Two applications of shape-based morphology: blood vessels segmentation and a generalization of constrained connectivity

Yongchao Xu, Thierry Géraud, Laurent Najman

To cite this version:
Yongchao Xu, Thierry Géraud, Laurent Najman. Two applications of shape-based morphology: blood vessels segmentation and a generalization of constrained connectivity. International Symposium on Mathematical Morphology, May 2013, Uppsala, Sweden. pp.386-397. hal-00798625

HAL Id: hal-00798625
https://hal.archives-ouvertes.fr/hal-00798625
Submitted on 9 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Two applications of shape-based morphology: blood vessels segmentation and a generalization of constrained connectivity

Yongchao Xu$^{1,2}$, Thierry Géraud$^1$, Laurent Najman$^2$

$^1$ EPITA Research and Development Laboratory (LRDE), France
\{yongchao.xu, thierry.geraud\}@lrde.epita.fr
$^2$ Université Paris-Est, LIGM, Équipe A3SI, ESIEE, France
l.najman@esiee.fr

Abstract. Connected filtering is a popular strategy that relies on tree-based image representations: for example, one can compute an attribute on each node of the tree and keep only the nodes for which the attribute is sufficiently strong. This operation can be seen as a thresholding of the tree, seen as a graph whose nodes are weighted by the attribute. Rather than being satisfied with a mere thresholding, we propose to expand on this idea, and to apply connected filters on this latest graph. Consequently, the filtering is done not in the space of the image, but on the space of shapes built from the image. Such a processing, that we called shape-based morphology \cite{30}, is a generalization of the existing tree-based connected operators. In this paper, two different applications are studied: in the first one, we apply our framework to blood vessels segmentation in retinal images. In the second one, we propose an extension of constrained connectivity. In both cases, quantitative evaluations demonstrate that shape-based filtering, a mere filtering step that we compare to more evolved processings, achieves state-of-the-art results.

1 Introduction

Mathematical morphology, as originally developed by Matheron and Serra \cite{23}, proposes a set of morphological operators based on structuring elements. Later, Salembier and Serra \cite{21}, followed by Breen and Jones \cite{3}, proposed morphological operators based on attributes, rather than on elements. Such operators rely on transforming the image into an equivalent representation, generally a tree of components (e.g., level sets) of the image: such trees are equivalent to the original image in the sense that the image can be reconstructed from its associated tree. Filtering then involves the design of a shape attribute that weights how much a node of the tree fits a given shape. Two different approaches for filtering the tree (and hence the image) have been proposed: the more evolved approach consists in pruning the tree by removing whole branches of the tree, and is easy to apply if the attribute is increasing on the tree (i.e., if the attribute is always stronger for the ancestors of a node). This process is depicted in the black path in Fig. 1.
However, most shape attributes are not increasing. When the attribute is not increasing, three strategies have been proposed (min, max, Viterbi; see [22] for more details). They all choose a particular node on which to take the decision, and remove the whole subtree rooted at this node. While it may give interesting results in some cases, it does not take into account the possibility that several relevant objects can have some inclusion relationship, meaning that they are on the same branch of the tree (e.g., a ring object in a tree of shapes, see Fig. 3(a)).

In the simplest approach, one simply removes the nodes of the tree for which the attribute is lower than a given threshold [27]. Such a thresholding does not take into account the intrinsic parenthood relationship of the tree. Moreover it is often impossible to retrieve all expected objects with one unique threshold. Fig. 2 shows the evolution of a shape attribute, the circularity, along two branches of the tree of shapes [15]. The light round shape and the dark one are both meaningful round objects compared to their context. However, their attribute values are very different. In order to obtain the light one, a higher threshold is applied, but some non-desired shapes appear, the ones in the background in Fig. 2.f.

The founding idea of shape-based morphology is to apply connected filters on the space of all the components of the image, such space being structured into a graph by the parenthood relationship (i.e., the neighbors of a node are its children and its parent). This process is illustrated by the black+red path of Fig. 1. This surprising and simple idea has several deep consequences that were first exposed in [30], where it is shown that this framework encompasses the usual attribute filtering operators [30]. Novel connected filters based on non-increasing criterion can also be proposed. When the first tree $T$ is respectively a Min-tree or a Max-tree, such filters are new morphological lower or upper levelings. When the first tree $T$ is the tree of shapes, we introduce a novel family of self-dual connected filters that we call morphological shapings. The proposed framework can also be used to produce extinction-based [28] saliency maps [18].
In general, shape-based morphological operators provide better results than threshold-based approaches. As stated above, when we want to process both upper and lower level sets at the same time, we use as $\mathcal{T}$ the tree of shapes, and the such created operator is called a morphological shaping. In both Fig. 2 and Fig. 3 the attribute $A$ is the circularity. The result of the shaping on Fig. 2a, shown in Fig. 2g, looks indeed better than the one of Fig. 2f. In Fig. 3 we compare this extinction-based self-dual shaping approach with a variant of the state-of-the-art thresholding approach [27]. When the threshold of $A$ is low, some objects do not appear (Fig. 3b). To get all the expected objects, we have to set a high threshold; however, in this case, too many unwanted objects are present (Fig. 3c). With our shaping filter, all the expected objects can be found, as depicted in Fig. 3d.

**Fig. 2.** (a) Evolution of “circularity” on two branches of $\mathcal{T}$; (b to e): Some shapes; (f) Attribute thresholding; (g) Shaping.

**Fig. 3.** Comparison of extinction-based shapings with attribute thresholding.
In this paper, we propose to detail how the framework can be used for two different types of applications. The rest of this paper is organized as follows. An application of our proposed shape-based upper levelings to blood vessels segmentation in retinal images is explained in Section 2. In Section 3, we detail how to extend the constraint connectivity framework first introduced in [24]. Finally we discuss and conclude in Section 4.

2 Blood vessels segmentation in retinal images

Blood vessels segmentation is a very important task in retinal images analysis. Unlike classical linear filters, connected operators perfectly preserve the location and the shape of the contours, which provides a motivation for using them.

Many existing methods work on the green channel of the color retinal image. To improve the visibility of the blood vessels, for each color retinal image $f_c$, a black top-hat transform is applied to the green channel $f_g$. When a mask of eye fundus is available, we combine it with the black top-hat $f_t$. We thus obtain an image $f_i$ in which the blood vessels are visible: indeed, the main structures of the blood vessels are present in the Max-tree $T$ of $f_i$ (the connected components of the upper level sets of $f_i$, [22]).

For each connected component represented by some node $N_k$ of the Max-tree $T$, we compute a shape attribute $A_e$ characterizing the blood vessels, which are usually long and thin structures. The attribute used here is the elongation $A_e$:

$$A_e(N_k) = |N_k|/(\pi \times l_{max}^2),$$  \hspace{1cm} (1)

where $|.|$ denotes the cardinality and $l_{max}$ denotes the length of the largest axis of the best fitting ellipse for the connected component represented by $N_k$. Since the blood vessels are long and thin, nodes having a low attribute $A_e$ correspond to the blood vessels.

The core of this application is the filtering of the Max-tree $T$. A mere thresholding of the elongation $A_e$ is not sufficient, often giving unwanted objects (noise). However, a very low thresholding value $t_{min}$ on $A_e$ ensures that thresholded nodes are blood vessels. Those initial extracted nodes are used as seeds in the sequel. We then apply a morphological filtering with a depth criterion: using the Min-tree $\mathcal{T}$ of the node-weighted graph $(T, A_e)$, we only preserve the nodes that have a certain depth $d_0$ in $\mathcal{T}$ and that furthermore contain the seeds. The connected components represented by the preserved nodes are considered as the segmented blood vessels. The whole process is one of the many variants of shape-based upper levelings [30].

An example of this blood vessels segmentation methods is given in Fig. 4. As compared with the manually segmented blood vessels segmentation (Fig. 4f), the elongation-based upper leveling (Fig. 4e) correctly extracts most of the blood vessels.

We have tested this specific shape-based upper leveling on the Digital Retinal Images for Vessel Extraction (DRIVE) database [5], [25] and on the STructured Analysis of the Retina (STARE) database [26], [8]. DRIVE is a database assembled in the Netherlands from a diabetic retinopathy screening program. It
includes 40 color fundus images of 584 × 565 pixels, captured using a 45° field-of-view fundus camera. The 40 color images are divided into a training set and a test set, both containing 20 images. For the training images, a single manual segmentation of the vasculature is available. For the test cases, two manual segmentations are available: one is used as gold standard; the other one can be used to compare computer generated segmentations with those of an independent human observer. The complete database contains seven pathological cases (four in the test set and three in the training set). The STARE database contains 20 images captured using a TopCon TRV-50 fundus camera at 35° FOV, and digitized to 700 × 605 pixels, 8 bits per RGB channel. A manual segmentation is available for each image of the database. Masks of the eye fundus, derived from the matched spatial filter [8], are also available. Note that, among the 20 images, 10 images are abnormal.

Fig. 5 and 6 show four segmentation results respectively from the DRIVE database and the STARE database. Qualitatively, most of the blood vessels are correctly extracted, although some noise points at the end of the vessels are also kept, and some very thin blood vessels are missed.

Quantitative assessment is based on three performance measurements named respectively sensitivity, specificity and accuracy [25]. Sensitivity measures the true positive rate (TPR), specificity measures the true negative rate (TNR), and accuracy measures the rate of pixels correctly classified. These measurements are
**Fig. 5.** Illustration of four segmented blood vessels from the 20 test retinal images of the DRIVE database. Top: Input color retinal images; Bottom: Corresponding segmented results. (White pixels: true positive; Black pixels: true negative; Blue pixels: false positive; Red pixels: false negative.)

**Fig. 6.** Illustration of four segmented blood vessels from the 20 test retinal images of the STARE database. Top: Input color retinal images; Bottom: Corresponding segmented results. (White pixels: true positive; Black pixels: true negative; Blue pixels: false positive; Red pixels: false negative.)
<table>
<thead>
<tr>
<th>Method</th>
<th>TPR</th>
<th>TNR</th>
<th>Accuracy</th>
<th>TPR</th>
<th>TNR</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd Expert</td>
<td>0.7761</td>
<td>0.9725</td>
<td>0.9473 (0.0048)</td>
<td>0.8949</td>
<td>0.9390</td>
<td>0.9354 (0.0171)</td>
</tr>
<tr>
<td>mendonça [14]</td>
<td>0.7344</td>
<td>0.9764</td>
<td>0.9452 (0.0062)</td>
<td>0.6996</td>
<td>0.9730</td>
<td>0.9440 (0.0142)</td>
</tr>
<tr>
<td>Staal [25]</td>
<td>0.7193</td>
<td>0.9773</td>
<td>0.9441 (0.0057)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Niemeijer [19]</td>
<td>0.6793</td>
<td>0.9801</td>
<td>0.9416 (0.0065)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Our</td>
<td>0.6924</td>
<td>0.9779</td>
<td>0.9413 (0.0078)</td>
<td>0.7149</td>
<td>0.9749</td>
<td>0.9471 (0.0114)</td>
</tr>
<tr>
<td>Zana [31]</td>
<td>0.6696</td>
<td>0.9769</td>
<td>0.9377 (0.0078)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Al-Diri [1]</td>
<td>-</td>
<td>-</td>
<td>0.9258 (0.0126)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Jiang [9]</td>
<td>0.6478</td>
<td>0.9625</td>
<td>0.9222 (0.0070)</td>
<td>-</td>
<td>-</td>
<td>0.9513</td>
</tr>
<tr>
<td>Perez [12]</td>
<td>0.7086</td>
<td>0.9496</td>
<td>0.9181 (0.0240)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hoover [8]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.6751</td>
<td>0.9567</td>
<td>0.9267 (0.0090)</td>
</tr>
</tbody>
</table>

Table 1. Benchmark of different blood vessels segmentation approaches on DRIVE and the STARE database.

Defined below as:

\[
\text{sensitivity} = TPR = \frac{TP}{P} = \frac{TP}{TP + FN} \quad (2)
\]

\[
\text{specificity} = TNR = \frac{TN}{N} = \frac{TN}{TN + FP} \quad (3)
\]

\[
\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN}.
\]

where \(TP\) stands for true positive, \(FP\) for false positive, \(TN\) for true negative and \(FN\) for false negative.

A benchmark of different approaches (including ours) is provided in Table 1. It shows the good performance of our proposed elongation-based upper leveling for both databases. In the case of the DRIVE database, our result is slightly under the best results given by the method of Mendonça [14]. Note also that the approaches of Staal [25] and Niemeijer [19] are supervised approaches. On the STARE database, our proposed method performs also very well, and is even better than the method of Mendonça [14]. Both methods give results that are very close to one of the second human observer. Table 2 shows that our method is more robust than others, in the sense that it performs equivalently on both abnormal and normal images.

Last, note that the proposed elongation-based upper leveling is only a “simple” filtering step, whereas other approaches are more complicated. Besides, our process is not complete, since further post-processing can improve the results.

## 3 Extending constrained connectivity

From an algorithmic point of view, constrained connectivity [24] is the application of an increasing criterion (e.g., the range) on the Min-tree \(T\) of the minimum spanning tree (MST) [16,17]. In the litterature, this tree \(T\) has been called the
<table>
<thead>
<tr>
<th>Method</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Normal cases</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd human observer</td>
<td>0.9646</td>
<td>0.9236</td>
<td>0.9283 (0.0100)</td>
</tr>
<tr>
<td>Shape-based upper leveling</td>
<td>0.7178</td>
<td>0.9802</td>
<td><strong>0.9493 (0.117)</strong></td>
</tr>
<tr>
<td>mendonca [14]</td>
<td>0.7258</td>
<td>0.9791</td>
<td>0.9492 (0.0122)</td>
</tr>
<tr>
<td>Hoover [8]</td>
<td>0.6766</td>
<td>0.9662</td>
<td>0.9324 (0.0072)</td>
</tr>
<tr>
<td><strong>Abnormal cases</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd human observer</td>
<td>0.8252</td>
<td>0.9544</td>
<td>0.9425</td>
</tr>
<tr>
<td>Shape-based upper leveling</td>
<td>0.7120</td>
<td>0.9696</td>
<td><strong>0.9447 (0.0106)</strong></td>
</tr>
<tr>
<td>mendonca [14]</td>
<td>0.6733</td>
<td>0.9669</td>
<td>0.9388 (0.0150)</td>
</tr>
<tr>
<td>Hoover [8]</td>
<td>0.6736</td>
<td>0.9472</td>
<td>0.9211 (0.0091)</td>
</tr>
</tbody>
</table>

Table 2. Benchmark of vessels segmentation methods (STARE images - Normal versus abnormal cases).

α-tree [20]. In this section, we apply our framework with a non-increasing criterion \( A \) derived from a popular work [6]. Precisely, we use for \( T \) a binary Min-tree of the MST [17]. The dissimilarity used for the MST is the maximal distance of the red, blue, and green channels taken independently.

The attribute \( A \) is derived from [6], where the authors propose a region merging process that follows the edges of the MST by increasing order of the weights (dissimilarity). When an edge \( \{x, y\} \) is considered, they search for the regions \( X \) and \( Y \) that respectively contain the points \( x \) and \( y \). The regions \( X \) and \( Y \) are merged if

\[
\text{Diff}(X, Y) < \min\{ \frac{k}{|X|}, \frac{k}{|Y|} \},
\]

where \(|\cdot|\) denotes the cardinality, \( \text{Diff}(X, Y) \) is the minimum weight of the edge connecting the two regions \( X \) and \( Y \), \( \text{Int}(X) \) is the largest weight in the MST of the region \( X \), and \( k \) is a parameter favouring the merging of small regions (a large \( k \) causes a preference for larger components). However, \( k \) is not a scale parameter in the sense of the causality principle: as shown in [7], a contour present at a scale \( k_1 \) is not always present at a scale \( k_2 < k_1 \). The merging criterion defined by Eq. [5] depends on the parameter \( k \) at which the regions \( X \) and \( Y \) are observed. So let us consider the attribute \( A \) as the \( k \) defined by

\[
k = \max\left\{ \left( \text{Diff}(X, Y) - \text{Int}(X) \right) \times |X|, \left( \text{Diff}(X, Y) - \text{Int}(Y) \right) \times |Y| \right\}.
\]

That is to say, for each node \( N_k \), let \( N_{k1}^c \) and \( N_{k2}^c \) be the two children of \( N_k \) in \( T \), then the attribute \( A \) for node \( N_k \) is given by

\[
A(N_k) = \max\left\{ \left( \text{Diff}(N_{k1}^c, N_{k2}^c) - \text{Int}(N_{k1}^c) \right) \times |N_{k1}^c|, \left( \text{Diff}(N_{k1}^c, N_{k2}^c) - \text{Int}(N_{k2}^c) \right) \times |N_{k2}^c| \right\}.
\]

The maxima of the attribute \( A \) correspond to meaningful regions. We thus compute a Max-tree \( \mathcal{T} \) on the node-weighted graph \((\mathcal{T}, \mathcal{A})\), with which we can
Table 3. Preliminary region benchmarks on the BSDS500.

<table>
<thead>
<tr>
<th>Method</th>
<th>GT Covering</th>
<th>Prob. Rand. Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ODS</td>
<td>OIS</td>
</tr>
<tr>
<td>FH [6]</td>
<td>0.43</td>
<td>0.53</td>
</tr>
<tr>
<td>Guimarães [7]</td>
<td>0.46</td>
<td>0.53</td>
</tr>
<tr>
<td>Ours</td>
<td>0.50</td>
<td>0.57</td>
</tr>
</tbody>
</table>

filter this graph. The contours of the flat zones of one level of filtering of \((T^T, A)\) provides a segmentation of the original image. Computing all levels of filtering produces a hierarchy of segmentations. A good representation of such a hierarchy is a saliency map \([18,10]\), which is equal to the sum of all the contours of every filtered image. An efficient algorithm for computing the saliency map in our framework, based on the notion of extinction value \([28]\), will be provided in an extend version of this paper. The saliency map defines a duality between closed, non-self intersecting weighted contours and a hierarchy of regions. Low levels of the hierarchy correspond to weak contours, thus to over-segmentations. High levels of the hierarchy correspond to strong contours, hence to under-segmentations. Moving between levels gives a continuous trade-off between these two extremes. A given level can be seen as an observation scale at which we consider the image.

We have tested our extended constrained connectivity framework on the Berkeley Segmentation Dataset BSDS500 \([2]\), an extension of the BSDS300 \([13]\). The dataset consists of 500 natural images divided into 200 test images, 200 images for training, and 100 validation images, together with human annotations. Each image is segmented by an average of five different subjects. Fig. 7 and 8 show some saliency maps computed on some images from the BSDS500 dataset. Two evaluation schemes are also provided by the authors. In the first one, the same fixed threshold level (observation scale) is used for all saliency maps in the dataset; we refer to it as the optimal dataset scale (ODS). In the second one, we evaluate the performance using an image-dependent threshold for each saliency map; we refer to this choice as the optimal image scale (OIS).

Quantitative evaluation is performed using the region-based performance measurements described in \([2]\), in terms of Ground-Truth Covering criterion and Probabilistic Rand Index. Here, we compare our results with the graph-based image segmentation (Felz-Hutt) \([6]\), and with another method named hierarchical graph based image segmentation (Guimarães et al.) \([7]\), also relying on the same criterion popularized by \([6]\). The comparison is given in Table 3. Our method ranks first, for both the optimal dataset scale (ODS) and for optimal image scale (OIS).

4 Discussion and conclusion

This paper has presented two applications of shape-based morphology, a generalization of existing tree-based connected operators. The first application uses one of the many shape-based upper levelings. Although such a filter is but a "simple" filtering step, it gives results almost as good as the second human observer
in the case of blood vessels segmentation. The second application is an extension of the constrained connectivity framework to non-increasing constraints. A quantitative evaluation based on a criterion given in [6] shows that our approach compares favorably to previous works.

The potential of shape-based morphology is tremendous. In this paper, we have hardly skimmed the surface of what this theory has to offer to the scientific community. In [29], we have also used this framework to achieve object segmentation on the tree of shapes. The key idea was to propose an efficient context-based energy estimator whose minima correspond to meaningful objects. Some other criterions that can be used are, for example an energy derived from the number of false alarms [4] or some snake energy [10]. A practical problem is that many minima of such energies do not correspond to meaningful components of the input image. In [29], a morphological closing in the space of shapes helps to filter those spurious minima.

Implementations of shape-based filters are easy thanks to the open-source Milena library [11]. A demonstration is available online from [http://olena.lrde.epita.fr/ICPR2012](http://olena.lrde.epita.fr/ICPR2012) (see also [http://olena.lrde.epita.fr/ICIP2012](http://olena.lrde.epita.fr/ICIP2012)). More applications will be studied in a forthcoming paper. Properties of those filters, such as conditions for idempotence, will be also studied.

**Acknowledgements.** This work received funding from the Agence Nationale de la Recherche, contract ANR-2010-BLAN-0205-03 and through “Programme d’Investissements d’Avenir” (LabEx BEZOUT n°ANR-10-LABX-58).
Fig. 8. Additional hierarchical segmentation results on the BSDS500. From top to bottom: Input mage, saliency map, and segmentations at the optimal dataset scale (ODS) and at the optimal image scale (OIS).

References


