MULTIVARIATE IMAGE SEGMENTATION USING LAPLACIAN EIGENMAPS

Ioannis Tziakos, Nikolaos Laskaris and Spiros Fotopoulos

Electronics laboratory, Department of Physics, University of Patras
Patras 26500, Greece
phone: +302610 997287, fax: +302610 997456, email: spiros@physics.upatras.gr
web: ellab.physics.upatras.gr

ABSTRACT

We are exploring the novel technique of *Laplacian Eigenmaps (LE)* [1] as a means of improving the clustering-based segmentation of multivariate images. A computationally efficient scheme, taking advantage of the ability of LE-algorithm to learn the actual *manifold* of the multivariate data, is introduced. After embedding the local image characteristics in a high-dimensional feature space, the skeleton of the intrinsically low dimensional manifold is reconstructed. A low-dimensional map, in which the variations in the local image characteristics are presented in the context of global image variation, is then computed. The non-linear projections on this map serve as inputs to the *fuzzy c-means* algorithm boosting its clustering performance significantly. The final segmentation is produced by a simple labelling scheme that works pixelwise. The experimental results using RGB-images were very promising and showed that robustness to noise and generic character are the main advantages of our method.

1. INTRODUCTION

Image segmentation is a key process in vision. All animals, including humans, use segmentation to identify objects and places. With this information they can understand, move and interact with their environment. Although living organisms can easily segment an image into coherent objects, computer systems need to devote a lot of computational power to approximate the same task.

In the machine vision literature, numerous algorithms have appeared. Usually, each one algorithm is addressing 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 or multi-thresholding schemes [2],[3]. Other approaches are characterized as region-based ones [4], and appear in the form of region-growing or region split-and-merge techniques. The list of novel or hybrid methods that use every possible combination of low-level, mid or high-level futures of image goes on, since segmentation is an extremely active area of research [e.g.5,6]. Nowadays, with the advents in remote sensing technology and the extensive use of medical imaging modalities giving rise to vector fields, techniques for multivariate image segmentation are a prerequisite for image understanding. Despite the continuous efforts, a generally applicable method is still missing. The family of clustering-based techniques is perhaps the most popular one due to their flexible character. In these techniques, using a proper feature extraction step, pixels are mapped in a feature space (usually) a partitional-clustering algorithm is employed to identify distinct groups. The selection of features, the 'curse of dimensionality' and the definition of number of clusters are the most important problems the user has to deal with. While the last one can be overcome via simple experimentation, there is no straightforward solution for the first two, since they are counteracting. The 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 apparatus and the current machine-vision segmentation algorithms. It is inherent in the biological systems the ability to recover the low-dimensional structure when confronted with stimuli lying in high dimensional spaces [7]. Motivated by this fact, a few techniques have recently appeared in the computational literature. Trying to imitate the perceptual *manifold leaning* from high-dimensional data, e.g. [8].

It was among the main objectives of this work to try to incorporate such a learning technique for the benefit of image segmentation. To provide the segmentation scheme with a universal character and in order to alleviate the problem of feature-selection, we propose an initial embedding of local image characteristics in a high-dimensional feature space, where LE-algorithm acts in order to describe, parameterise and visualize the learned data-manifold. The nonlinear dimensionality reduction technique produces a sketch of the manifold in a reduced feature space. Based on this sketch the data manifold can be partitioned efficiently and, consequently, the image regions can be classified accordingly.

By releasing the condition for a 'good' 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 will be discovered by the LE-technique and used for segmenting the image plane.

The LE-technique is described in section 2, where the exact algorithmic-steps (as incorporated in this study) have been also included. Our segmentation technique is presented in section 3, while the results from the application to RGB-images have been included in section 4. Finally, a short discussion on the methodology and the future implications is provided in section 5.

2. THE LAPLACIAN EIGENMAP METHODOLOGY

The LE-technique is an application of Spectral Graph Theory. Given a set of N multivariate observations embedded as vectors $\boldsymbol{x}_1, \boldsymbol{x}_2,...,\boldsymbol{x}_N$ in R^l (l>>1), a weighted graph G is built over the endpoints of these vectors. It consists of N nodes, one for each point and a set of edges connecting neighbouring points. Consider the problem of mapping the weighted graph G to a line so that connected points stay as close as possible. If two points are close enough, then there is an edge between them. Let $\boldsymbol{y}=\{y_1,y_2,...y_N\}^T$ be such a map. A reasonable criterion for choosing a "good" map is to minimize the following objective function,

$$\sum_{ij}\left(y_i - y_j\right)^2 \cdot W_{ij}$$

where W is the weight matrix defined as follows:

$$W_{ij} = \left\{ \begin{matrix} e^{-\frac{\|x_i - x_j\|^2}{t}} & if \quad i,j \quad are \quad connected \\ 0 & otherwise \end{matrix} \right\}$$

The objective function with our choice of weights W_{ij} incurs a heavy penalty if neighbouring points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are mapped far apart. Therefore, **minimising** it is an attempt to ensure that if $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are close then y_i and y_j are close as well. It turns out that the minimization problem reduces to finding,

$$\arg\min_{y^T D y = 1} y^T L y$$

where $L=D-W$ is the *Laplacian matrix*. D is a diagonal weight matrix such that its entries are column (or row, since W is symmetric) sums of W, $D_{ij} = \sum_j W_{ij}$ .Laplacian is symmetric, positive semidefinite matrix which can be thought of as an operator on functions defines on vertices of G.

The algorithmic procedure is formally stated below:

- [Constructing the adjacency graph]. Nodes i and j are connected by an edge if $\left\|x_i - x_j\right\|^2 \le \varepsilon$ where the norm is the usual Euclidean norm in $\Re^l$ and $\varepsilon \in \Re$.

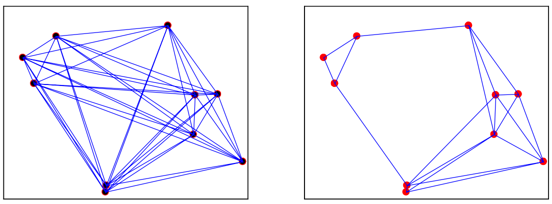

Figure 1: An example graph in R^2, before and after the operation for defining the neighbouring-relationships.

- [Defining the weights] Select the heat-kernel parameter $t \in \Re$.
- [Eigenmap] Assume the graph G, constructed above, is connected; otherwise proceed with this step for each connected component. Compute eigenvalues and eigenvectors for the generalized eigenvector problem:

$$Ly = \lambda Dy \quad (1)$$

Let $\mathbf{y}_0,...,\boldsymbol{y}_{k-1}$ be the solutions of eq.(1), ordered according to their eigenvalues. We leave out the eigenvector $\boldsymbol{y}_0$ corresponding to eigenvalue 0 and use the next m eigenvectors for embedding in m-dimensional Euclidean space.

$$x_i \rightarrow (y_1(i),...,y_m(i))$$

where $\boldsymbol{y}_j(i)$ is the i^{th} element of eigenvector $\boldsymbol{y}_j$.

3. THE SEGMENTATION TECHNIQUE

Without sacrificing generality, and since via the feature extraction step we aimed at representing the local image characteristics, we adopted the following simple step. Using a regular grid, the multivariate image was first partitioned into non-overlapping square patches. Then, for each patch a representing feature-vector was formed by concatenating the vectorial measurements corresponding to the pixels included in the specific patch. In this way, a correspondence between specific image-plane locations (the centre pixel of each patch) and feature space locations (the endpoint of each feature vector) is established. This apparent spatial undersampling was motivated by computational economy, recovered by the final labelling procedure and fully justified by the obtained results.

In summary, our segmentation technique includes the following steps:

- A given p-variate image is first divided in to n x n blocks, where n is an odd number (see an example in Fig. 2*)*.
- Blocks are reshaped into vectors of $l=pn^2$ dimensions (representing them in a high-dimensional feature space).
- We construct the adjacency graph using the ε-neighbourhoods variant of the LE-algorithm. As radius ε, the minimum value that leaves the graph connected is used. The graph is weighted using as heat kernel parameter t the 70% of the maximum distance in the adjacency graph (gives in general better results).
- A map of the blocks, in m<<l dimensions, is produced (i.e. in reduced feature space). Usually m is set to 2 or 3 enabling visualization of the clustering tendencies (see Fig.3.) and therefore facilitating the definition of the number of clusters.
- *Fuzzy C-means* is applied to the 'projections' of the blocks in the reduced space. In this way, the blocks are partitioned into a user-defined number of classes.
- To infer the classification of all the individual pixels from the classification of the formed blocks, a simple label-assignment scheme was followed. The class-label of the corresponding block was assigned to each pixel at the central location. For each one of the rest of

the pixels, an equally-sized block is formed. By comparing this block with the spatially adjacent ones that have already classified in the previous algorithmic step, we identify the most similar one and assign the same label.

The algorithm requires three parameters: the block size [nxn], the reduced space dimension m, and the number of requested classes C. No other values are needed to be defined.

4. EXPERIMENTAL RESULTS

We extensively applied the proposed segmentation technique to RGB-images and compared its performance with respect to well known clustering-based techniques. Due to space limitations, we refrained from including a list of quantitative measures and decided to include some typical examples that will make clear the involved steps and enable the direct justification of our proposal.

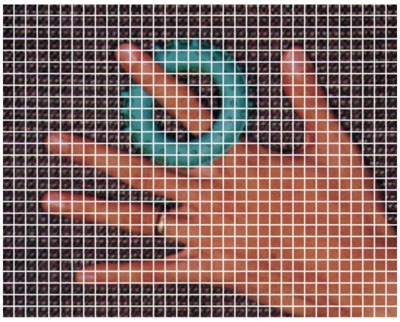

Figure 2: Test image divided in [9x9] blocks

We first discuss the performance of the new technique when applied to the color-image of Fig.2 that shows a hand and a hoop over a textured background (with the regular grid used for the definition of the blocks overplotted). The specific image seems to be a very "simple" one (i.e. two single objects are included). However, the combination of color-information (needed to differentiate between the two objects) and textural information (needed to distinguish the two objects form the homogenous background) makes this image a very intricate input to a segmentation routine. Apparently, clustering on a pixel-by-pixel basis is not an effective procedure.

On the contrary, our segmentation technique, with its first steps acting on blocks, encompasses easily the distinct local characteristics and incorporates both color-layout and textural-features in a straightforward manner. In Fig.3, we have included the map of the [9x9] blocks as produced by the LE-algorithm.

The coordinates of the depicted points in the 2-d reduced feature space, were used as inputs to the Fuzzy C-means algorithm which split the points into 3 groups (denoted using different colors). The labels of these points defined the labelling of the corresponding blocks, as can be seen in the included inset of Fig.3 where the blocks have been colored accordingly. The spatial resolution of this image is (9x9)-times lower than that of the original image, since these labels strictly correspond to the pixel at the centre of each block. With the last step of our technique, labels were assigned to the rest of the pixels. The final, properly-labelled image is shown in Fig.4, where the average color of each group has been used in order to denote to corresponding label.

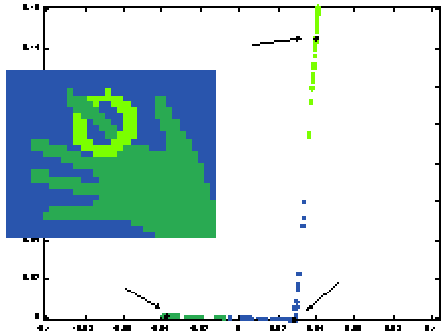

Figure 3: Clusters in reduced feature space and labels of the corresponding blocks.

Figure 4: Refined Segmentation

Since, the successfully resulted segmentation could have been, simply, due to the 'clever' feature extraction step (definition of square image patches), which incorporates at once a variety of local characteristics (i.e. color and texture), we tried to apply the same clustering algorithm directly in the high-dimensional feature space (in other words, to bypass the LE-step). Figure 5, depicts the produced low-resolution, labelled image and the corresponding refined segmentation.

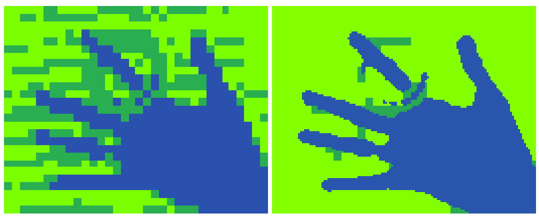

Figure 5:Original block labelling (left) and Refined segmentation (right) after the direct application of Fuzzy C-means.

The previous example makes clearly evident, that the manifold-learning achieved by the LE-algorithm (along with the subsequent dimensionality reduction) is what actually constitutes our segmentation technique efficient. Next, we discuss briefly the behaviour of the technique in the presence of noise

We are including a typical example in which the performance of our technique is tested when different types of noise is added to the original image. The image of a woman

(265x120 pixels), is divided into blocks of 5x5 pixels and the segmentation is performed (based on a 2-d map) by setting the number of clusters C to 3. The visual comparison (via Fig.6) with the results from noisy images gives us an indication about the robustness against noise. Without the noise we have a fairly good grouping of the background, the clothes and the skin. The noise categories in this test included gaussian, salt-and- pepper and speckle noise.

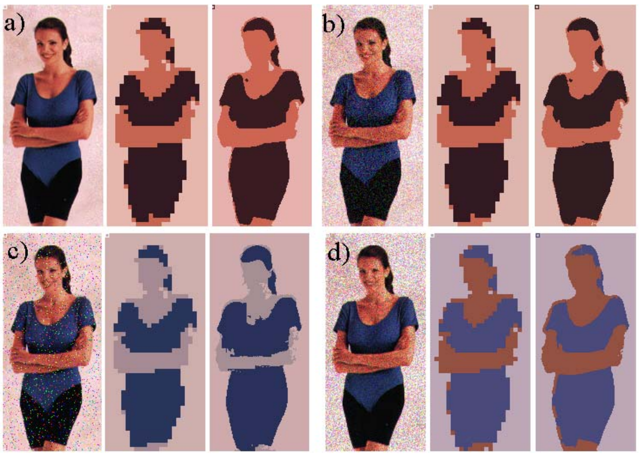

Figure 6:In each panel the input image, the initial block labelling and the refined segmentation appear from left to right. Panel a) corresponds to the original non-corrupted image, while b),c) and d) to the same image contaminated, correspondingly, with gaussian, salt-&-pepper, and speckle noise.

It is clear from the above figure that the introduced segmentation technique shows high resistance to these kinds of noise. It is inherent to the nonlinear character of the LE-algorithm the insensitivity to additive noise. As someone could expect, the added-noise has the form of random excursions from the 'true' manifold. The involved local-operations of the LE-algorithm act on the perturbed manifold as a restoration mechanism. Dimensions of random variation (noise) can easily be distinguished from dimensions of systematic variation (signal), as it is often considered in eigenalysis.

5. DISCUSSION

A novel feature-based segmentation scheme for multivariate images has been introduced, with the main advantage the avoidance of the usual complications related with the tuning of the involved parameters. The basic characteristics of our technique are:

- It can afford working with numerous features and is very flexible regarding the type of extracted features. Different feature selection procedures are expected to 'converge' in terms of results, since the 'true' underlying data-manifold is only one and an efficient learning algorithm is utilized for its discovery.
- It presents noise insensitivity, due to the nonlinear character of the LE-algorithm
- It is computational efficient since the resulting laplacian is very sparse. Furthermore the computation of only a few eigenvectors (corresponding to the smallest eigenvalues) in eq.(1) is necessary.

It should be mentioned, that our method bears a similarity with the well-known technique of Normalized-Cuts [6]. Both are graph-theoretic approaches and based on Spectral Graph Theory. The great difference lies in the generic character of our method, which emphasises -for the first time in the literature- the possibility of working in high-dimensional spaces.

Finally we need to stress that, we restricted (within here) the application only to RGB-images since these images enabled the direct visual judgment, by the reader, of the obtained results. The proposed technique is directly applicable to other types of multivariate data, like satellite or MRI images. With slight modifications the technique is expected to contribute to image understanding and that it might act as an intermediate tool for image analysis purposes like data fusion

Our current research efforts revolve around the improvement of the method 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 that similar feature vectors from different spatial locations. Among the feature objectives is the experimentation with different form of information representation, for instance to work with colour spaces like HSV, or the HMDD proposed in MPEG7.

ACKNOWLEDGEMENTS

This work was partially supported by a grant from the Greek GSRT (ENTER 2001).

REFERENCES

[1] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation", *Neural Computation,* col.15, pp.1373-1396, 2003.

[2] P. Sagoo, S. Soltani and A. Wong, "A survey of threholding techniques", *Computer Vision, Graphics, and Image Processing*, CVGIP-41, pp.233-260, 1988.

[3] N. Papamarkos, C. Strouthopoulos and L. Andreadis, "Multithresholding of color and gray-level images through a neural network technique", *Image and Vision Computing*, vol. 18, pp. 213-222, 2000.

[4] S. Makrogiannis, G. Economou, and S. Fotopoulos, "Region oriented compression of color images using fuzzy inference and fast merging", *Pattern Recognition,* Vol.35, pp. 1807-1820, 2002.

[5] A. Jain and Y. Zhong, "Page segmentation using texture analysis", *Pattern Recognition* , vol 5, pp. 169-184, 1992.

[6] J. Shi and J. Malik., "Normalized cuts and image segmewntation", IEEE Trans. PAMI, vol 22(8), pp.888-905, 2000.

[7] H. Seung Jain and D. Lee, "The manifold way of perception", *Science*, vol. 290, pp.2268-2269, 2000.

[8] J. Tenenbaum, V. de Silva, J. Langford,"A global geometric framework for nonlinear dimensionality reduction", *Science*, vol. 290, pp.2319-2223, 2000.