GRAPH-BASED MORPHOLOGICAL PROCESSING OF MULTIVARIATE MICROSCOPY IMAGES AND DATA BASES*

O. Lézoray

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

ABSTRACT

The extension of lattice based operators to manifolds is still a challenging theme in mathematical morphology. In this paper, we propose to explicitly construct complete lattices and replace each element of a manifold by its rank suitable for classical morphological processing. Manifold learning is considered as the basis for the construction of a complete lattice. The whole processing of multivariate functions is expressed on graphs to have a formalism that can be applied on images, region adjacency graphs, and image databases. Several examples in microscopy do illustrate the benefits of the proposed approach.

Index Terms— Multivariate, Mathematical Morphology, Graphs.

1. INTRODUCTION

Mathematical Morphology (MM) is a nonlinear approach to image processing that relies on a fundamental structure, the complete lattice $\mathcal{L}$ [1]. A complete lattice is a nonempty set equiped with an ordering relation, such that every non-empty subset $\mathcal{K}$ of $\mathcal{L}$ has a lower bound $\wedge \mathcal{L}$ and an upper bound $\vee \mathcal{K}$. With the complete lattice theory, it is possible to define morphological operators for any type of data once a proper ordering is established [2]. Then morphological operators are represented as mappings between complete lattices in combination with matching patterns called structuring elements. If Mathematical Morphology is well defined for binary and gray scale images, there exist no general accepted extension that permits to perform basic operations on multivariate data since there is no natural ordering on vectors. Several orders have been reported in literature but they are reduced to considering one specific type of images (color [3] or tensor images [4]). As a consequence, if mathematical morphology has been very popular for the segmentation of gray scale microscopy images, this is not the case with multivariate images. In this paper, we propose an approach to tackle this aspect. To this aim, a graph-based formalism for the morphological processing of multivariate images is presented.

2. RANK TRANSFORM

A multivariate function is considered as a mapping $f : \Omega \to \mathbb{R}^p$ where p is the dimensionality of the vectors. One way to define an ordering relation between vectors is to use a transform [5] $h : \mathbb{R}^p \to \mathbb{R}^q$, with $q \ll p$ followed by a conditional ordering on each dimension of $\mathbb{R}^q$. Then, $\forall (\mathbf{x}_i, \mathbf{x}_j) \in \mathbb{R}^p \times \mathbb{R}^p, \mathbf{x}_i \leq \mathbf{x}_j \Leftrightarrow h(\mathbf{x}_i) \leq h(\mathbf{x}_j)$. From this, it is easy to show the following equivalence (complete lattice on $\mathbb{R}^p$)$\Leftrightarrow$(bijective application $h : \mathbb{R}^p \to \mathbb{R}^q$)$\Leftrightarrow$(rank transform on $\mathbb{R}^p$) [6]. This implies that, to induce a complete lattice, the vectors' values are not important but only their position in the lattice: this corresponds to a rank transform defined by the mapping $h : \mathbb{R}^p \to \mathbb{N}$. Ordering comparisons involved in morphological operations are performed directly on ranks and one obtains a common framework valid for data of arbitrary dimensions.

3. MATHEMATICAL MORPHOLOGY ON GRAPHS

A graph is a couple $G = (V, E)$ where V is a finite set of vertices and E is a set of edges included in a subset of $V \times V$. Two vertices are adjacent if the edge $(u, v) \in E$. $u \sim v$ denotes the set of vertices u connected to the vertex v via the edges $(u, v) \in E$. A graph is weighted if it is associated with a weight function $k : E \to \mathbb{R}^+$ satisfying $k(u, v) > 0$ if $(u, v) \in E$, and $k(u, v) = 0$ if $(u, v) \notin E$. We now introduce several definitions. The neighborhood set of vertices $\mathcal{N}(G, v)$ of a vertex v is defined as: $\mathcal{N}(G, v) = \{u \in V : (u, v) \in E\} \cup \{v\}$. The set of edges $\mathcal{A}(G, v)$ connecting any vertices in $\mathcal{N}(G, v)$ is defined as: $\mathcal{A}(G, v) = \{(u, w) \in E : u \in \mathcal{N}(G, v), w \in \mathcal{N}(G, v)\}$. A structuring element $\mathcal{S}(G, v)$ at a given vertex v is a sub-graph of G defined as: $\mathcal{S}(G, v) = (\mathcal{N}(G, v), \mathcal{A}(G, v))$. With these definitions, the erosion ϵ of a function f on a graph G at a vertex v is defined by: $\epsilon(G, f, v) = \{f(u) : h(f(u)) = \wedge h(f(w)), w \in \mathcal{N}(G, v)\}$. If we compare this definition with the usual definition of an erosion, the structuring element is directly expressed by the graph topology and the lattice is defined by the use of the rank transform h. For the case of images, this definition is equivalent to the classical algebraic case. Indeed, for images, one considers grid graphs

(one vertex per pixel) and vertices are then connected according to the chosen structuring element. However, our formulation is more general since it can be applied on graphs of the arbitrary topologies. Similar definitions can be found in [7]. With this definition, the graph topology never changes, but only vectors associated to vertices. We can reformulate the erosion as a *contracting* erosion that modifies the graph topology. To that aim, we define the erosion at a vertex v in terms of vertex preservation: $\epsilon_\mathcal{V}(G, f, v) = \{u : h(f(u)) = \wedge h(f(w)), w \in \mathcal{N}(G, v)\}$. Then, one can define the vertex erosion $\epsilon_\mathcal{V}(G, f)$ and the edge erosion $\epsilon_\mathcal{E}(G, f)$ of a graph as: $\epsilon_\mathcal{V}(G, f) = V \cap \{\epsilon_\mathcal{V}(G, f, v), \forall v \in V\}$ and $\epsilon_\mathcal{E}(G, f) = \{(u, v) \in E, u \in \epsilon_\mathcal{V}(G, f), v \in \epsilon_\mathcal{V}(G, f)\}$. Finally a contracting erosion $\epsilon_C(G, f)$ is an operation that produces a new graph $(\epsilon_\mathcal{V}(G, f), \epsilon_\mathcal{E}(G, f))$ that is a sub-graph of G. Similar definitions apply for dilation.

4. COMPLETE LATTICE LEARNING

In the previous definitions of Mathematical Morphology on Graphs, the complete lattice is assumed to be known and expressed by the rank transform h. However, the construction of such a rank transform is a difficult problem. To perform this, we consider manifold learning methods that enable to perform dimensionality reduction. This is equivalent to a rank transform when the dimension of the projected is space is equal to one. Graph-based methods have recently emerged as powerful tools for nonlinear dimensionality reduction. Among the existing methods, we consider Laplacian Eigenmaps [8]. Let $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\} \in \mathbb{R}^p$ be a set of n initial vectors. Manifold learning aims at searching for a new representation $\{\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n\}$ with $\mathbf{y}_i \in \mathbb{R}^n$. From a neighborhood graph G built from the initial data set, an adjacency matrix W is considered and weighted by a Gaussian kernel $W_{ij} = e\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2}\right)$. To have a parameterless Gaussian kernel, σ is estimated by $\sigma = \max_{v \in V, u \sim v} \|f(v) - f(u)\|$. Then, one seeks to minimize $\frac{1}{2} \sum_{ij} W_{ij} \|\mathbf{y}_i - \mathbf{y}_j\|_2 = Tr(\mathbf{Y}^T \Delta \mathbf{Y})$ with $\Delta = D - W$ that represents the un-normalized Laplacian (D is the degree matrix). The solution of the previous minimization problem can be found by solving $\Delta \mathbf{y} = \lambda D \mathbf{y}$. The eigenvectors of this equation corresponding to the smallest non zero eigenvalues form the manifold representation. To perform a complete lattice learning with manifold learning, a vertex is associated to each input vector data and a neighborhood graph is constructed. Then, we consider only the first non-zero eigenvector of the obtained Manifold representation and re-arrange the initial vectors increasingly according to their value in the first non-zero eigenvector: this defines the rank transform. Manifold learning, although being attractive, is a time consuming step for the complete lattice construction when the amount of data is large: complexity is $O(n^3)$. To overcome this, several strategies can be considered that rely on the same

idea: to reduce the size of the data on which the complete lattice construction is performed. We propose two strategies in the sequel.

4.1. Data Quantization

A first strategy can consist in reducing the input data size by Vector Quantization (VQ). Given an initial data set of size n, VQ: $\mathbb{R}^p \to \mathbb{R}^p$ is applied to construct a codebook $C : \mathbb{N} \to \mathbb{R}^p$ and an encoder $I : \mathbb{R}^p \to \mathbb{N}$. An index $h : \Omega \to \mathbb{N}$ can be deduced from D and I by applying $h(x) = I(f(x))$ to each vector $f(x) = \mathbf{x}$ of the original data set. The initial data set can be reconstructed with loss from the index and the codebook by $C(h(x))$: the obtained data set is an approximation of the initial data set with only 2^k elements. The codebook being of reduced size, one can apply manifold learning on the complete graph associated to the codebook. This enables to construct the complete lattice (the ordering of the codebook) and to define the rank transform (obtained with the function h).

4.2. Local lattice learning

A second strategy can consist in performing locally the complete lattice creation. The rank transform h is defined on subgraphs of the initial graph: the structuring elements $\mathcal{S}(G, v)$. This comes to define the rank transform only on a reduced set of vertices: $\mathcal{N}(G, v)$. The manifold learning is therefore applied on the data set $\{f(u), u \in \mathcal{N}(G, v)\}$. With this strategy, the complete lattice is not available for the whole manifold but only one sub-manifold defined, for a local processing, by $\mathcal{S}(G, v)$.

5. RESULTS

In this Section, we illustrate the two above-mentioned strategies for complete lattice learning.

5.1. Vector Quantization with Manifold Learning

We illustrate the use of vector quantization with manifold learning for morphological processing. Figure 1 illustrates this principle on a color image ($f : \mathbb{Z}^2 \to \mathbb{R}^3$) represented by a $(k^2 - 1)$-adjacency graph that means using a $k \times k$ square structuring element. The image is quantized into 512 colors and the obtained codebook is re-ordered by Manifold Learning to construct the complete lattice of the 512 colors. A rank image is created by assigning to each pixel its rank on the complete lattice of the codebook. Then, morphological operations are performed on the rank image and the final color images are obtained by reconstruction with the codebook defining the lattice. As it can be seen in Figure 1, the induced morphological operations enable an accurate processing of the image. To show that our formalism is easily applicable to any multivariate image, Figure 2 presents a morphological

processing of a multispectral image ($f : \mathbb{Z}^2 \to \mathbb{R}^{20}$) with a 1024 codebook. The segmentation of the image is performed with a watershed on the morphological gradient of an Alternate Sequential Filter of the rank image.

5.2. Local Manifold Learning

We illustrate the use of local manifold learning for morphological processing. First, we consider the morphological processing of Region Adjacency Graphs (RAG). From a cytological microscopic image (Figure 3(a)), a partition is constructed (Figure 3(b)) by labeling connected components obtained from a k-means clustering with k = 4. To the obtained partition, a RAG can be associated where each vertex represents a region and edges model adjacency relations between regions. To perform morphological operations on such a graph, one needs to define the feature vectors associated to vertices and the distance used to compare these features. We have used here a Mahalanobis distance and $f : V \to \mathbb{R}^{3\times3}$ that represents the variance-covariance matrix associated to each region. Several morphological contracting operations are then applied successively: one erosion and two dilations (Figures 3(c)-3(e)). Since these operations are contracting ones, the number of vertices is reduced at each operation. Figure 3(f) presents the original image with boundaries of Figure 3(e) superimposed. Such processing on a RAG is a simple alternative to region merging.

Finally, we consider the morphological processing of image manifolds that represent high dimensional real-world data. A database of cytological cellular images is used. This database contains color images of cells of different sizes that belong to 18 different classes. To each cell is associated a region map that delineates its nuclear boundary. For visualization purposes, we only consider the class of dystrophic mesothelials (38 cells in this category). One problem with such a database is that the images of cells have different sizes. Therefore, we consider the 64-colors quantized color histogram of each cell (only inside the nucleus) and we have $f : V \to \mathbb{R}^{64}$ that associates a color histogram to each vertex. To model this image manifold, a k-nearest neighbor graph is constructed (k = 7). The Earth Mover Distance (EMD) [9] is used to compare histogram feature vectors. For this image manifold, Morphological processing is applied: two erosions and two contracting erosions (Figure 4). The two erosions simplify the image manifold while maintaining its size. Therefore, a same feature vector can be associated to different vertices and simplification acts as a suppression of outliers. When the two erosions are contracting morphological operations, the manifold size is decreased and few representative images have been retained. To better understand the behavior of such contracting erosions, the surviving images of Figure 4(c) are shown with red borders on a graphical representation of the graph associated to Figure 4 (this projection is obtained with Laplacian Eigenmaps). One can
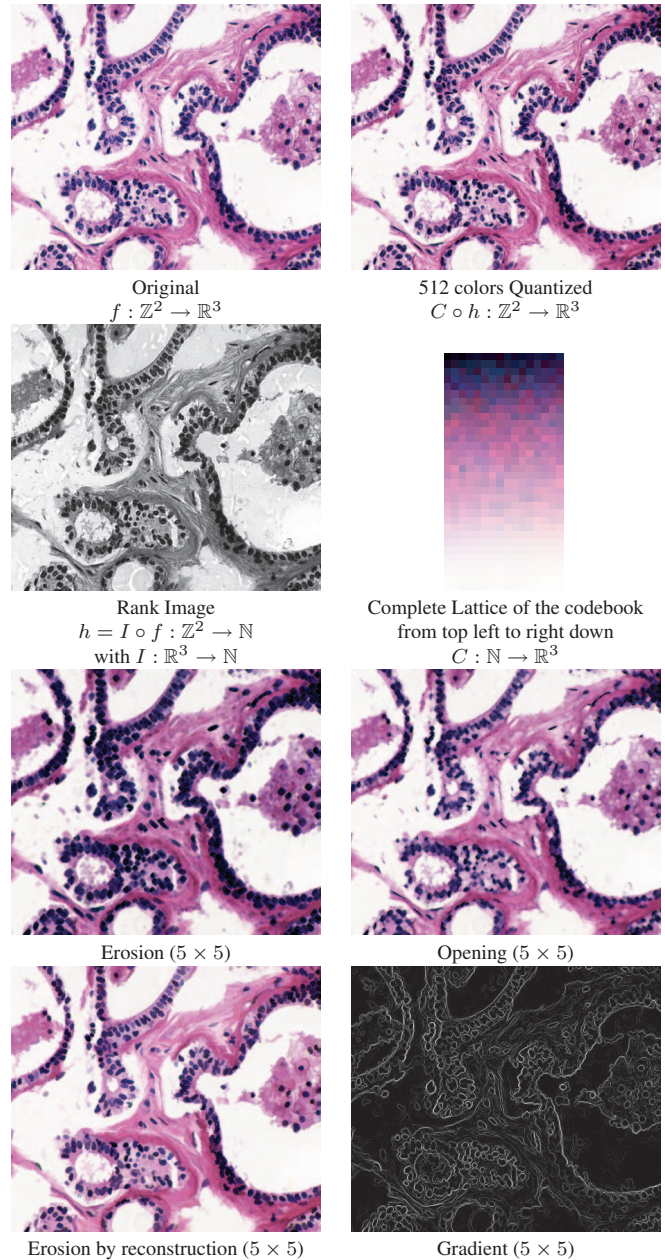

Original
$f : \mathbb{Z}^2 \to \mathbb{R}^3$

512 colors Quantized
$C \circ h : \mathbb{Z}^2 \to \mathbb{R}^3$

Rank Image
$h = I \circ f : \mathbb{Z}^2 \to \mathbb{N}$
with $I : \mathbb{R}^3 \to \mathbb{N}$

Complete Lattice of the codebook
from top left to right down
$C : \mathbb{N} \to \mathbb{R}^3$

Erosion (5×5)

Opening (5×5)

Erosion by reconstruction (5×5)

Gradient (5×5)

Fig. 1. Processing examples with a rank image obtained from Manifold Learning with Vector Quantization.

Channel 1   Channel 10   Rank Image

ASF (7 iterations with a square)   MM gradient (3 × 3 square)   Watershed
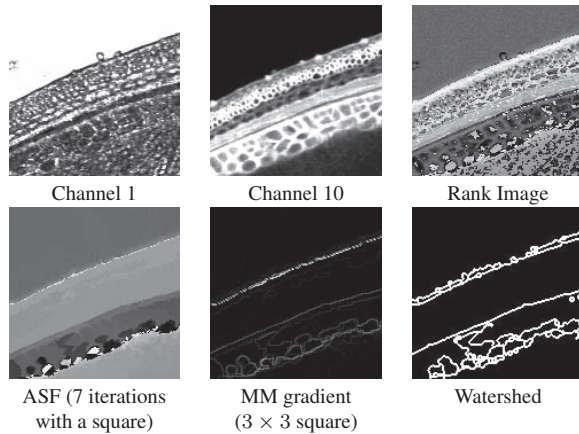
Fig. 2. MM Processing example of a 20-channels multi-spectral barley grain image.

(a) Original Image   (b) Labeled partition   (c) ϵ_C

(d) $\delta_C(\epsilon_C)$ (e) $\delta_C(\delta_C(\epsilon_C))$ (f) boundaries of (e)
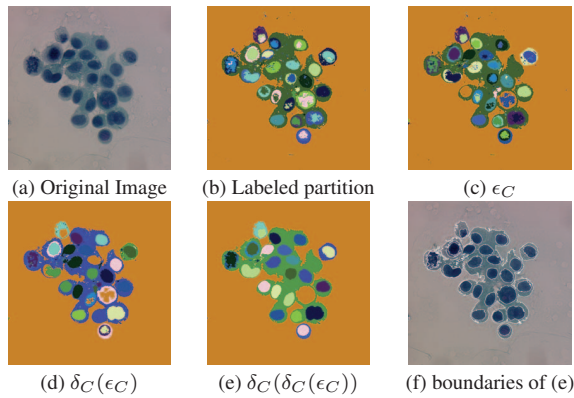
Fig. 3. Morphological operations on the region adjacency graph of a cytological image segmentation.

see that the surviving images correspond to the most representative elements of the manifold. The processing is interesting for extracting relevant items of image data bases.

6. CONCLUSION

In this paper we presented a method that enables the use of morphological operations on multivariate functions defined on graphs. Morphological operators relying on a complete lattice, the latter is dynamically constructed by manifold learning with Laplacian Eigenmaps. To have a computationally efficient solution, the manifold learning is performed either locally or after data compression. Several examples have illustrated the interest of such a family of operators for the morphological multivariate processing of images and databases in microscopy.
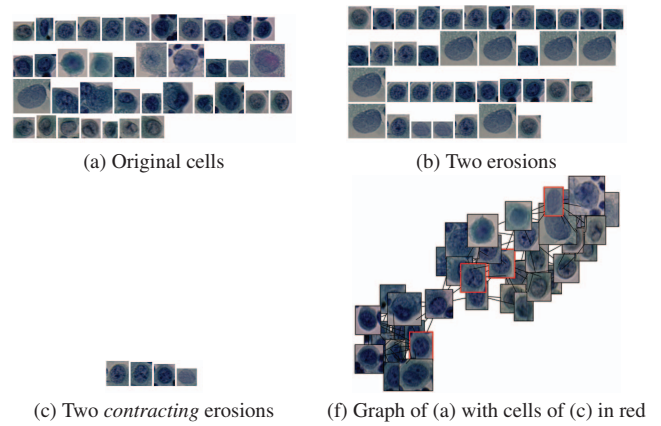

(a) Original cells   (b) Two erosions

(c) Two *contracting* erosions   (f) Graph of (a) with cells of (c) in red

Fig. 4. Morphological processing of cellular cytology image data set.

7. REFERENCES

[1] C. Ronse, "Why mathematical morphology needs complete lattices," *Signal Processing*, vol. 21, no. 2, pp. 129–154, 1990.

[2] E. Aptoula and S. Lefèvre, "A comparative study on multivariate mathematical morphology," *Pattern Recognition*, vol. 40, no. 11, pp. 2914–2929, 2007.

[3] J. Angulo, "Unified morphological color processing framework in a lum/sat/hue representation," in *Proceedings of the 7th International Symposium on Mathematical Morphology*, 2005, pp. 387–396.

[4] B. Burgeth, A. Bruhn, N. Papenberg, M. Welk, and J. Weickert, "Mathematical morphology for matrix fields induced by the loewner ordering in higher dimensions," *Signal Processing*, vol. 87, no. 2, pp. 277–290, 2007.

[5] J. Goutsias, H.J.A.M. Heijmans, and K. Sivakumar, "Morphological operators for image sequences," *Computer Vision and Image Understanding*, vol. 62, no. 3, pp. 326–346, 1995.

[6] O. Lezoray, C. Charrier, and A. Elmoataz, "Rank transformation and manifold learning for multivariate mathematical morphology," in *Proceedings of EUSIPCO*, 2009.

[7] H. Heijmans, P. Nacken, A. Toet, and L. Vincent, "Graph morphology," *Journal of Visual Communication and Image Representation*, vol. 3, no. 1, pp. 24–38, 1992.

[8] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, no. 6, pp. 1373–1396, 2003.

[9] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *International Journal of Computer Vision*, vol. 40, no. 2, pp. 99–121, 2000.