Color image segmentation using Laplacian eigenmaps

Ioannis Tziakos
Queen Mary University of London
Department of Electronic Engineering
United Kingdom

Christos Theoharatos
University of Patras
Department of Physics
Electronics Laboratory
Patras 26500, Greece
E-mail: htheohar@upatras.gr

Nikolaos A. Laskaris
Aristotle University of Thessaloniki
Department of Informatics
Artificial Intelligence and Information Analysis Laboratory
Thessaloniki 54124, Greece

George Economou
University of Patras
Department of Physics
Electronics Laboratory
Patras 26500, Greece

Abstract. *The novel technique of Laplacian eigenmaps (LE) is studied as a means of improving the clustering-based segmentation of color images. Taking advantage of the ability of the LE algorithm to learn the actual manifold of the multivariate data, a computationally efficient scheme is introduced. After embedding the local image characteristics, extracted from overlapping regions, in a high-dimensional feature space, the skeleton of the intrinsically low-dimensional manifold is constructed using spectral graph theory. Using the LE-based dimensionality reduction technique, a low-dimensional map is computed in which the variations of the local image characteristics are presented in the context of global image variation. The nonlinear projections on this map serve as inputs to the Fuzzy C-Means (FCM) algorithm, boosting its clustering performance significantly. The final segmentation is produced by a simple labeling scheme. The application of the presented approach to color images is very encouraging and illustrates the effectiveness of the performance over alternative methods.* © *2009 SPIE and IS&T.*
[DOI: 10.1117/1.3122369]

1 Introduction

One of the key processes in computer vision and image understanding is image segmentation, which tries to solve the problem of partitioning an image into smaller disjoint homogeneous regions that share similar attributes. In the literature, numerous algorithms have appeared over the years. Usually, each one addresses a specific facet of the segmentation problem (e.g., the incorporation of textural characteristics for object recognition). Some approaches are based on global features like the color histogram and employ thresholding[1] or multithresholding schemes.[2] Other techniques are characterized as region-based ones and appear in the form of region-growing or region split-and-merge methods.[3] Nowadays, with the advents in remote sensing technology and the extensive use of medical imaging modalities that give rise to vector fields, techniques for multivariate image segmentation are a prerequisite for image understanding.

The family of clustering-based techniques is perhaps the most popular approach toward color image segmentation, due to their flexible character.[4] In general, clustering techniques can be categorized into (1) hierarchical-based techniques that try to discover structures and, thereafter, substructures in a recursive manner,[5] and (2) partition-based techniques that try to obtain a single partition of data based on the optimization of an appropriate objective function.[6] In the latter case, in the fields of which lies the presented approach, using a proper feature extraction step, pixels are mapped into a feature space and a clustering algorithm—such as Fuzzy C-Means (FCM; e.g., Ref. 7), K-means or adaptive K-means (e.g., Refs. 8 and 9), Expectation Maximization (e.g., Ref. 10), and Self-Organizing Map (e.g.,

Paper 08054RR received Apr. 5, 2008; revised manuscript received Feb. 12, 2009; accepted for publication Mar. 10, 2009; published online Apr. 21, 2009.

1017-9909/2009/18(2)/023004/10/$25.00 © 2009 SPIE and IS&T.

Ref. 11), etc.—is employed to identify distinct groups. The selection of features, the "curse of dimensionality,"[12] and the definition of a number of clusters are the most important problems the user has to deal with. While the latter problem can be overcome via simple experimentation, there is no straightforward solution for the first two, since they counteract each other. In addition, a user is not allowed to include as many features as possible, since the emptiness of the constructed feature space obscures clustering. Here lies a striking difference between the procedures taking place in the human perceptual mechanism and the current machine-vision segmentation algorithms. Inherent in the biological systems is the ability to recover the low-dimensional structure when confronted with stimuli lying in high-dimensional spaces.[13]

Motivated by this fact, a few techniques have recently appeared in the computational literature that try to imitate the perceptual manifold learning from high-dimensional data.[14–16] The incorporation of such a learning technique can infuse substantial benefit to color image segmentation. In our previous work,[17] we used such a technique to segment images in a supervised manner. In order to improve the segmentation scheme (or result) and alleviate the problem of feature selection, we incorporate an initial embedding of local image patches (i.e., vectors) representing overlapping image regions in a high-dimensional feature space. The Laplacian Eigenmap (LE) algorithm[18] is utilized so as to discover, parameterize, and visualize the learned data manifold. This nonlinear dimensionality reduction technique reveals the structure of the manifold in a reduced feature space, validated by the substantial success of classifying points using spectral graph partitioning and clustering methods.[19–21] Based on this sketch, the data manifold can be partitioned efficiently. Consequently, the image regions can be classified properly using a well-known clustering algorithm like the Fuzzy C-Means (FCM)[22] combined with a clustering validity index[23] to select the optimum number of clusters. By realizing the requirement for a simple feature selection step, we suggest the use of a generic strategy that incorporates, in a long feature vector, different region characteristics like illumination changes, texture, color layout, etc. The "true" dimensions of image variation can be discovered by the LE technique and used for segmenting the image plane. Promising results on color images demonstrate the potentials of the proposed segmentation approach.

2 The Laplacian Eigenmap

2.1 *Overview*

Given a set of high-dimensional data sampled from a low-dimensional manifold, a large family of algorithms has been presented in the literature that utilizes the top or bottom eigenvectors of an appropriately constructed (dis)similarity matrix, in order to find and compute a faithful (nonlinear) embedding. The existing algorithms, linear and nonlinear ones,[24] can reveal low-dimensional manifolds that are not detected by classical methods such as principal component analysis (PCA).[25]

The proposed segmentation technique is organized around the nonlinear manifold learning and low-dimension data representation abilities of LE. The locality-preserving character of the LE algorithm makes it relatively insensitive to outliers and degradation imposed by the presence of noise. In addition, it is not prone to short circuiting, as only the local distances inside a neighborhood are used. In their work, Belkin and Niyogi[18] argued that by trying to preserve local information in the embedding, the algorithm implicitly emphasizes the natural clusters in the data. More particularly, the LE has close connections to spectral clustering algorithms developed in learning and computer vision (e.g., the approach of normalized cuts by Shi and Malik[19]). In this sense, dimensionality reduction and clustering can be considered as two sides of the same coin. In contrast, global methods like Ref. 14 do not reveal any sufficient tendency to cluster (they mainly represent the data structure within a feature space in which all data are embedded and utilize this structure for identifying similar patterns), since the optimization criterion pursued by this algorithm is the preservation of all pairwise geodesic distances between points. The basic mathematical notion of LE can be summarized as follows.

Given a set of N multivariate observations embedded as vectors, $x_1, x_2, \ldots, x_N \in \Re^p$ $(p \gg 1)$, a weighted adjacency graph $\mathbf{G}$ is built over the endpoints of these vectors. It consists of N nodes and a set of edges connecting neighboring points. The embedding map can emerge by computing the eigenvectors of the graph Laplacian. Let us consider the problem of mapping the weighted graph $\mathbf{G}$ to a line so that connected points stay as close as possible. If two points x_i and x_j (with $i, j = 1, 2, \ldots, N$) are close enough, then there exists an edge that connects them. Let $\mathbf{y} = (y_1, y_2, \ldots, y_N)^T$ be such a map. A reasonable criterion for choosing a "good" map is to minimize the objective function:

$$\sum_{i,j} \|y_i - y_j\|^2 W_{ij}, \tag{1}$$

where the weight W (which will be explained in the following sections) incurs a heavy penalty if neighboring points x_i and x_j are mapped far apart.[18] Therefore, minimizing is an attempt to ensure that if points x_i and x_j are close, then y_i and y_j are close as well. It turns out that the minimization problem reduces to finding

$$\arg\min_{\substack{\mathbf{y}^T D\mathbf{y}=1 \\ \mathbf{y}^T D\mathbf{1}=0}} \mathbf{y}^T L\mathbf{y}, \tag{2}$$

where $L = D - W$ is the Laplacian matrix, and D is a diagonal weight matrix whose entries are column (or row, since W is symmetric) sums of W, i.e., $D_{ii} = \Sigma_j W_{ij}$. Matrix D provides a "measure of importance" on the vertices of the graph, in a way that the bigger the value D_{ii} is, the more "important" is that vertex. In the previous equation, the constraint $\mathbf{y}^T D\mathbf{y} = 1$ removes an arbitrary scaling factor in the embedding, while the condition $\mathbf{y}^T D\mathbf{1} = 0$ can be thought of as removing a translation invariance in $\mathbf{y}$, with $\mathbf{1}$ being a constant function taking 1 at each vertex.[18]

Shi and Malik[19] showed that the solution to the general problem of mapping the weighted graph $\mathbf{G}$ to m dimensions is provided by the matrix of eigenvectors corresponding to the lowest eigenvalues of the generalized eigenanalysis problem:

$$L\mathbf{y} = \lambda D\mathbf{y}. \quad (3)$$

The core algorithm is very simple. The solution reflects the intrinsic geometric structure of the manifold. It does, however, require a search for neighboring points in a high-dimensional space. This makes LE an appealing algorithm since computational effectiveness and efficiency (i.e., speed), although usually neglected, is of great importance for every algorithmic scheme especially when the goal is to segment various images for real-life applications.

The LE technique produces a mapping that preserves the locality of the original lattice on the manifold. It is worthwhile noticing that an isometric embedding that preserves global distances, such as that attempted by Isomap, is theoretically possible only when the surface of the manifold is smooth and convex. In that case, the curvature tensor is zero, which is the case in manifolds like the "Swiss roll."[14] However, a classical result shows that even for a 2-D sphere (or any part of a sphere), no distance-preserving mapping into the plane can actually exist. Additionally, nongraph dimensionality reduction algorithms tend to scale badly with the number of training vectors. To compensate for the large number of vectors, landmark tricks are used,[15] but this procedure contradicts with the assumption of a dense manifold sampling needed to approximate the manifold structure.

These features led to the adoption of LE as a promising candidate to assist and improve the color image segmentation problem. This selection is justified by the experimental results (presented later), which show a clear improvement of the segmentation procedure in the reduced space, produced by LE, over the resulting segmentation using the original high-dimensional feature vectors.

2.2 *The LE Steps*

The individual steps that comprise the algorithmic procedure used in this work are summarized here:

1. Construct the adjacency graph $\mathbf{G}$ using the k nearest neighbors ($k \in N$). Nodes i and j are connected by an edge if i is among the k nearest neighbors of j and vice-versa.
2. Choose the weights W using the heat kernel ($t \in \Re$):

$$W_{ij} = \begin{cases} e^{-\|x_i - x_j\|^2/t} & \text{if } i,j \text{ are connected} \\ 0 & \text{otherwise} \end{cases}. \quad (4)$$

3. Compute the N eigenvectors of the generalized eigenvector problem in Eq. (3). Eigenvectors $\mathbf{y}_{[N]}$ are sorted in ascending order according to their eigenvalues $\lambda_{[N]}$,

$$\begin{matrix} \mathbf{y}_0, \mathbf{y}_1, \cdots, \mathbf{y}_{N-1} \\ 0 = \lambda_0 \leqslant \lambda_1 \leqslant \cdots \leqslant \lambda_{N-1} \end{matrix}. \quad (5)$$

4. The resulting reduced space of m dimensions, $\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{N-1} \in \Re^m$ $(m \ll p)$ is obtained from the eigenvectors (excluding $\boldsymbol{y}_0$ corresponding to eigenvalue 0) by following the simple scheme:

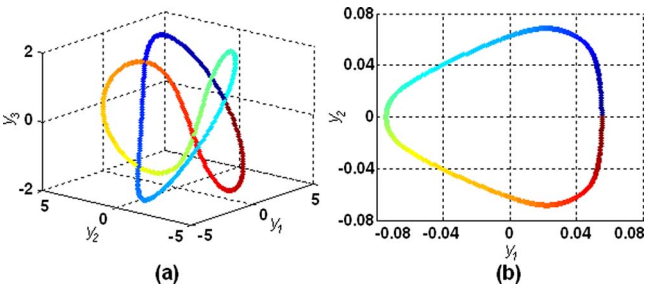

Fig. 1 (a) The trefoil knot manifold of 539 points. (b) Unfolded manifold in two dimensions using the LE algorithm.

$$x_i \rightarrow [\mathbf{y}_1(i), \mathbf{y}_2(i), \ldots, \mathbf{y}_m(i)], \text{ with } i = 1,2,\ldots,N, \quad (6)$$

where the convention $\boldsymbol{y}_j(i)$ stands for the selection of the i'th component of eigenvector $\boldsymbol{y}_j$.

A visual example of 539 vectors sampled from a trefoil knot in three dimensions, is presented in Fig. 1(a). In this case, the underlying manifold is a 1-D curve, but due to the fact that the loop is closed, it can be represented in a Euclidean space only by using a 2-D closed loop. The application of LE to this manifold using $m=2$ dimensions as well as the average node-to-node distance in the resulting graph (i.e., the heat kernel parameter), produces Fig. 1(b), which clearly indicates the manifold learning and unfolding abilities of the algorithm. The derived dimensionality reduction is not just a projection into a 2-D space, but as the coloring of vectors reveals, the whole sequence of points on the manifold corresponds exactly to the one appearing in the reduced space.

3 FCM-Based Clustering Algorithm

The projection of data into lower dimensions uncovers the structure of the data manifold for identifying similar patterns. Based on this sketch, the need to cluster the projected data arises so as to classify the image regions, taking into careful consideration the edge pixels that are not yet explicitly classified. Thus, we are highly interested in using a well-known clustering algorithm and efficiently treating the problem of cluster validation in the context of effective data partitioning. In the core of the procedure lies the Fuzzy C-Means (FCM) clustering algorithm,[22] along with a recently introduced validity index V_{WSJ} (Ref. 23). Due to its beneficial concept of assigning membership grading, FCM is considered to perform quite effectively when dealing with outliers. This is very important in practice, since several validity indices proposed in the literature are based on it.

Mathematically, the FCM-based clustering algorithm can be presented using the following notations. Let $X = \{x_k\}_{k=1}^N \subset R^p$ be a set of vectors, with N being the number of samples to be clustered and p the dimension of the input space. The goal is to estimate a partition matrix $U = [u_{ik}]_{C \times N}$, $1 \leqslant i \leqslant C$, and $1 \leqslant k \leqslant N$ that defines its segmentation into a number of $C \ll N$ homogeneous clusters. In the previous equation, $u_{ik} \in [0,1]$ denotes the fuzzy membership matrix of vector k in a region i. Each row in matrix U

represents the membership grade of all vectors belonging to a distinct cluster region, while the elements of U are subject to the usual conditions in fuzzy clustering:

$$\sum_{i=1}^{C} u_{ik} = 1, \quad \forall\, k; \quad 0 < \sum_{k=1}^{N} u_{ik} < N, \quad \forall\, i; \quad \sum_{i=1}^{C}\sum_{k=1}^{N} u_{ik} = N. \tag{7}$$

The FCM algorithm derives C cluster centroids V_i in R^p, by minimizing the dispersion of feature attributes between the center of region i and vector k to determine the pixel membership as follows:

$$J_m = \sum_{i=1}^{C}\sum_{k=1}^{N} u_{ik}^m d_{ik}^2 = \sum_{i=1}^{C}\sum_{k=1}^{N} u_{ik}^m \|X_k - V_i\|^2, \tag{8}$$

where $m \in (1,\infty)$ is a weighted parameter[22] that controls the fuzziness in the classification process, the distance function $\|\cdot\|$ stands for the standard Euclidean norm, and the matrix V denotes the set of centroids tabulated in the form $V = [v_{ik}]_{C\times p} = [V_1, V_2, \ldots, V_C]$. After a proper initialization of the partition matrix, local minimization of the function J_m is accomplished by iteratively adjusting the values of u_{ik} and V_i according to the following two formulae:

$$u_{ik} = \left[\sum_{j=1}^{C}\left(\frac{d_{ik}}{d_{jk}}\right)^{2/m-1}\right]^{-1}, \quad \forall\, i,k$$

$$\text{and} \quad V_i = \frac{\sum_{k=1}^{N} u_{ik}^m X_k}{\sum_{k=1}^{N} u_{ik}^m}, \quad \forall\, i. \tag{9}$$

Each time this double-step procedure is calculated, the evolved matrix is compared with its former version and the algorithm terminates when the changes become negligible. The termination can also be executed based on the comparison of successive J_m values.

In most clustering algorithms, including FCM, the number C of clusters comprising a dataset is a user-defined parameter and, indeed, a difficult one to set in practical applications. The definition of the optimal number of classes C_0 is generally considered an important issue that deserves a careful treatment, since this number has to be tailored to the dataset at hand.[26] Several validity indices, specifically devoted to fuzzy clustering, have been reported in the literature, and their relative success seems to be case dependent.[27] Early indices such as the partition coefficient and classification entropy make use of membership values only and possess the advantage of easy computation. Nowadays, it is widely accepted that a better definition of a validity index must consider both the compactness within each cluster and the separation between distinct clusters. In Ref. 23, Suna *et al.* proposed a new validity index that performs efficiently even when clusters tend to overlap each other, based on a linear combination of compactness and separation. The validity index follows the formula

$$V_{WSJ}(U,V,C) = Scat(C) + \frac{Sep(C)}{Sep(C_{\max})}, \tag{10}$$

where $Scat(C)$ and $Sep(C)$ represent the compactness and separation of the obtained clusters respectively. A cluster number that minimizes $V_{WSJ}(U,V,C)$ is considered to be the optimal value for estimating the number of clusters present in the data.

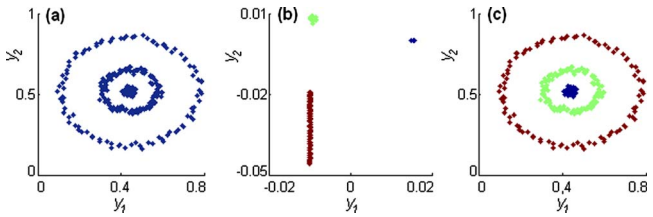

Fig. 2 Graphical example of the utilized procedure: (a) original dataset, (b) dataset after the application of LE, and (c) final clustering result: the FCM plus the validity index result in three well-separated classes.

By fusing together the three tools defined earlier (i.e., LE, FCM, and validity index), a spectral graph-clustering algorithmic scheme is formed that emphasizes the data structure of the manifold. The basic steps that should be followed are simple and straightforward. First, the LE-based dimensionality reduction algorithm is applied, having formerly defined the parameters for the graph weights as well as the number of the resulting feature-space dimensions. Next, the FCM algorithm is employed in order to acquire the membership matrix U and the cluster centers V for C_0 groups estimated from the utilized validity index V_{WSJ}. Last, the selected clustering result is obtained as the one for which the validity index is optimal.

A simple example of the algorithmic performance, when confronted with data that lie on a manifold, is shown in Fig. 2. The original dataset consists of 299 vectors (illustrated as points) in $\Re^2$. The incorporation of the LE-step clearly separates the distinct modules from each other, forming dense clusters in the 2-D reduced space, as illustrated in Fig. 2(b). The final clustering outcome emerges from the FCM plus the validity index application, resulting in the three well-separated clusters shown in Fig. 2(c).

We need to emphasize here that the mapping from the original space onto a space spanned by the first m eigenvectors defines a *diffusion map* of the derived reduced space. According to Nadler *et al.*,[28] diffusion distances in the original space are the same as Euclidean distances in the image of the diffusion map, and consequently, this justifies the use of Euclidean distances in the image of the diffusion map. For this purpose, the FCM procedure (or any other analogous approach such as K-Means, etc.) can be directly applied in the reduced feature space for clustering the involved data.

4 Segmentation Technique

The final segmentation is effected by selecting the local image characteristics in a simple form, without sacrificing our generality. Using a rectangular grid, each multivariate image is first partitioned into overlapping square blocks, with every block being a $n\times n$ RGB matrix. In general, the block size should be large enough to capture statistical regularities in image parts. For each block, a representative feature vector x_i of size $3n^2$ is formed by concatenating the vectorial measurements corresponding to the pixels included in the specific image patch—that is; $x_i = (R_{ii}, G_{ii}, B_{ii})$, with $i = 1,2,\ldots,n$. By doing so, a correspon-

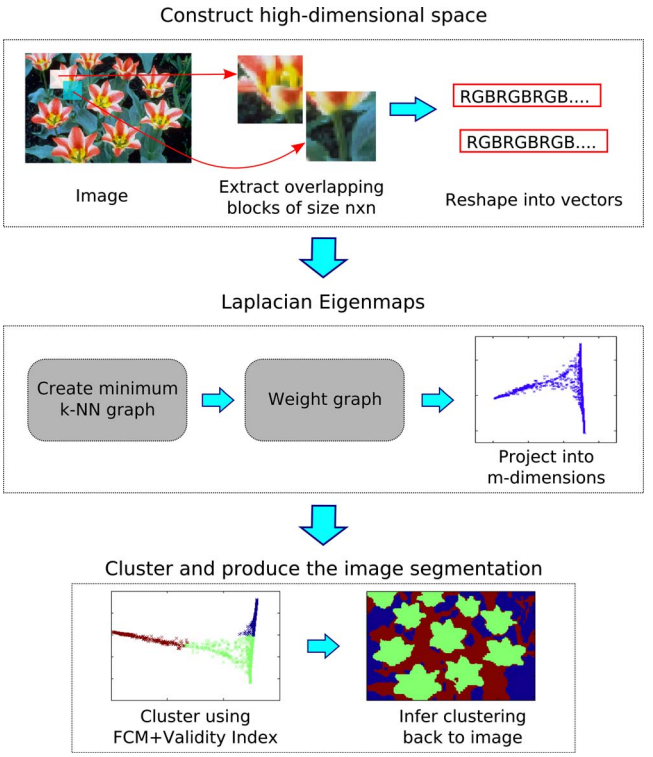

Fig. 3 Block diagram of the proposed approach. Overlapping blocks are extracted from the image and then flattened into vectors and projected with LE. The result is clustered using the FCM plus the validity index, and the clustered representation is referred back to the image.

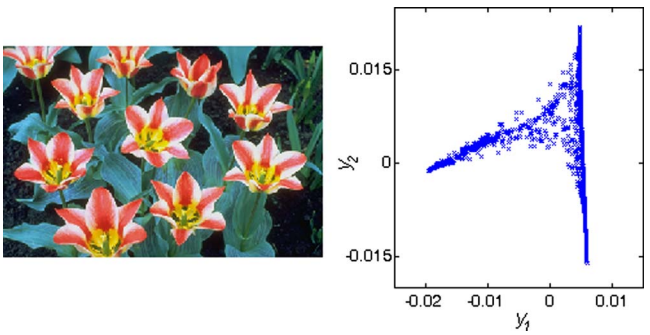

Fig. 4 Left: Sample RGB image Tulips (758×512 pixels). Right: Reduced 2-D space. Each vector represents a block of 8×8 pixels with 2 pixels overlapping.

dence between the specific image-plane locations (i.e., the spatial domain) and feature space locations (i.e., the endpoint of each feature vector) is established. This apparent spatial undersampling is motivated by computational economy, refined by the final labeling procedure and fully justified by the obtained results.

The block diagram in Fig. 3 illustrates graphically the proposed approach. In summary, our segmentation technique includes the following sequential steps:

1. The color image is first divided into a restricted number of $n \times n$ blocks. The image is resized so that the block partitioning fits entirely into the image plane. This is dictated by the need to keep the feature space free from blocks that carry "misleading" information, since the quality of the feature space is of great importance. In this way, blocks are reshaped into vectors of $3n^2$ dimensions (representing them in a high-dimensional feature space).
2. The adjacency graph is then constructed using the k-Nearest Neighbors (NN) variant of the LE algorithm. The parameter k is the minimum value that leaves the graph connected. The procedure of finding k is based on an iterative procedure that creates a graph using an increasing number of nearest neighbors until the resulting graph is connected. The final graph is weighted by the heat kernel W_{ij} provided in Sec. 2.2, where the parameter t is selected as the average distance of the resulting fully connected graph. The value is acceptable since the representations produced by LE are not very sensitive to this value of t, assuming small values for k.
3. A map of the blocks, in $m \ll p$ dimensions, is produced (i.e., in reduced feature space). Parameter m varies in the range of $[2-15]$, usually setting its value to 2 or 3 in order to be able to visualize the clustering tendencies (see for example Fig. 4) and therefore facilitate the definition of the final number of clusters.
4. The FCM plus the validity index combination is applied to the "projections" of the blocks in the reduced space. In this way, the blocks are partitioned into classes.
5. To infer the classification of all the individual pixels from the classification of the formed blocks, a simple label-assignment scheme is followed. The class label of the corresponding block is assigned to each pixel of that block. The pixels at the overlapping regions, where there is a transition between diverse classes, are not classified.

Due to the overlapping nature of the blocks, thick areas of unclassified pixels exist in the resulting segmentation map (which will be explained and illustrated in Sec. 5). These pixels belong to adjacent patches and should be assigned to different classes. In order to classify them, a small patch is originally formed around each pixel. This patch is afterward compared to the original ones from which the pixel in question belong. Last, each pixel is assigned to the most similar class. This procedure results to the final segmentation map.

The segmentation method described above is exemplified in Figs. 4 and 5. More particularly, in Fig. 4, the original RGB image Tulips is presented along with the projected feature vectors in 2-D, with each vector (illustrated using a blue dot label) representing a block of 8×8 pixels with an overlap of two pixels wide. Following the procedure described earlier, the image is segmented into multiple regions. The final segmentation, after the block relaxation step, is computed and illustrated in Fig. 5 using different numbers of distinct classes in the resulting image partitioning. Figure 5(a) illustrates the output image having three separate segments, while Fig. 5(b) presents the image having two segments, respectively.

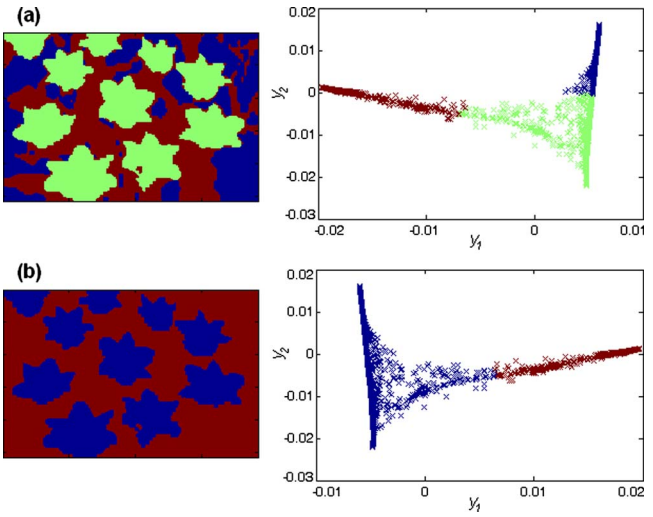

Fig. 5 Final segmentation after the block relaxation step, partitioning the reduced feature space into (a) three separate classes (right panel), thus producing three distinct segments in the segmented image (left panel), and (b) two separate classes (right panel), thus producing two segments in the segmented image (left panel).

5 Experimental Setup

The algorithmic procedure described earlier requires the setup of two distinct parameters: the block size $n \times n$ and the dimensions m of the reduced space. By assigning global values, which are justified by the following experimental results, a good segmentation map is obtained. The proposed segmentation technique was applied to several RGB images from the Berkeley Segmentation Dataset (BSD).[29] Results generated with FCM and Normalized Cuts (Ncuts) algorithms are also shown for comparison. The following examples consider various aspects of the introduced color image segmentation methodology, including a sort discussion for justifying the incorporation of the LE algorithm.

5.1 *Justifying the Need for LE*

The successfully resulting segmentation could have been attributed simply to the "clever" feature extraction step (i.e., definition of square image patches), which incorporates at once a variety of local characteristics (i.e., color and texture). Thus, the need to justify the use of LE arises. The test is simply to apply the same clustering algorithm directly in the high-dimensional feature space (in other words, to bypass the LE-step). Working in high-dimensional feature space where the data lay on a manifold makes distances shorter.[30,31] Figure 6 depicts the preceding proposition test by revealing the actual (geodesic) distance

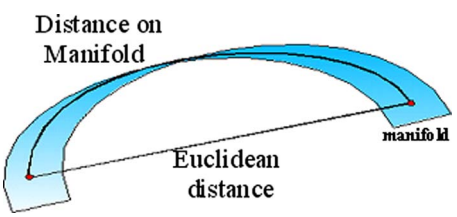

Fig. 6 Geodesic distance on a manifold in regard to the Euclidean distance.

on the presented manifold in regard to the Euclidean one. It is evident that the segmentation benefits from the LE-step since it discovers the intrinsic geodesic manifold, which in turn benefits the FCM-clustering process.

It is generally true that the distances acquired from the high-dimensional space are sometimes misleading, because the intrinsic relation between them, as well as the geodesic configuration, is ignored. In this case, the involved actions based on these distances do not correspond to the true structure of the data,[32] and thus FCM fails to partition the regions correctly. On the other hand, the LE works with local distances where the error is negligible. The resulting low-dimensional representation retains the information for the relative placing of the feature vectors on the manifold structure.[13] In addition, FCM stumbles upon the "curse of high dimensionality," at which due to the increased number of dimensions of the feature vectors and their finite number, the resulting space is very sparse and empty.[12] Likewise, data density is close to zero, and vectors show up as a small density fluctuation. This makes every classification attempt very difficult and error prone.

The final reason for using LE is dictated by the nature of the feature vectors extracted from the image, using the predefined block-partitioning scheme. These vectors are characterized from the RGB color-coding of each block facet, which is generally not perceptually uniform. The LE seems to overcome the apparent difficulty and discover the hidden reduced dimensionality, since the constructed graph actually connects (neighboring) blocks with small variations in color layout.

In terms of time complexity, the bottleneck of the approach is the graph construction. A brute force approach would need $O(N^2)$, where N is the number of vectors. An approximate neighborhood search using dedicated data structures[33] could perform this more efficiently. The resulting graph is sparse, and thus the generalized eigensolver needs to compute only a few eigenvectors. The overhead of using LE is only a fraction compared to traditional isometric embeddings and other spectral clustering techniques.

5.2 *Experimental Results and Comparisons*

To provide proof of the benefits of the proposed framework, we compared the LE-based algorithm with Normalized Cuts[19] to a set of images from the Berkley dataset; the results are illustrated in Figs. 7 and 8. Figure 7 shows results for a number of images having a low number of clusters in the final segmentation map (i.e., less than three), while Fig. 8 presents the outcome for images containing more than three clusters. All images in both figures are divided into 8×8 blocks (i.e., $n=8$) with 2-pixel overlap. The formed long vectors are projected in a manually selected number of dimensions (usually from 2 to 5), which reveals the manifold structure of the extracted features and the underlying distribution of the given image. The final number of classes is mainly estimated using the validity index. On the other hand, regarding the Normalized Cuts (Ncuts) based image segmentation presented in columns (d) and (e) of both Figs. 7 and 8, the number of resulting segments is predefined (i.e., manually set) for a low and high number of segments, respectively. In both columns, all

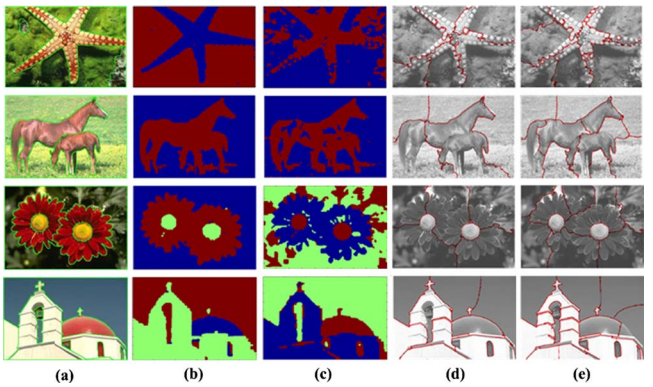

Fig. 7 Segmentation results for images coming from the Berkeley dataset[29] having fewer than (or equal to) three clusters in the final segmentation map: (a) ground-truth hand segmentation, along with the segmentation results from different methodologies; (b) with and (c) without the LE-step; as well as the Normalized Cuts result with (d) low and (e) high number of final segments.

Ncuts-based segmentations were estimated after careful experimentation, selecting the final number of segments that produce the best perceivable result.

From the illustrated results, we can perceive that the proposed framework performs more adequately for images containing some amount of texture information. This can be attributed to the efficient feature extraction technique of incorporating square image patches that utilizes, automatically, both color and texture characteristics. What is also interesting to notice is that the concept of classes of regions instead of segments is easier to relate visually in the image. In this way, selecting to segment the image into a number of visual regions tends to segment the image into complete objects. In contrast, the segments emerging from the Normalized Cuts do not follow this concept. For example, the third row of Fig. 7 depicts flowers, for which a logical segmentation would be a foreground–background segmentation. Our algorithm generates such a segmentation result [Fig. 7(b)], while this is not the case for Normalized Cuts. One of the possible explanations for this effect is the fact

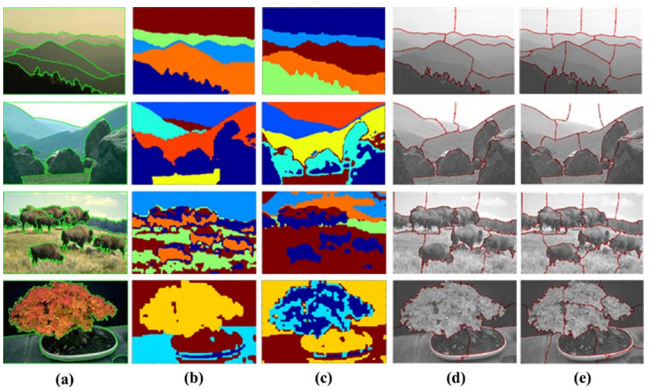

Fig. 8 Segmentation results for images coming from the Berkeley dataset[29] having more than three clusters in the final segmentation map: (a) ground-truth hand segmentation, along with the segmentation results from different methodologies; (b) with and (c) without the LE-step; as well as the Normalized Cuts result with (d) low and (e) high number of final segments.

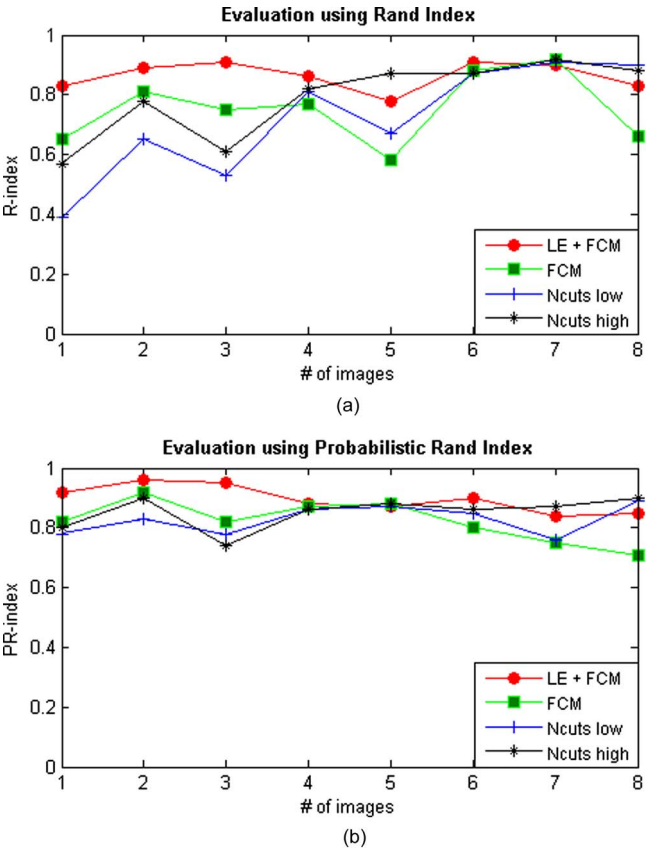

Fig. 9 Calculation of the Rand index (a) and the probabilistic Rand index (b) for the eight images of Figs. 7 and 8.

that the Normalized Cuts algorithm starts from an image boundary, which is not always the segmentation case for a number of images. Additionally, in its standard version, Normalized Cuts is able to function using the grayscale images, while the presented technique is not limited to the number of channels used. Thus, the LE-based method can function with the full RGB color data fusing together the additional information.

In order to quantitatively evaluate the proposed segmentation technique, two indices were utilized in our experiments: the Rand index (R-index) and the probabilistic Rand index (PR-index).[34] R-index is a nonparametric test that works by counting pairs of pixels that have compatible label relationships in the two segmentations to be compared. The calculation is executed between the different results produced with the preceding segmentation techniques and the ground-truth images of the human-segmented dataset. On the other hand, the PR-index is a generalization of the R-index that allows comparison of a segmentation result with multiple ground-truth images through soft nonuniform weighting of pixel pairs as a function of the variability in the ground-truth set. Figure 9 demonstrates the evaluation results of R- and PR-indices (upper and lower panels, respectively) for the set of eight images segmented in Figs. 7 and 8 using the different segmentation algorithms. In both panels, values corresponding to numbers 1 to 4 on the horizontal axis are attributed to the evaluation of the four seg-

Table 1 Comparison of the segmentation results for a subset of the Berkeley dataset (30 images), containing a different number of clusters in the final segmentation map, using the R-index and the PR-index.

Number of classes	Evaluation index	FCM+LE	FCM	Ncuts low	Ncuts high
2	R	0.85	0.67	0.59	0.71
	PR	0.93	0.84	0.81	0.86
3	R	0.87	0.69	0.65	0.75
	PR	0.91	0.81	0.79	0.83
4	R	0.80	0.66	0.72	0.75
	PR	0.85	0.73	0.80	0.81
5	R	0.81	0.63	0.70	0.80
	PR	0.84	0.75	0.79	0.87
6	R	0.77	0.59	0.71	0.75
	PR	0.81	0.70	0.79	0.80

mentation results provided in Fig. 7, while values corresponding to numbers 5 to 8 are attributed to the evaluation of those provided in Fig. 8.

To provide a more comprehensive evaluation of the proposed technique, we experimentally demonstrated the performance using a subset of 30 images coming from the Berkeley dataset. The images were selected so as to contain a different number of classes (according to the utilized validity index) in the final segmentation map. For all images, the R- as well as the PR-index were estimated, and the results were averaged for images having the same number of classes. Table 1 presents the results provided using the proposed segmentation technique, in comparison to the FCM- and Ncuts-related methods. Regarding Normalized Cuts, we followed the same procedure as before by estimating two different segmentations, one for lower and one for higher number of final segments, trying to produce the best possible perceivable outcome. Last, the segmentation results using all methods are illustrated in Fig. 10, which summarizes all pairwise measurements presented in Table 1. The left group of bars contains the evaluation using the Rand index, while the right group contains the results of the probabilistic Rand index. For both settings, our proposed segmentation scheme achieves highest accuracy, outperforming the FCM-based and the Ncuts-based segmentation approaches. To conclude, the experimentation phase revealed the apparent superiority of the introduced methodology using both qualitative attributes and quantitative measurements. The results using LE are more consistent and provide better segmentations closer in context to humans.

5.3 *Justifying the Edge-Blocking Effect*

During the partitioning of each color image into a certain number of blocks, the edge-blocking effect arises. In this direction, blocks including an edge that exist between different objects, or objects and the background, should be treated efficiently. The corresponding vectors originating from the feature extraction step and their mapping into a low-dimensional space should not be classified into a class of their own, i.e., one that contains a selected portion of various objects in the image. The number of those artificially created classes may be significant enough, depending each time on the "richness" of color information of the image under consideration. In that case, an increased possibility of transition classes (i.e., edge classes) can be observed.

In Fig. 11(a), the depicted color image visualizes a hand with a large blue ring on a fabric (textured background). Its segmentation is produced in Fig. 11(b), using an 8 $\times$ 8 pixels block partitioning with an overlap of 2 pixels

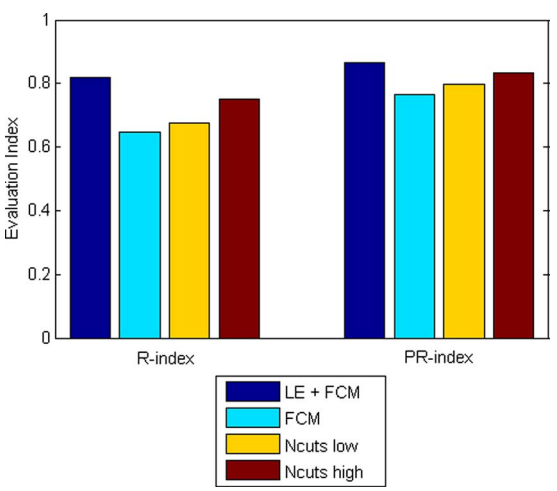

Fig. 10 Segmentation results on a subset of 30 images of the Berkeley dataset, using both R-index and PR-index.

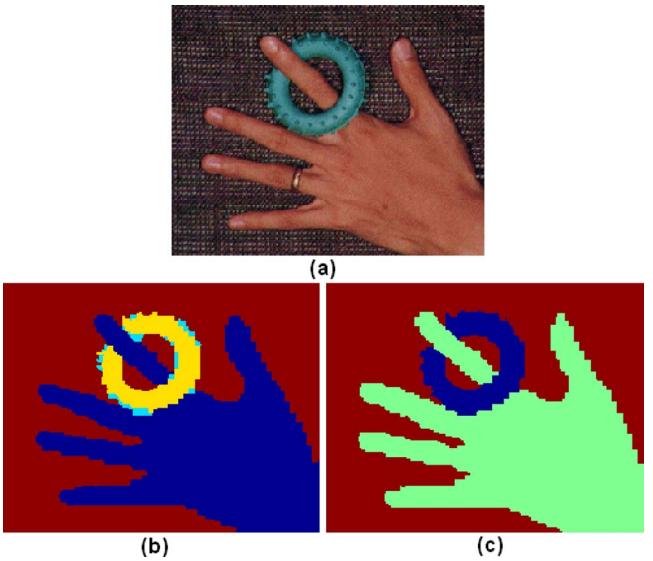

Fig. 11 There are three main classes of the RGB image in (a). However, the edge-clusters create an erroneous fourth class (around the ring) in (b). Using the validity index, the image is properly segmented in (c).

and a predefined class parameter for FCM of four classes in the reduced space of $m=2$ dimensions. One way to get rid of the erroneously produced edge clusters is to reveal the exact number of classes that are present in the preceding image. Although this is generally a difficult task, in this case where the real classes of the image (with high cardinality) are well separated from the artificial edge clusters, it is considered a rather easy proposition. The fuzzy validity index V_{WSJ} is invoked,[23] which has been found suitable when classes tend to overlap each other. The number of classes and the final partition in the reduced space is computed by using the combination of FCM plus the validity index algorithm, presented in Fig. 11(c). The validity index results in a considerable optimization of the algorithm, defining the number of classes properly.

6 Conclusions and Future Works

A novel feature-based segmentation scheme for color images is introduced in this study, with the main advantage being the avoidance of the usual complications related to the tuning of several involved parameters. The presented work explores the idea of using Laplacian eigenmaps, a manifold learning algorithm, to aid the color image segmentation problem. The produced segmentation maps reveal the existence of manifolds in an image, the structural information of which is taken into account in order to improve the resulting segmentation. The basic characteristics of our technique can be summarized as follows:

- It can work with numerous features and is exceptionally flexible regarding the type of extracted features. The flexibility of the approach is justified by the fact that different features can be used instead of the RGB pixels (like DCT). As long as we can construct a graph based on the extracted blocks, regardless of the features used to calculate the similarity between them, the algorithm can be applied unaltered.
- It is computationally efficient since the resulting Laplacian is very sparse. Furthermore, the computation of only a few eigenvectors (corresponding to the smallest eigenvalues) in Eq. (3) is necessary.
- The LE-algorithm receives blockwise inputs instead of pixelwise inputs that are utilized in other similar approaches, due to the "clever" feature extraction scheme.

It should be mentioned, that our method bears a similarity with the well-known technique of Normalized Cuts.[19] Both are graph-theoretic approaches and based on spectral graph theory. The great difference lies in the generic character of our method, which emphasizes the possibility of working in high-dimensional spaces.

Our current research efforts revolve around the improvement of the method toward a fully unsupervised segmentation algorithm. This requires estimating[35] the manifold dimensionality instead of manually selecting it. Additional improvement is expected by incorporating spatial information in the feature extraction step, so that similar feature vectors nearby in the image plane are more prone to be grouped together than similar feature vectors from different spatial locations. Furthermore, experimentation with different forms of information representation—for instance, to work with perceptual color spaces (e.g., HSV or CIE L*a*b*)—are on the agenda. Last, alternative ways to cluster the projected data need to be explored, since the images of points belonging to the same cluster were experimentally observed to be distributed angularly around straight lines (e.g., Figs. 4 and 5), while classical algorithms like FCM assume approximately round clusters. To conclude, although several factors have to be studied in depth, the preceding idea seems to work well and improves the clustering-based segmentation of color images.

References

1. P. K. Sagoo, S. Soltani, A. K. C. Wong, and Y. C. Chen, "A survey of thresholding techniques," *Comput. Vis. Graph. Image Process.*, **41**(2), 233–260 (1988).
2. N. Papamarkos, C. Strouthopoulos, and L. Andreadis, "Multithresholding of color and gray-level images through a neural network technique," *Image Vis. Comput.* **18**(3), 213–222 (2000).
3. S. Makrogiannis, G. Economou, and S. Fotopoulos, "A region dissimilarity relation that combines feature-space and spatial information for color image segmentation," *IEEE Trans. Syst., Man, Cybern., Part B: Cybern.* **35**(1), 44–53 (2005).
4. R. Xu and D. I. I. Wunsch, "Survey of clustering algorithms," *IEEE Trans. Neural Netw.* **16**(3), 645–678 (2005).
5. A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM Comput. Surv.* **31**(3), 264–323 (1999).
6. M. Filippone, F. Camastra, F. Masuli, and S. Rovetta, "A survey of kernel and spectral methods for clustering," *Pattern Recogn.* **41**(1), 176–190 (2008).
7. S. Chen and D. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," *IEEE Trans. Syst., Man, Cybern., Part B: Cybern.* **34**(4), 1907–1916 (2004).
8. T. Kanungo, D. M. Mount, N. Netanyahu, C. Piatko, R. Silverman, andA. Y. Wu, "An efficient k-means clustering algorithm: analysis and implementation," *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 881–892 (2002).
9. C. Chinrungrueng and C. H. Sequin, "Optimal adaptive k-means algorithm with dynamic adjustment of learning rate," *IEEE Trans. Neural Netw.* **1**(6), 157–169 (1995).
10. C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: image segmentation using expectation-maximization and its application to image querying," *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(8), 1026–1038 (2002).
11. J. Lazaro, J. Arias, J. L. Martin, A. Zuloaga, and C. Cuadrado, "SOM segmentation of grayscale images for optical recognition," *Pattern Recogn. Lett.* **27**(16), 1991–1997 (2006).

12. J. Costa and A. O. Hero, "Geodesic entropic graphs for dimension and entropy estimation in manifold learning," *IEEE Trans. Signal Process.* **52**(8), 2210–2221 (2004).
13. H. S. Jain and D. D. Lee, "The manifold ways of perception," *Science* **290**(5500), 2268–2269 (2000).
14. J. Tenenbaum, V. de Silva, and J. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science* **290**(5500), 2319–2322 (2000).
15. V. de Silva and J. Tenenbaum, "Global versus local methods in nonlinear dimensionality reduction," *Adv. Neural Inf. Process. Syst.*, **15**, 721–728 (2003).
16. S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science* **290**(5500), 2323–2326 (2000).
17. I. Tziakos, N. A. Laskaris, and S. Fotopoulos, "Multivariate image segmentation using Laplacian eigenmaps," in *Proc. European Signal Processing Conference (EUSIPCO)*, pp. 945–948 (2004).
18. M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Comput.* **15**, 1373–1396 (2003).
19. J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(8), 888–905 (2000).
20. A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: analysis and an algorithm," *Adv. Neural Inf. Process. Syst.*, **14**, 849–856 (2002).
21. C. Fowlkes, S. Belongie, F. Chung, and J. Malik, "Spectral grouping using the Nystrom method," *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(2), 214–225 (2004).
22. J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, pp. 43–93, Plenum Press, New York (1981).
23. H. Suna, S. Wanga, and Q. Jiang, "FCM-based model selection algorithms for determining the number of clusters," *Pattern Recogn.* **37**(10), 2027–2037 (2004).
24. M. Brand, "Charting a manifold," *Adv. Neural Inf. Process. Syst.*, **15**, 961–968 (2003).
25. I. T. Jolliffe, *Principal Component Analysis*, Springer-Verlag, New York (1986).
26. N. A. Laskaris and S. P. Zafeiriou, "Beyond FCM: graph-theoretic post-processing algorithms for learning and representing the data structure," *Pattern Recogn.* **41**(8), 2630–2644 (2008).
27. W. Wang and Y. Zhang, "On fuzzy cluster validity indices," *Fuzzy Sets Syst.* **158**(19), 2095–2117 (2007).
28. B. Nadler, S. Lafon, R. R. Coifman, and I. G. Kervekidis, "Diffusion maps, spectral clustering, and eigenfunctions of Fokker-Planck operators," *Adv. Neural Inf. Process. Syst.*, **18**, 955–962 (2006).
29. D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. Int. Conf. Computer Vision*, vol. 2, pp. 416–423 (2001).
30. S. Jaiyen and C. Lursinsap, "Non-Euclidean self-organizing classification using natural manifold distance," in *Proc. IEEE Conf. Neural Networks*, vol. 1, pp. 799–802 (2004).
31. N. Vasconcelos and A. Lippman, "A multiresolution manifold distance for invariant image similarity," *IEEE Trans. Multimedia* **7**(1), 127–142 (2005).
32. J. B. Tenenbaum, "Mapping a manifold of perceptual observations," *Adv. Neural Inf. Process. Syst.*, **10**, 682–687 (1998).
33. S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Wu, "An optimal algorithm for approximate nearest neighbor searching," *J. ACM* **45**(6), 891–923 (1998).
34. R. Unnikrishnan, C. Pantofaru, and M. Hebert, "Toward objective evaluation of image segmentation algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(6), 929–944 (2007).
35. B. Kegl, "Intrinsic dimension estimation using packing numbers," *Adv. Neural Inf. Process. Syst.*, **15**, 681–688 (2003).

Ioannis Tziakos received a BSc degree in physics in 2002 and an MSc degree in electronics and computer science in 2005 from the Physics Department, University of Patras (UoP), Greece. He is currently pursuing his PhD degree as a research student at the School of Electronic Engineering and Computer Science of Queen Mary University of London, United Kingdom. His main research interests include dimensionality reduction and applications, graph theory, and image and video segmentation.

Christos Theoharatos received a BSc degree in physics in 1998, an MSc degree in electronics and Computer science in 2001, and a PhD degree in image processing and multimedia retrieval in 2006, all from the University of Patras (UoP), Greece. He has actively participated in several national research projects and is currently working as a Postdoctoral researcher at the Electronics Laboratory (ELLAB), Electronics and Computer Division, Department of Physics, UoP. Since 2002, he has been working as a tutor at the degree of lecturer in the Department of Electrical Engineering, Technological Institute of Patras. His main research interests include pattern recognition, multimedia databases, image processing and computer vision, data mining, and graph theory.

Nikolaos A. Laskaris received both an MSc degree in medical physics (1995) and a PhD degree in biomedical signal processing (1998) from Patras University, Greece. He joined the Laboratory for Human Brain Dynamics in the Brain Science Institute (BSI) of RIKEN in Japan for four years (1999 to 2003), where he worked mostly as a neuroengineer. He received a grant from the Greek General Secretariat of Research and Technology (ENTER Program) for working as an associate researcher at the Electronics Laboratory of the Physics Department, Patras University (2003 to 2004), where he was involved in the development of neuromorphic signal/image processing techniques. Currently, he is a faculty member in the Informatics Department of Aristotle University of Thessaloniki. His current research interests include computational intelligence, machine vision, soft computing, data mining, and nonlinear dynamics and their applications in biomedicine and neuroscience.

George Economou received a BSc degree in physics from the University of Patras. (UoP), Greece, in 1976; an MSc degree in microwaves and modern optics from University College, London, in 1978; and a PhD degree in fiber optic sensor systems from the University of Patras in 1989. He is currently an associate professor at the Electronics Laboratory (ELLAB), Department of Physics, UoP, where he teaches at both undergraduate and postgraduate level. He has published papers on nonlinear signal and image processing, fuzzy image processing, multimedia databases, data mining, and fiber-optic sensors. He has also served as referee for many journals, conferences, and workshops. His main research interests include signal and image processing, computer vision, pattern recognition, and optical signal processing.