OBJECT REPRESENTATION USING COLOUR, SHAPE AND STRUCTURE CRITERIA IN A BINARY PARTITION TREE

Christian Ferran Bennström and Josep R. Casas

Signal Theory and Communications Department, Technical University of Catalonia UPC-Campus Nord, C/ Jordi Girona, 1-3, 08034, Barcelona E-Mail: {cferran,josep}@gps.tsc.upc.es

ABSTRACT

Binary Partition Trees (BPTs) are a well known technique used for region-based image representation and analysis. BPTs are usually created as a result of a merging process based on homogeneity properties, such as colour or motion. In this paper, we present a BPT creation technique based on a general merging algorithm, where the homogeneity criteria are neither low-level (pixel oriented, intra-region), nor high-level features (object oriented, semantics), but a combination of several criteria including regionbased and structural features such as shape and partial-inclusion. We are thus combining intra-region homogeneity (e.g colour-based) with inter-region homogeneity (structural), with the long term aim of bridging the gap in region-based image analysis from low-level features to a higher level interpretation of the image by the intermediate description of the image structure (which we call "syntactic visual features"). Syntactic visual features are geometric relationships among regions based on shapes and the spatial configuration of neighboring regions in the image and can be found by structure analysis (or syntactic analysis). The aim of this work is to present how the addition of combined pixel-region and structural features leads to better object BPT representation.

1. INTRODUCTION

The Binary-Partition-Tree (BPT) is a region-based image representation, usually obtained from a segmentation procedure. This region-based image representation allows performing image filtering, segmentation, information retrieval, visual browsing, etc. Image analysis techniques based on the BPT representation aim to obtain every object in the image as a single node of the tree, where ideally its children nodes are parts of the object.

The segmentation procedure is performed by iteratively merging pairs of neighbour regions according to certain homogeneity properties. Object homogeneity properties are used to merge regions in a given order (merging order), i.e. the regions with the most similar colours are merged first (order relation). Therefore, the definition of object is based on an homogeneity criterion and is directly related to the merging order. The notions of *merging order*, *region model and merging criterion* and its application for the creation of BPTs are explained in [1] and are briefly presented in section 2. This iterative process is bottom-up: starts either at the pixel level (every single pixels is assumed to be a region) or at the region level (using an initial over-segmented partition) and continues by merging regions until it gets a single region representing the whole image support. In this procedure regions representing different parts of the objects progressively merge into single regions (nodes in the tree) ideally representing every object in the image. We propose the following homogeneity properties defining an object: colour, motion, shape and structure.

Colour and motion have been studied in [2] with good results, however these homogeneity criteria have some drawbacks for object extraction. Colour cannot be used to characterise as a single region complex objects composed of several colour-homogeneous parts. Motion can only be used to characterise an object when object and background have different motion models, which is not always the case. Eventhough the segmentation algorithm is based on a Region-Adjacency-Graph (RAG) when the homogeneity criterion is colour or motion, the merging cannot assess structural region relationships for complex objects, such as inclusion. Two different approaches are possible for region-based image analysis to extract objects from an oversegmented partition at this stage. One approach is to introduce heuristics or specific object information based on a model of the object to be extracted or recognised as a single region. In [3], pre-defined object shapes are sought among the contours represented across the different levels of a BPT. A second approach consists in extracting meaningful features from the available regions without restricting the analysis to specific objects. In [4], a prove of concept has been carried out concluding that colour-based criteria have limitations in a general context and that structural features (such as partial-inclusions among regions) help to overcome this limitation in some particular cases.

The aim of this paper is to present some examples validating the use of combined pixel-region and structural features. For this purpose we present results based on the combination of colour, shape and structure. It is important to notice that structural regionrelationships or syntactic features are defined and computed over more than one single region, and that the syntactic features proposed are general and non-specific to any particular object (such as colour homogeneity is not specific). The notions related to syntactic image analysis were described in section [5].

The paper is organised as follows: Section 2 presents the general merging algorithm based on a combination of colour, shape and structure. The BPT creation and interpretation described in 3 is used to compare image analysis results based on individual features and on combined approaches in section 4. Finally, we will draw our conclusions and mention the plans for future work in section 5.

This material is based upon work partly supported by the IST program of the EU in the project IST-2000-32795 SCHEMA and by the grant TIC2001-0996 of the Spanish Government.

2. REGION-ORIENTED SEGMENTATION

The work presented in this paper is an extension introduced to include syntactic features in the General Merging Algorithm described in [2]. Syntactic features are defined over regions. Therefore, in a first stage the original image is partitioned with a colourbased segmentation algorithm as described in [2]. This algorithm is used to provide an over-segmentation of the image with N regions, typically $N \in [10, 100]$. It is important to avoid the merging of regions belonging to different objects at this stage. This is because the second stage is also a merging stage, and will not introduce any new contours. Therefore, the final segmentation will just be a subset of the contours represented in the partition resulting from the first stage. This partition is used as input for a second stage, where colour, shape, syntactic feature or the combination of them is used as merging criteria. The result of the second stage is a new partition with M regions, where $M \in [1, N]$. In this paper, we focus in each individual homogeneity criterion by separating them in order to analyse the influence of colour, shape and structure in the segmentation process.

Next section presents the second stage algorithm.

2.1. Graph-based region merging algorithm

Graphs are widely used for image analysis and representation. Indeed, they offer a rich and compact representation for structural relationships. In this paper we are interested in spatially region configurations, which are very well synthesised using the RAG. A RAG is a set of nodes and links. Nodes represent regions in the image, and links connect each pair of neighboring regions. RAGs are suited for an spatial structured description of the object parts. The merging algorithm is applied on pairs of neighbourhood regions by evaluating the homogeneity criteria. Moreover, each node models the feature of the region and each link models the result of merging the two nodes it connects. Merging of neighboring regions is done in the RAG by a three step procedure 1) removing the link connecting those regions, 2) merging the two associated nodes in order to create a new merged region with the model specified in the link and 3) updating the links to the neighboring regions.

According to [2], every concept will be associated to a different data structure resulting in the following entities:

- **Region model** \mathcal{M}_R : concept related to the region (*R*) representation.
- **Merging order** $O(R_i, R_j)$: concept related to the region homogeneity definition, that is, the likelihood that two regions have to be merged.
- Merging criterion $C(R_i, R_j)$: concept related to the termination criteria.

All RAG nodes are firstly initialised with their corresponding region model. Then, links are placed in a hierarchical queue and extracted following the priority given by the merging order. The merging criterion is boolean and states that two nodes will be merged only if the criterion is true. The termination criterion is, in this paper, the desired number of final regions (M) and it is therefore omitted in the description of the visual features. When two nodes are merged, the queue is updated. This is an iterative procedure. Note that the merging order defines which regions should be merged first. Therefore, the introduction of higher level primitives for the merging process, such as syntactic criteria, has to be done by expressing them properly through the merging order. When the desired object can be modelled by colour, shape or syntactic homogeneity criteria, the segmentation may result in a single region representing the object at some stage of the merging process. The following visual features have been used with the aim of assessing object homogeneity properties:

2.2. Colour-based model

Segmentation in the colour space is performed using the following values within the *general merging algorithm*,

- \mathcal{M}_{R_i} : Region is modelled by the mean colour value of the pixels belonging to the *i*-th region (R_i) .
- $O(R_i, R_j)$: Using a zero-order model the colour homogeneity order is defined as follows $O_{colour}(R_i, R_j) = ||M_{R_i} - M_{R_iUR_j}||_2 + ||M_{R_j} - M_{R_iUR_j}||_2, ||.||_2$ is the \mathcal{L}_2 norm and A_{R_k} is the area of $R_k, k = i, j$.

Note that only colour information is taken into account, therefore, starting from individual pixels, this merging process yields partitions whose regions are homogeneous in colour (flat zones). Merging order is normalised to the interval [0, 1] by dividing by the maximum variation in gray level 256^2 .

2.3. Shape-based model

Shape could be added weighting O_{colour} , as in [2], with the region area. However shape is in this paper represented by the relative area of neighbour image regions and introduced as an independent criterion.

 \mathcal{M}_{R_i} : Region is modelled by its area A_{R_i} .

 $O(R_1, R_2)$: The Merging order is defined as

$$O_{area}(R_1, R_2) = 1 - \frac{\min(A_{R_1}, A_{R_2})}{\max(A_{R_1}, A_{R_2})}$$

This homogeneity criterion is based on relative areas aims to merge regions with similar area. This feature will never lead to semantic entities, however it is of interest in addition to more complex criteria. By definition, this criterion is normalised to [0, 1].

2.4. Syntactic-based model

The partial-inclusion homogeneity criterion was described in [4] This criterion is computed as the percentage of a region R_i included into the convex-hull of the neighboring region R_j . For this purpose the region is modelled by its boundary, see [6], allowing fast computation of the region convex-hull. The notions of merging order and criteria are as follows,

 \mathcal{M}_{R_i} : The region is modelled by its boundary (position function).

 $O(R_i, R_j)$: The merging order is the percentage of R_i partiallyincluded in the convex-hull of R_j ($CH(R_j)$), given by

$$O_{qinc}(R_i, R_j) = 1 - \frac{A_{CH(R_j) \cap R_i}}{A_{R_i}}$$

Note that for each two neighbour regions, the minimum merging order of $O_{qinc}(R_i, R_j)$ and $O_{qinc}(R_j, R_i)$ have to be considered. The lowest merging order is reached when a region is totally included into the other, as opposite to adjacent regions (0% of partial-inclusion) which have the highest value.

2.5. Combined, colour, shape and syntactic-based model

For object extraction, one may want to consider together the three defined criteria: colour, shape and syntactic features, and to allow the combination of the previous merging orders during the iterative merging process.

 \mathcal{M}_{R_i} : General model including colour, area and region contour.

 $O(R_i, R_j)$: Colour, area and partial-inclusion homogeneity computed as a weighted sum,

$$O_{combined} = \sum_{n} \omega_n O_n(R_i, R_j)$$

Where O_n is O_{colour} , O_{area} and O_{qinc} , which are obtained using the previous expressions and $\sum_n \omega_n = 1$.

The aim of combining the different criteria is to merge, at each iteration, the most *meaningful* pair of regions for the defined criteria.

To illustrate colour, area and partial-inclusion homogeneity, a merging process is performed using each property individually. The original image is shown in Fig. 1(a) and the contours of the partition resulting from the first stage (N = 8) can be seen on Fig. 1(b). This is the initial partition for the second stage. Each row of Fig. 2 shows 6 iterations of the 3 different merging processes using colour, area and partial-inclusion. For example, Fig. 2 shows, in the first row, that the initial iteration for the second stage using colour similarity is merging regions 4 and 5, whereas using area similarity regions 1 and 2 are merged initially, and using partial-inclusion regions 7 and 8 are the first to be merged. For each criterion the merging process is performed in the same way, and can be observed moving downwards each column of images.



Fig. 1. (a) Original image. (b) Initial partition Contours (N = 8).

The cost of merging two nodes of the tree during the merging process is defined as the instant merging order. This cost has been tracked for all the BPT mergings in the second stage and is shown in Fig. 3. In (a) the individual cost curves are represented while in (b) the combined costs are shown. Each curve represents the cost at each iteration associated to merging regions with each individual homogeneity criterion. As we can see, after a colour-based merging, regions are modelled by its average and, therefore, the cost of a new merging increases smoothly. However, area-cost curves tend to abruptly increase because the regions with similar area are first merged and at the end of the process the merging applies to the regions with very different areas are merged. Partial-inclusion cost is unstable because after merging two regions the resulting region tend to be convex and therefore it is not overlapping any other region convex-hull. When combined, the overall cost tends to be more robust if the weight of the area is small.

In this section we presented the extension of the *general merging algorithm* with a new region model, merging order and merging criterion within a syntactic approach. This is the necessary base for undertaking the combined BPT-based object representation.



Fig. 2. (a) Colour, (b) Area and (c) Partial-Inclusion mergings.



Fig. 3. (a) Individual costs. (b) Combined costs.

3. BPT-BASED OBJECT REPRESENTATION

We will take advantage of the BPTs capability of tracking the sequence of mergings performed. The BPT construction is based on the merging order and the region model, whereas the merging criterion is not analysed because mergings are performed until a single node (root of the tree) represents the whole image support (M = 1). The merging algorithm presented in section 2 with a trivial colour, shape or structure merging criterion leads to a BPT representation. The BPT presents a compromise between representation accuracy and processing efficiency. That means that only a subset of all possible mergings in the RAG are represented in the BPT, i.e. only those mergings selected by the priorities set in the merging order are actually done.

Within the BPT framework, leaves of the tree are the regions belonging to the initial colour partition. The remaining nodes represent the regions resulting from the successive mergings according to the homogeneity criterion. For example, under a colour criterion, the algorithm will favour those mergings of nodes representing regions with similar colour. Therefore, the BPT will be composed of nodes representing sets of pixels with similar colour, and those nodes will be our potential object candidates, disregarding objects composed of several non-homogeneous colour parts. When homogeneity is the area, those nodes representing regions with similar area will be priorized. Potential object candidates present in the BPT will then be nodes whose region of support has different area than their neighboring regions. In a similar way, using partial–inclusion homogeneity, will result in a BPT with nodes representing object candidates which tend to be convex.

Next section presents some results comparing single homogeneity criterion and combined criteria for object representation.

4. RESULTS

Fig. 4 presents an original image from the America's Cup oversegmented with N = 40 regions generated with a colour-based segmentation, in the first stage, such as the described in [2]. As we can see, several regions can be merged in the second stage in order to obtain the left boat. However, this is not possible with the right one, because not all the boundary of the right boat are present in the initial partition.



Fig. 4. (a) Original image. (b) Initial partition contours (N = 40).

Fig. 5(a) shows the BPT associated to the colour-based merging process for the second stage. The colour homogeneity criterion merges regions in the sea with regions in the boat. The image in Fig. 5(b), corresponds to the region represented by the black node on the right. The white areas in the image, correspond to regions represented by the two bold nodes on the left of the BPT.



Fig. 5. (a) Colour-based BPT and (b) right black node in the tree.

The combination of colour, shape and syntactic features have been used to build the BPT represented in Fig. 6(a). In this case the homogeneity criteria have been weighted with 0.2, 0.3 and 0.5, respectively. With these weights, a region included (or partiallyincluded) into another region with similar area and colour will be merged before any other pair of neighbour regions. In Fig. 6(b), the region represented by the black BPT node corresponds to the boat. This illustrates that the use of combined homogeneity criteria is able to generate nodes in the BPT which are better object candidates overcoming the limitations of individual homogeneity criterion.



Fig. 6. (a) BPT with combined criteria. (b) Black BPT node.

5. CONCLUSIONS

In this paper we have extended the *general merging algorithm* to allow the combination of intra-region and inter-region homogeneity criterion for the creation of BPTs. Moreover, some promising results have been presented. However, there are several issues which remain as topics for future research. For example, the initial partition size, the merging order coefficient weights, their combination, their evolution along the merging process or the criteria definition to make them increasing. Also, the evaluation of this representation with object extraction tools, such as in [3]. Therefore, we plan to extend this work with more syntactic features and other merging order functions.

6. REFERENCES

- L. Garrido. "Hierarchical Region Based Processing of Images and Video Sequences: Application to Filtering, Segmentation and Information Retrieval." *PhD Thesis UPC*, Barcelona, Spain, April 2002.
- [2] L. Garrido and P. Salembier. "Region based analysis of video sequences with a general merging algorithm." *In European Signal Processing Conference (EUSIPCO)*, Rhodes, Greece, September 1998.
- [3] O. Salerno, M. Pardàs, V. Vilaplana, F. Marqués. "Object Recognition Based on Binary Partition Trees". *In IEEE Int. Conference on Image Processing (ICIP 2004)*, Singapore, October 24-27, 2004.
- [4] C. Ferran Bennström, J.R. Casas. "Binary-Partition-Tree Creation using a Quasi-Inclusion Criterion". In Eighth Int. Conference on Information Visualisation (IV'04), London, England, July 2004.
- [5] Rafael C.Gonzalez and Michael G. Thomason. "Syntactic Pattern Recognition. An introduction." Addison-Wesley Publishing Company, Inc, 1978. ISBN 0-201-02930-8
- [6] P.J. van Otterloo. "A Contour-Oriented Approach to Shape Analysis." *Prentice-Hall*, 1991.