Ultimate Attribute Opening Segmentation with Shape Information

Jorge Hernández and Beatriz Marcotegui

Mines ParisTech
CMM- Centre de morphologie mathématique
Mathématiques et Systèmes
35 rue St Honoré 77305-Fontainebleau-Cedex, France
{ hernandez, marcotegui }@cmm.ensmp.fr

Abstract. In this paper, a method for morphological segmentation using shape information is presented. This method is based on a morphological operator named ultimate attribute opening (UAO). Our approach considers shape information to favor the detection of specific shapes. The method is validated in the framework of two applications: façade analysis and scene-text detection. The experimental results show that our approach is more robust than the standard UAO.

1 Introduction

Segmentation is a fundamental problem in image analysis to distinguish between objects of interest and "the rest". It creates a partition of the image into disjoint and uniform regions, according to some features such as gray value, color, or texture [1]. An overview of morphological segmentation is presented by Meyer in [2] where a unified framework for supervised or unsupervised, multi-scale or single scale, color or grayscale and 2D or 3D images is introduced. Furthermore, a new morphological operator, named ultimate opening (*UO*) [3], has been increasingly used as a powerful segmentation method due to its various advantages (non-parametric operator, segmentation of contrasted structures, intrinsically multi-scale, etc). This morphological operator can be used for shape analysis by associating a granulometry function. Retornaz and Marcotegui have proposed and implemented ultimate attribute opening (*UAO*) [4], where attribute opening (*AO*) was introduced by Breen and Jones [5].

In this paper, we present an integrated approach for image segmentation based on *UAO* combined with shape constraints. In contrast to using only grayscale values to locate regions with the *UAO*, the proposed method uses a similarity function defined through a prior knowledge of the shapes. This similarity function is based on the characteristics of connected components *CC*s in images (shapes). Urbach *et al.* defined in [6] vector-attribute filters based on shape descriptors and a dissimilarity measure. A threshold is required in order to filter out *CC*s different from the prior one. In contrast, our method detects the most contrasted shapes similar to the prior one. Our results are shown on real applications: façade analysis and scene-text detection.

The paper is organized as follows. In Section 2, *UAO* are presented. Section 3 describes *UAO* with shape information. In Section 4, two applications are shown and the advantages of our method are illustrated. Finally, conclusions are drawn in Section 5.

2 Ultimate Attribute Opening

Ultimate opening (*UO*), closing by duality, has been introduced by Beucher in [3]. This is a non-parametric method and a non-linear scale-space based on morphological numerical residues to extract (*CCs*). Several applications have been developed: automatic localization of text [7] and façade segmentation [8].

2.1 Ultimate Opening

The ultimate opening θ analyzes the difference between consecutive openings. This operator has two significant outputs for each pixel x from an input image I: the maximal difference between openings (Residue, $R_{\theta}(I)$) and the opening size, when the maximal residue is generated $(q_{\theta}(I))$. The equations describing the UO evolution are written as:

$$R_{\theta}\left(I\right) = \max\left(r_{\lambda}\left(I\right)\right), \, \forall \lambda \geq 1$$

$$\text{with } r_{\lambda}\left(I\right) = \gamma_{\lambda}\left(I\right) - \gamma_{\lambda+1}\left(I\right)$$

$$q_{\theta}\left(I\right) = \max\left(\lambda\right) + 1 : \lambda \geq 1, \, r_{\lambda}\left(I\right) = R_{\theta}\left(I\right) \wedge R_{\theta}\left(I\right) > 0$$

$$(1)$$

where, γ_{λ} is an opening of size λ .

2.2 Attribute Opening

A binary attribute opening, defined by Breen and Jones [5], consists in a connected opening associated with a given increasing criterion T, based on attributes. This criterion is used to keep or discard CCs. Gray attribute openings γ^T can be defined as follows:

$$\gamma^{T}(I(x)) = \max\left(h|x \in \Gamma^{T}(\mathcal{J}_{h}(I))\right) \tag{2}$$

where, $\mathcal{J}_h(I)$ is the threshold at level h of I, and Γ^T is a binary attribute opening. If κ_{CC} is a CC attribute, λ is a scalar value and T(CC) criterion is $\kappa_{CC} \geq \lambda$, an attribute opening γ^T can be denoted as γ_{λ} . AO and UAO are connected operators. They can only merge input flat zones but never cut them. As stated by Salembier in [9], the Max-Tree is a suitable image representation to compute connected operators.

2.3 Max-Tree

The Max-Tree, by duality Min-Tree, was introduced by Salembier [10] as a structure for computing connected operators. It is a multi-scale image representation and a hierarchical structure in which the nodes C_h^k represent k'th binary connected components of $\mathcal{J}_h(I)$. The root node corresponds to the whole image. The leaf nodes correspond to the image maxima. The links between the nodes describe the inclusion relationship of the binary connected components. Once the Max-Tree is created, several attributes can be estimated, allowing an efficient computation of attribute openings.

On Max-Tree, an attribute opening γ_{λ} removes a node C_h^k when its attribute $\kappa_{C_h^k}$ is smaller than a parameter λ . Fig. 1(a) illustrates a synthetic image and its corresponding Max-Tree (Fig. 2(a)). Fig. 1(d) shows the attribute opening results, with λ increasing

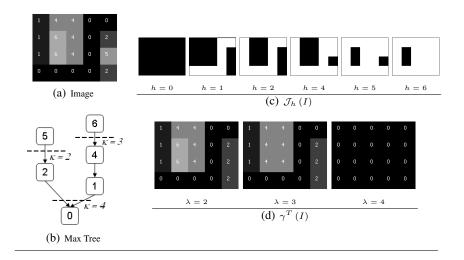


Fig. 1. (a) Original image, (b) $\mathcal{J}_h(I)$, (c) Max-Tree and (d) height openings $\gamma_T(I)$, where $T: \kappa_{C_h^k} \geq \lambda$.

values. The attribute κ used is height (y-extent) of C_h^k . These openings are obtained by pruning the tree at the corresponding dot-lines of Fig. 2(a).

Fabrizio in [11] implemented a fast $U\!AO$ based on Max-Tree taking advantage of the structure. The residue of each removed node in a given γ_{λ} is computed by the difference between its gray level and the gray level of its first ancestor with a different attribute. Fig. 2 shows intermediate steps of $U\!AO$ computation. A height opening with $\lambda=2$ removes the node C_5^0 , producing a residue of $\gamma_1-\gamma_2=5-2=3$. Then, a height opening with $\lambda=3$ removes C_6^0 with a residue of $\gamma_2-\gamma_3=6-4=2$. Finally, a height opening with $\lambda=4$ generates a residue for tree $C\!C\!s$, C_2^0 , and C_1^0 and C_4^0 . These residues are compared to current R_{θ} and only pixels with a larger residue are updated.

 $r_{\lambda}\left(C_{h}^{k}\right)$ computation on the Max-Tree is summarized in an iterative procedure as follows:

$$r_{\lambda}\left(C_{h}^{k}\right) = \begin{cases} h - h' & \kappa_{C_{h}^{k}} \neq \kappa_{C_{h'}^{k'}} \\ h - h' + r_{\lambda}\left(C_{h'}^{k'}\right) & \text{otherwise} \end{cases}$$
(3)

where, $C_{h'}^{k'}$ is the parent node ($C_h^k \subset C_{h'}^{k'}$ and h' < h). The residue is calculated for a parent node and it is propagated to all children. Every child compares his residue with its parent, and R_θ keeps the maximum value between them. Each child becomes parent and repeats the process. q_θ is the child attribute $\kappa+1$ when the maximal residue is generated.

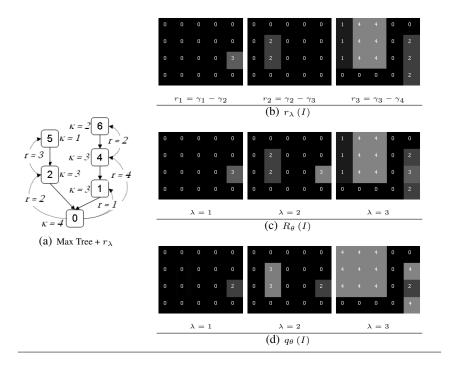


Fig. 2. (a) Max-Tree and residue $r_{\lambda}\left(C_{h}^{k}\right)$ and (b) height openings $R_{\theta}\left(I\right)$ and $q_{\theta}\left(I\right)$

2.4 Masking Problem

In spite of this non-parametric operator capacity to segment the most contrasted structures, it shows a problem of *blindness* named as "masking". The UO is an operator without memory and it only saves the last maximum residue. Hence, when a structure is nested in others, it may be masked by a bigger residue. For example, in Fig. 3(a), the masking problem is shown by applying the ultimate attribute (height) opening. The image has three nested shapes: a rectangular shape (dimensions 120×40), a square shape (dimensions 30×30) and a circle one (diameter 90). These structures are enclosed in a minimum gray value bounding box.

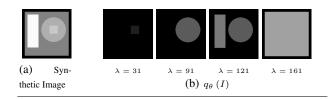


Fig. 3. (a) Original image, (b) $q_{\theta}\left(I\right)$ of intermediate results after openings $(\lambda+1)$

The square is the first shape found ($\lambda=31$), and then this shape is masked by a circle filtered by opening $\lambda=91$. Before the height opening $\lambda=161$ is applied, two shapes have been detected; nevertheless, in this opening, an important residue masks the relevant information. In order to solve the masking problem, we propose to use shape information, exploiting a prior knowledge of the image and preserving specific shapes before being masked.

3 Ultimate Attribute Opening with Shape Information

In contrast to employing only grayscale values to locate regions with the *UAO*, the proposed method uses a similarity function based on the characteristics of connected components in images (shapes).

3.1 Shape Information

In this Section, we define our shape similarity measure. The shape definition has been widely studied in the literature. Charpiat *et al.* [12] note Ω any shape, i.e. any regular bounded subset of D, and Γ or $\partial\Omega$, its boundary, a smooth curve of \mathbb{R}^2 . In our context, we are interested in comparing two different shapes, and their similarity measure. Many different definitions of the similarity functions ψ () between two shapes (Ω_i, Ω_j) have been proposed in the computer vision literature.

In order to compare two shapes, we propose to define a similarity function via shape descriptors κ_{Ω} . In our method, we propose to use the simplest shape descriptors: geometric features (height, width, etc) and their relations (fill ratio, circularity, moments, etc). We define $\psi_{\kappa}\left(\Omega_{i},\Omega_{j}\right)$ by the use of Eq. 4, where τ_{κ} is the similarity threshold of attribute κ .

$$\psi_{\kappa}\left(\Omega_{i},\Omega_{j}\right) \leftarrow \begin{cases} 1 & \kappa_{\Omega_{i}} = \kappa_{\Omega_{j}} \\ \left[0,1\right] & \left|\kappa_{\Omega_{i}} - \kappa_{\Omega_{j}}\right| < \tau_{\kappa} \\ 0 & \text{otherwise} \end{cases}$$
(4)

We have defined similarity functions with only one attribute κ . In practice, several measures are used to describe a prior shape in a same application. So, we expand the comparison between two shapes Ω_i, Ω_j to several attributes by using a simple multiplication function of similarity functions as follows: $\psi_{\forall \kappa} = \prod_{\forall \kappa \in \Omega} \psi_{\kappa} \left(\Omega_i, \Omega_j \right)$.

All these possible shape attributes and similarity functions can be utilized to give an advantage over specific shapes in a segmentation process. Nevertheless, we must be careful with the selection of measures, because of the following reasons:

- Computing Time: Measures are computed for each tree node. To keep a reasonable computing time, we have used the simplest shape attributes, because they can straightforwardly and accurately be estimated during the Max-Tree construction.
- Robustness of position, scale, and rotation invariant (PSRI): UO may be intrinsically PSRI. However, the invariance of the proposed method also depends on the chosen measures. For example, if we choose fill ratio attribute, the new operator

will only be *PSI*, besides if we choose compactness descriptor, it will be fully invariant.

3.2 Definition

We propose to consider a shape factor function $f\left(\Omega_{i},\Omega_{ref}\right)$ to a reference shape Ω_{ref} within the residue computation (Eq. 5). In that way, the residue of a Ω_{i} similar to Ω_{ref} is artificially increased. Thus, masking becomes more difficult. As the UAO is computed by using Max-Tree, each tree node C_{h}^{k} corresponds to a Ω_{i} and, $f\left(\Omega_{i},\Omega_{ref}\right)$ is denoted by $f\left(C_{h}^{k}\right)$.

$$r_{\lambda}^{\Omega} \leftarrow f\left(\Omega_{i}, \Omega_{ref}\right) r_{\lambda}$$
 (5)

The factor function $f\left(C_h^k\right)$ is related to the similarity function $\psi_{\forall \kappa}$ of C_h^k as follows: $1+\alpha\psi_{\forall \kappa}$. An offset of 1 is added in order to switch to standard UO when the similarity function is equal to 0 (r_λ^Ω becomes r_λ). As well, a multiplicative factor α is used to control the influence of the shape factor with respect to the gray level. Hence, $1+\alpha$ represents the maximum value that the function may reach.

Finally, function $f_{C_h^k}$ is stored on an image $F_{\theta}^{\Omega}(I)$ when the maximal residue $(R_{\theta}(I))$ is generated. With this information, we modify the original expression of UO Eq. 1 by Eq. 6.

$$R_{\theta}^{\Omega}(I) = \sup \left(r_{\lambda}^{\Omega}(I) \right)$$

$$F_{\theta}^{\Omega}(I) = f\left(C_{h}^{k} \right)$$

$$q_{\theta}^{\Omega}(I) = \max \left(\lambda \right) : \lambda \ge 1$$

$$R_{\theta}^{\Omega}(I) = r_{\lambda}^{\Omega}(I) ,$$

$$R_{\theta}^{\Omega}(I) > 0$$

$$(6)$$

3.3 Example on Synthetic Image

Now, we test our approach on a synthetic image (Fig. 3(a)) to analyze it. First, we try to favor rectangular shapes. For this purpose, we use as a shape metric the fill ratio $\Upsilon_{\Omega} = \frac{A_{\Omega}}{Abbox_{\Omega}}$, where (A_{Ω}) is the shape area and $(Abbox_{\Omega})$ is the bounding box area. The ratio lies in the range [0,1]; where, if the value is close to 1, it means that the shape corresponds to a rectangular polygon without rotation. We suppose that Ω_{ref} is a rectangular shape without rotation, i.e. $\Upsilon_{\Omega_{ref}} = 1$. Then, we have imposed area limit to validate the shape. We utilize a maximum area limits (90 % of image area A_{I}) to reject the largest regions. This factor function is translated into Eq. 7. Fig. 4(b) presents the result on a synthetic image. In this case, the masking problem is solved and the three shapes are segmented. The importance of limits $\psi_{2}\left(C_{h}^{k}\right)$ is remarkable in this example, because the masking opening, size = 161, shape reaches a high factor $\Upsilon_{C_{h}^{k}} \approx 1$. Using the limits, these shapes have a factor equal to 1, thus $r_{\lambda}^{\Omega} = r_{\lambda}$.

Even though Υ mainly favors rectangular structures, the circular shape factor is high enough to unmask it. If we want to favor only rectangular shapes, we modify $\psi_1\left(C_h^k\right)$ by a narrow function. For example, we change $\Upsilon_{C_h^k}$ in Eq. 7 by $\Upsilon_{C_h^k}$ cube value as shown in Eq. 8.

$$f\left(C_{h}^{k}\right) = 1 + \alpha\psi_{1}\left(C_{h}^{k}\right)\psi_{2}\left(C_{h}^{k}\right)$$

$$\psi_{1}\left(C_{h}^{k}\right) = \Upsilon_{C_{h}^{k}},$$

$$\psi_{2}\left(C_{h}^{k}\right) = \begin{cases} 1 A_{C_{h}^{k}} < 90\%A_{I} & (7) \\ 0 & \text{otherwise} \end{cases}$$

$$\psi_{1}\left(C_{h}^{k}\right) = \begin{cases} 1 A_{C_{h}^{k}} < 90\%A_{I} & (8) \\ 0 & \text{otherwise} \end{cases}$$

$$\psi_{1}\left(C_{h}^{k}\right) = \begin{pmatrix} 1 A_{C_{h}^{k}} < 90\%A_{I} & (8) \\ 0 & \text{otherwise} \end{cases}$$

$$\psi_{1}\left(C_{h}^{k}\right) = \begin{cases} 1 A_{C_{h}^{k}} < 90\%A_{I} & (8) \\ 0 & \text{otherwise} \end{cases}$$

$$\psi_{1}\left(C_{h}^{k}\right) = \frac{4\pi A_{C_{h}^{k}}}{\left(L_{C_{h}^{k}}\right)^{2}},$$

$$\psi_{2}\left(C_{h}^{k}\right) = \begin{cases} 1 A_{C_{h}^{k}} < 90\%A_{I} & (9) \\ 0 & \text{otherwise} \end{cases}$$

$$\psi_{2}\left(C_{h}^{k}\right) = \begin{cases} 1 A_{C_{h}^{k}} < 90\%A_{I} & (9) \\ 0 & \text{otherwise} \end{cases}$$

Fig. 4. Synthetic image segmentation (Fig. 3(a)) using UAO with shape constraints, κ is height.

Another shape factor example is implemented to favor circular shapes. In circle detection, the most frequently used metric is circularity. The metric is the shape area (A_{Ω}) ratio to a circle area having the same perimeter (L_{Ω}) : $\frac{4\pi A_{\Omega}}{(L_{\Omega})^2}$. Eq. 9 shows factor function by using the circularity expression. Fig. 4(d) and Fig. 4(f) confirm that rectangular shapes and circular shapes are segmented, respectively.

4 Applications

The aim of the proposed method is to improve *UAO* segmentation results avoiding masking problems. In order to demonstrate the performance of our method, we illustrate two segmentation applications: façade image analysis and scene-text detection. We show all databases and result tests on the following web site: http://cmm.ensmp.fr/~hernandez/results/UOSC/testsegmentationOUSC.html.

4.1 Façade Image Analysis

Initially, we employed an ultimate attribute opening to segment façade images. For the façade structure detection, a height attribute of the CCs bounding box is used. Fig. 5(b) shows an example of an ultimate attribute opening using a color gradient of Fig. 5(a). In the example, all internal structures are masked in the segmentation process because the contrast between the sky and the building façade is bigger than the contrast between the wall and the windows. Most urban images contain sky information; for this reason their UAO segmentation is highly affected by the masking problem.

In façade images, windows and doors have particular features. Mayer and Reznik [13] describe that most windows are at least partially rectangular and the height-width

ratio (\aleph) of a window generally lies between 0.20 and 5. With these features, we can define a similarity function for the internal structures of façades.

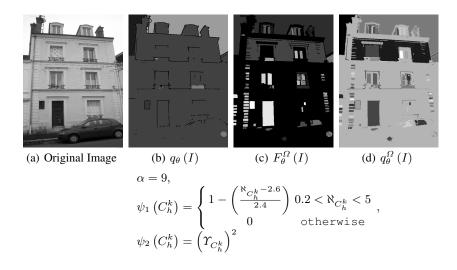


Fig. 5. (a) Original image: façade, (b) $q_{\theta}\left(I\right)$ of UAO , (c)-(d) F_{θ}^{Ω} , $q_{\theta}\left(I\right)$ of UAO with shape information, κ is height. Factor function: $f\left(C_{h}^{k}\right)=1+\alpha\psi_{1}\left(C_{h}^{k}\right)\psi_{2}\left(C_{h}^{k}\right)$. Color images in the web site.

In the case of the synthetic image, α was selected as $\max/2$, however after several tests with façade images, the contrast between the sky and the façade is about ten times bigger than the one between the wall and the windows. For this reason, we have chosen $\alpha=9$, i.e. $1+\alpha=10$. The first shape attribute is \aleph , we have used a dynamic function centred on $\aleph_{\Omega_{ref}}=2.6$ and the limits have been defined on $\tau_{\aleph}=\pm 2.4$. The second shape attribute is Υ because we suppose that windows are partially rectangular ($\Upsilon_{\Omega_{ref}}=1$). We have employed a square value of $\Upsilon_{C_h^k}$ as a dynamic function to penalize non-rectangular shapes or rectangular shapes with holes.

The segmentation result on the façade image is illustrated in Fig. 5. Our approach shows a better segmentation because interest structures appear (seven over nine windows/doors). Nevertheless, some structures which are neither windows nor doors become visible on the image segmentation, such as bricks (Fig. 5(d)).

4.2 Scene-Text Detection

The text in a scene is linked to the semantic context of the image and constitutes a relevant information for content-based image indexation [7]. In several cases, the text on images is at least partially placed on a surface of different color such as: placards, posters, etc; and favors the visibility of letters. But this surface is also contrasted in comparison to its surrounding (Fig. 6(a)). When we utilize *UAO*, characters may be masked by the contrast of the signboard with its surroundings (Fig. 6(b)).

In the same way, we propose some text features to define shape information. Based on histograms of approximately 2000 analyzed characters , letter features are described as follows: 1- the range of height/width ratio mostly lies between 0.4 and 2, 2- the range of Υ falls approximately between 0.4 and 0.8 and 3- the biggest height and width of a character is 1/3 of the image height and width respectively. We have used a similarity function analogous to the façade application. For \aleph and Υ metrics, the dynamics functions are centred on $\aleph_{\varOmega_{ref}}=1.2$ and $\varUpsilon_{\varOmega_{ref}}=0.6$, respectively.

In Fig. 6, the example image shows the results of the text detection using *UAO* with shape constraints. The results from this preliminary study indicate that the proposed method is superior to the classical *UAO* segmentation.

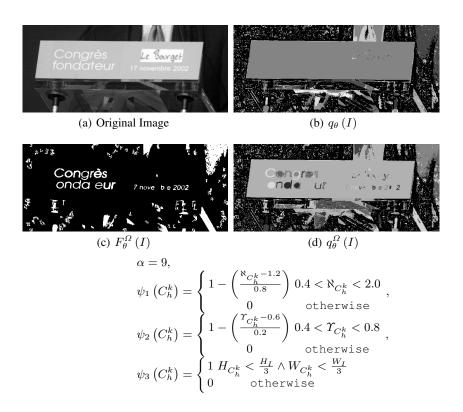


Fig. 6. (a) Original image: placard, (b) $q_{\theta}(I)$ of $U\!AO$, (c)-(d) F_{θ}^{Ω} , $q_{\theta}(I)$ of $U\!AO$ with shape information, κ is height. Factor function: $f\left(C_h^k\right) = 1 + \alpha \psi_1\left(C_h^k\right) \psi_2\left(C_h^k\right) \psi_3\left(C_h^k\right)$. Color images in the web site.

5 Discussion and Future Work

In this paper we introduce a segmentation method based on ultimate opening with shape constraints. Our approach exploits a prior knowledge to define shape information. The method can be combined with all types of similarity shape functions, thanks to the independence between the shape method computation and the classical UAO process. UAO provides two pieces of information, contrast $(R_{\theta}^{\Omega}(I))$ and size $(q_{\theta}^{\Omega}(I))$. The proposed method provides a third interesting piece of information: the shape factor image $(F_{\theta}^{\Omega}(I))$, that conveys a shape similarity measure with a reference shape. We store this factor when the maximal residue is generated.

The proposed method has been validated in two applications of structure extractions from façade and text images. The proposed method produces much better segmentation results than the standard UAO. In façade example (Fig. 5(c)), we can see that one window with shutters (down - right) is not detected, even if this shape is similar to a rectangle. The reason is its low contrast making the shape factor insufficient to avoid the masking effect. Fig. 6(c) illustrates factor images for text example. Several letters are not detected. "m" case is not favored ($f(\Omega_i, \Omega_{ref}) = 1$) because its \aleph is outside the fixed limits. In "f" and "t" cases, the factor value is not big enough to unmask them. On the other hand, many noise CCs are valued with a high factor function and they are still masked thanks to their low contrast.

In the future, we will analyze in details the factor function of the segmentation algorithm in order to validate the performance on a larger databases. The detection process is the first step in computer vision problems. We intend to apply a machine learning process using shape features (shape and color descriptors) to classify regions in both applications. As well, machine learning techniques could be considered as a factor function into the presented method.

6 Acknowledgment

The work reported in this paper has been performed as part of Cap Digital Business Cluster Terra Numerica project.

References

- Gonzalez, R.C., Woods, R.E.: Digital Image Processing (3rd Edition). Prentice-Hall, Inc., Upper Saddle River, NJ, USA (2006)
- Meyer, F.: An overview of morphological segmentation. International Journal of Pattern Recognition and Artificial Intelligence 15(7) (2001) 1089–1118
- 3. Beucher, S.: Numerical residues. Image and Vision Computing 25(4) (2007) 405-415
- Retornaz, T., Marcotegui, B.: Ultimate opening implementation based on a flooding process (2007) The 12th International Congress for Stereology.
- Breen, E.J., Jones, R.: Attribute openings, thinnings, and granulometries. Computer Vision and Image Understanding 64(3) (1996) 377–389
- Urbach, E.R., Boersma, N.J., Wilkinson, M.H.: Vector-attribute filters. In: ISMM '05,Proceedings of the seven International Symposium on Mathematical Morphology, Paris (18-20 April 2005) 95–104
- Retornaz, T., Marcotegui, B.: Scene text localization based on the ultimate opening. In: ISMM '07,Proceedings of the eigth International Symposium on Mathematical Morphology. Volume 1. (October 2007) 177–188

- 8. Hernández, J., Marcotegui, B.: Segmentation of façade images using ultimate opening. Technical report, CMM-Mines ParisTech (2007)
- 9. Salembier, P., Serra, J.: Flat zones filtering, connected operators, and filters by reconstruction. IEEE Transactions on Image Processing 4 (1995) 1153–1160
- Salembier, P., Oliveras, A., Garrido, J.L.: Anti-extensive connected operators for image and sequence processing. IEEE Transactions on Image Processing 7 (1998) 555–570
- 11. Fabrizio, J., Marcotegui, B.: Fast implementation of the ultimate opening. In: ISMM '09,Proceedings of the nine International Symposium on Mathematical Morphology. (August 2009)
- 12. Charpiat, G., Faugeras, O., Keriven, R.: Approximations of shape metrics and application to shape warping and empirical shape statistics. Foundations of Computational Mathematics **5**(1) (2005) 1–58
- 13. Mayer, H., Reznik, S.: Building façade interpretation from image sequences. In: CMRT05. Volume XXXVI. (2005) 55-60