Pattern Recognition 42 (2009) 1113 – 1125

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Graph-based tools for microscopic cellular image segmentation

Vinh-Thong Ta*, Olivier Lézoray, Abderrahim Elmoataz, Sophie Schüpp

Université de Caen Basse-Normandie, GREYC CNRS UMR 6072, ENSICAEN, 6 Boulevard Maréchal Juin, F-14050 Caen Cedex, France

ARTICLE INFO

Article history:
Received 30 November 2007
Received in revised form 29 September 2008
Accepted 23 October 2008

Keywords:
Cytological and histological images
Pathology
Weighted graphs
Image processing tools
Discrete regularization
Fast image processing
Automatic and interactive segmentation schemes

ABSTRACT

We propose a framework of graph-based tools for the segmentation of microscopic cellular images. This framework is based on an object oriented analysis of imaging problems in pathology. Our graph tools rely on a general formulation of discrete functional regularization on weighted graphs of arbitrary topology. It leads to a set of useful tools which can be combined together to address various image segmentation problems in pathology. To provide fast image segmentation algorithms, we also propose an image simplification based on graphs as a pre processing step. The abilities of this set of image processing discrete tools are illustrated through automatic and interactive segmentation schemes for color cytological and histological images segmentation problems.

1. Introduction

During the last decade, pathologists and biologists have witnessed major advances in computer hardware and software technologies. The first major impact of these technologies was in the management of text information and has led to medical knowledge management systems. Nowadays, traditional knowledge management systems in the medical field, predominantly focus on medical knowledge and problem solving like diagnosis, prognosis, therapy planning, image processing with image classification, and teaching or practical medical knowledge. Among all these fields, image processing systems have become a very promising one. The main reason is that the advances in imaging technologies have led to an important increase of the manipulation and use of digital images.

In recent years, computer aided image processing and analysis systems have played a significant part in quantitative pathology. Pathologists traditionally make a diagnostic decision by viewing a specimen and measuring various diagnostically important attributes of an isolated object such as size, shape, darkness, color and texture. This is a complex process. Digital images can be analyzed to extract feature information, processed to be enhanced, archived to help diagnostic efforts or transmitted to facilitate the communication between pathologists. These components can be integrated within information management systems to make the basis for a workstation for diagnosis. Image processing techniques are of special interest for pathologists because they can be more reliable than classical assessment by eye screening using poorly defined criteria that cannot be reliably reproduced. Such imaging technologies are likely to improve the productivity and diagnostic ability of pathologists. In particular, the detection of the rare cellular cancer events is a key problem to address. However, to perform quantitative analysis directly on images, quantitative data have to be extracted from the latter. To this end, it is inevitable to perform image segmentation interactively or automatically. Once trying to assist the pathologist in his work, massive quantities of data have to be processed and evaluated, an automation of the segmentation step is absolutely necessary to take care of the reproducibility of the results, barring therefore the problem of heavy work load for human experts.

In microscopic cellular imaging, the objective of segmentation is the extraction of cellular or tissue components. This problem is a difficult problem due to the large variations of the features present in such structures. There are several strategies for segmenting images in literature. For instance, methods based on histogram analysis, pixel classification, morphological operators or partial derivative equations (PDEs) approaches can be mentioned. Their performances depend largely on the type of images to be processed and on a priori knowledge relative to the object features. PDEs and morphological segmentations are very popular methods for segmenting medical images. PDEs-based methods are very effective tools which enable to perform a lot of different image processing tasks under an unified formalism. Malladi and Sethian [1], have

* Corresponding author. Tel.: +33 231452706; fax: +33 231452698.
E-mail addresses: vinhthong.ta@unicaen.fr (V.-T. Ta), olivier.lezoray@unicaen.fr (O. Lézoray), abder.elmoataz@greyc.ensicaen.fr (A. Elmoataz), sophie.schupp@unicaen.fr (S. Schüpp).
URL: http://www.info.unicaen.fr/~vta (V.-T. Ta).

0031-3203/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2008.10.029

proposed a geometrical model based on PDEs for filter and segment histological images [2]. Many other PDEs-based resolution schemes have been presented so far in literature (see for instance [3,4], and references therein for more details on variational methods). Moreover, recent data sets analysis and machine learning methods have been developed. They are based on graph Laplacian diffusion processes and have been used to perform data sets classification [5,6], dimensionality reduction [7] or interactive image segmentation based on label diffusion methods [8,9]. If we consider an image as a set of pixels, graph Laplacian classification is difficult to use, due to the great mass of data to analyze. Then, a data reduction or image simplification can be performed to provide a faster resolution [10]. Indeed, microscopic image processing is mainly required for quantitative analysis of masses of images and the used image processing methods have to be fast and with few parameters for an easy model selection. This is not really the case with PDEs. Alternatively, unregularized methods have been proposed [11]. Morphological methods are also very popular for processing microscopic images [12] with an object oriented approach (i.e. a priori knowledge on the objects to extract is used).

In this work, we use our recently proposed discrete regularization framework based on weighted graph [13] to address the image segmentation problem. This framework is inspired by continuous regularization and data dependent function analysis methods. It provides a common formulation for a wide range of applications in various domains. In this paper, we focus on microscopic cellular image segmentation and classification problems. Microscopic cellular image segmentation is an application dependent task and no general scheme can be given. Therefore, we propose a set of graph-based tools to address cellular extraction problems. These tools constitute a framework. Within this framework, a large variety of operations can be performed, combined or derived to produce a specific segmentation scheme for a given problem. This framework leads to a family of linear and nonlinear filters, and provides label-based diffusion processes for image automatic and interactive segmentation. One strong specificity of the proposed framework is to use graphs as a discrete modeling of images at different levels (pixels or regions) and different component relationships (grid graph, proximity graph, etc.). Working on graphs of arbitrary structure, our framework leads to a set of flexible tools for image segmentation, image regularization or clusters extraction. To provide a fast segmentation scheme, we also propose a discrete image simplification inspired by an approach based on the generalized Voronoi diagram [14]. This simplification scheme can be viewed as an image/graph reduction. The main purpose of this paper is not to solve a particular class of cytology or histology problems but to show how, with our graph-based tools, we can provide particular schemes to address several classes of problems in microscopic images segmentation.

The rest of this paper is organized as follows. In Section 2 we study and attempt to model microscopic imaging problems. We describe in Section 3 the set of graph-based tools by formulating the discrete regularization framework, the image simplification by discrete energy partition and the label propagation approach. Section 4 shows applications for color cytological and histological image segmentation in pathology. It also describes segmentation schemes (combining the above mentioned graph-based tools) to perform specific segmentation tasks and compare a particular scheme with other methods. Finally, Section 5 summarizes this paper and proposes future works.

2. Modeling of microscopic imaging problems in pathology

Pathology is roughly composed of two sections: cytology and histology. For both these sections, the visual inspection of cellular specimens and histological sections through a light microscope plays an important part in clinical medicine and biomedical research.

2.1. Color cytology

Cytology literally means the study of cells. It studies morphological features of human body fluid cells which are put on a glass slide and stained. The study of the modification of the main cellular components (nucleus and cytoplasm) is the ground of the cytological study. The morphological features of cells are visually evaluated by cytotechnologists and cytopathologists and these features involve several notions including size, shape, color, texture and topography. The interaction between nucleus and cytoplasm is also of interest: the position of the nucleus in the cytoplasm, the nucleus–cytoplasm area ratio, the position of nucleols in the nucleus, the color and the granularity of the cytoplasm. Figs. 1(a)–(c) present several different cytological images from different body fluids with different stainings and magnifications.

2.2. Color histology

Histology is defined as the anatomical study of the microscopic structure of tissues. In a clinical setting, histology is used to analyze disease states at a cellular level by means of light and/or electron microscopy, histochemistry and immunochemistry. Histology studies cells which are grouped in big and complex structures: tissues. The latter cannot be only characterized by the properties of individual cells such as color staining intensity or expression of specific proteins but also by the geometric arrangement of cells and by the topographic relationships between cells. Figs. 1(d)–(f) present several histological images from different organ tissues with different stainings and magnifications.

2.2.1. Image object oriented modeling

As discussed above, image processing methods are of high interest to provide Image Decision Guided Systems (IDGS) to perform prognostic, diagnostic and early detection of cancer [15]. For the case of problems in pathology, image processing can be used for several tasks: quantification of a cellular content (DNA, proteins, color), recognition and sorting of cellular types, extraction of cellular groupings or clusters and of their topographic relations. In terms of image processing problems, the objectives concern segmenting or analyzing objects at different levels: cellular or cellular grouping level. Any cytological or histological images have common properties and can be described by an image object oriented modeling [16]. First of all, the images have to be acquired in color. This is essential to follow the exact visual way the pathologist follows to evaluate microscopic images. For a microscopic image, one can always divide it into two parts: the background and the rest of the image, namely the objects to be extracted. The background is always close to a given color and is usually homogeneous even if some debris or artifacts can occur (some mucus for instance). The rest of the image is composed of the elements of interest for the pathologists. In this work, we use an image object oriented modeling of the elements of an image: they can be classified and characterized in terms of size, shape, color and texture. The objectives of the segmentation can appear at two object levels: cellular or cellular grouping. For each one of these levels, different configurations can be established. At the cellular level, one can find isolated, touching or overlapping cells. At the cellular grouping level, one can find groupings (groups of cells) or clusters (groups of groupings). The cellular grouping level can be seen as an upper level of the cellular level: cells $\subset$ groupings $\subset$ clusters. Fig. 2 summarizes this image object oriented modeling of microscopic imaging problems in pathology. Before trying to conceive any image segmentation technique, a pool of representative images has

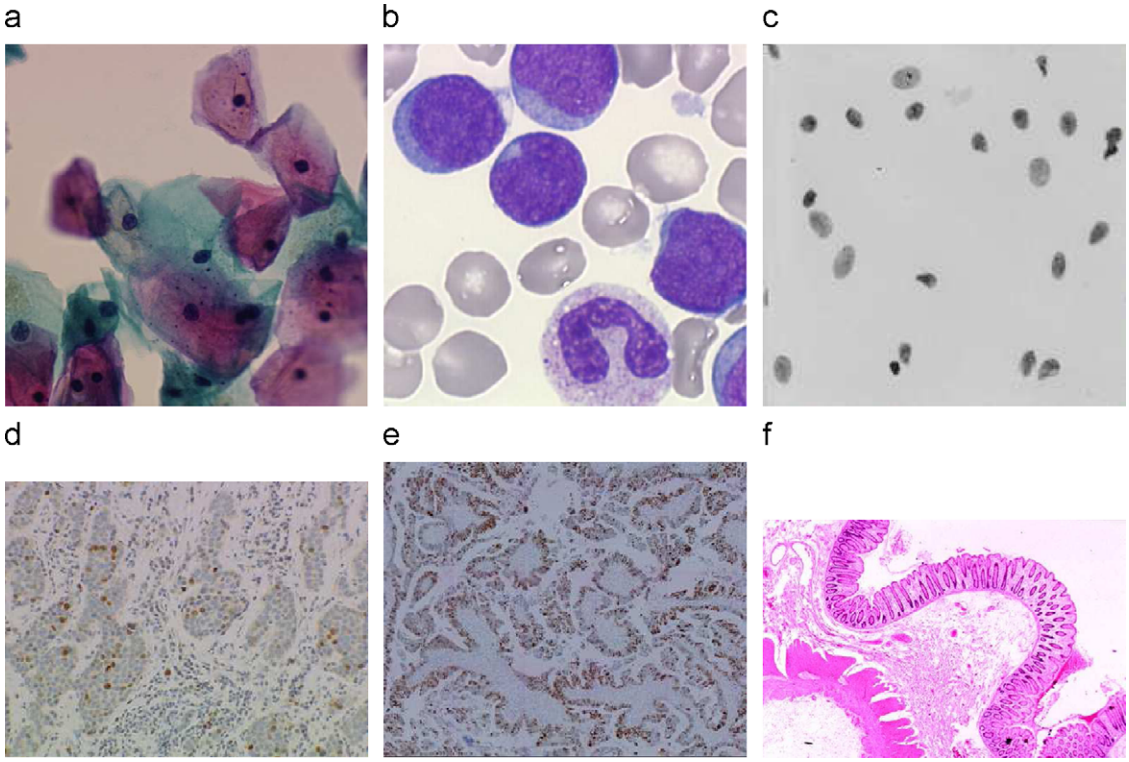

Fig. 1. Body fluid images with different stainings and magnifications. First row: color cytological images. Second row: color histological images. (a) Gynecology (Papanicolaou, ×20). (b) Hematology (H&E, ×100). (c) DNA (Feulgen, ×63). (d) Breast (Immuno-staining, ×33). (e) Breast (Feulgen, ×33). (f) Colon (High magnification, ×66).

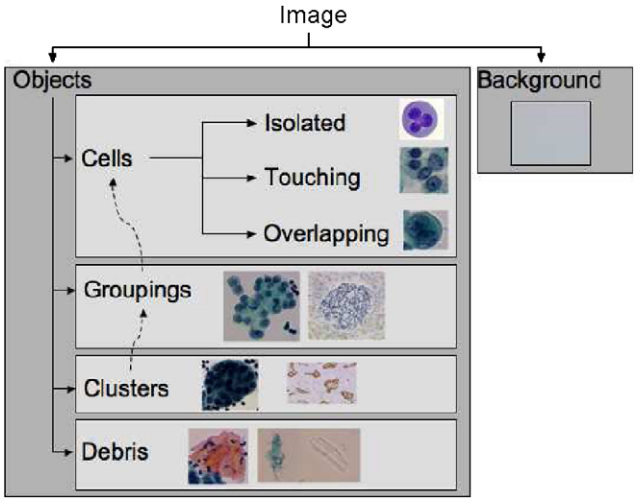

Fig. 2. Object oriented modeling of microscopic imaging problems in pathology.

to be built to have a modeling of the problem which is as close as possible to reality. Once this database has been constructed, a segmentation strategy can be conceived.

3. Graph-based segmentation framework

Image segmentation, in particular the case of microscopic cellular images, is an application dependent task and no general scheme or rule can be applied. In this work, we propose a set of graph-based tools to address the microscopic cellular image segmentation problems. One strong specificity of these tools is to use graphs of arbitrary topology as a discrete modeling of images at different levels. It leads to a set of flexible methods that can be combined or derived to produce a specific segmentation scheme for image segmentation, image regularization or clusters extraction. In the sequel, after recalling basic definitions on graphs, we present our discrete regularization framework. From this framework, a label propagation method for image automatic or interactive segmentation is derived. Moreover, based on the generalized Voronoi diagram, a discrete pre-segmentation approach is proposed to simplify images and to allow faster segmentation methods.

3.1. Graphs as a discrete modeling of images

3.1.1. Preliminaries on graphs

A graph is a structure used to describe a set of objects and the pairwise relations between those objects. The objects are called *vertices* and a link between two objects is called an *edge*. A weighted graph $G=(V,E,w)$ is composed of a finite set $V=\{u_1,\ldots,u_N\}$ of N vertices, a set of edges $E\subset V\times V$, and a *weight function* $w: V\times V\to\mathbb{R}^+$. An edge of E, which connects two *adjacent neighbor* vertices u and v, is noted (u,v). In the rest of this paper, the notation $v\sim u$ means that vertex v is an adjacent neighbor of vertex u. We assume that the graph G is simple, connected and undirected (see in [17] for details on these notions). This implies that the weight function w is symmetric i.e. $w(u,v)=w(v,u)$ if $(u,v)\in E$ and $w(u,v)=0$ otherwise.

Let $\mathscr{H}(V)$ be the Hilbert space of real valued functions on the vertices of a graph. Each function $f: V\to\mathbb{R}$ of $\mathscr{H}(V)$ assigns a real value $f(u)$ to each vertex $u\in V$. The function f forms a finite N-dimensional space and can be thought as a column vector $f=[f(u_1),\ldots,f(u_N)]^{\mathrm{T}}$. By analogy with functional analysis in continuous spaces, the integration of f over the graph, is noted $\int_V f=\sum_V f$. Similarly, let $\mathscr{H}(E)$ be the Hilbert space of real valued

functions defined on the edges of the graph. These two spaces are endowed with the usual inner products: $\langle f,g\rangle_{\mathscr{H}(V)}=\sum_{u\in V}f(u)g(u)$ with $f,g\in\mathscr{H}(V)$, and $\langle F,G\rangle_{\mathscr{H}(E)}=\sum_{u\in V}\sum_{v\sim u}F(u,v)G(u,v)$, where $F,H\in\mathscr{H}(E)$.

3.1.2. Representation of images by graphs

Any discrete domain can be modeled by a graph. In image processing, this structure is commonly used to represent digital image. In machine learning community, graphs are usually used to represent data sets and relationships between data points. Many typical structures can be quoted:

- *Grid graphs* [18] which are natural structures corresponding to the definition of digital images: vertices represent pixels and edges represent pixel adjacency relationship.
- *Region adjacency graphs* (*RAG*) [19] which provide very useful and common ways of describing the structure of a picture: vertices represent regions and edges represent region adjacency relationship.
- *Proximity graphs* [20], for instance the k-nearest neighbor graph, where each vertex is associated with a set of k close vertices depending on a similarity criterion. For a given vertex $u\in V$, if we consider all the vertices $V\backslash\{u\}$ as the vertex u neighborhood then the associated proximity graph is the *fully connected graph*.

Graph structures are extremely useful and occur naturally while processing digital images. A graph can be associated with any color image representation according to the definition of a distance or a similarity. In that case, processing images is reduced to processing graphs. The weight function measures the similarity between two vertices of the graph. When $w(u,v)\to 0$, the two vertices u and v are dissimilar. To weight a graph, the following standard Gaussian weight function can be used. For two vertices $u,v\in V$ and an edge $(u,v)\in E$, the edge weight is $w(u,v)=\exp(-|f(u)-f(v)|_2^2/\sigma^2)$ where σ is a scaling parameter living on the data and depends on the application, and $|\cdot|_2$ is the $\mathscr{L}_2$-norm. The topology of the graph depends on the problem under consideration: grid graphs for image simplification/segmentation, RAGs for image segmentation/analysis. In this paper, the proposed experiments are performed with different graph topologies to show the flexibility and the behavior of our graph-based tools.

3.2. Discrete tools on graphs

In this section, we introduce our discrete tools on graphs. These tools constitute a set of methods and provide a large variety of operations. These operations can be combined to formulate microscopic cellular image segmentation or classification schemes.

3.2.1. Discrete regularization

Discrete operators on graphs: Let us recall some basic definitions. We consider that a graph $G=(V,E,w)$ and a function $f\in\mathscr{H}(V)$ are given. The *weighted difference operator* $d:\mathscr{H}(V)\to\mathscr{H}(E)$ of a function f on an edge (u,v) linking two vertices $u,v\in V$ is defined as

$$d(f)(u,v)=\sqrt{w(u,v)}(f(v)-f(u)). \tag{1}$$

This operator leads us to define the *directional derivative* of f, over an edge (u,v), as $\partial_v f(u)=d(f)(u,v)$. Then, the *weighted gradient operator* $\nabla_w f(u)$ is defined as

$$\nabla_w f(u)=(\partial_v f(u):v\sim u)^{\mathrm{T}}=(\partial_{v_1}f(u),\dots,\partial_{v_k}f(u))^{\mathrm{T}}.$$

This operator corresponds to the *local variation* of the function f at the vertex u and measures the regularity of f in the adjacent neighborhood $v_1,\dots,v_k$ of the vertex u. Hence, the $\mathscr{L}_2$-norm of the weighted gradient operator is

$$|\nabla_w f(u)|_2=\left[\sum_{v\sim u}(\partial_v f(u)^2)\right]^{1/2}=\left[\sum_{v\sim u}w(u,v)(f(v)-f(u))^2\right]^{1/2}.$$

Then, the *weighted p-Laplace operator* $\Delta_w^p f(u)$ at vertex u is defined as

$$\Delta_w^p f(u)=\sum_{v\sim u}\gamma(u,v)(f(v)-f(u))$$
$$\text{where } \gamma(u,v)=w(u,v)(|\nabla_w f(u)|_2^{p-2}+|\nabla_w f(v)|_2^{p-2}). \tag{2}$$

Clearly, in the case where $p=1$ and 2, we have the definitions of the standard graph curvature $\Delta_w^1 f=\kappa f$ and graph Laplace $\Delta_w^2 f=\Delta f$ operators. More details on these definitions can be found in our previous works [13,21].

Discrete regularization framework: To regularize a function $f^0\in\mathscr{H}(V)$ using the p-Laplacian (Eq. (2)), we consider the following general variational problem on graphs:

$$\min_{f\in\mathscr{H}(V)}\left\{E_w(f,f^0,\lambda,p)=R_w(f,p)+\frac{\lambda}{2}|f-f^0|_2^2\right\}. \tag{3}$$

The first term, $R_w(f,p)$, is the regularizer and is defined as, with $0<p<+\infty$:

$$R_w(f,p)=\frac{1}{p}\sum_{u\in V}|\nabla f(u)|_2^p=\frac{1}{p}\sum_{u\in V}\left[\sum_{v\sim u}w(u,v)(f(v)-f(u))^2\right]^{p/2}. \tag{4}$$

The second term is the fitting term. $\lambda\geqslant 0$ is a fidelity parameter called the Lagrange multiplier which specifies the trade-off between the two competing terms. Both terms of E_w in Eq. (3) are strictly convex function of f [18]. By standard arguments in convex analysis, this optimization problem has a unique solution for $p=1$ and 2 which satisfies, for all $u\in V$:

$$\frac{\partial}{\partial f(u)}E_w(f,f^0,\lambda,p)=\Delta_w^p f(u)+\lambda(f(u)-f^0(u))=0. \tag{5}$$

Eq. (5) can be viewed as the discrete analogue of the Euler–Lagrange equations. Using the p-Laplacian formulation (Eq. (2)) in Eq. (5), the optimization problem solution is also the solution of the following system of equations. For all $u\in V$,

$$\left(\lambda+\sum_{v\sim u}\gamma(u,v)\right)f(u)-\sum_{v\sim u}\gamma(u,v)f(v)=\lambda f^0(u).$$

To approximate the solution of the minimization (3), we can linearize this system of equations and use the Gauss–Jacobi method to obtain the following iterative algorithm:

$$\begin{cases}f^{(0)}(u)=f^0(u),\\ f^{(t+1)}(u)=\dfrac{\lambda f^0(u)+\sum_{v\sim u}\gamma^{(t)}(u,v)f^{(t)}(v)}{\lambda+\sum_{v\sim u}\gamma^{(t)}(u,v)},\end{cases} \tag{6}$$

where $\gamma^{(t)}(u,v)$ is the γ function (in Eq. (2)) at the iteration step t. At each iteration of the algorithm, the value of f at step $(t+1)$, for a vertex u, only depends on two quantities: the original value f^0 and the sum of weighted local variation of the existing values in the neighborhood of u. By using different formulations of w and different values of p, a family of linear and nonlinear filters is obtained. Indeed, when $p=2$ and $w(u,v)=1$ one obtains the linear diffusion on graphs. When $p=1$ and $w(u,v)=1$ one recovers the TV digital filter [18]. The reader can note that this isotropic regularization corresponds to the weighted discrete transcription of the regularization functional

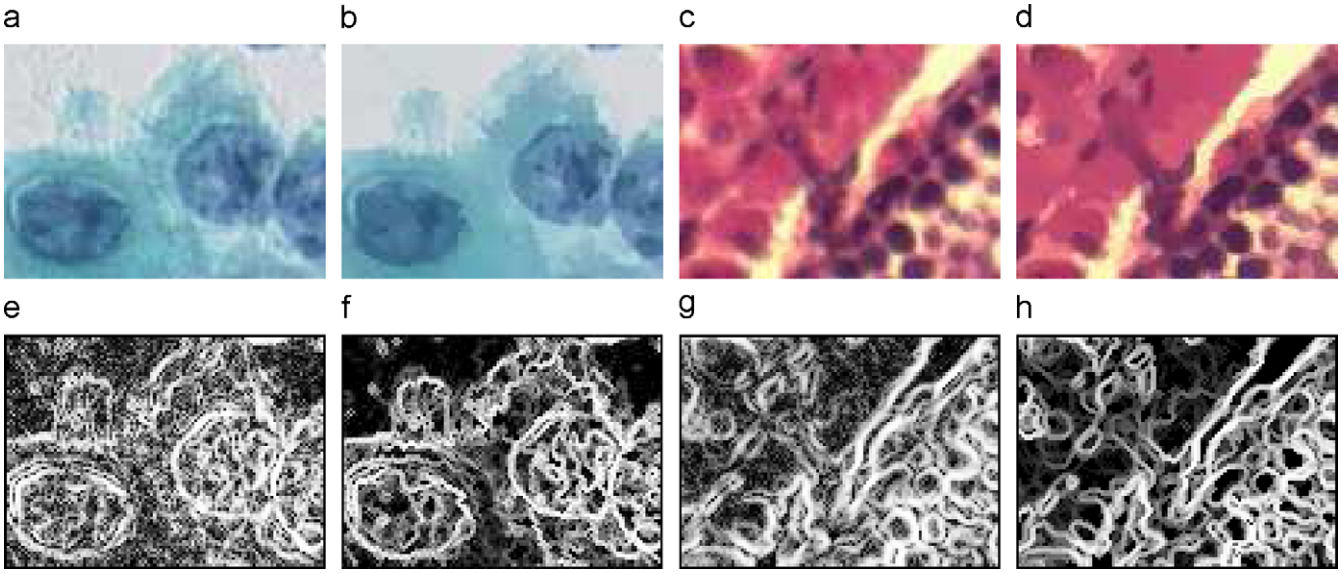

Fig. 3. Color cytological and histological image filtering. (a) and (c) Original images. (b) and (d) Their filtered image (nonlinear filtering with $p=1$). (e) and (g) Gradient images from original ones. (f) and (h) Gradient images from filtered ones.

in the continuous case. The interested reader can refer to [13,21] for more details on the formulation and the connections with other formalisms. Moreover, in [22], we have extended this discrete isotropic regularization to a discrete anisotropic regularization framework for image and data processing.

Image filtering: Through the values of the p parameter, the discrete regularization (5) describes a family of linear and nonlinear filters. This image filtering/denoising can be viewed as an image simplification that can ease a seed extraction step. Fig. 3 shows a nonlinear ($p=1$) filtering on noisy cytological and histological images represented by an 8-adjacency grid-graph. This image filtering enhances image components, as shown in Figs. 3(b) and (d) as compared to original ones in Figs. 3(a) and (c). We have the same effect on the gradient images, Figs. 3(f) and (h) as compared to Figs. 3(e) and (g).

3.2.2. Discrete energy partition

Recent interactive image segmentation approaches have become increasingly popular in the image processing community. They reformulate image segmentation tasks into semi-supervised classification approaches by label propagation methods. They are usually based on image pixel diffusion strategies. Thus, if we consider large images, these methods are difficult to apply due to the great mass of data to analyze. To avoid this computational restriction and to provide a fast image segmentation, we propose to use a simplified version of the image in place to work with the pixel-based representation. One possible representation is to construct a fine partition of the image and to consider neighborhood relations between the obtained regions. Many methods have been proposed to construct image fine partition in literature. For instance, see [23] for methods based on graph cuts algorithm or [24] for the ones based on morphological operators. In this work, we propose to use a graph-based method inspired by an approach based on the generalized Voronoi diagram [25]. This method can be viewed as an image simplification or a graph reduction. This data reduction can lead to further faster algorithm convergence and to fast image processing schemes.

We consider a graph $G=(V,E,w)$ to be given. Let $f\in\mathscr{H}(V)$ be a function defined on V and $P_G(u,v)$ be the set of paths connecting two vertices $u,v\in V$. A *path* $\rho(u,v)$ is a sequence of vertices $(u_1,\dots,u_m)$ such as $u=u_1$ and $v=u_m$ with $(u_i,u_{i+1})\in E$ and $i=1,\dots,m-1$. We define the pseudo-metric $\delta: V\times V\to\mathbb{R}$ to be

$$\delta(u,v)=\min_{\rho\in P_G(u,v)}\sum_{i=1}^{m-1} d(f)(u_i,u_{i+1})$$

$$=\min_{\rho\in P_G(u,v)}\sum_{i=1}^{m-1}\sqrt{w(u_i,u_{i+1})}(f(u_{i+1})-f(u_i)), \tag{7}$$

where df is the difference equation defined in (1).

Given a set of K seeds $S=\{s_i\}\subseteq V$, where $i=1,\dots,K$, the *energy* $\delta_S: V\to\mathbb{R}^n$ induced by the metric δ for all the seeds of S can be expressed as

$$\delta_S(u)=\min_{s_i\in S}\delta(s_i,u)\quad\forall u\in V.$$

The *influence zone* z (also called Voronoi region) of a given seed $s_i\in S$ is the set of vertices which are closer to s_i than to any other seeds with respect to the metric δ. It can be defined, $\forall j=1,\dots,K$ and $j\neq i$, as

$$z(s_i)=\{u\in V:\delta(s_i,u)\leqslant\delta(s_j,u)\}.$$

Then, the *energy partition* of G, for a given set of seeds S and a metric δ, is the set of influence zones $Z(S,\delta)=\{z(s_i),\forall s_i\in S\}$.

For a given graph G, to find the energy partition corresponds to seek a minimal cost path over G. Among the graph algorithms dedicated to this shortest path problem, the Dijkstra algorithm can be applied. Using Fibonacci heap structure, the amortized time complexity to obtain such partition is $\mathscr{O}(E+V\log V)$.

Let $G=(V,E,w)$ be the grid graph associated to a given image f where each vertex u of V corresponds to a pixel of f. With these definitions, from a set of seeds which constitutes a set of source vertices and a pseudo-metric, we can obtain the exact energy partition which considers the total variation of df along a path with $w(u,v)=1$. Then, the energy partition of G represents an approximation of the image by assigning a model for each influence zone of Z (where each zone corresponds to an obtained image region). A simple model can be a mean or a median value of each influence zone. Finally, the RAG can be constructed by considering the dual representation (the Delaunay graph [26]) of the obtained energy partition where each vertex of the RAG corresponds to an image region and the edges, the region adjacency relationships.

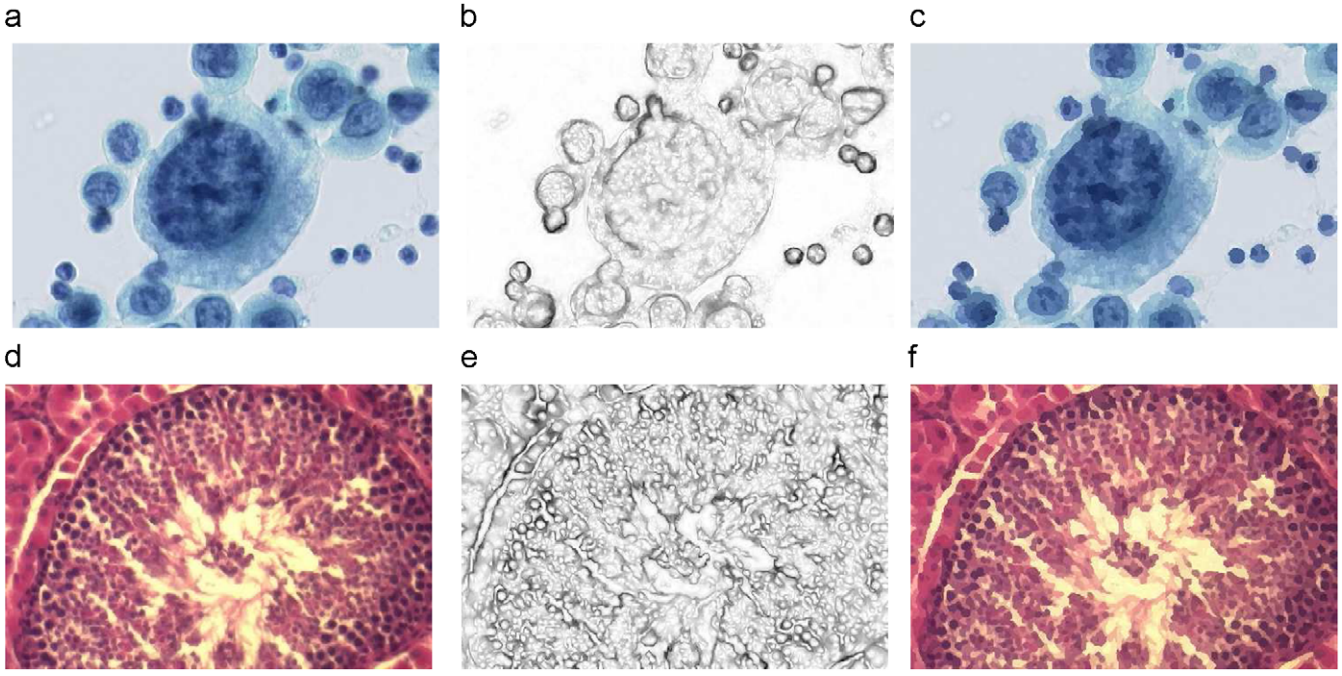

Fig. 4. Discrete energy partitions on cytological and histological images. (a) and (d) Original images of size 480×320. (b) and (e) Energy images. (c) and (f) Reconstructed images from influence zones with region mean color as model; the specified value corresponds to the number of final influence zones and to the percentage of reduction as compared to the original image. The partition images computing time is about 1 s including seeds extraction step. (a) and (d) 134400 pixels, (b) and (e) energy images, (c) 7132 zones (95% of reduction) and (d) 2691 zone (98% of reduction).

To segment a microscopical image, a classical method is to use a region growing approach starting from seeds (such as a watershed for instance). Identically, our discrete energy partition can be used as a segmentation method. To obtain a set of seeds, several methods can be used. In image processing, a common approach is to use the image local extrema (minima or maxima) in a fixed search window. Other schemes, based on machine learning approaches have also been proposed to extract the initial seeds [12,27,28].

Image simplification for fast segmentation: Fig. 4 illustrates the application of graph-based energy partition to simplify cytological and histological images represented by an 8-adjacency grid graph. In the proposed example, we use both local image minima and maxima as the initial set of seeds. This example shows the efficiency of the discrete energy partition approach to simplify the original images while also respecting the image components. One can note the significant data reduction. The original image (Figs. 4(a) and (d)) have 134 400 components (image pixels). After applying our method, the obtained number of influence zones corresponds approximatively to less than 5% of the original ones (Figs. 4(c) and (f)). The reconstructed images (Figs. 4(c) and (f)) are obtained from the corresponding partition where the pixel values of each region of the partition are replaced by the mean pixel color value in the original image of its regions. Finally, as shown by the energy images (Figs. 4(b) and (e)), our method has the ability to respect object structure by preserving edge information. The computing time to obtain the energy partitions in Fig. 4 is about 1 s including the time to compute the image extrema. The computation is performed on a standard Linux computer equipped with quadri 2.4 GHz Intel Xeon processors and 16 GB of RAM.

In Section 4, we will show the application of discrete energy partition to segment cellular components.

3.2.3. Discrete semi-supervised clustering

Numerous automatic segmentation schemes have been proposed in literature and they have shown their efficiency. But, sometimes, automatic segmentation results are not accurate when images are more complex. Meanwhile, recent interactive image segmentation approaches have been proposed. They reformulate image segmentation tasks into semi-supervised classification approaches by label propagation strategies [8,9,29]. Other applications of these label diffusion methods can be found in [5,30]. Our previously presented discrete regularization framework (Section 3.2.1) can be naturally adapted to address this learning problem for semi-supervised segmentation. The adaption of our regularization framework leads to a clustering method. In the sequel, we use the terms of semi-supervised classification and segmentation as the same procedure but additional computation steps must be performed to obtain final results. Indeed, a final classification is obtained by a class membership probabilities estimation. To obtain the final image segmentation, a labeling of connected image components must be performed.

Problem formulation: Let $V = \{u_1, \ldots, u_N\}$ be a finite set of data, where each data u_i is a vector of $\mathbb{R}^m$. Let $G = (V, E, w)$ be a weighted graph such as data are connected by an edge of E. The semi-supervised clustering of the set V consists in grouping the set V into k classes where the number of k classes is given. For this, the set V is composed of labeled and unlabeled data. The objective is then to estimate the unlabeled data from labeled ones.

Let $C = \{c_i\}_{i=1,\ldots,k}$ be the set of initial *labeled* vertices and let $V \backslash C$ be the initial *unlabeled* vertices (the whole set of vertices except the labeled ones). This situation can be modeled by considering k initial label functions (one per class) $f_i^0 : V \to \mathbb{R}$, with $i = 1, \ldots, k$. For a given vertex $u \in V$, if u is initially labeled then

$$f_i^0(u) = \begin{cases} +1 & \text{if } u \in c_i, \\ -1 & \text{otherwise.} \end{cases} \tag{8}$$

If u is initially unlabeled (i.e. $u \in V \backslash C$) then $f_i^0(u) = 0$, the vertex clustering is accomplished by k regularization processes by estimating resultant function $f_i : V \to \mathbb{R}$ for each i th class. Using our proposed

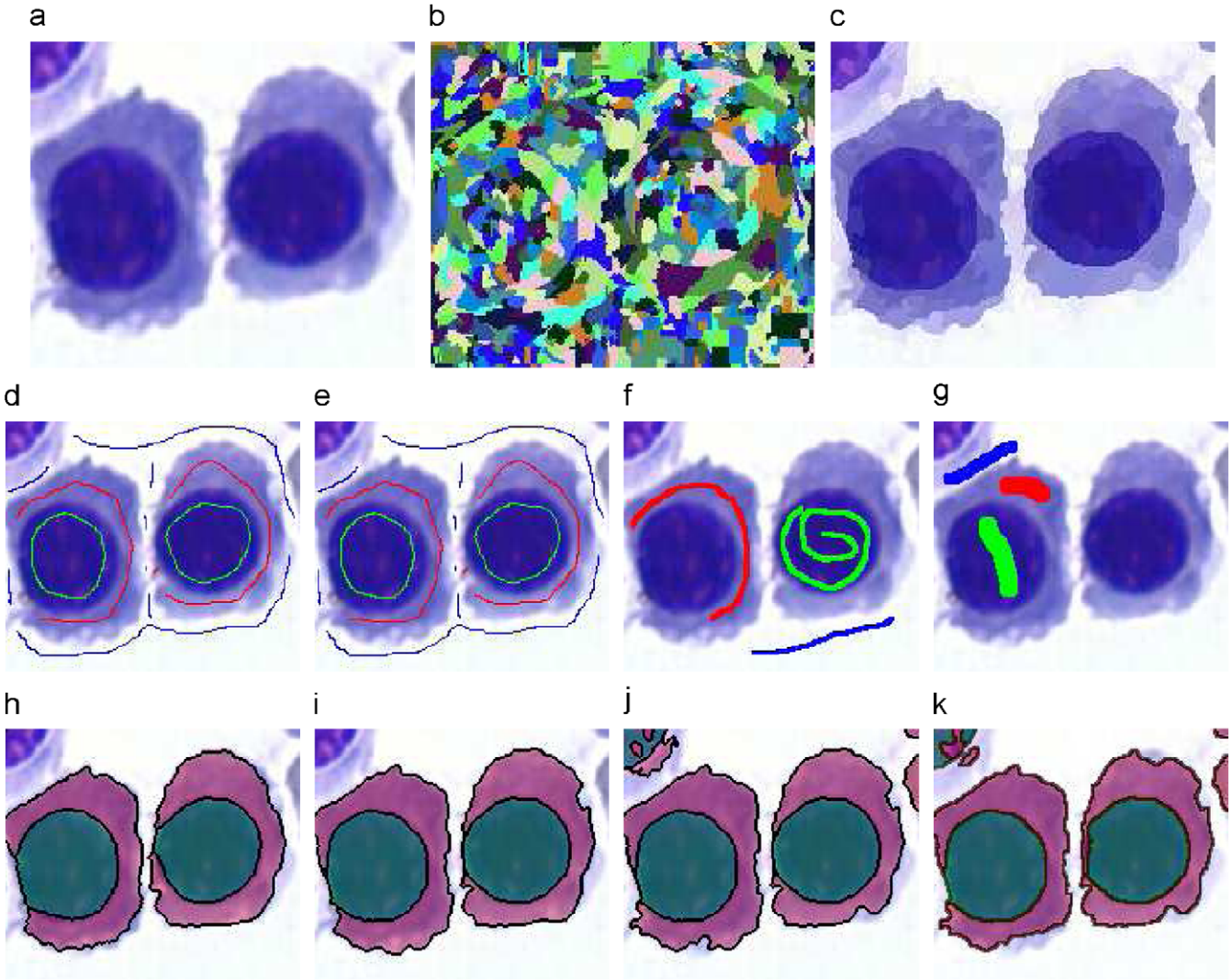

Fig. 5. Semi-supervised image segmentation with $p=2$, $\lambda=1$, t iterations for different graph topologies and user input strokes. First row: (a) original image of size 152×181, (b) discrete energy partition image, (c) reconstructed images from partition (mean color model). Second row: user input labels. Third row: original image with the obtained segmented regions superimposed: cytoplasm (red), nuclei (green) and region boundaries (black); the segmentation is performed with the specified iteration steps t and the corresponding computation time. The images (h), (i), (j) and (k) are, respectively, obtained from label images (d), (e), (f) and (g). Graph topologies used to obtain results (h) *grid graph* associated with original image, (i) RAG associated with simplified image, (j) and (k) *fully connected graph* associated with simplified image. (a) 27512 pixels, (b) 639 zones (98% of reduction), (c) reconstructed image, (d)–(g) original+labels (h) $t=50$ (11s), (i) $t=5$ (<1s), (j) $t=2$ (<1s) and (k) $t=2$ (<1s).

discrete regularization framework, this is formalized as follows:

$$\min_{f_i\in\mathscr{H}(V)}\left\{R_w(f_i,p)+\frac{\lambda}{2}|f_i(u)-f_i^0(u)|_2^2\right\},$$

where the first term $R_w(f_i,p)$ is the one defined in Eq. (4). We use the discrete diffusion process (Eq. (6)) to compute each minimization. At the end of the label propagation processes, the class membership probabilities can be estimated and the final classification can be obtained for a given vertex $u\in V$ by the following formulation. For all $i\in 1,\ldots,k$, we have

$$\operatorname*{argmax}_i\left\{f_i(u)\Big/\sum_i f_i(u)\right\}.$$

Any other formulations can be used (for instance see [31]). To obtain a final image segmentation, a connected image components labeling can be performed on classified elements.

Fast semi-supervised image segmentation: Image semi-supervised segmentation are usually based on image pixel diffusion strategies. The drawback of this method is that when the considered image is large, the label propagation method is inefficient due to the great mass of data to analyze. To avoid this computational problem, we propose to simplify the image (i.e. to reduce the graph structure) by applying our discrete energy partition method (Section 3.2.2). Fig. 5 illustrates the proposed label propagation method. It shows segmentation of cytological image into three classes (nuclei, cytoplasm and background) with different graph structures. This example compares the computation time and the segmentation results between a pixel-based grid graph and two region-based proximity graphs (the RAG and the fully connected graph). It also shows the robustness of our approach regarding the initial user input. For the cases of proximity graphs, the computing times include the graph construction itself. Moreover, due to the size of proposed example, the computing time for the energy partition construction can be ignored. All the results are obtained with a standard Linux computer

equipped with quadri 2.4 GHz Intel Xeon processors and 16 GB of RAM.

Fig. 5(h) is the semi-supervised segmentation result obtained from the initial labels (Fig. 5(d)) and an 8-adjacency grid graph as original image (Fig. 5(a)) representation. One can observe the number and the precise location of the initial labels, in particular the necessary labels between the two cells. Fig. 5(b) is a simplified version (98% of reduction) of the original image (Fig. 5(a)) obtained by our discrete energy partition approach. Fig. 5(c) is a reconstructed image from the partition. The same method has been used to obtain such image in Section 3.2.2 with Fig. 4 by using mean color model. With this simplified version, we construct two proximity graphs associated with this image: the RAG and the fully connected graph. (The RAG is constructed as described at the end of Section 3.2.2.) Fig. 5(i) shows the segmentation result obtained from the RAG with the same initial labels (Fig. 5(e)) as in the grid graph case. We can observe that the two results (Figs. 5(i) and (h)) are similar but in the RAG-based segmentation case, the computation time is significantly reduced. Figs. 5(j) and (k) show the segmentation result obtained from the fully connected graph. If we consider an image as a set of pixels, it is clear that this approach cannot be applied due to the computation time. But, if we consider the simplified version of images, this method becomes an efficient one and we can quote interesting properties to use this structure.

- The graph contains all the image information in the weighted edges and therefore the regularization process only needs a minimal number of iterations to reach the algorithm convergence.
- A minimal number of labels is needed to obtain correct results as compared to the case of the grid graph or of the RAG. In Figs. 5(f) and (g), only one nucleus and one cytoplasm are marked, and there is no separating label between the two cells.
- An interesting property is that the objects can be quickly labeled in the same class, even if they are not spatially adjacent or close. In Figs. 5(j) and (k), the two main nuclei and cytoplasm are segmented even if there are no initial labels. Moreover, the label diffusion process has also found the two pieces of cytoplasm on the left and the piece of cells on the top-left corner of the image.

Moreover, these two examples show the robustness of our approach. Indeed, Figs. 5(j) and (k) show similar results with two different user input labels (Figs. 5(f) and (g)).

4. Microscopic cellular imaging applications

In this section, we illustrate the abilities of the presented graph-based methods to address color cytological and histological segmentation problems. In the sequel, various schemes are proposed to extract cellular components in such images. These schemes are obtained by combining the previously described discrete tools together or in addition with classical approaches (such as morphological operation or fuzzy clustering). They are also performed either automatically or interactively (i.e. user guideline segmentation).

All the following results are obtained with a standard Linux computer equipped with quadri 2.4 GHz Intel Xeon processors and 16 GB of RAM.

4.1. Color cytology applications

In this section, we consider color cytological images. All the images, in the next applications, are from serous cytology. The images are from a database of digitized cell images, collected from pleural and peritoneal effusions with different pathologies [15]. In this class of images, the cytoplasm and the nuclei are, respectively, colored in green and blue. Both cytoplasm and nuclei have to be segmented. We have the following object organization in an image: Image = ((cells = (nucleus $\subset$ cytoplasm)) $\cup$ background). In the sequel, all the schemes are based on our discrete tools. We propose both automatic and interactive segmentation approaches.

4.1.1. Automatic color cytological image segmentation by discrete energy partition (methodology: image filtering by regularization + seeds extraction and classification + segmentation by discrete energy partition)

Fig. 6 shows an automatic color cytological image segmentation by discrete energy partition. First, the image (represented by an 8-adjacency grid graph) is simplified by regularization with $p=2$. This corresponds to a linear diffusion on a grid graph. The aim of this image simplification is to ease the seeds extraction. On this simplified image, the minima of the gradient image are extracted (Fig. 6(b)). The gradient is computed with the Di Zenzo formulation [32]. Minima are then classified into three classes by a fuzzy c-means clustering. These three classes, respectively, correspond to the nuclei (Fig. 6(c)), the background (Fig. 6(d)), and the cytoplasm (Fig. 6(e)). Using these labeled minima as vertex sources, a localization is performed by discrete energy partition which leads to the complete segmentation of the image (Fig. 6(f)).

4.1.2. Automatic color cytological image segmentation by label propagation on a MST (methodology: image simplification by discrete energy partition + seeds extraction and labeling + segmentation by label diffusion processes)

Fig. 7 shows an automatic color cytological image segmentation by label propagation on minimum spanning tree (MST). First, from the original image (Fig. 7(a)) we compute a simplified image by discrete energy partition (Fig. 7(b)) and construct the MST. An MST is a well-known structure in graph theory and gives a connected graph with no cycles and it is also sparse (for N data points, it has only $N-1$ edges). We use the discrete energy partition set of seeds as initial labels of our label diffusion approach. Seeds are classified into three classes by a classical K-means algorithm (based on region color features). Result of this seeds classification is shown in Fig. 7(c). One can note that this classification is not perfect and some nuclei seeds have been classified as cytoplasm ones. Figs. 7(d)–(g) show, respectively, the obtained region maps and the region boundaries for different values of λ parameter of the proposed label diffusion algorithm. When λ is positive value, the label diffusion is highly oriented by initial labels. When the λ parameter is null, the algorithm has the ability to modify initial labels classification. Fig. 7(f) shows label modification effects on classification results. Bad initial cytoplasm labels are changed to background label and conversely. This interesting property resides in the graph topology and in our label diffusion formulation. One can note that there are differences between a simple region merging algorithm and our automatic classification scheme. Indeed, the proposed method does not need any merging criterion to assign a label to a region: it is implicitly done by the weighted graph representation. Moreover, our method does not need any stopping criterion: the final result is obtained when the algorithm reaches convergence.

4.1.3. Interactive color cytological image segmentation by label propagation on a fully connected graph (methodology: image simplification by discrete energy partition + segmentation by user labels diffusion processes)

Fig. 8 shows an interactive color cytological image segmentation by label propagation on a fully connected graph. First, from the original image (Fig. 8(a)) we compute a simplified image by discrete energy partition. With this simplified image (Fig. 8(b)) we construct the fully connected graph. As previously mentioned in the example Fig. 5, the fully connected graph provides interesting properties for

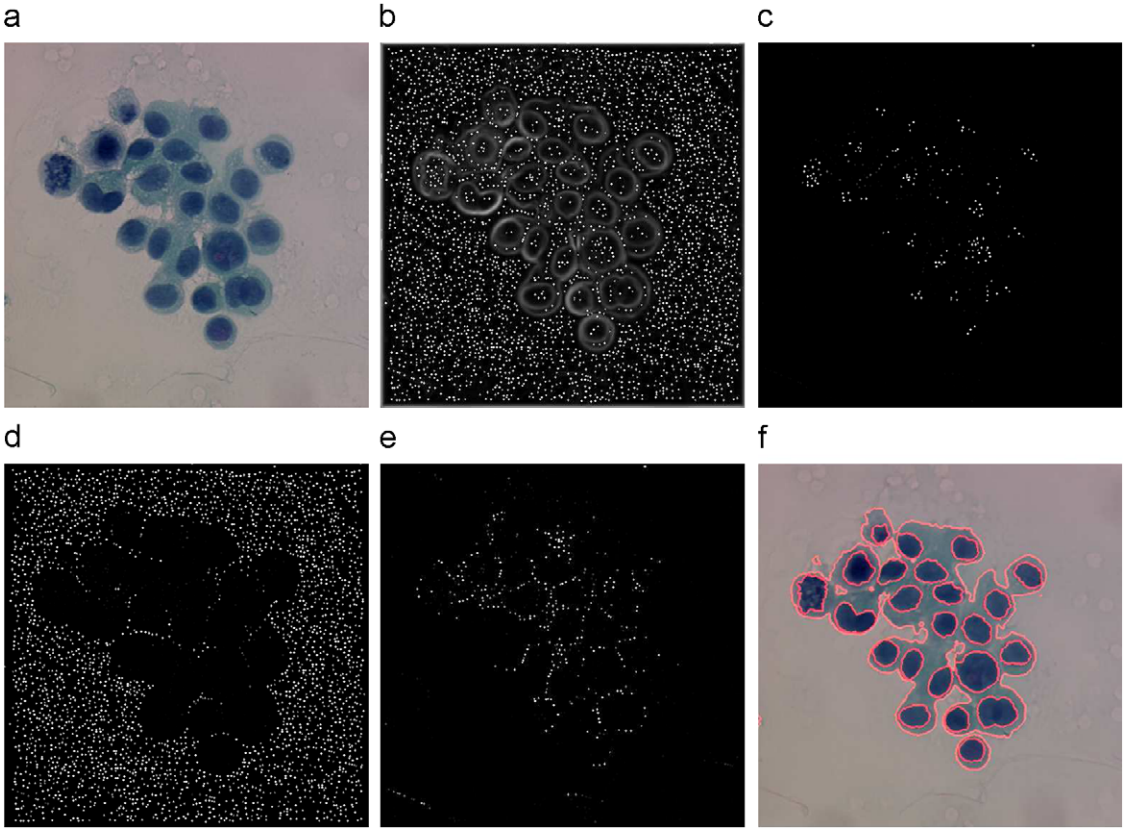

Fig. 6. Automatic color cytological image segmentation based on 8-adjacency grid graph and discrete energy partition. (a) Original image. (b) Gradient minima. (c)–(e) Nuclei, background and cytoplasm minima classified with a fuzzy clustering. (f) Segmentation result with found region boundaries superimposed in red.

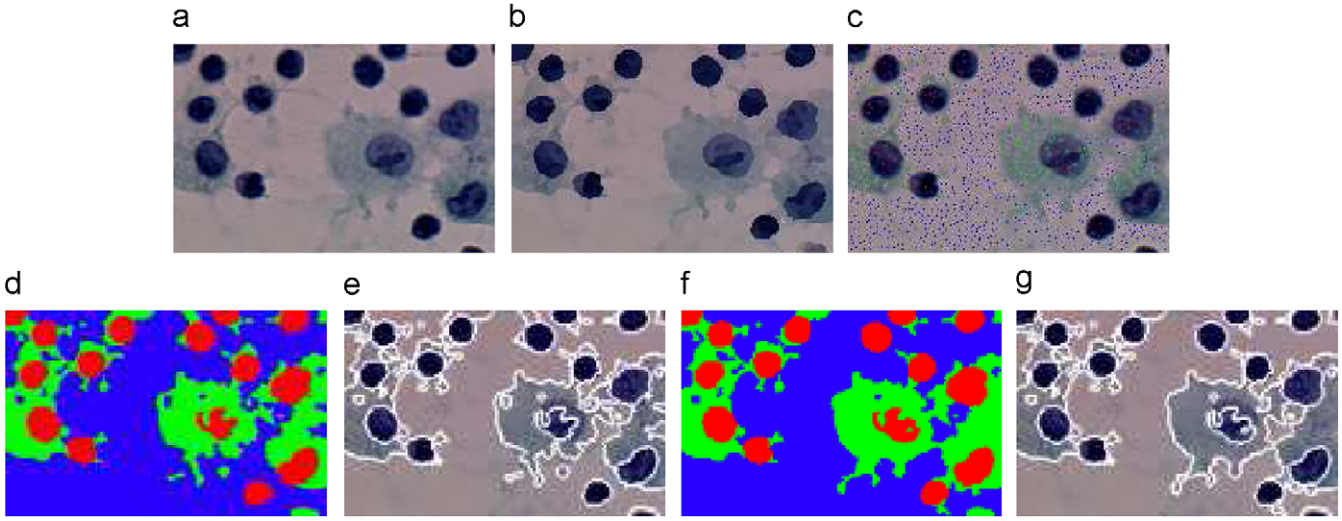

Fig. 7. Automatic cytological segmentation by label diffusion processes based on an MST and λ parameter. (a) Original image. (b) Energy partition reconstructed image with region mean color as model. (c) Original image with labeled seeds superimposed; the seeds were classified by K-means ($k=3$) algorithm. (d) and (f) Obtained regions map with the specified λ parameter. (e) and (g) Original image with the obtained region boundaries superimposed in white. (a) 26000 pixels, (b) 936 zones (96% of reduction) (c) original+labels, (d) segmentation $\lambda=1$, (f) segmentation $\lambda=0$.

label propagation: a minimal number of user initial labels is needed, as shown in Fig. 8(c) where only one nucleus is marked; and a fast fully nonlocal propagation where the graph topology allows a non-spatially adjacent labeling, as shown in Fig. 8(d) where all the nuclei are finally segmented.

4.2. Color histology applications

Classical color histological images contain background and cellular objects. The segmentation problem of this class of images consists in extracting clusters of abnormal nuclei. These clusters contain

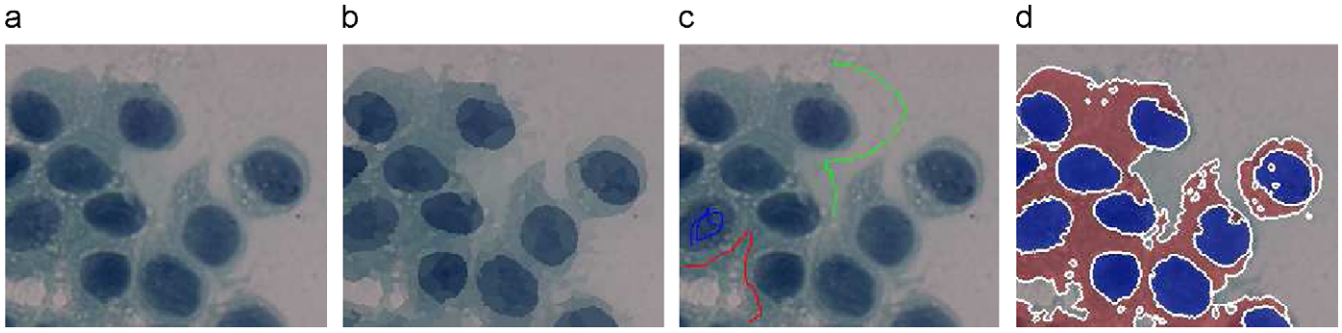

Fig. 8. Semi-supervised cytological image segmentation based on fully connected graph. (a) Original image of size 198×206. (b) Reconstructed energy partition image with mean color from original image. (c) Original image with initial labels superimposed. (d) Original image with the obtained regions superimposed (red for the cytoplasm, blue for the nuclei, and white for region boundaries).

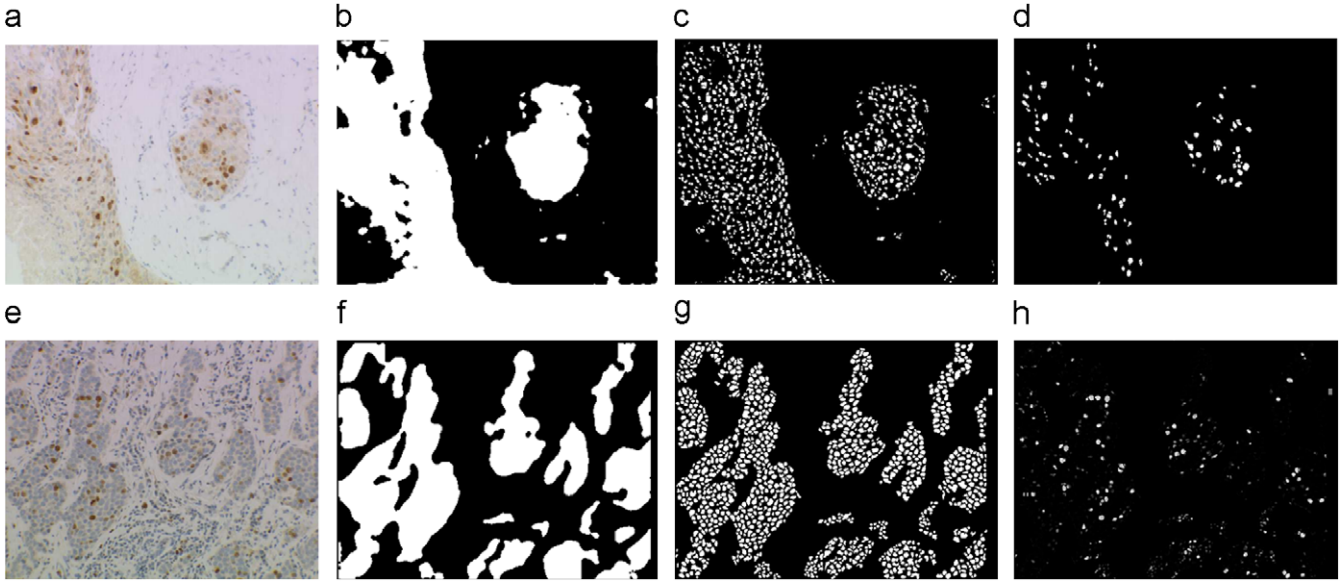

Fig. 9. Automatic segmentation of two color histological images. (a) and (e) Original color histological images. (b) and (f) Clusters extraction. (c) and (g) Clusters' nuclei. (d) and (h) Clusters' marked nuclei.

nuclei which have been colored or not by a chemical marker. In all the next applications, the images are from a database of breast tissues digitized images, collected from different pathologies and marked with a Ki67 marker. We have the following organization in a color histological image: Image = (Background ∪ Nuclei ∪ (Marked Nuclei ∪ Unmarked Nuclei) ⊂ Clusters).

4.2.1. Automatic color histological image segmentation with image simplification by regularization (methodology: image filtering by regularization + clusters *extraction and classification)*

Fig. 9 shows an automatic color histological image segmentation with image simplification by regularization on two images. To extract clusters of abnormal cells, the image (represented by an 8-adjacency grid graph) is simplified by regularization with $p = 1$. This corresponds to a nonlinear diffusion on a grid graph. As for color cytology, gradient minima are classified into two classes by fuzzy clustering and a localization is performed (Figs. 9(b) and (f)). Once clusters of abnormal cells have been extracted, the corresponding regions in the image are the only ones used for further analysis. In these regions, a simplification by regularization is performed (with $p = 2$ and $w(u, v) = 1$) and a residual analysis of the simplified image is accomplished to extract nuclei. Residual analysis consists in seeking zero-crossing zones of the difference between the original image and the simplified one (this provides an approximation of the Laplacian). Extracted zones constitute seeds and a localization is then performed by discrete energy partition (Figs. 9(c) and (g)). Once these nuclei inside the clusters have been extracted, a fuzzy clustering is performed to split them into two classes: marked and unmarked (Figs. 9(d) and (h)).

4.2.2. Interactive color histological image segmentation by label diffusion processes on a RAG (methodology: image filtering by regularization + image *simplification by discrete energy* partition + segmentation *by user labels diffusion processes)*

Fig. 10 shows an interactive color histological image segmentation by label propagation on a RAG. This application illustrates the advantage of user guideline segmentation. When images become much more complex, automatic segmentation results are not always accurate and a user's correction becomes necessary to obtain the desired result. In this application, first we filter the original image (Fig. 10(a)) by regularization on an 8-adjacency grid graph to ease the seeds extraction for the discrete energy partition method. The image filtering result is shown in Fig. 10(b). With this simplified image, we perform our discrete energy partition to obtain a significant reduced

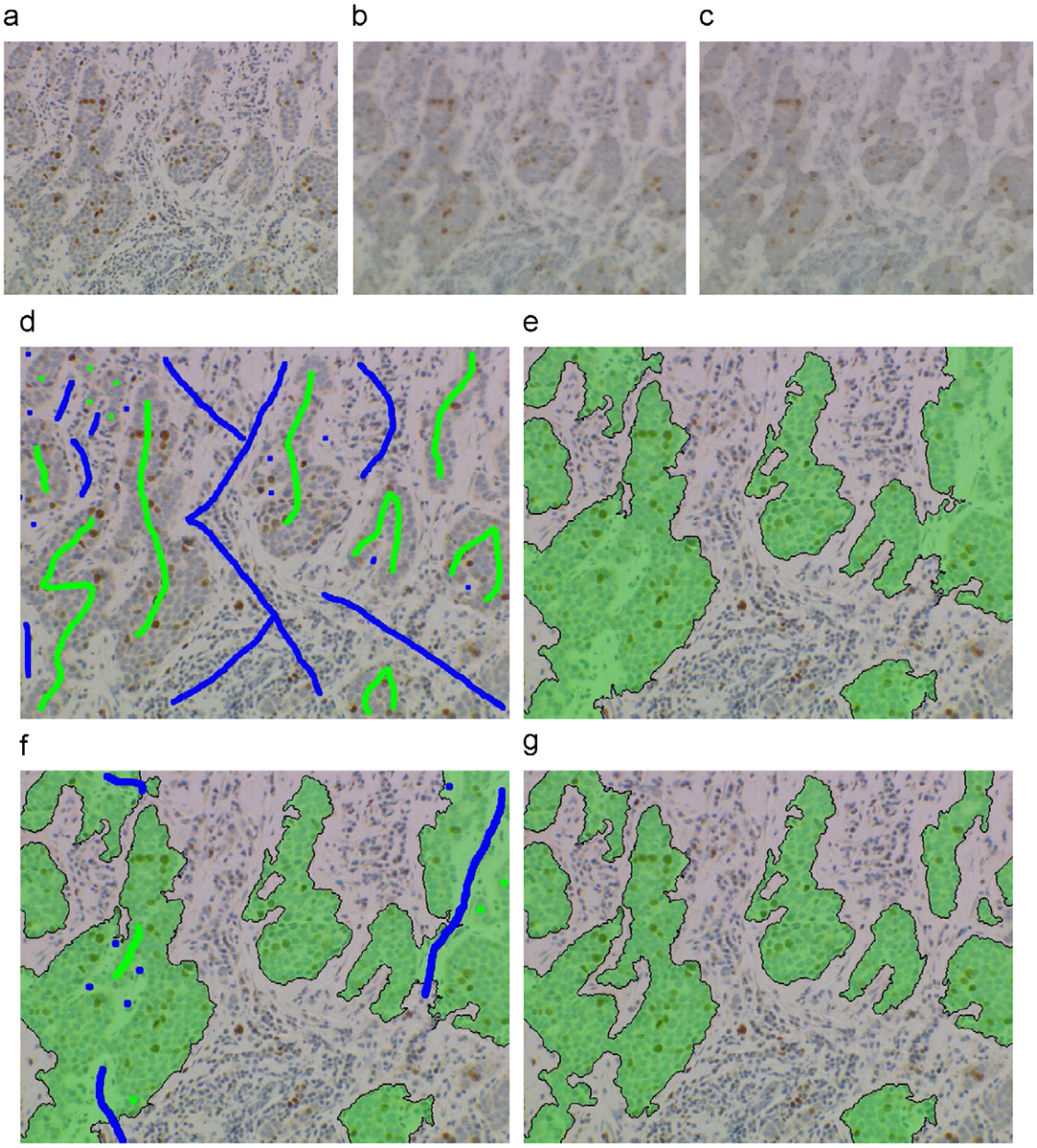

Fig. 10. Semi-supervised (interactive) color histological image segmentation with pre-processing step. (a) Original image of size 555 × 770. (b) Filtered image from original. (c) Image of reconstructed energy partition with mean color from filtered image. (d) Original image with initial labels superimposed. (e) Original image with the obtained regions superimposed (green for the clusters and black for cluster boundaries). (f) User additional labels. (g) Original image with the obtained final regions superimposed (green for the clusters and black for cluster boundaries).

version of the original image (Fig. 10(c)). A RAG is constructed with this reduced image where each vertex corresponds to an influence zone with region mean color as model. Our label diffusion process is performed with this graph and the initial user labels (Fig. 10(d)). The obtained segmentation is shown in Fig. 10(e). One can note that the segmentation is not correct, the initial labels are not enough precise to separate the clusters. Fig. 10(f) shows the user additional labels to correct the segmentation and Fig. 10(g) shows the corrected final classification result.

4.3. Comparisons with others methods

In this section, we compare one of our proposed schemes with other segmentation methods proposed by literature. The chosen scheme is the fast interactive image segmentation. This scheme is compared with the following methods.

- Two approaches based on *K*-means and Bayesian classifications proposed by Lézoray et al. [12].
- Two approaches based on pixel and subwindow random tree classifications proposed by Dumont et al. [33].

For more details on these methods, the interested reader can refer to the corresponding articles.

The test set is composed of 10 images from serous cytology [12]. Cells are colored by the international standard coloration of Papanicolaou. Two classes are considered: nuclei pixels and other pixels. Fig. 11 shows three examples of the test set with the expert

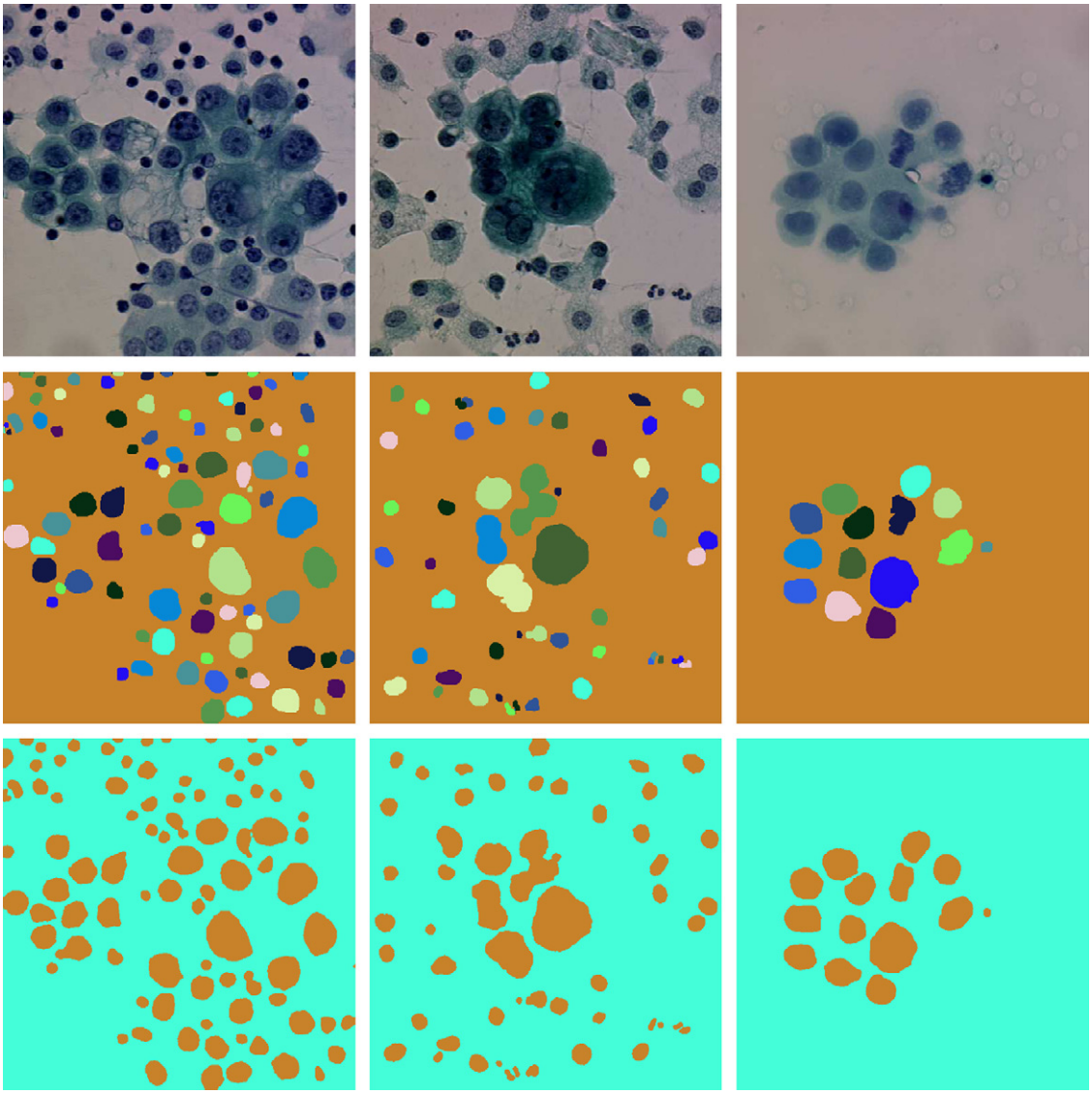

Fig. 11. Three examples of the test data set images. First row: original images. Second row: expert ground truth. Third row: segmentation results with our interactive scheme.

Table 1
Global classification rates and comparison with Dumont et al. approaches [33]

	Pixel based (%)	Subwindow based (%)	Our scheme (%)
R_0	95.93	96.39	**98.41**

Table 2
Classification rates and comparison with Lézoray et al. approaches [12]

	K-means (%)	Bayesian (%)	Our scheme (%)
R_1	88.7	**97.53**	92.31
R_2	98.65	95.40	**99.15**
R_3	93.67	**96.47**	95.73

manually segmented ground truth (second row) and the segmentation results of our interactive scheme (third row). To compare classification schemes, we use four common classification rates based on well-classified pixels. Tables 1 and 2 show classification rate comparisons between our interactive scheme and two pixel-based schemes relying on random decision trees proposed by Dumont et al. [33] and two classification models (K-means and Bayesian) proposed by Lézoray et al. [12]. Best accuracies are face bolded. Classifications rates R_0, R_1, R_2 and R_3 are expressed as follows:

$$R_0 = \frac{\text{number of pixels well classified}}{\text{number of pixels of the image}},$$

$$R_1 = \frac{\text{number of nuclei pixels well classified}}{\text{number of nuclei pixels of the image}},$$

$$R_2 = \frac{\text{number of background pixels well classified}}{\text{number of background pixels of the image}},$$

$$R_3 = (R_1 + R_2)/2.$$

Results in Table 1 show that our approach has the best classification accuracies as compared to methods proposed by Dumont et al. [33]. Table 2 presents classification accuracies per class for the three methods under consideration. Our approach clearly outperforms K-means but is less accurate than Bayesian classification for nuclei pixels. These results can be interpreted as follows. One of the advantages of our interactive scheme is that even if the segmentation is not correct, the user can modify his input labels until

obtaining the desired results. The consequence is that we can obtain better results as compared to an automatic scheme since rates are very close. Regarding the background pixel and the nuclei pixel classification rates, one can note that Bayesian tends to over segment the nuclei and our approach tends to subsegment them.

5. Conclusion

In this paper, we have considered a framework of graph-based tools for microscopic images segmentation. These discrete tools can be combined together or in addition with classical approaches to formulate specific schemes for the segmentation of a particular class of cellular images. Through specific strategies, we have shown the efficiency of our methodology to address color cytological and histological images segmentation problems in pathology. Moreover, the proposed framework is sufficiently general to be applied to any type of microscopic images. We have proposed, on the one hand, efficient automatic and interactive user guideline segmentation techniques; and, on the other hand, fast segmentation algorithms via image simplification and/or graph topologies. The proposed discrete tools use graphs as an underlying representation and a unified formulation. This strong specificity implies that our graph-based tools can be easily adapted to be used with any type of discrete data which can be represented by a graph. Hence, it leads to an interesting ongoing work: to use the same tools to analyze, categorize, recognize the segmented cellular objects in huge image databases.

Acknowledgments

This work was partially supported under a research grant of the ANR Foundation (ANR-06-MDCA-008-01/FOGRIMMI) and a doctoral grant of the Conseil Régional de Basse-Normandie and of the Cœur et Cancer association in collaboration with the Department of Anatomical and Cytological Pathology from Cotentin Hospital Center.

References

[1] R. Malladi, J. Sethian, A unified approach to noise removal, image enhancement, and shape recovery, IEEE Transactions on Image Processing 5 (11) (1996) 1554–1568.
[2] U. Adiga, R. Malladi, R. Fernandez-Gonzalez, C.O. de Solorzano, High-throughput analysis of multi-spectral images of breast cancer tissue, IEEE Transactions on Image Processing 15 (8) (2006) 2259–2268.
[3] T. Chan, J. Shen, Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods, SIAM, Philadelphia, PA, 2005.
[4] G. Aubert, P. Kornprobst, Mathematical Problems in Image Processing. Partial Differential Equations and the Calculus of Variations, second ed., Applied Mathematical Sciences, vol. 147, Springer, Berlin, 2006.
[5] M. Belkin, P. Niyogi, V. Sindhwani, Manifold regularization: a geometric framework for learning from labeled and unlabeled examples, Journal of Machine Learning Research 7 (2006) 2399–2434.
[6] D. Zhou, B. Schölkopf, Discrete Regularization, Semi-Supervised Learning, Adaptive Computation and Machine Learning, MIT Press, Cambridge, MA, USA, 2006, pp. 221–232 (Section 3.13).
[7] S. Lafon, A.B. Lee, Diffusion maps and coarse-graining: a unified framework for dimensionality reduction, graph partitioning, and data set parameterization, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (9) (2006) 1393–1403.
[8] F. Wang, C. Zhang, H.C. Shen, J. Wang, Semi-supervised classification using linear neighborhood propagation, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Conference Proceedings, vol. 1, IEEE Computer Society, Silver Spring, MD, 2006, pp. 160–167.
[9] L. Grady, Random walks for image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (11) (2006) 1768–1778.
[10] V.-T. Ta, O. Lézoray, A. Elmoataz, Graph based semi and unsupervised classification and segmentation of microscopic images, in: The 7th IEEE International Symposium on Signal Processing and Information Technology, ISSPIT, 2007, pp. 1177–1182.
[11] S. Schüpp, A. Elmoataz, M.-J. Fadili, D. Bloyet, Fast statistical level sets image segmentation for biomedical applications, in: Scale-Space, 2001, pp. 3380–3888.
[12] O. Lézoray, H. Cardot, Cooperation of color pixel classification and color watershed: a study for microscopic images, IEEE Transactions on Image Processing 11 (7) (2002) 783–789.
[13] A. Elmoataz, O. Lézoray, S. Bougleux, Nonlocal discrete regularization an weighted graphs: a framework for image and manifolds processing, IEEE Transactions on Image Processing 17 (7) (2008) 1047–1060.
[14] P. Arbeláez, L. Cohen, Energy partitions and image segmentation, Journal of Mathematical Imaging and Vision 20 (1–2) (2004) 43–57.
[15] O. Lézoray, A. Elmoataz, H. Cardot, A color object recognition scheme: application to cellular sorting, Machine Vision and Applications 14 (3) (2003) 166–171.
[16] A. Renouf, R. Clouard, M. Revenu, How to formulate image processing applications, in: International Conference on Computer Vision Systems, 2007, p. 10.
[17] R. Diestel, Graph Theory, Graduate Texts in Mathematics, vol. 173, Springer, Berlin, 2005.
[18] T. Chan, S. Osher, J. Shen, The digital TV filter and nonlinear denoising, IEEE Transactions on Image Processing 10 (2) (2001) 231–241.
[19] A. Trémeau, P. Colantoni, Regions adjacency graph applied to color image segmentation, IEEE Transactions on Image Processing 9 (4) (2000) 735–744.
[20] U. von Luxburg, A tutorial on spectral clustering, Statistics and Computing 17 (4) (2007) 395–416.
[21] O. Lézoray, S. Bougleux, A. Elmoataz, Parameterless discrete regularization on graphs for color image filtering, in: M. Kamel, A. Campilho (Eds.), Image Analysis and Recognition, Lecture Notes in Computer Science, vol. 4633, Springer, Berlin, 2007, pp. 46–57.
[22] V.-T. Ta, S. Bougleux, A. Elmoataz, O. Lézoray, Nonlocal anisotropic discrete regularization for image, data filtering and clustering, Technical Report hal-00187165, Université de Caen Basse-Normadie, GREYC, HAL, October 2007.
[23] X. Ren, J. Malik, Learning a classification model for segmentation, in: International Conference on Computer Vision, 2003, pp. 10–17.
[24] F. Meyer, An overview of morphological segmentation, in: 7th International Workshop on Combinatorial Image Analysis, 2000, pp. 119–142.
[25] P. Arbeláez, L. Cohen, A metric approach to vector-valued image segmentation, International Journal of Computer Vision 69 (1) (2006) 119–126.
[26] B. Delaunay, Sur la sphère vide, Izvestia Akademii Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk 7 (1934) 793–800.
[27] S. Derivaux, S. Lefevre, C. Wemmert, J. Korczak, On machine learning in watershed segmentation, in: IEEE International Workshop on Machine Learning in Signal Processing (MLSP), 2007, pp. 187–192.
[28] I. Levner, H. Zhang, Classification-driven watershed segmentation, IEEE Transactions on Image Processing 16 (5) (2007) 1437–1445.
[29] A. Sinop, L. Grady, A seeded image segmentation framework unifying graph cuts and random walker which yields a new algorithm, in: ICCV, 2007, pp. 1–8.
[30] D. Zhou, B. Schölkopf, Regularization on discrete spaces, in: Pattern Recognition, Proceedings of the 27th DAGM Symposium, Lecture Notes in Computer Science, vol. 3663, Springer, Berlin, 2005, pp. 361–368.
[31] O. Chapelle, B. Schölkopf, A. Zien, Semi-Supervised Learning, MIT press, Cambridge, MA, 2006.
[32] S. Di Zenzo, A note on the gradient of a multi-image, Computer Vision, Graphics, and Image Processing 33 (1) (1986) 116–125.
[33] M. Dumont, R. Marée, P. Geurts, L. Wehenkel, Random subwindows and multiple output decision trees for generic image annotation, in: 6th Annual Machine Learning Conference of Belgium and The Netherlands, 2007.

About the Author—VINH-THONG TA received the M.S. degree in artificial intelligence and algorithmic in 2002. Since 2006, he is a Ph.D. student in computer science at the Université de Caen Basse-Normandie, France.
His research interests concerns graph-based variational and morphological models for image and data processing.

About the Author—OLIVIER LEZORAY received the M.S. and doctoral degrees, both in computer science, from the Université de Caen Basse-Normandie, France, respectively, in 1996 and 2000.
From 1999 to 2000, he was an Assistant Professor with the Computer Science Department, Université de Caen Basse-Normandie. Since 2000, he has been an Associate Professor with the Cherbourg Institute of Technology in the Communications, Networks, and Services Department. His research concerns color image segmentation and filtering (graph-based variational and morphological methods) and machine learning techniques for image mining (neural networks and support vector machines).

About the Author—ABDERRAHIM ELMOATAZ is a full-time Professor of computer science at the Computer Science Department, Université de Caen Basse-Normandie, France. His research concerns graph-based and discrete variational methods for image and data processing.

About the Author—SOPHIE SCHUPP is an Associate Professor with the Computer Science Department, Université de Caen Basse-Normandie, France.
Her research concerns PDEs-based variational models for image segmentation and data clustering.