Chapter 7

Conclusions

7.1 Contribution

Our objective in this thesis has been the discussion of region based processing of images and video sequences using hierarchical representations. Tree based structures have been chosen in front of graph-bases structures for the representation since the former are inherently hierarchical and allow implementing efficient and complex techniques on it due to its fixed structure. The processing of these hierarchical structures has been based on pruning techniques, that is, techniques that remove several subtrees of the structure based on an analysis algorithm applied on the nodes of the tree structure. Moreover, the processing algorithms developed in this thesis have been restricted to a framework in which the input and the output of the processing are pixel based representations. Let us now briefly recall the main contributions of this thesis.

Two representations have been developed in our work: the *Max-Tree*, and the *Binary Partition Tree*. The former was created to address the efficient implementation of anti-extensive connected operators, ranging from classical ones (for instance, area filter) to new ones (such as the motion filter). The latter was developed initially to overcome some of the drawbacks imposed by the Max-Tree, but has demonstrated to be a structure useful for a rather large range of applications.

The *general processing scheme* in which our algorithms are enclosed is divided in two stages: in the first stage, the tree representation is constructed using the image or video sequence, whereas in the second step, the tree is processed and analyzed to decide how the tree has to be pruned.

The process of constructing the tree representation from the pixel based representation can be viewed as the process where the abstraction from pixel to region is performed. In the case of the Max-Tree, construction is based on structuring in the level sets of a gray-level image, whereas in the case of the Binary Partition Tree, the *General Merging Algorithm* has been

used. The tree construction is, from our experience, the most computationally expensive task. *Efficient algorithms* based on the intensive use of hierarchical queues have been developed in order to enable fast construction of the tree.

Pruning techniques applied on the Max-Tree lead to anti-extensive operators, whereas self-dual operators are obtained on the Binary Partition Tree if the tree is created in a self-dual manner. In this work, pruning techniques have been targeted on filtering (connected operators), segmentation and content based image retrieval techniques. In the case of filtering and segmentation, the pruning strategy is based on performing an analysis of the nodes of the tree by measuring a specific criterion on each of the associated regions. Examples of available criteria discussed in this work are based on area, contrast or *motion*. Other criteria, such as the *marker & propagation* based or *rate-distortion* based segmentation, have also been studied. The decision on the elimination or preservation is usually taken on each node by means of a simple threshold. A solution to the *non-increasingness* of some criteria based on an optimization algorithm has been presented and discussed. Furthermore, as discussed in this work, the analysis and decision steps are very efficient in terms of time and memory consumption. As a result, segmentation and filtering processing techniques, based on tree representations, may be used in an intensive manner.

The Binary Partition Tree has also been used for content based image retrieval. In fact, this tree structure offers a way to represent the image at different levels of resolution. As a result, the description can be done at different scales of resolution. The description has been done using only low level descriptors, namely color and geometry descriptors. *Single and multiple region retrieval* has been developed. From our experimental results, the quality and relevance of the retrieved results depend on whether a good segmentation has been obtained, which is one of the main difficulties in image analysis. However, as pointed out in the future research directions, multiple region query may be used to overcome this problem.

7.2 Future research

Within the differents parts discussed in this thesis the following research work may be devised:

Creation

- Max-Tree: as it has been seen in Sec. 3.4.2 the Max-Tree representation may have a high number of nodes for a fixed image. For certain applications and frameworks it may be interesting to have a tree with a reduced number of nodes. Techniques that allow constructing the tree with a reduced complexity seems an interesting problem to tackle. These techniques may be based, for instance, on optimization strategies that minimize the number of nodes of the tree given the distortion between the original image and its associated tree representation.

- Binary Partition Tree: in the work developed up to now, motion information between two frames is used to obtain more meaningful nodes (see Sec. 4.3.3). Instead of using two frames of the video sequence, a motion analysis through the whole sequence for each region may improve the creation of the Binary Partition Tree. Other types of criteria that may help creating the tree in a more robust way should also be studied.

Processing

- Filtering and segmentation: in this work the analysis algorithm is mainly based on the assessment of a criterion independently on each node and the use of a threshold to decide which nodes have to be removed or preserved. New methods based on the analysis of several nodes at a time for processing may be thus devised.

  New types of filters using a combination of time and geometry descriptors may be devised. For instance, one may think of a filter that removes from a sequence all small regions whose global life span is less than a fixed threshold on the number of frames. These type of filtering may be interesting for content based coding applications.

- Content based image retrieval: the results of the content based retrieval engine may be enhanced by adding more descriptors to describe each region. Similarity based on fuzzy sets, as discussed in Sec. 6.6.7, may also improve the retrieval results.

  Up to now, the hierarchy of the tree has not been really exploited to code in its nodes the region descriptors in an efficient manner. Efficient coding of descriptors is obtained if descriptors can be computed in a recursive manner (see Sec. 6.4). The recursivity has been exploited to code the color histogram in an efficient manner. But, how does shape or texture behave under the union of two regions ? In other words, given the shape or texture descriptor of two neighboring regions, is it possible to compute or, at least approximate, the corresponding descriptor of the union of both regions ?

  Multiple region search may overcome some of the problems posed of the tree creation through segmentation. In fact, the Binary Partition Tree is created by means of a segmentation algorithm and thus it is difficult to ensure that complex regions are represented as a single region in the tree. In this case single region query would not work properly since the latter method needs to have the objects of interest to be represented as nodes in the tree. Future search work may study how to use multiple region search to be able to search for a fixed single object whose region of support has been divided in several nodes during the tree construction: the user may only define the whole region of support to search for and it would be the task of the search engine to decompose properly the query into subregions to perform the search.

  Relevance feedback also is an interesting topic of future research. In particular, relevance feedback may be used to restruct the structure of the trees included in the database

in order to ensure that relevant objects are represented as nodes in the tree (object formation by learning [100]). A second topic of interest in relevance feedback is to study the possibility of using some of the relevant objects marked by the user as query to improve the results, in addition of properly updating the weights.

- Non-pruning operations: this work has focused on pruning operations. Further research work may focus on operations on nodes which do not lead to pruning operations. For instance, in [71] a method to restructure the tree in order to get a representation that reflects more clearly the image structure is presented.

As can be seen, there are still many different topics to be covered. However, the main objective of all of them is to enable easier manipulation and acess to the vast amount of information available today.