Tessella-oriented segmentation and guidelines estimation of ancient mosaic images

Lamia Benyoussef
Stéphane Derrode
Institut Fresnel (CNRS UMR 6133)
GSM Group and École Centrale Marseille
Technopôle de Château-Gombert
8, rue Frédéric Joliot Curie
13451 Marseille Cedex 20, France
E-mail: lamia.benyoussef@ec-marseille.fr

Abstract. *Automatic segmentation and analysis of ancient mosaic images can help archaeologists and experts build digital collections and automatically compare mosaics by means of image database indexing and content-based retrieval tools. However, ancient mosaics are characterized by low contrast colors and irregular tessella shape, orientation, and positioning, making automatic segmentation difficult. We propose a tessella-oriented strategy whose first step consists of isolating tessellas from their cemented network by computing the watershed transformation of a criterion image generated to exhibit the cement network as watershed crests. Then a simple k-means algorithm is used to classify tessellas and segment mosaic images with more accuracy than with a pixel-oriented strategy. Additionally, we propose a method to automatically obtain the main directional guidelines of mosaics by estimating tessella orientation. This is done by minimizing a contextual energy computed from gray-level means of neighboring tessellas and the orientation of their borders. Several examples of cartographies show the effectiveness of the method. © 2008 SPIE and IS&T.* [DOI: 10.1117/1.3013543]

1 Introduction

The aim of this work is (1) to analyze ancient mosaic images and (2) to characterize their structure and color by means of automatic processing tools. The final goal is to detect and localize objects with a semantic meaning such as animal, human, and object in a complex mosaic scene. This can help archaeologists and experts in their historical and artistic studies, especially the analysis of ancient mosaicist styles i.e., *opus musivum* and *opus vermiculatum*[1]). Such tools can also be of interest (1) for museums in order to categorize mosaics and to draw up a digital inventory of their collection [2] and (2) for computed-aided generation of old-style mosaic images from a master image (see Ref. 3 for an overview of digital mosaic frameworks and references cited therein). A first attempt to propose a content-based and image retrieval system dedicated to ancient mosaic images has been presented in Ref. 4. In this work, efforts have been focused on pattern recognition aspects, by using an invariant description of semantic objects present in scenes using a Fourier-Mellin transform. [5,6] Semantic objects are isolated using statistical segmentation and morphological operators, but extraction remains a difficult task inherent to the way mosaics are built.

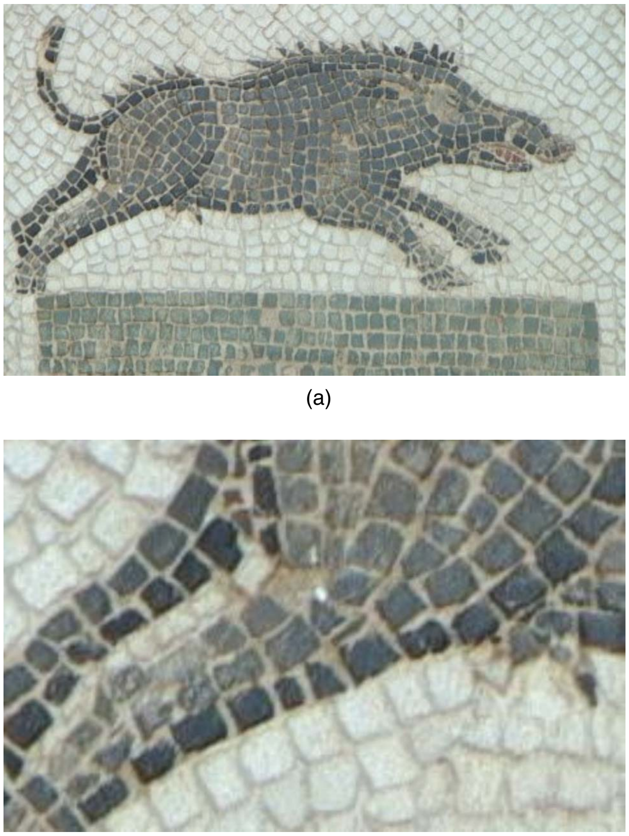
Mosaics are made of colored tiles, called tessera or tessella, usually formed in the shape of a cube of materials separated by a cement joint. Smart and judicious use of orientation, shape, and size of tessellas characterize the artwork style and exhibit the "general flow" of the mosaic chosen by the mosaicist. Figure 1 shows a typical example of a mosaic image to be processed. These kinds of images show specific difficulties inherent to their age and artwork style:

- Tessellas of ancient mosaics are characterized by pastel colors, with low contrast. Color information is not discriminant, and gray-level values are generally sufficient to describe color dynamics in such an image.
- The shapes of tessellas are irregular, from square shapes to polygonal shapes. Their positioning and orientation are not aligned according to a rectangular grid.
- The positioning of tessellas makes the joint appear as an irregular network with numerous interconnections throughout the mosaic. Network intensity, mainly middle gray, is not uniform through the image because of tessella shadows due to nonflat mosaic surfaces and snapshot acquisition angle.

These particularities make segmentation methods based on pixel values inefficient. Indeed, pixels associated to the cement network interfere with and introduce confusion in the classification process. Hence, the strategy under which this work was conducted is to consider that tessellas are indivisible entities with an almost uniform gray-level value. So the first stage is to extract tessellas from the cement network. In Sec. 2, we present a strategy adapted to the mosaic network specificity. It is based on the watershed transformation of a particular criterion image built from the original image in order to exhibit the cement network as watershed crests. At this point, mosaic images are considered tessella-oriented and not pixel-oriented, i.e., all processing is applied on tiles and not pixels. Hence, it was

Paper 08006RR received Jan. 17, 2008; revised manuscript received Aug. 28, 2008; accepted for publication Sep. 25, 2008; published online Dec. xx, xxxx.

1017-9909/2008/17(4)/1/0/$25.00



(a)

(b)

Fig. 1 Detail of an ancient mosaic showing a boar (a) with a close-up of its hind legs (b). This image will be used later as a guiding thread for the algorithm illustration.

easy to obtain a robust segmentation of mosaic images by using a simple tessella-based k-means algorithm, which outperforms the classical pixel-based one.

Since orientation of tiles has a strong visual influence on the overall perception of the mosaic and also in order to facilitate delimitation of semantic objects in the mosaic scene, we propose a simple and efficient way of estimating the main orientations of tessellas (allowing us to exhibit directional guidelines of mosaics) in ancient mosaic images in Sec. 3. The proposed approach consists of minimizing a contextual energy computed from the mean-gray values of neighboring tessellas and the orientation of their borders. Conclusions and further works are presented in Sec. 4.

2 Tessella Extraction for Mosaic Segmentation

The extraction of a network in an image is a recurrent problem, especially for road extraction from aerial photos[7,8] or for vascular network segmentation from angiographies.[9,10] Several approaches have been proposed. Methods based on contour extraction are widely used and mainly rely on the assumption that the network pixels and neighboring ones have different gray levels in order to compute gradients. But methods based on high-pass filters, such as Harris's corner detector, highlight pixels belonging to the network, not connected components. Higher-level processing detects lines with varying widths.[11,12] Strategies that track the entire network from a starting point[13,14] are difficult to justify in our case study due to the high number of intersections in a typical mosaic network.

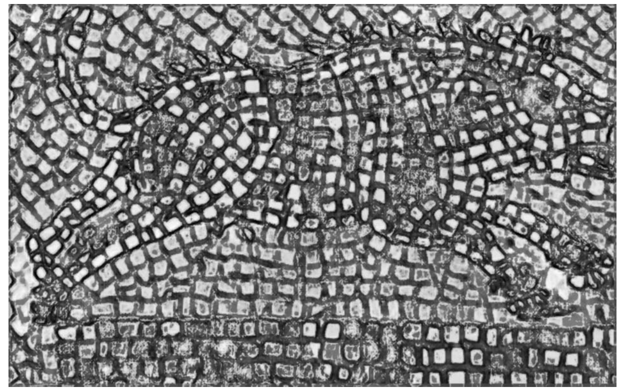
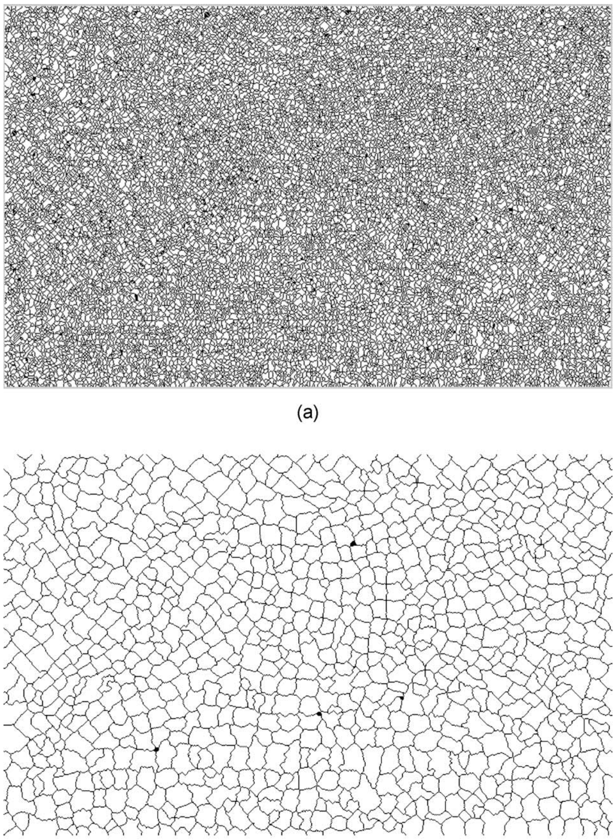


Fig. 2 Criterion image obtained from the boar image in Fig. 1.

Numerous methods based on Markov modeling[15,16] or active contours[17,18] have also been proposed. These methods are quite efficient but time consuming. In the case of mosaics, these methods are not suited because of the high density of the network to be extracted in images. In Ref. 16, a Markov model is applied on a graph of adjacency crests, detected by a watershed transformation (WT) applied on a criterion image. This criterion image, computed from the original image, exhibits the potential of each pixel to belong to the network.

Among those methods, the WT approach appears interesting for mosaic images since this method is a good compromise between low-level methods (contour detection) and approaches by energy minimization (Markov model or active contours), which are unworkable due to cement network complexity in mosaic images. To work well, the WT needs to be computed on a criterion image that shows tessellas as catchment basins and the network as crests. But the network, mainly middle-gray-valued, is sometimes darker than the tessellas and sometimes lighter in the same image. Hence, for each pixel in the image, we study the gray-level profile around it according to four directions (0 deg, 45 deg, 90 deg, and 135 deg) and compare them to two templates characteristic of the two situations, i.e., dark network and light tessellas and light network and dark tessellas. The value of a pixel in the criterion image is the minimum value among the eight values. If this value is high, we face a somewhat flat profile that indicates a pixel inside a tessella.

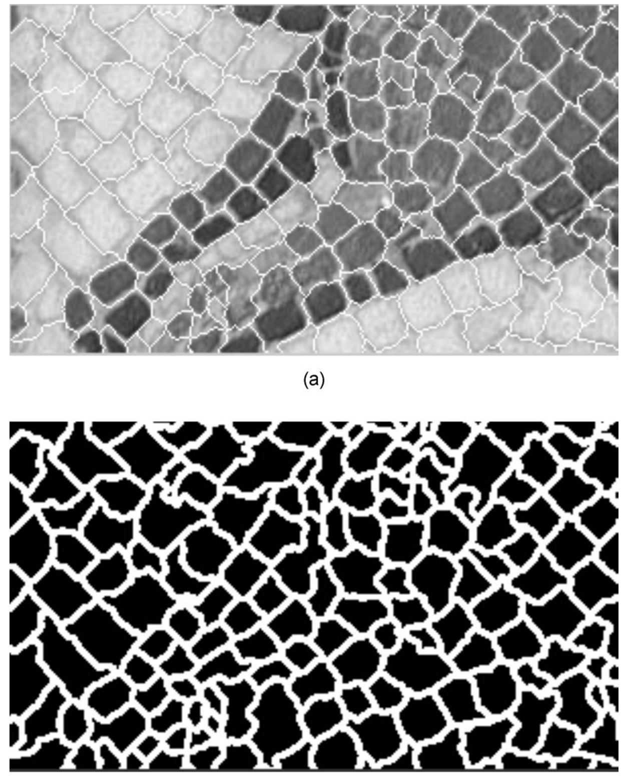
Figure 2 shows the criterion image obtained by applying the method to a boar image. As can be seen in this example, the network appears dark. However, tessellas are not uniform in texture and show local gray-level crests that should be deleted before WT in order to avoid oversegmentation. Following Ref. 16, we first compute an area closing[19] of the criterion image that gives fewer minima while retaining crest locations. The WT result is illustrated in Fig. 3. The crest contours now correctly represent the network, which is confirmed by the close-up shown in Fig. 4(a). To determine the width of the network (and not only a one-pixel skeleton, as is done by WT), which varies through the image, a simple threshold is applied on neighboring pixels of



(a)

(b)

Fig. 3 Extraction of tessellas from the criterion image in Fig. 2, without (a) and with (b) area closing operator.



(a)

(b)

Fig. 4 (a) Result of tessella extraction on the close-up in Fig. 1(b), and (b) network/tessellas classification.

crests; a pixel is aggregated to the crest if its gray value differ by no more than 10% from the skeleton mean-gray value. The result of applying such a threshold is shown in Fig. 4(b).

For segmentation, we are now able to consider only the tessellas of the mosaic, not its network. Moreover, instead of using all pixels from the tessellas, and since tessellas are almost homogeneous in color, we can segment the image by using a tessella-oriented strategy: each tessella is characterized by one or more features used for classification. Simple examples of features are mean gray-level value or variance of the tessella, number of pixels in the tessella, etc. For our application, a simple k-means algorithm on the mean gray-level value of tessellas was sufficient to get a nice segmentation, as illustrated in Fig. 5. This result can be compared with a classical pixel-based k-means strategy. A second segmentation example is given in Fig. 6.

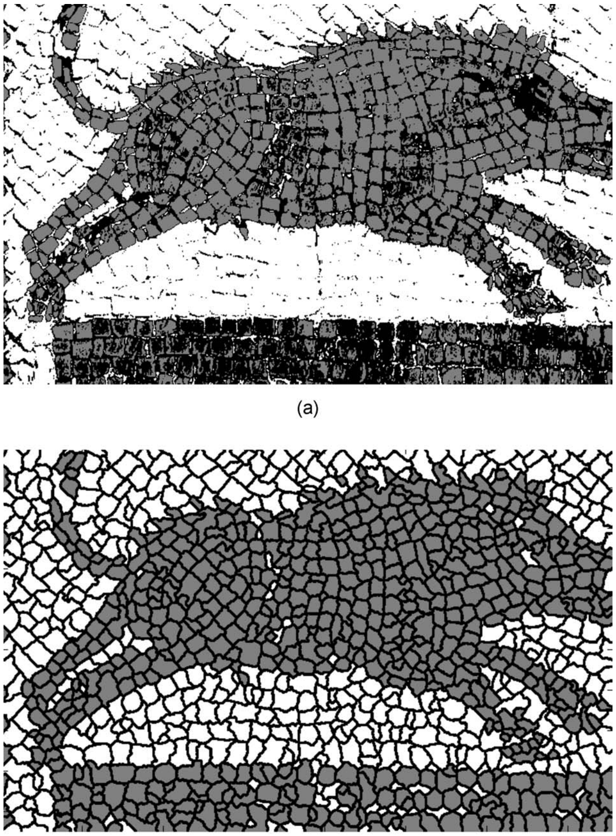
Note: The entire processing is based on two parameters: (1) the length l of the profile to compute the criterion image and (2) the area closing threshold s. These parameters can be set proportional to the mean tessella size α, which is almost constant in a mosaic. Parameter l should be greater than 2α for the profiles to fit at least two tessellas, and parameter s should be less than α^2 to avoid small tessellas being deleted by the morphological operator. In our experiments, values $l=3\alpha$ and $s=\alpha^2/2$ give good results. Coefficient α depends on the image zoom and can be either set by an operator or estimated automatically on a small uniform part of the mosaic for example.

3 Tessella Orientation Estimation for Directional Guidelines Detection

Ancient mosaicists avoided aligning their tiles according to rectangular grids. Indeed, such grids emphasize only horizontal and vertical lines and may distract the observer from seeing the overall picture. Hence, mosaicists placed tiles in order to emphasize the strong edges of the subject to be represented, influencing the overall perception of the mosaic. Thus, organization and positioning of tessellas are therefore interesting information for experts since they emphasize the main directional guidelines chosen by the artist. This information is of crucial interest for mosaic-dedicated applications such as content-based retrieval of mosaic elements or region-based mosaic image compression.

To get directional guidelines, one can first think of using the principal axes of an ellipse-equivalent shape of each tessella, using well-known formulae based on geometrical moments (minor and major axes). However, ancient mosaic tessellas are not box- or regular-shaped, and principal axes quickly appear not robust enough. One major drawback of such a method is that it does not take into account information about neighboring tessellas, which is of great importance for regularization and for recovering the main guidelines that emphasize the general flow of a mosaic.

We propose an energy-based contextual algorithm for retrieving main directional guidelines in a mosaic. The energy to be minimized is constructed by using two key fea-



(a)

(b)

Fig. 5 Segmentation of the mosaic image in Fig. 1 with a pixel-based strategy (a), and the tessella-based strategy proposed here (b), using a k-means algorithm with two classes.

tures: the mean-gray value and the border directions of each tessella. The optimization is done either by gradient descent or by simulated annealing.

3.1 *Methodology*

We denote by N the number of tessellas detected in the mosaic. Each tessella i is represented by

- Its barycenter (x_i, y_i) computed on the support Ω_i of i:

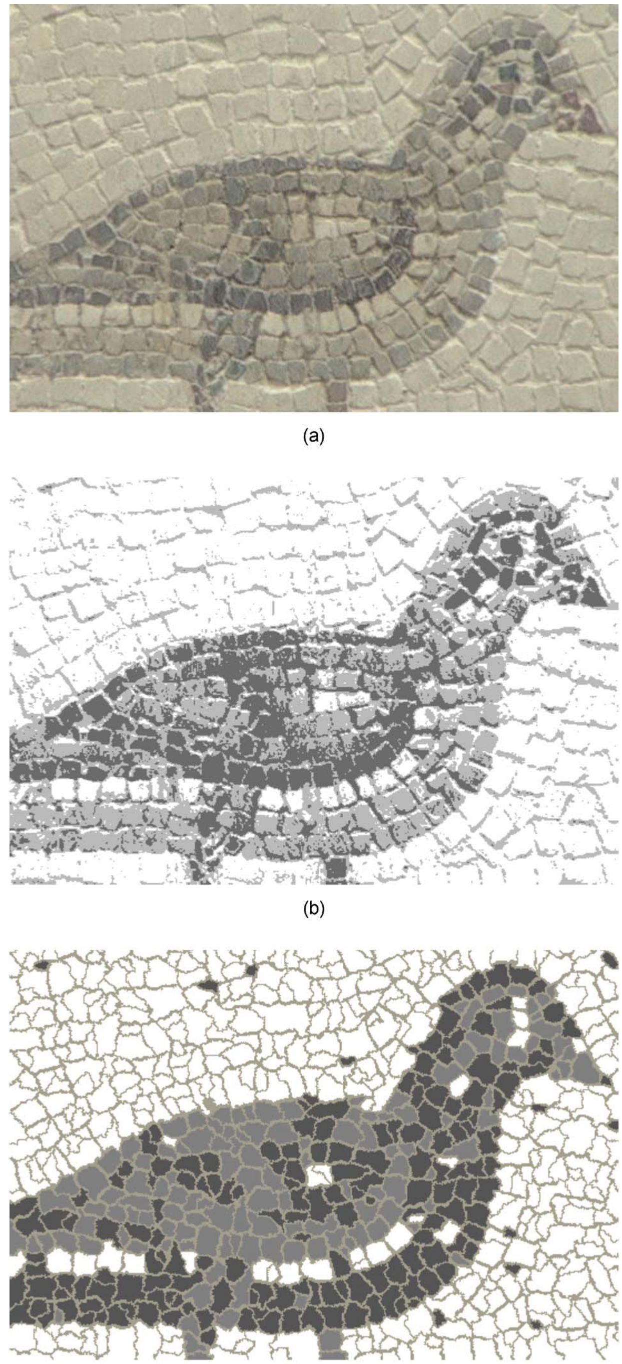
$$(x_i, y_i) = \left(\frac{m_{1,0}}{m_{0,0}}, \frac{m_{0,1}}{m_{0,0}}\right)$$

$$\text{with } m_{p,q} = \int\int_{\Omega_i} x^p y^q f_i(x,y)\, \mathrm{d}x\, \mathrm{d}y.$$

- The list of its neighboring tessellas: $\mathcal{V}_i = \{v_{i,1}, \ldots, v_{i,T_i}\}$. A neighbor is a tessella that shares at least one pixel with i.

It should be noted that the number of neighbors T_i is different from one tile to the other since tessellas are not organized according to a regular grid.

Each tessella i is characterized by an energy of configuration that links itself to each of its neighbor $v_{i,t} \in \mathcal{V}_i$. This energy, denoted by $E_{i,t}, t \in [1, \ldots, T_i]$, is the sum of two complementary terms:



(a)

(b)

(c)

Fig. 6 Segmentation of a mosaic representing a bird (a), with a pixel-based strategy (b), and with a tessella-based strategy (c), using a k-means algorithm with three classes.

- The first term Q is based on the mean-gray value of tessellas. It is proportional to the sum of the difference of gray-level means (1) between i and $v_{i,t}$ and (2) between i and the symmetrical tessella of $v_{i,t}$ with respect to i. This feature favors alignment of tessellas with low contrast, which is a characteristic of directional guidelines.
- The second term R is based on the orientation of tessella contours. We compute the histogram of the ori-

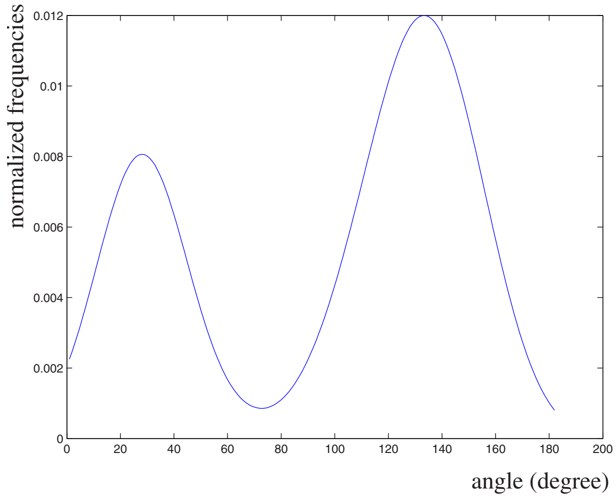


Fig. 7 Plot of the regularized and normalized histogram of the contour orientation of one tile from the boar mosaic in Fig. 1.

entation of segments constituting the contour of tessella i. This histogram is regularized using a Gaussian kernel, the result of which is illustrated in Fig. 7. It should be noted that the two modes at approximatively 90 deg to each other correspond to the two ambiguous orthogonal main directions of a square-shaped tile. It is then possible to estimate the pdf at angle $\alpha_{i,t}$ given by the barycenter of i and that of $v_{i,t}$.

Terms Q and R are normalized to belong to range $[0, 1]$. We can then initialize the "main direction" of a tile, i.e., the direction of the neighboring tessella that gives the highest $Q+R$ value:

$$t_{i,max} = \arg \max_{t \in [1,\ldots,T_i]} E_{i,t}.$$

The energy of a tessella is then defined as $C_i=(2-E_{i,t_{i,\max}})+\lambda V_i$, with λ a weighting factor set manually. Term V_i is defined as

$$V_i = \frac{1}{T_i} \sum_{t=1}^{T_i} |\alpha_{i,t_i,max} - \alpha_{t,t_t,max}|,$$

which is the normalized sum of the absolute difference between the main direction of tile i and the main direction of its neighbor t.

We try next to minimize the mosaic energy, defined as the sum of C_i for all tessellas in the mosaic. This is done by selecting the tessella i that gives the highest value for V_i. To reduce the contribution of this tessella, we try another main direction and recompute the mosaic energy. At that point, two strategies have been tested:

- Deterministic framework [gradient descent (GD)] If the mosaic energy reduces, then the new main direction is validated; otherwise, another main direction is tested. When all directions for this tile have been tested, we repeat the process for the next tile with high V_i value.
- Stochastic framework [simulated annealing (SA)]: A configuration that gives a higher mosaic energy can be validated according to the simulated annealing principle.[20] This strategy allows us to search for the global minimum, which cannot be reached with the previous strategy because the mosaic energy function is not convex.

The process is iterated until the mosaic energy is almost constant. The cartography of tessella orientation consists of the main direction of each tile at the last iteration.

3.2 *Experimental Results*

Figure 8 illustrates the application of the tessella orientation methodology on the close-up of the boar image in Fig. 1. From the initial configuration of tessellas [Fig. 8(a)], we get the final configuration [Fig. 8(b)] using simulated annealing (SA) for optimization. Figure 8(c) shows the evolution of the computed energy during iterations of both the GD algorithm and the SA algorithm. As expected, SA reaches a lower minimum than GD, but to the detriment of numerous additional iterations (150 for SA versus 50 for GD). Indeed, GD searches for a local minimum and is highly dependent on the initial configuration, whereas SA is expected to reach the global minimum due to its stochastic nature.

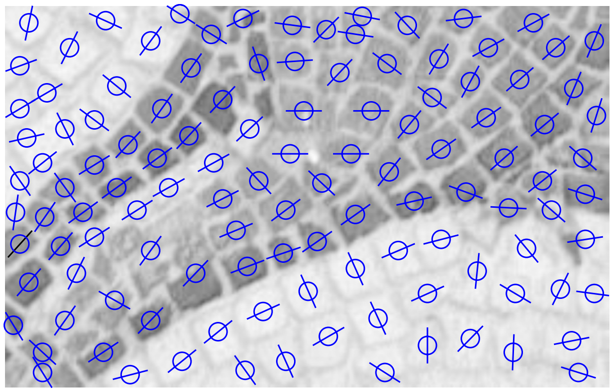
The tessella cartography obtained with SA optimization is very satisfying when visually compared to the main directional guidelines of the mosaic. This is especially true for regions at the borders between classes. A second example of cartography is proposed in Fig. 9. Once again, the tessella orientation estimation methodology, which makes use of contextual information, gives regularized results that emphasize the mosaic guidelines. Nevertheless, confusion can be found in areas with homogeneous colors and where tessellas are square-shaped. Indeed, for those kinds of tessellas, two orthogonal directions are equally probable, which generally gives ambiguous results. However, these areas of uniform color are of limited interest for object-based scene applications, such as mosaic pattern recognition.

Figure 10 shows a failure case in tessella orientation estimation. Indeed, the result will not allow us to find the main directional guidelines in the mosaic, which can be more easily observed in Fig. 6(a). The main reason comes from an overdetection of tessellas from the extraction step. This behavior is observed in mosaics built with tessellas of different sizes (e.g., large tessellas for the background and small ones for objects or details). Hence, the shape of extracted tiles does not correspond to the shape of tessellas, and the orientation is corrupted, showing no particular guideline in the mosaic.

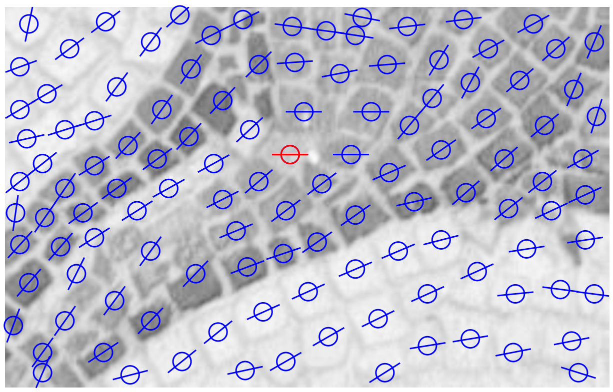
Note: For all experiments, the weighting factor λ has been set to 1. A study not reported here showed the low impact of λ value on the results.

4 Conclusion

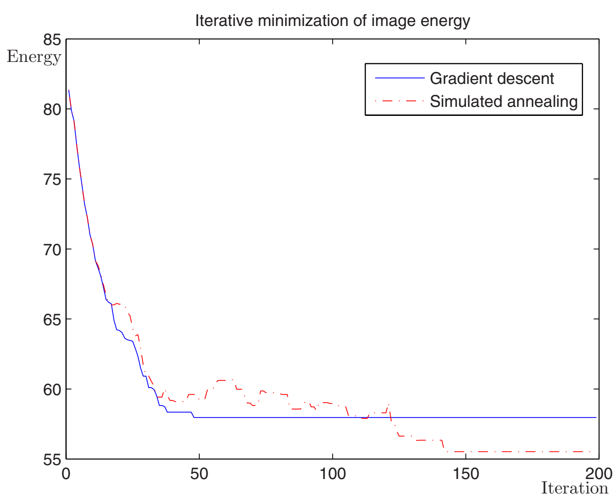
In this work, a method for analyzing the structure and color of ancient mosaic images has been presented, based on a tile-oriented strategy. To extract tessellas from the cement network, we applied a watershed transform (WT) on a criterion image computed from the original image. The criterion image was generated in order to exhibit the cement



(a)



(b)

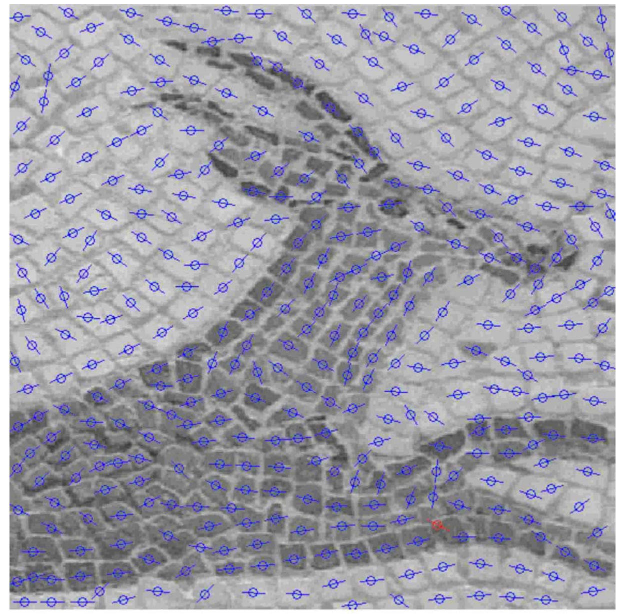


(c)

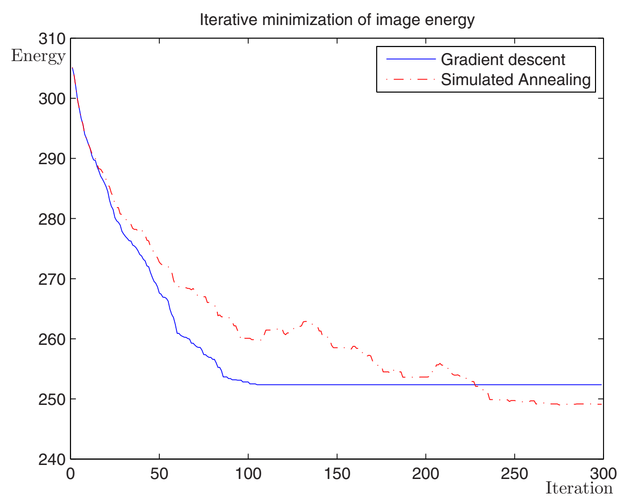
Fig. 8 Cartography of tessellas orientation. Each tessella is characterized by its center of mass (circle) and its orientation (segment crossing the circle).

network as watershed crests and each tessella as a catchment basin. Then, from the individual tiles, we were able to compute a tessella-based k-means classification, using the mean-gray value as a feature to characterize tessellas. Results of segmentation are distinctly of higher quality than those obtained from a pixel-based k-means strategy.

Then, in order to help archaeologists understand mosaic structure and mosaicists' way of working, and also to facilitate extraction of individual semantic objects in complex



(a)



(b)

Fig. 9 Another example of tessella orientation cartography.

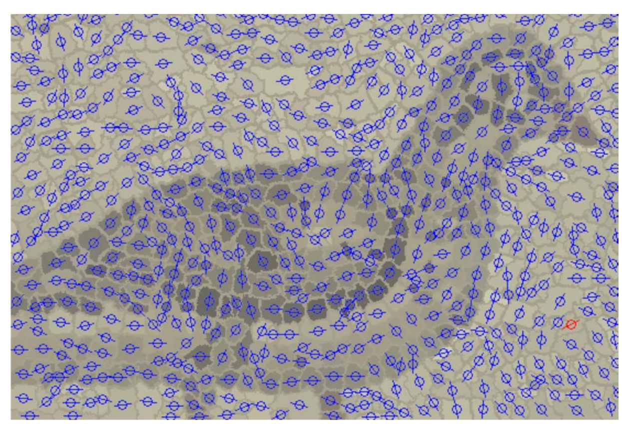


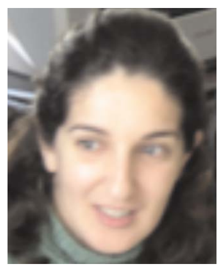
Fig. 10 Result of guidelines detection for the 'bird' mosaic in Fig. 6.

mosaic scenes, we proposed a method to estimate the main directional guidelines of tessellas in mosaics. To that goal, each tessella is described by a contextual energy computed from the mean-gray value and the main directions of tile borders. The minimum energy is searched for by means of the simulated annealing (SA) algorithm. Results showing cartographies of tessella orientation are very interesting because most of the searched guidelines are retrieved, especially those at the borders between objects and scene background.

This processing is the first step toward a system devoted to the indexation and retrieval of semantic objects in mosaic images, which should help archaeologists to compare mosaics from different sites or built at different dates. Future works will include a tessella-based invariant description of objects to enable the comparison between images of mosaics taken at different zooms and orientations, for example. Image compression of ancient mosaics, using a tessella-based coding strategy, is also an interesting perspective, e.g., for a quick look at a remote catalog using the Internet.

References

1. Mosaic Art in Vitreous Glass, Millefiori, Tesserae Mosaics by Shelby Glass Studio, www.mosaic-tileart.com/mosaic.html, 2006.
2. Example: *Musée des Antiquités Nationales*, Saint-Germain-en-Laye, France. www.culture.gouv.fr/culture/arcnat/vienne/en/. *Musée National du Bardo*, Tunisia, www.di.com.tn/museebardo/.
3. S. Battiato, G. D. Blasi, G. M. Farinella, and G. Gallo, "Digital mosaic frame works: an overview," *Comput. Graph. Forum* **26**, 794–812 (2007).
4. M. M'hedhbi, R. Mezhoud, S. M'hiri, and F. Ghorbel, "A new content-based image indexing and retrieval system of mosaic images," in *Third Int. Conf. on Information and Communication Technologies: From Theory to Applications (ICTTA'06)*, Damascus, Syria, pp. 1715–1719 (2006).
5. F. Ghorbel, "A complete invariant description for gray-level images by the harmonic analysis approach," *Pattern Recogn. Lett.* **15**, 1043–1051 (1994).
6. S. Derrode and F. Ghorbel, "Robust and efficient Fourier-Mellin transform approximations for invariant gray-level image description and reconstruction," *Comput. Vis. Image Underst.* **83**, 57–78 (2001).
7. H. Mayer, S. Hinz, U. Bacher, and E. Baltsavias, "A test of automatic road extraction approaches," in *Photogrammetric Computer Vision (PCV'06)*, W. Förstner and R. Steffen, Eds., Vol. 3, pp. 209–214, ISPRS Commission III, Bonn, Germany (2006).
8. J. B. Mena and J. A. Malpica, "An automatic method for road extraction in rural and semi-urban areas starting from high resolution satellite imagery," *Pattern Recogn. Lett.* **26**, 1201–1220 (2005).
9. A. C. S. Chung, "Image segmentation methods for detecting blood vessels in angiography," in *IEEE Int. Conf. on Automation, Robotics, and Computer Vision (ICARCV'06)*, Singapore, pp. 1424–1429 (2006).
10. P. J. Yim, G. Boudewijn, C. Vasbinder, V. B. Ho, and P. L. Choyke, "Isosurfaces as deformable models for magnetic resonance angiography," *IEEE Trans. Med. Imaging* **22**, 875–881 (2003).
11. M. A. Fischler, J. M. Tenenbaum, and H. C. Wolf, "Detection of road and linear structures in low resolution aerial images using multi-source knowledge integration techniques," *Comput. Graph. Image Process.* **15**(3), 201–223 (1981).
12. C. Steger, "An unbiased detector of curvilinear structures," *IEEE Trans. Image Process.* **20**, 113–125 (1998).
13. A. Baumgartner, S. Hinz, and C. Wiedemann, "Efficient methods and interfaces for road tracking," in *Photogrammetric Computer Vision (PCV'02)*, Vol. XXXIV, pp. 28–31, ISPRS Commission III, Graz, Austria, B:28 (2002).
14. O. Wink, W. Niessen, and M. Viergever, "Multiscale vessel tracking," *IEEE Trans. Med. Imaging* **23**, 130–133 (2004).
15. F. Tupin, H. Maitre, J. F. Mangin, J. M. Nicolas, and E. Pechersky, "Detection of linear features in SAR images: application to road network extraction," *IEEE Trans. Geosci. Remote Sens.* **36**(2), 434–453 (1998).
16. T. Géraud and J. B. Mouret, "Fast road network extraction in satellite images using mathematical morphology and Markov random fields," *EURASIP J. Appl. Signal Process.* **2004**(16), 2503–2514 (2004).
17. C. M. van Bemmel, L. J. Spreeuwers, M. A. Viergever, and W. J. Niessen, "Level-set-based artery-vein separation in blood pool agent CE-MR angiograms," *IEEE Trans. Med. Imaging* **22**, 1224–1234 (2003).
18. M. Rochery, I. H. Jermyn, and J. Zerubia, "Higher order active contours," *Int. J. Comput. Vis.* **69**, 27–42 (2006).
19. L. Vincent, "Grayscale area openings and closings: their applications and efficient implementation," in *Proc. EURASIP Workshop on Mathematical Morphology and Its Applications to Signal Processing*, Barcelona, Spain, pp. 22–27 (1993).
20. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science* **220**, 671–680 (1983).



Lamia Benyoussef received her electronics and informatics engineering degree from the Department of Computer Science of the University of Sciences, Tunis, Tunisia, in 1996. She received her PhD degree in signal and image processing from Bordeaux University, France in 2004. From 2003 to 2007, she was with the College of Higher Education of Sciences and Technics of Tunis (ESSTT) and the GRIFT Department of Ecole Nationale des Sciences de l'Informatique, Tunisia, as an associate professor. Since September 2007, she has been an associate professor in the GSM Group, Institut Fresnel (CNRS UMR 6133), Marseille, France. Her research interests include group theoretical image representation, registration methods, and Markov models for segmentation.

Stéphane Derrode received his telecommunication engineering degree from Télécom Lille 1. France, in 1995, and his PhD degree from the University of Rennes I, France, in 1999. From 1999 to 2001, he worked as a research engineer at the Ecole Nationale Supérieure des Télécommunications de Bretagne (ITI Department), Brest, France. Since September 2001, he has been with the Ecole Centrale Marseille where he is currently an associate professor in the GSM Group, Institut Fresnel (CNRS UMR 6133). His research interests include invariance, group theoretical image representation, and Markov models for pattern recognition, image indexing, and segmentation.