ULTIMATE OPENING SEGMENTATION WITH SHAPE CONSTRAINTS

Jorge Hernández and Beatriz Marcotegui
Mines ParisTech
CMM- Centre de morphologie mathématique
35 rue St Honoré 77305-Fontainebleau-Cedex, France
email: { hernandez,marcotegui }@cmm.ensmp.fr

ABSTRACT
In this paper, a novel method for image segmentation with shape constraints is presented. This method is based on a morphological operator named ultimate opening (UO). The UO is an operator based on numerical residues. Our approach introduces 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 UO. Lastly, we conclude and present some perspectives of this work.

KEY WORDS
Ultimate opening, mathematical morphology, shape information, façade analysis, scene-text detection.

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.

We present an integrated approach for image segmentation based on UO combined with shape constraints. In contrast to employing only grayscale values to locate regions with the UO, the proposed method uses a similarity function based on the characteristics of connected components in images (shapes). This similarity function has been defined through a prior knowledge of the shapes in images.

This paper is organized as follows. In Section 2, some basic concepts of UO are presented and shape information is introduced. Section 3 describes our method that introduces shape information into UO definition. In Section 4, two applications are shown and the advantages of our method are illustrated. Finally, conclusions are drawn in Section 6.

2 Basic Notions

2.1 Ultimate 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 connected components (CCs). Retornaz and Marcotegui have proposed and implemented ultimate attribute opening (UAO) [4], where AOs were introduced by Breen and Jones [5]. Three applications have been developed using UO: Image analysis to measure the granulometry of rocks [6], automatic localization of text [7] and façade segmentation [8].

2.1.1 Definition

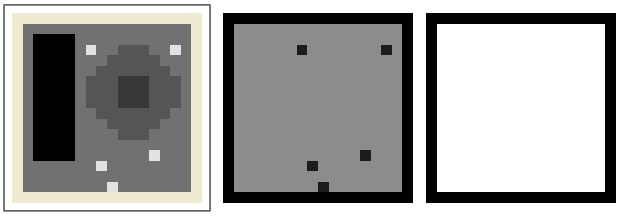
The ultimate opening θ analyzes the difference between two consecutive openings. This operator has two significant outputs for each pixel x from an input image I: the maximal difference between openings (Residue, $R_\theta(I)$) and the opening size, when the maximal residue is generated ($q_\theta(x)$).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

$$\begin{aligned} R_\theta(I) &= sup\left(r_\lambda(I)\right), \quad \forall \lambda \geq 1 \\ &\text{with} \quad r_\lambda(I) = \gamma_\lambda - \gamma_{\lambda+1} \\ q_\theta(x) = max(\lambda) : \lambda \geq 1, \quad & r_\lambda(x) = R_\theta(x) \wedge R_\theta(x) > 0 \end{aligned} \tag{1}$$

where, γ_λ is an opening of size λ.

In spite of the capacity of this non-parametric operator to segment the most contrasted structures, it presents a problem of *blindness* named as "masking". In Fig. 1, the masking problem is shown using a synthetic image by applying an ultimate closing. The image has three internal shapes (two rectangular shapes and one circle), and five

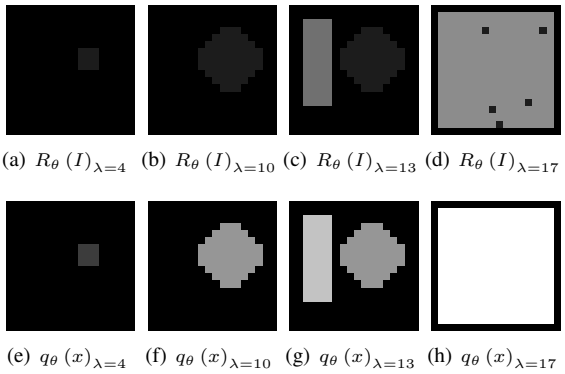
noise points. These structures are enclosed in a maximum gray value bounding box (Fig. 1(a)).



(a) Synthetic Image (b) $R_\theta(I)$ (c) $q_\theta(x)$

Figure 1. UAC Example: problem of masking

Fig. 1(b) and Fig. 1(c) show the ultimate height opening results (q_θ and R_θ) of Fig. 1(a). Fig. 2 illustrates some intermediate results after the closing of sizes (λ): 4, 8, 10, 17. The square of height 3 is the first shape found (Fig. 2(e)), and then this shape is masked by a circle filtered by height closing of size 10 (Fig. 2(f)). Before the last closing size, 17, two shapes have been detected (Fig. 2(g)); nevertheless, in the last size closing an important residue masks the relevant information (Fig. 2(h)) and noise points are visible in the residue image (Fig. 2(d)). 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.



(a) $R_\theta(I)_{\lambda=4}$ (b) $R_\theta(I)_{\lambda=10}$ (c) $R_\theta(I)_{\lambda=13}$ (d) $R_\theta(I)_{\lambda=17}$

(e) $q_\theta(x)_{\lambda=4}$ (f) $q_\theta(x)_{\lambda=10}$ (g) $q_\theta(x)_{\lambda=13}$ (h) $q_\theta(x)_{\lambda=17}$

Figure 2. UAC Example: Intermediate results

2.2 Shape Information

The shape definition has been largely studied in the literature. Charpiat et al. [9] note Ω any shape, i.e. any regular bounded subset of D, and Γ or $\partial\Omega$, its boundary, a smooth curve of R^2. In our context, we are interested in comparing two different shapes, and their similarity measure. Many different definitions of the similarity functions $\psi()$ between two shapes (Ω_i, Ω_j) have been proposed in the computer vision literature. We propose to use the simplest shape features for comparison: height, width, area, etc. and their relations (Example: Eq. 2).

$$\psi(\Omega_i, \Omega_j) \leftarrow abs\left(Height_{\Omega_i} - Height_{\Omega_j}\right) \quad (2)$$

More complex shape descriptors such as Fourier moments or those based on multiple characteristics of polygon form (e.g. the length of the major and minor internal axes, or the set of possible straight lines included within the polygon) can be used to describe shape information. As well, distance notions can be employed to compare two shapes (Hausdorff distance Eq. 3, etc.)

$$\psi(\Omega_i, \Omega_j) \leftarrow d_H(\Gamma_i, \Gamma_j) = \max\left\{\sup_{x\in\Gamma_i} d_{\Gamma_j}(x), \sup_{x\in\Gamma_j} d_{\Gamma_i}(x)\right\}$$
$$where \quad d_\Gamma(x) = \inf_{y\in\Gamma} d(x,y) \quad (3)$$

All these possible similarity functions can be utilized to give an advantage over specific shapes in a segmentation process. We will describe that point in Section 3.

3 Ultimate Opening with Shape Constraints

The basic UO computes the residue using only grayscale values. We propose to introduce a shape similarity function ($f(\Omega_i, \Omega_{ref})$) to a reference shape Ω_{ref} within the residue computation (Eq. 4). In that way, the residue of a CC - Ω_i similar to Ω_{ref} is artificially increased, thus its masking becomes more difficult.

$$r_\lambda^\Omega = f(\Omega_i, \Omega_{ref})\, r_\lambda \quad (4)$$

In equation 4, the function $f(\Omega_i, \Omega_{ref})$ is related to the similarity function $\psi(\Omega_i, \Omega_j)$ via a weighted function $w\{\}$ and a multiply factor α as describes Eq. 5. The weighted function is used to centre, normalize and specify shape measures. $1.0+\alpha$ represents the maximum value that the function reaches. In addition, a set of limits of shape Ω_i is considered as shape information. If Ω_i feature lies within the range of limits, the residue is advantaged, otherwise it is not favored.

$$f(\Omega_i, \Omega_j) = \begin{cases} 1.0 + \alpha w\{\psi(\Omega_i, \Omega_j)\} & \forall \Omega_i \in limits \\ 1.0 & otherwise \end{cases} \quad (5)$$

Additionally, shape factor ($f(\Omega_i, \Omega_j)$) is stored on an image $F_\theta^\Omega(x)$ when the maximal residue ($R_\theta(x)$) is generated. With this information, we modify the original expression of UO Eq. 1 by Eq. 6.

$$\left.\begin{array}{l} R_\theta^\Omega(I) = \sup\left(r_\lambda^\Omega(I)\right) \\ F_\theta^\Omega(x) = f(\Omega_i, \Omega_{ref}) \\ q_\theta^\Omega(x) = \max(\lambda) : \lambda \geq 1 \end{array}\right\}, \quad R_\theta^\Omega(I) = r_\lambda^\Omega(I), \quad R_\theta^\Omega(I) > 0 \quad (6)$$

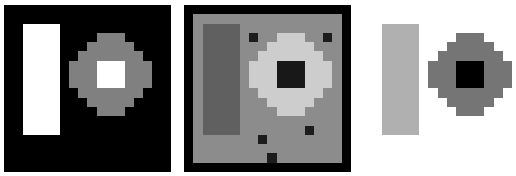
With our approach, UO can be combined with almost all types of shape measures, thanks to the independence

between the computation of shape information and the classical process of UO. Nevertheless, we must be careful with the selection of measures, because these measures will normally be computed each time after each opening, and it will require more computing time. To keep a reasonable computing time, we have used the simplest attributes of shapes.

Now, we test the example of a synthetic image (Fig. 1(a)) to analyze our approach. In regular shapes (e.g., circle or square), the most common measures are based on the perimeter (L_Ω), area (A_Ω) and bounding box area ($Abbox_\Omega$) ratios of shape Ω. Firstly, we have considered the case of fill ratio Fr as shape metric, that is $\frac{A_\Omega}{Abbox_\Omega}$. 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. Then, we have imposed maximum and minimum area limits to validate the shape. Finally, we have used a linear function as weighted function. This shape information is translated into Eq. 7.

$$f\left(\Omega_i, \Omega_{ref}\right) = \begin{cases} 1.0 + \alpha Fr_{\Omega_i} & \forall \Omega_i : 2\%A_I < A_{\Omega_i} < 90\%A_I \\ 1.0 & otherwise \end{cases} \tag{7}$$

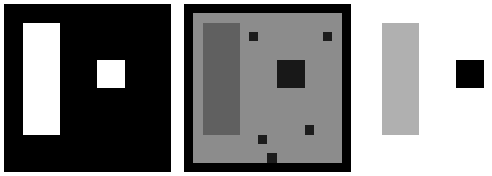
where, $\alpha = \max(I)/2$, $max = 255$. We suppose that Ω_{ref} is a rectangular shape without rotation, i.e. $Fr_{\Omega_{ref}} = 1$. We utilize two area limits (2% and 90 % of image area A_I) to reject the smallest and the largest regions. Fig. 3 presents the result of our approach on a synthetic image. In this case, the masking problem is solved and the three shapes are segmented. The importance of limits is remarkable in this example, because noise points and the last closing shape reach a high factor $Fr_{\Omega_i} \approx 1$. Using limits, these shapes have a factor equal to 1.0.



(a) $F_\theta^\Omega(x)$ (b) $R_\theta^\Omega(I)$ (c) $q_\theta^\Omega(x)$

Figure 3. Segmentation of a synthetic image using UAC with shape constraints

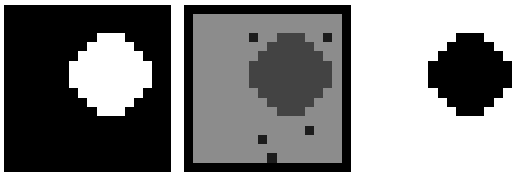
If we want to segment rectangular shapes, we modify Fr_{Ω_i} in Eq. 7 by cube value of Fr_{Ω_i}. Another example of shape factor is implemented to give preference to circle shapes. In circle detection, the most frequently used metric is the circularity. The metric is the ratio of the shape area to a circle area having the same perimeter ($\frac{4\pi A_\Omega}{(L_\Omega)^2}$). Eq. 8 shows factor information after changing Fr_{Ω_i} in Eq. 7 by circularity expression. Fig. 4 and Fig. 5 that confirm rectangular shapes and circular shapes are segmented, respectively.



(a) $F_\theta^\Omega(x)$ (b) $R_\theta^\Omega(I)$ (c) $q_\theta^\Omega(x)$

Figure 4. Segmentation of rectangular shapes of a synthetic image using UAC with shape constraints

$$f\left(\Omega_i, \Omega_{ref}\right) = \begin{cases} 1.0 + \alpha \frac{4\pi A_\Omega}{(L_\Omega)^2} & \forall \Omega_i : 2\%A_I < A_{\Omega_i} < 90\%A_I \\ 1.0 & otherwise \end{cases} \tag{8}$$



(a) $F_\theta^\Omega(x)$ (b) $R_\theta^\Omega(I)$ (c) $q_\theta^\Omega(x)$

Figure 5. Segmentation of circular shapes of a synthetic image using UAC with shape constraints

4 Applications

The aim of the proposed method is to improve UO segmentation results avoiding masking problems. In order to demonstrate the performance and to test our method, we illustrate two segmentation applications: facade image analysis and scene-text detection.

4.1 Façade Image Analysis

Several approaches of urban environment modeling have focused on coarse modeling, for instance: polyhedral representation, main walls, roof planes and ground planes. Nevertheless, the last research issues try to analyze building façade in real images. This analysis extracts and reconstructs windows, doors and ornaments to provide rich information of the buildings and to add realism for visualization. Our goal is the automation of the façade interpretation from images; especially to facilitate the extraction of semantic/grammatical information. We are focused on a segmentation procedure behind the façade modeling [10, 11] to detect/extract structural objects, mainly windows.

Initially, we have employed ultimate an attribute opening to segment façade images. For the façade structure detection, height attribute of the CCs' bounding box is analyzed. Fig. 6 shows two examples of UAO using a color gradient. In the first example, the operator shows an acceptable segmentation of façade internal structures (Fig. 6(b)). On the second 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.

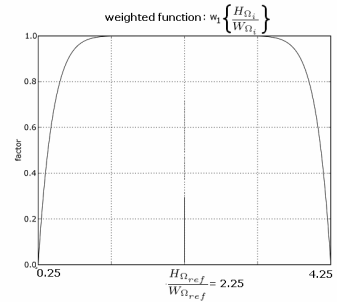
(a) Original Image (b) $q_\theta(x)$ (c) $R_\theta(I)$

(d) Original Image (e) $q_\theta(x)$ (f) $R_\theta(I)$

Figure 6. Example of UAO on façade image

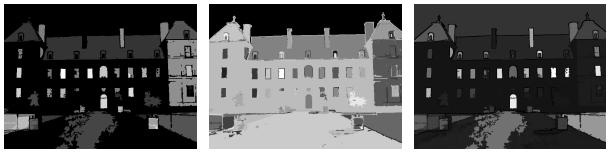
In façade images, windows and doors have particular features. Mayer and Reznik [12] describe that most windows are at least partially rectangular and the height-width ratio (HWr) of a window lies generally between 0.25 and 5. With these features, we can define a similarity function for the internal structures of façades.

$$f(\Omega_i, \Omega_{ref}) = \begin{cases} 1.0 + \alpha w_1\left\{HWr_{(\Omega_i,\Omega_{ref})}\right\} w_2\left\{Fr_{\Omega_i}\right\} & \forall \Omega_i \in limits \\ 1.0 & otherwise \end{cases} \quad (9)$$

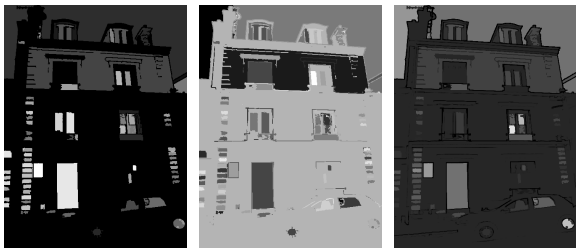


In the case of the synthetic image, α was selected as $max/2$, however after several tests with façade images, the contrast between the sky and the façade is ten times bigger than the one between the wall and the windows, for this reason, we have chosen $\alpha = 9$. Next, we have associated a weighted function to each shape metric. The first metric is HWr and we have used a weighted function centred on $HWr_{\Omega_{ref}} = 2.25$ and equal to $1 - x^{100}$, value 100 allows a large range in the limits of HWr. The limits have been defined on $HWr_{\Omega_{ref}} \pm 2.0$. Fig. 7(d) shows the weighted function utilized in the segmentation process. The second metric is Fr because we suppose that windows are partially rectangular $(Fr_{\Omega_{ref}} = 1)$. We have employed a square value of Fr_{Ω_i} as weighted function to penalize non- rectangular shapes or rectangular shapes with holes.

The results of shape - constraints segmentation on façade images are illustrated in Fig. 7. Our approach presents a better segmentation in both cases. In the first case, where the masking problem is less important, some "new" windows have been found. Likewise, interest structures appear in the second case (seven over nine windows/doors). Nevertheless, some structures which are neither windows nor doors become visible on image segmentation, for example flowerpots Fig. 7(b) and bricks Fig. 7(e).



(a) $F^\Omega_\theta(x)$ (b) $q^\Omega_\theta(x)$ (c) $R^\Omega_\theta(I)$



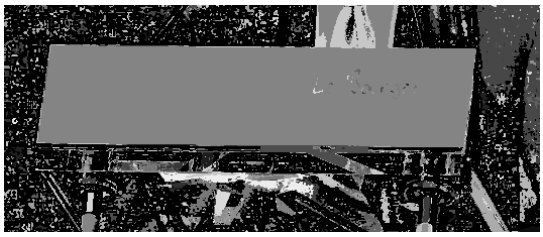
(d) $F^\Omega_\theta(x)$ (e) $q^\Omega_\theta(x)$ (f) $R^\Omega_\theta(I)$

Figure 7. Example of UAO with shape constraints on façade image

4.2 Scene-Text Detection

The text present 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 the surface of different colors such as: placards, posters, etc; and favoring the visibility of letters. But this surface is also contrasted in comparison to its surrounding (Fig. 8(a)). When we utilize UAO, characters may be masked by the contrast of the signboard with its surroundings (Fig. 8(b)).

(a) Original Image



(b) $q_\theta(x)$

(c) $R_\theta(I)$

Figure 8. Example of UAO on text image

(a) $F_\theta^\Omega(x)$

(b) $q_\theta^\Omega(x)$

(c) $R_\theta^\Omega(I)$

Figure 9. Example of UAO with shape constraints on text image

In the same way, we propose some features of text to define shape information as follows:

- The range of height/width ratio lies mostly between 0.8 and 2.
- The range of Fr falls approximately between 0.5 and 0.9.

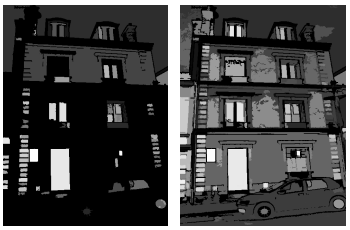
We have utilized a similarity function analogous to façade application (Eq. 9). For HWr and Fr metrics, the same weighted window has been used and it is centred on $HWr_{ref} = 1.4$ and $Fr_{ref} = 0.7$, respectively.

In Fig. 9, the example image shows the results of the text detection using UO with shape constraints. The results from this preliminary study indicate that the proposed method is superior to the classical UO segmentation.

5 Discussion

UO provides two pieces of information, contrast ($R_\theta^\Omega(x)$) and size ($q_\theta^\Omega(I)$). The proposed method provides a third interesting piece of information: the shape function ($F_\theta^\Omega(x)$), that conveys a shape similarity measure with a reference shape. We store this factor when the maximal residue is generated.

In façade example (Fig. 10(a)), we can see that one window with shutter (down - right) is not detected, even if this shape is similar to a rectangle. The reason is the low contrast and the shape factor is not sufficient to avoid the masking effect. Fig 10(b) shows the maximum of shape factor ($\widehat{F_\theta^\Omega}(x)$) during the UO process. We observe that the factor of the window with shutter is higher than the shape factor of third floor windows ($5.7 > 2.4$). In fact, the window is still masked because of its low contrast.



(a) $F_\theta^\Omega(x)$ (b) $\widehat{F_\theta^\Omega}(x)$

Figure 10. Factor analysis façade example

Fig. 11 illustrates factor images for text example. Several letters are not detected. "m" case is not valued ($f(\Omega_i, \Omega_{ref}) = 1.0$) because it is outside of the weighted windows. In "f" and "t" cases, the factor value is not big enough to unmask them. On the other hand, many CCs of noise are valued with a factor function and they are still masked thanks to their low contrast.

(a) $F_\theta^\Omega(x)$

(b) $\widehat{F_\theta^\Omega}(x)$

Figure 11. Factor analysis text example

6 Conclusions and Future Work

In this paper we present a novel 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 computation of shape method and the classical process of UO. The proposed method has been validated in two applications of structure extractions from façade and text images. This method produces much better segmentation results than the standard UO.

In the future, we will analyze in details the shape factor information of the segmentation algorithm in order to improve the performance on large databases. The detection process is the first step in computer vision problem. We intend to apply a machine learning process using regional features (shape and color descriptors) to classify regions in both applications. As well, this machine learning could be used as shape factor into the proposed method.

7 Acknowledgements

This work is supported by by the Pôle de Competitivé Cap Digital through TerraNumerica project.

References

[1] Hieu T. Nguyen and Qiang Ji, "Improved watershed segmentation using water diffusion and local shape priors," pp. 985–992, 2006.

[2] Fernand Meyer, "An overview of morphological segmentation," *IJPRAI*, vol. 15, no. 7, pp. 1089–1118, 2001.

[3] Serge Beucher, "Numerical residues," *Image Vision Computing*, vol. 25, no. 4, pp. 405–415, 2007.

[4] Thomas Retornaz and Beatriz Marcotegui, "Ultimate opening implementation based on a flooding process," 2007, The 12th International Congress for Stereology.

[5] Edmond J. Breen and Ronald Jones, "Attribute openings, thinnings, and granulometries," *Computer Vision and Image Understanding*, vol. 64, no. 3, pp. 377–389, 1996.

[6] Souhaïl Outal, *Quantification par analyse d'images de la granulométrie des roches fragmentées : amélioration de l'extraction morphologique des surfaces, amélioration de a reconstruction stéréologique*, Ph.D. thesis, CMM/GEOSCIENCES - École Mines Paris, 2006.

[7] Thomas Retornaz and Beatriz Marcotegui, "Scene text localization based on the ultimate opening," in *International Symposium on Mathematical Morphology, 8 (ISMM)*, October 2007, vol. 1, pp. 177–188.

[8] Jorge Hernandez and Beatriz Marcotegui, "Segmentation of façade images using ultimate opening," Tech. Rep., CMM-Mines ParisTech, 2007.

[9] Guillaume Charpiat, Olivier Faugeras, and Renaud Keriven, "Approximations of shape metrics and application to shape warping and empirical shape statistics," *Foundations of Computational Mathematics*, vol. 5, no. 1, pp. 1–58, 2005.

[10] K. Leinemann J. Benner, A. Geiger, "Flexible generation of semantic 3d buildings models," in *First International Workshop on Next Generation 3D City Models*, 2005.

[11] Pascal Müller, Gang Zeng, Peter Wonka, and Luc Van Gool, "Image-based procedural modeling of façades," pp. 85–93, 2007.

[12] Helmut Mayer and Sergiy Reznik, "Building façade interpretation from image sequences," in *CMRT05*, 2005, vol. XXXVI, pp. 55 – 60.