F_t and , R_j denotes region j_{th} of frame F_{t+1}. $F_t(M(R_i))$ denotes region R_i in frame F_t while $F_{t+1}(M(R_j))$ denotes region R_j on frame F_{t+1} .

The initial population is unique and randomly generated in the solution space. The lowest leaf is a very small region and there are a lot of regions at the lowest level. The randomly generated chromosomes are engineered to select larger regions at higher levels nearer the root first to produce faster solutions.

At each iteration of the GA the quality of the chromosome relative to the target gradually improves. Crossover and mutation are two operators which aid the solution and play an important role in modifying the chromosome in order to search for non-local solutions. Crossover is an attempt to improve the fitness of the chromosome genes by combining the genes of two high ranking chromosomes. A roulette mechanism is chosen in our algorithm. Each time 2 chromosome parents are chosen, the crossover mechanism is applied until a better fitness value is achieved compared to both of their parents or a maximum crossover is achieved. In the case of maximum crossover the process is aborted and the original chromosome survives to the next generation.

Mutation is another operation which achieves fitter chromosomes in the subsequent generations. This operation is done every time a random value is greater than the mutation threshold. Mutation will execute until a better chromosome is achieved or the maximum number of mutation attempts (m) is reached. The mutation rate parameter is an empirical value in this algorithm. Mutation is aborted when the fitness value falls below that of the original chromosome. If this is the case, the original chromosome will live on in the next generation. Mutation can be achieved either by random selection of a gene to be changed from 1 to 0 or 0 to 1or a bitwise operation shifting left or right. The flowchart of the proposed algorithm is illustrated in figure 2.

V. Experiment Result

In our experiments the genetic algorithm is configured according to table 1 below. The results of the first experiment are shown in figure 3. A branch is selected as the target and represents the helmet object of the foreman to the left side of the BPT. The solution is found in the next frame on the right hand side of the BPT. The solution is found on the 3rd iteration of the GA. The solution is a single branch which starts from the circled node in the right side BPT.

Table 1 Genetic Algorithm Parameter for the experiment

No	Parameter	Value
1	Population Size	100
2	Crossover rate	0.7
3	Mutation rate	0.1
4	Max Generation	40
5	Min fitness	500
6	Max crossover attempt	10
7	Max Mutation attempt	10
8	Max solution branch	100

The second experiment is conducted to test the ability of the proposed algorithm to deal with solutions that are across several branches. In the real world this happens for example when an object is covered by another object or an object is broken into smaller pieces.

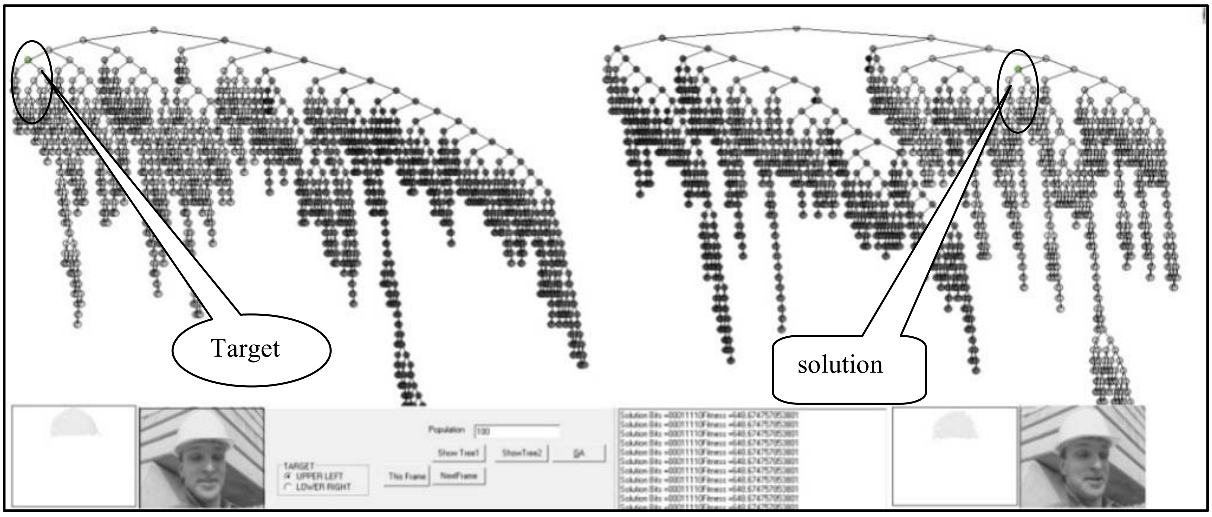

Figure 3 Second Exeperiment Result one to Many Braches Target

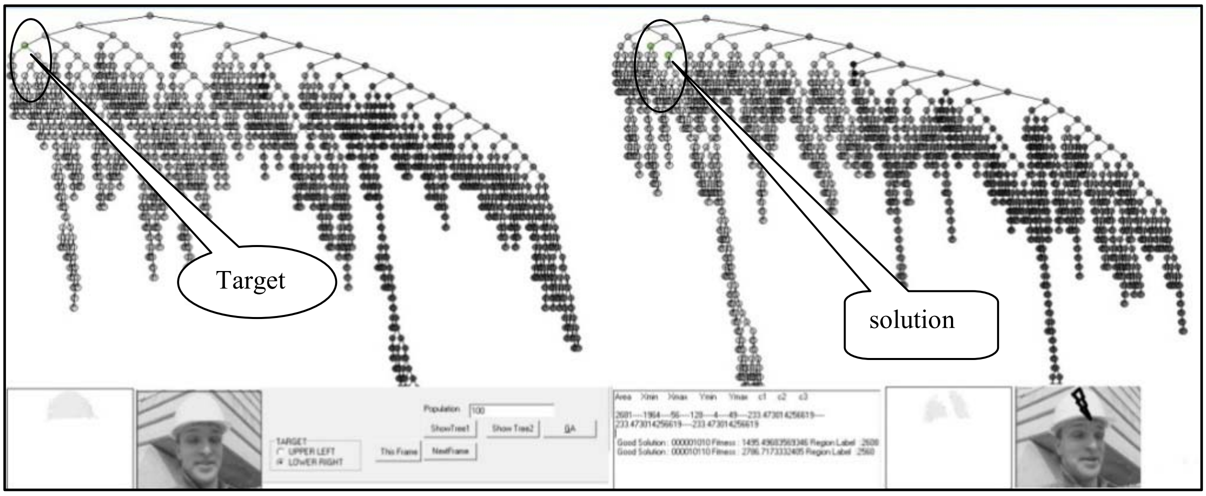

Figure 4 First Exeperiment Result one to one Branch Target

The result of the second experiment is depicted in the figure 4. A single branch on the left side BPT that represents a helmet is selected. There is a small difference in the second frame, the object has become divided into two separated areas but the GA has still found a good solution. The solution is found in two different branches which are node numbers 2608 and 2560 with their leaves. This experiment shows that the algorithm can deal with occlusion.

VI. Conclusion

This paper presents the application of a genetic algorithm to solve the problem of temporal correlation of objects inside a binary tree representation. This problem is identified and robustly solved in this paper and represents the first attempt to solve an otherwise ill posed problem. The algorithm can successfully identify regions inside a binary partition tree and find their best match according to a cost function in subsequent frames. The algorithm is robust and can cope with objects either within a single branch or separated into a number of branches. Further the measure of the fitness function provides a measure of the quality of the object match. This algorithm can also find a target object in the presence of translation, rotation and occlusion.

Future work based on this algorithm will include: Event recognition in a video shoot and key frame identification. The quality of the initial pre segmentation can be improved by the inclusion of texture, edge and other features in the region modelling.

Acknowledgement

This work has been supported by Directorate General of Higher Education (DHGE) Scholarship, Indonesian Ministry of Education. The authors would like to thank Huihai Lu for the legacy code previous research in the multimedia group the University of Essex.

References

[1] Philippe Salembier and Luis Garrido, "Binary Partition Tree as an Efficient Representation," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 9, no. 4, pp. 561-576, 2000.

[2] Michael Affenzeller, Stefan Wagner, Stefan Winkler, and Andreas Beham, *[Genetic Algorithm and Genetic Prgramming:modern concept and practical Application]*, P Borne et al., Eds. New York: CRC Pess, 2009.

[3] Huihai Lu, John C Woods, Mohammed Ghanbari, "Binary Partition Tree for Semantic Object Extraction and Image Segmentation," *IEEE Trans. On Circuit and Systems for Video Technology*, vol. 17, no. 3, pp. 378-383, 2007.

[4] Camilo C Dorea, Montse Pardàs, and Ferran Marques, "Trajectory Tree as an Object-Oriented Hierarchical Representation for Video," *IEEE Trans. on Circuit and Systems for Video Technology*, vol. 19, no. 4, 2009.

[5] Ghanbari Shirin, John C Woods, and Simon M Lucas, "BPT Using Multi-Dimensional Information for Semi-Automatic Content Retrieval," in *Conference for Visual Media Production*, 2009, pp. 169-175.

[6] Zoran Stejic, Yafusumi Takama, and Kauro Hirota, "Genetic Algorithm based relevance feedback for image retrieval using local similarity patterns ," *Information Processing and Management* , vol. 39, pp. 1-23, 2003.

[7] Mantas Paulinas and Andrianus Usinkas, "A Survey of Genetic Algorithms Application For Image Enhancement and Segmentation," *Information Technology and Control*, vol. 36, no. 3, pp. 278-284, 2007.

[8] M Sarfraz, Mehmood El Hasan, and M Iqbal, "Object recognition using Fourier Descriptor and Genetic Algorithm," in *International Conference of Soft Computing and Pattern Recognition* , 2009, pp. 318-323.

[9] Chih-Chin Lai and Ying-Chuan Chen, "A User-Oriented Image Retrieval System Based on Interactive Genetic Algorithm," *IEEE Transaction On Instrumentation and Measurement*, vol. 60, no. 10, pp. 3318-3325, 2011.

[10] Hun-Woo Yoo and Sung Bae Cho, "A Video Scene Retrieval With Interactive Genetic Algorithm ," *Multimedia Tools Application*, vol. 34, no. 3, pp. 317-336, 2007.