# REGION BASED SEGMENTATION USING THE TREE OF SHAPES

*Juan Cardelino, Gregory Randall*\*

IIE, Universidad de la República
Montevideo, Uruguay
{juanc,randall}@fing.edu.uy

*Marcelo Bertalmío, Vicent Caselles*

Dep. de Tecnologia, Universitat Pompeu Fabra
Barcelona, Spain
{marcelo.bertalmio,vicent.caselles}@upf.edu

## ABSTRACT

The tree of shapes is a powerful tool for image representation which holds many interesting properties. There are many works in the literature that use it for image segmentation, but most of them use only boundary information along the level lines. In many real images this is not enough to achieve a good segmentation, and region information must be introduced. In this work we present a novel region-based segmentation algorithm using the tree of shapes. The approach taken consists in the selection of relevant level-lines according to region based descriptors computed from their interior. We describe a region using the histogram of its features and we select interesting regions by identifying parts of the tree with an *homogeneous* histogram. The main contribution of this work is the joint use of histograms and suitable metrics between them, with the powerful representation of the tree of shapes. This allows us to handle complex region models and thus improves on previous works which were only able to deal with piecewise constant models. We validate our approach with real images and we obtain results which are favorably compared with some well known related approaches.

## 1. INTRODUCTION

Many authors (see [1]) claim that in natural images, the significative edges correspond with segments of level-lines. In addition, the use of level lines provides a complete representation of the image, in the sense that it can be completely re-

---

G. Randall is also with the Institute for Mathematics and its Applications (IMA), Universiy of Minnesota, Minneapolis, USA.

constructed from its level lines. This representation is contrast invariant and generates a structured representation of the image known as the tree of shapes [2].

Let us mention several works in the literature that use level lines for image segmentation. The notion of topographic map as boundaries of level-sets was proposed in [1] and a simplification of it was performed by choosing level-lines joining $T$-junctions. In [3] Morel et al. define the notion of *significative* level line (a well contrasted level-line) and propose a parameter-less algorithm to compute those level-lines. In a later work, Cao et al. [4] improved the model of Morel et al. [3] by using a local concept of *significative* level-line. Finally, Cao [5] proposes an algorithm that selects level lines using the *good continuation* principle, which favors the detection of smooth curves. In [6] Pardo proposes a segmentation algorithm that selects those level lines matching certain criteria: contrast, smoothness, unions and area.

Perhaps the approach most related to ours is the work by Ballester et al. [7]. They propose the selection of the level lines that minimize the well known Mumford-Shah functional restricted to the tree of shapes. Finally, in [8] the authors use an alternative and independent formulation of the tree of shapes, the so-called *Monotonic Tree*. They define structural elements (some subsets of the tree) from which they extract certain features and classify them according to those features. Afterwards, these elements are grouped in the feature space in order to obtain an interpretation of the image.

The rest of the paper is organized as follows. Our algorithm is described in section 2. Section 3 shows the experimental results. Finally, in section 4 we give some conclusions and future work.

## 2. OUR ALGORITHM

Despite of the aforementioned works, little effort was carried out towards the selection of level lines using region based information. To the best of our knowledge the only work that deals with this problem is [7]. However in that work the image model is piecewise constant and only gray
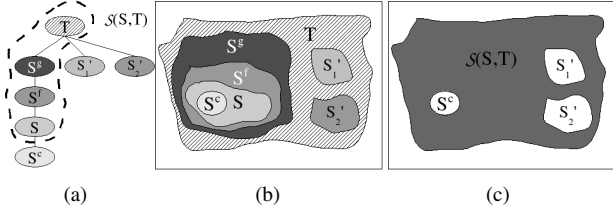
**Fig. 1**. Definition of terms used. (a) Tree of Shapes. (b) Original gray-level image. (c) Pixels of the segment $\mathcal{S}(S,T)$ (dark gray).



(a) Stable segments $\lambda = 0.003$.

(b) Our result, after merging $\mu = 0.09$.

(c) Ballester et al. $\gamma = 1600$.

(d) Cao et al.

**Fig. 2**. Comparison of results over the *Ram* image.

level information is used. The approach taken here consists in the selection of relevant level-lines according to region based descriptors computed from their interior. We describe a region using the histogram of its features (gray-level or color), and we select interesting curves by identifying segments of branches[1] with an *homogeneous* or *stable* histogram. We consider normalized histograms as an estimation of the probability density function (PDF) of the region. In order to determine when a region is homogeneous we use standard dissimilarity measures on its PDF. The main contribution of this work is the joint use of histograms and PDF metrics between them with the powerful representation of the tree of shapes. The use of PDFs and their dissimilarity measures allows us to handle complex region models and thus improves on previous works which take piecewise constant approximations for regions. In addition, the use of region information allows us to model complex images where boundaries are not well defined. We validate our approach with real images and compare the results with well known related approaches.

### 2.1. The tree of shapes

Before we describe the proposed algorithm, we review some basic concepts and notation, for further details we refer to [1, 2]. Given an image $I : \Omega \subset R^2 \to R$, we define the lower and upper level-sets of $I$ as:

$$L_\delta = \{x \in \Omega : I(x) \leq \delta\} \tag{1}$$

$$U_\eta = \{x \in \Omega : I(x) \geq \eta\} \tag{2}$$

According to Mathematical Morphology, level-sets give a complete and contrast invariant representation of the image. Indeed, we can reconstruct the image from (1) or (2) by:
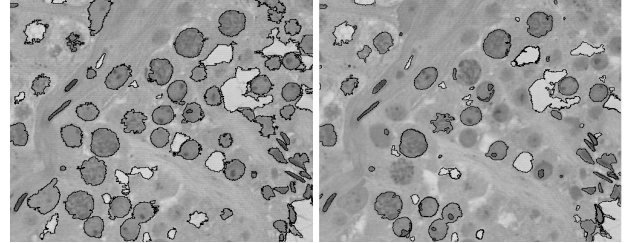
$$I(x) = \inf \{\delta : x \in L_\delta\} = \sup \{\eta : x \in U_\eta\} \tag{3}$$

Observe that upper(lower) level-sets are ordered by inclusion: if $\alpha < \beta$ then $L_\alpha \subset L_\beta$ and $U_\beta \subset U_\alpha$. Moreover, the family of connected components of all the upper(lower) level-sets has a tree structure. These two trees

---

[1] A branch is a set of nodes going from a leave to a root of the tree, a segment is a subset of a branch (see fig. 1).
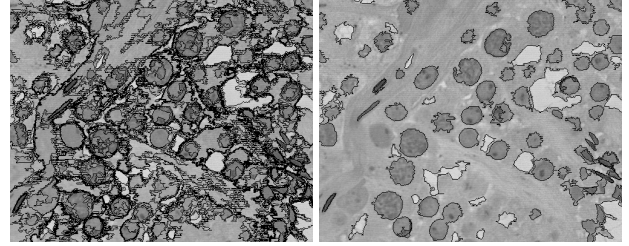
can be merged into a single one, the tree of shapes of an image [9]. In order to construct the tree the authors define the basic notion of shape: a shape is a connected component of an upper(lower) level-set where we have filled-in its holes. Thus any shape can be described by its boundary, which is called a level-line. In [9] Monasse et al. show an efficient algorithm to compute the tree of shapes, the so-called Fast Level Set Transform (FLST).

We denote a shape by $S$ and its child, parent and grandparent by $S^c, S^f$ and $S^g$ respectively (see fig. 1). We define a segment $\mathcal{S}(S,T)$ between two shapes $S$ and $T$, with $S \subset T$, as $\mathcal{S}(S,T) = T \setminus (\{\bigcup S_i' : S_i' \cap S = \emptyset\} \bigcup \{\bigcup S_i^c : S_i^c \subset S\})$, and the histogram of $I$ restricted to the segment as $h_{\mathcal{S}(S,T)}$.

The tree of shapes permits to encode the family of level lines of an image. However, not all the level lines are equally relevant and the main task is to simplify them, to obtain a smaller set of the relevant ones [4, 7, 6]. Our approach will be to look for regions determined by level lines with homogeneous or stable properties in its interior. In order to do so, we must define in which sense a region is homogeneous. Our algorithm has two stages: a first pass, where we traverse the tree computing stable segments (with respect to an homogeneity criterion) of the tree. In the second pass, we merge segments according to the same homogeneity criterion.
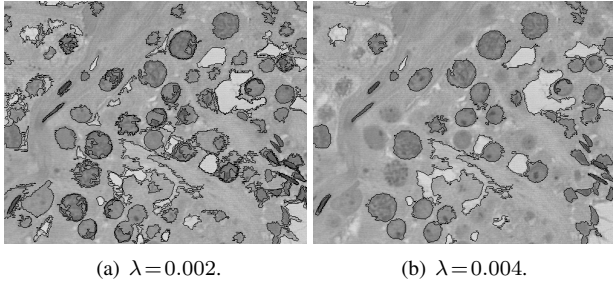
2

(a) $\lambda = 0.002$.  (b) $\lambda = 0.004$.

**Fig. 3**. Effect of the parameters on the results. *Ram* image.



(a) *Church* $\lambda = 0.004$.  (b) *Plane* $\lambda = 0.007$.

**Fig. 4**. Results of our algorithm over some tests images.

### 2.2. Identifying stable segments

In order to detect stable segments we start by identifying all the leaves (shapes without childs) of the tree. For each leaf $L$ we travel the corresponding branch upwards (towards the root) and we compute the *distance* between the histogram of two consecutive segments. More precisely if $T$ is a shape containing $L$, we say that the segment $\mathcal{S}(L, T)$ is stable if the distance between $h_{\mathcal{S}(L,T)}$ and $h_{\mathcal{S}(L,T^f)}$ is *big enough*, where $T^f$ is the parent shape of $T$ in the tree. A big distance value means that the new pixels of $T^f$ (those of $\mathcal{S}(L, T^f) \setminus \mathcal{S}(L, T)$) are different from those on the segment. So now we define a new segment starting on $T^f$ and make it grow upwards until another big change occurs. We repeat this procedure until we arrive to the root of the tree. Once all the segments corresponding to $L$ are computed, we remove them from the tree. The remaining after this substraction is a family of trees, so we repeat the procedure iteratively on them. When this stage finishes we get a partition of the image given by the set of stable segments identified. Due to the way that these segments are constructed, this partition is an over-segmentation of the image. We confirmed this fact experimentally with the results shown in figure 2(a). For this reason, we introduced a merging algorithm as a second stage.

### 2.3. Merging algorithm

Due to the inclusion property of the shapes on the tree, it only makes sense to merge two shapes that are direct relatives (parent and child). In order to determine what shapes to merge, we traverse the tree again and compute the pairwise distance between all consecutive segments. After that, we order the segments in decreasing order of distance and select to be merged those whose distance is lower than a threshold $\mu$. Then we merge the segments with the lowest distance. This merging defines a new segment, so we must update the distance value on the neighboring segments. Then the procedure is repeated again until no more segments can be merged.
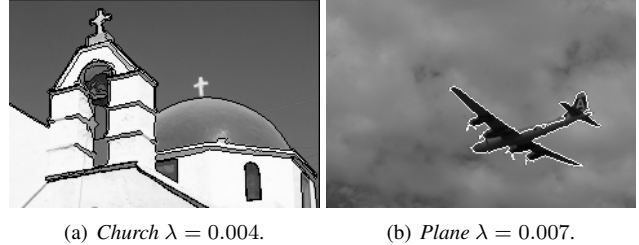
### 2.4. Histogram metrics

What is left to complete our algorithm is the definition of a suitable metric to compare histograms. For this purpose we consider the normalized histogram of a region as an empirical estimation of the conditional probability density function (PDF) in each region. According to this, we compare two segments $\mathcal{S}$ and $\mathcal{T}$ by means of the distance between their respective cumulative distribution functions (CDF), and compute that distance as the L1 norm of the difference between both CDFs. The results of all the experiments showed below use this measure. This is an uncommon choice compared to other well known PDF metrics, like the Battacharyya measure or the Kullback-Leibler divergence (for a review of these metrics we refer to [10]).The reason for not using them is that they are not suitable for the first stage: if the values of the new pixels added by the second segment $\mathcal{T}$ are different from those of the first segment, then the distance between the two segments does not depend on the actual value of the new pixels, but only on its number.

## 3. RESULTS

The only free parameter to be set is a threshold $\lambda$ to determine when a histogram changed enough to end a segment. There is also a threshold $\mu$ to determine when two segments must be merged together, but this parameter is fixed as a multiple of $\lambda$, and its not tuned for the examples. In the second stage objects are bigger, and since our distance measure depends on the area, it makes sense to use a bigger threshold but related to the first one[2]. In the future $\mu$ could be tailored by normalizing the measure taking into account the area of the objects.

With respect to the computational effort required, our algorithm takes 4 sec. to process the *Ram* image (400x500) on an AMD64 3400+ processor with non-optimized C code.

In figure 4 we show the results in two examples: *Church* and *Plane*. The objects are correctly detected but there are some errors. In the first case, the cross over the roof is lost and in the case of the plane the boundaries are not well localized, due to a smooth variation of the gray-level. Figure

---

[2]In our examples, $\mu = 30\lambda$

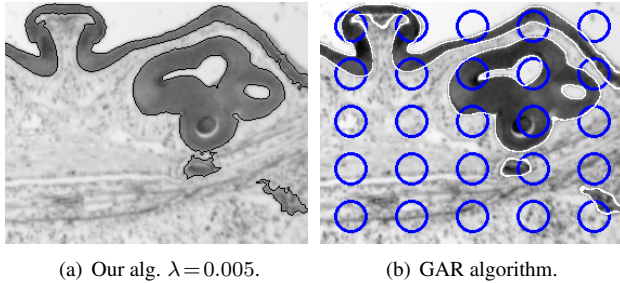(a) Our alg. $\lambda = 0.005$.      (b) GAR algorithm.

**Fig. 5**. Comparison of the results over the *Spider* image. (in (b) black: initial condition, GAR final results).

2 shows the results of the two stages of the algorithm over the *ram* image. As mentioned before, the first stage gives an over-segmented image (fig. 2(a)) and the second stage (fig. 2(b)) cleans up the results. The only errors are a couple of white zones and one light gray zone that are not detected. In addition some of the dark blobs are little over-segmented, but that is tolerable because they don't have a completely homogeneous interior. To show the effect of the parameter $\lambda$ over the segmentation results, we ran two examples varying its value. As figure 3 shows, $\lambda$ controls the size and the quantity of objects detected. A big value implies that the algorithm will merge many objects, so it will tend to lose some of them. When $\lambda$ is reduced we are able to detect more objects but also more false positives appear.

In figure 2 we compare our results with two related approaches. The first of them is Ballester et al. [7]. As we can see, their results are similar to those obtained by our algorithm. As a difference, our algorithm is much more simpler and runs an order of magnitude faster. Furthermore, our parameter has a clear interpretation as the amount of change allowed in an object, and their parameter ($\gamma$) does not. Finally, our algorithm is more robust when facing images where the relevant information is given by the distribution of the gray-level rather than by its mean value. The second approach (fig. 2(d)) is that by Cao et al. [4]. As they only use the minimum contrast along the boundaries, many shapes which have low contrast boundaries are lost. This confirms the importance of using region based information. The last comparison, shown in figure 4, is with the GAR algorithm [11] which is a region based active contour approach. The main difference between them is that the GAR approach uses a previously specified, fixed number of regions (not necessarily connected), and our algorithm does not. In this example both algorithms perform well, maybe our algorithm yields slightly better localized boundaries.

## 4. CONCLUSION AND FUTURE WORK

In this work we presented a novel segmentation approach based on the tree of shapes of an image. We introduced re-gion based descriptors based on the empirical PDFs of the pixels belonging to each shape. With these descriptors we performed a two pass algorithm based on pairwise comparisons in order to identify stable segments and merge them into meaningful regions. We also studied and tested suitable probabilistic metrics to perform this comparison. The results obtained by our algorithm are quite good and compare favorably with those presented in the literature.

This work can be extended in many directions. It would be interesting to find an automatic way to tailor our merging threshold $\mu$. In addition, the extension to multivalued images must be introduced, and tested over real images. Our approach is not a variational one, so it would be interesting to study how to re-formulate this algorithm as an energy minimization problem.

## 6. REFERENCES

[1] V. Caselles; B. Coll; J. Morel, "Topographic maps and local contrast changes in natural images," *Int. J. Comp. Vision*, , no. 33, pp. 5–27, 1999.

[2] C. Ballester, V. Caselles, and P. Monasse, "The tree of shapes of an image," *ESAIM-COCV*, pp. 1–18, 2003.

[3] A. Desolneux; L. Moisan; J-M. Morel, "Edge detection by helmholtz principle," *J. of Math. Imaging And Vision*, vol. 14, pp. 271–284, August 2001.

[4] F. Cao; P. Musé; F. Sur, "Extracting meaningful curves from images," *J. of Math. Imaging And Vision*, Dec. 2005.

[5] F. Cao, "Good continuations in digital image level lines," *ICCV*, vol. 1, pp. 440–448, 2003.

[6] A. Pardo, "Semantic image segmentation using morphological tools," *ICIP*, 2002.

[7] C. Ballester; V. Caselles; L. Igual, "Minimal morphological shape selection for segmentation and encoding," *(Preprint) J. of Math. Imaging and Vision*, 2004.

[8] Y. Song ; A. Zhang, "Analyzing scenery images by monotonic tree," *ACM Mult. Systems Journal*, 2002.

[9] P. Monasse and F. Guichard, "Fast computation of a contrast-invariant image representation," *IEEE Trans. Image Proc.*, 2000.

[10] D. Johnson; S. Sinanovic, "Symmetrizing the kullback-leibler distance," *IEEE Trans. on Information Theory (submitted)*, 2001.

[11] T. Brox; M. Rousson; R. Deriche and J. Weickert, "Unsupervised segmentation incorporating colour, texture, and motion," *10th Int. Conf. on Comp. Analysis of Images and Patterns,*, vol. 2756, pp. 353–360, August 2003.