

A multiresolution diffused expectation–maximization algorithm for medical image segmentation

Giuseppe Boccignone^{a,*}, Paolo Napoletano^a, Vittorio Caggiano^b, Mario Ferraro^c

^aNatural Computation Lab, DIIE-Università di Salerno, via Ponte Don Melillo, 1, 84084 Fisciano (SA), Italy

^bDipartimento di Informatica e Sistemistica, Università di Napoli Federico II, via Claudio, 21, 80125 Napoli, Italy

^cDipartimento di Fisica Sperimentale, Università di Torino, via P. Giuria, 1, 10100 Torino, Italy

Received 3 October 2005; accepted 3 October 2005

Abstract

In this paper a new method for segmenting medical images is presented, the multiresolution diffused expectation-maximization (MDEM) algorithm. The algorithm operates within a multiscale framework, thus taking advantage of the fact that objects/regions to be segmented usually reside at different scales. At each scale segmentation is carried out via the expectation–maximization algorithm, coupled with anisotropic diffusion on classes, in order to account for the spatial dependencies among pixels. This new approach is validated via experiments on a variety of medical images and its performance is compared with more standard methods.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Image segmentation; Expectation–maximization; Multiresolution

1. Introduction

Computer algorithms for segmentation, the partitioning of an image into meaningful regions, play a crucial role in the delineation of anatomical structures of interest in several biomedical imaging applications, such as diagnosis, localization of pathology, study of anatomical structures, treatment planning and computer integrated surgery. The process of image segmentation assigns pixels to regions defined by labels. For instance, in segmenting skin lesions, one label may be assigned to pixels within the lesion region, another label to pixels outside such region; in magnetic-resonance (MR) images three labels may be used to distinguish gray-matter, white-matter and cerebrospinal-fluid tissue. Labeling is performed requiring certain regularity constraints to be satisfied.

More formally, consider an image F defined on a domain Ω , then the segmentation problem is to determine the set of K regions $\mathcal{R}_k \subset \Omega$, $k = 1, \dots, K$, satisfying an homogeneity predicate \mathcal{H} , such that: (1) $\bigcup_{k=1}^K \mathcal{R}_k = \Omega$, with $\mathcal{R}_k \cap \mathcal{R}_l \neq \emptyset$, $k \neq l$; (2) $\mathcal{H}(\mathcal{R}_k) = \text{true}$, $\forall k$; (3) $\mathcal{H}(\mathcal{R}_k \cup \mathcal{R}_l) = \text{false}$, $\forall \mathcal{R}_k, \mathcal{R}_l$

adjacent. The first condition states that the partition has to cover the whole image and that any given pixel cannot belong to two regions; the second indicates that each region has to be uniform with respect to the predicate \mathcal{H} ; the last condition prevents two adjacent regions from being merged into a single region that satisfies \mathcal{H} .

A wide array of techniques, both for gray-level and color images, has been used in the past (for in-depth surveys, see [1–3]) either exploiting image-domain or feature-space based approaches but so far there is no satisfactory solution to image segmentation.

Image-domain based techniques try to account for feature-space homogeneity while ensuring spatial compactness, for instance by progressively growing image regions or either by subdividing and merging the regions according to a feature-based predicate \mathcal{H} (e.g. color similarity) [4]. Classical region growing, however, is not completely automated, since initial seed points must be given from which regions are grown. Split-and-merge techniques do not require seed points [5], but they may exhibit over-segmentation, with the occurrence of many small, disconnected regions; in order to promote significant regions, multiresolution schemes can be applied. Multiresolution provides an interesting strategy to carry out the refinement of segmentation [6,7], by operating on the image at different

* Corresponding author. Tel.: +39089964275; fax: +39089964218.

E-mail addresses: boccig@unisa.it (G. Boccignone), pnapoletano@unisa.it (P. Napoletano), vcaggian@unina.it (V. Caggiano), ferraro@ph.unito.it (M. Ferraro).

scales/resolutions, using for instance a quadtree, wavelets [7] or a pyramid structure [8]. Segmentation can also be obtained by detecting edges among regions [1], however edges cannot be used directly to define a region and further processing (edge linking, grouping etc.) is then required. In general such classical techniques have the critical limitation of using local criteria (pixel based). The homogeneity criterion has the function of deciding whether a pixel belongs to the growing region or not and the decision of merging is generally taken based only on the contrast between the evaluated pixel and the region. However, it is not easy to decide when this difference is small (or large) enough to take a decision. Limitations of purely local methods has led to the use of Markov random fields (MRF) models (in particular Gibbs MRF), which combine local spatial interactions, also incorporating edge information, with a global cost function (e.g., energy) [9]. Difficulties associated with MRFs are the proper selection of the parameters controlling the strength of spatial interactions, and that they require computationally intensive algorithms [10]. Deformable models [11] (e.g., active contours), partial differential equations (e.g., anisotropic diffusion [12]), and in general variational methods relying upon the minimization of some global energy function, either related to image discontinuities (edges) or to region uniformity, can be seen as deterministic counterparts of MRFs [13,14]. In general, active contours are sensitive to initial conditions and are really effective only when the initial position of the contour in the image is sufficiently close to the real boundary.

Feature-space based techniques work out in some feature space (e.g., color or texture) and solve the segmentation problem by finding clusters (regions) in such space or by finding the peaks of histograms (empirical distributions of the features); for instance, histogram thresholding is among the most popular technique [15]. The main limitations of this technique are that often, because of the noise, the profiles of the histograms are rather jagged giving rise to spurious peaks, and that it does not take into account the spatial characteristics of an image; furthermore the extension to color or multivalued images is not straightforward. Some of these problems can be solved by embedding connectivity information [16] or by resorting again to multiresolution representations, e.g., wavelets [17,18]. Clustering methods such as K-means, fuzzy c-means can be seen as unsupervised classifier methods [19], pattern recognition techniques that seek to partition (cluster) a feature space derived from the image [1]. Of particular interest for the work presented here, is that such clustering algorithms can be given a generalized form [20] known in the literature as the expectation–maximization (EM) algorithm, [21]. The EM algorithm assumes that pixel intensities are independent samples drawn from a linear combination of probability distributions, usually Gaussians [22] (finite mixture model, see Section 2 for details). In the E-step expectation–maximization computes the posterior probabilities that a pixel belongs to a certain class, whereas in the M-step computes maximum likelihood estimates of the parameters characterizing the distributions (that for Gaussian distributions, correspond to means and covariances), and mixing coefficients. Assignment to a class of

a pixel is eventually provided according to maximum posterior probability with respect to the given classes. Note that like thresholding and clustering algorithms, EM does not directly incorporate spatial modeling and can therefore be sensitive to noise and intensity inhomogeneities. To overcome such drawbacks several methods have tried to incorporate MRFs or, in general, a prior term in order to maximize a log-posterior probability instead of log-likelihood, thus leading to quite complex EM steps (see, for a discussion [23,24]). Recently, a diffused expectation–maximization (DEM) algorithm has been proposed for gray-level images [25], in which a diffusion step provides spatial constraint satisfaction.

Here we presents a twofold extension of the diffused expectation–maximization approach: the algorithm will be extended to deal with vector (i.e. color) images and furthermore it will operate in a multiresolution framework to take advantage of the fact that objects/regions to be segmented usually reside at different scales.

2. Method

2.1. Theoretical background

Let the plane of the image be formed by N pixels, and let f_i be the gray-level associated to the i th pixel; gray-levels are supposed to be drawn from a finite set and range from 0 to $M - 1$, so that there are M possible values. More formally, any gray-level image can be considered a set of N unlabeled samples $F = \{f_1, f_2, \dots, f_N\}$, on a 2D discrete support $\Omega \subseteq Z^2$, with $N = |\Omega|$.

The image segmentation problem can be defined in probabilistic terms as the problem of assigning a label k to each pixel i , given the observed data F , and where each label $k \in [1, \dots, K]$ defines a particular region/model. Different models are selected with probability $P(k)$, and a sample is generated with probability distribution $p(f_i|k, \theta)$ where $\theta = \{\theta_k, k = 1, \dots, K\}$ and θ_k is the vector of the parameters associated to label k . Thus $p(f_i|k, \theta)$ is the probability of f_i given the parameters of all models and the fact that we have selected model (label) k . Each image can be conceived as drawn from a mixture density, so that, for any pixel,

$$p(f_i|\theta) = \sum_{k=1}^K P(k)p(f_i|k, \theta); \quad (1)$$

terms $P(k)$ represent the mixing coefficients and it is usual to adopt the notation $\alpha_k \equiv P(k)$. The likelihood of the data is

$$\mathcal{L} = p(F|\theta) = \prod_{i=1}^N p(f_i|\theta). \quad (2)$$

For clarity's sake, we define two probability distributions $p(f_i)$ and $\pi(f)$; the former is the probability that, given a gray-level f it is assigned to pixel i , so that $\sum_i p(f_i) = 1$, whereas the latter is the probability that, given any pixel i it has gray-level f . Then $p(f_i)$ is a spatial distribution of the gray-levels across the image, whereas $\pi(f)$ is the probability to find a given gray-level f ,

irrespective of the position in the image. Image segmentation can be achieved by finding the set of labels that maximizes the likelihood $\mathcal{L} = \prod_{i=1}^N \sum_{k=1}^K P(k)p(f_i|k, \theta)$, or, equivalently, by using Eqs. (1) and (2)

$$\frac{1}{N} \log \mathcal{L} = \frac{1}{N} \sum_{i=1}^N \log \sum_{k=1}^K p(f_i|k, \theta)P(k). \quad (3)$$

By the weak law of large numbers and the ergodic theorem the right-hand side of Eq. (3) can be written as the expectation with respect to gray-level f

$$E \left[\log \sum_{k=1}^K p(f_i|k, \theta)P(k) \right] = \sum_{f=0}^{M-1} \left(\pi(f) \log \sum_{k=1}^K p(f_i|k, \theta)P(k) \right). \quad (4)$$

Simple manipulations lead to

$$\frac{1}{N} \log \mathcal{L} = \sum_{f=0}^{M-1} \pi(f) \log \pi(f) - \sum_{f=0}^{M-1} \left(\pi(f) \log \frac{\pi(f)}{\sum_{k=1}^K p(f_i|k, \theta)P(k)} \right). \quad (5)$$

Hence a straightforward maximization of $\log \mathcal{L}$ can be obtained by minimizing the second term of the last expression, namely the Kullback–Leibler (KL) divergence $D(\pi(f) \parallel \sum_{k=1}^K p(f_i|k, \theta)P(k))$ between distributions $\pi(f)$ and $\sum_{k=1}^K p(f_i|k, \theta)P(k)$, while holding the first term fixed. This is exactly what is performed by the classic EM algorithm [21] which minimizes the KL divergence between the manifold of the observed data and that of the true distribution.

In particular, when Gaussian distributions are used to model the density mixture of Eq. (1), the E- and M-steps can easily be computed in closed form. Define $h_{ik} = p(k|f_i, \theta)$; in the E-step these posterior probabilities are given by

$$h_{ik}^{(t)} = \frac{\alpha_k^{(t)} p(f_i|k, \mu_k^{(t)}, \sigma_k^{(t)})}{\sum_k \alpha_k^{(t)} p(f_i|k, \mu_k^{(t)}, \sigma_k^{(t)})} \quad (6)$$

while in the M-step, with $h_{ik}^{(t)}$ fixed, the parameters θ and mixing proportions $\alpha_k = P(k)$ that maximize $\log \mathcal{L}$, are obtained as

$$\alpha_k^{(t+1)} = \frac{1}{N} \sum_i h_{ik}^{(t)}, \quad \mu_k^{(t+1)} = \frac{\sum_i h_{ik}^{(t)} f_i}{\sum_i h_{ik}^{(t)}}, \quad \sigma_k^{(t+1)} = \frac{\sum_i h_{ik}^{(t)} [f_i - \mu_k^{(t+1)}]^2}{\sum_i \tilde{h}_{ik}^{(t)}}, \quad (7)$$

The E- and M-steps are iterated until $|\log \mathcal{L}^{(t+1)} - \log \mathcal{L}^{(t)}| < \varepsilon$.

Alternatively, one could attempt a multistep approach by iteratively minimizing the entropy $H(f) = -\sum_{f=0}^{M-1} \pi(f) \log \pi(f)$, while holding $p(F|k, \theta)$, $P(k)$ fixed, and then minimizing the divergence D , while keeping $H(f)$ fixed. Interestingly, from the segmentation standpoint, the problem can be reformulated in

a way which takes into account the spatial correlations among pixels, as follows.

Empirical approximations of probabilities $p(f_i)$ and $\pi(f)$ are given by

$$p_i = p(f_i) \simeq \frac{f_i}{f_{\text{tot}}} \quad \text{and} \quad f_{\text{tot}} = \sum_{i=1}^N f_i, \quad (8)$$

and

$$\pi(f) \simeq \frac{n_f}{N}, \quad (9)$$

and the relation $f_{\text{tot}} = \sum_{i=1}^N f_i = \sum_{f=0}^{M-1} n_f f$ holds. From these probabilities two different entropies can be derived, namely H and H_s , that capture different types of uncertainty related to the stochastic process of which the image is a sample; from $\pi(f)$ the entropy H can be defined as

$$H(f) = - \sum_{f=0}^{M-1} \pi(f) \ln \pi(f), \quad (10)$$

whereas a spatial measure of entropy is

$$H_s = - \sum_{i=1}^N p(f_i) \ln p(f_i) = E \left[\ln \frac{1}{p(f_i)} \right]. \quad (11)$$

By making use of (8) and (9) it is not difficult to prove that H_s increases when H decreases. Intuitively, when the image is uniform (all pixels have the same gray-level) H_s will be maximum, while H will be 0 since there is only one gray-level f with probability $p(f) = 1$. Thus a minimization of H corresponds to a maximization of H_s .

The main idea behind the DEM approach is that maximization should be attained in a way label assignment to a pixel depends on the labels in the pixel neighborhood; then, a process must be devised that takes into account spatial correlations. It has been proved [26] that $-H_s = \sum_{i=1}^N p(f_i) \log p(f_i)$ is a Liapunov functional decreasing under isotropic diffusion; however this result as such does not allow to select the optimal label. Note that neighboring pixels should have the same probability to be assigned a given label k and that labels at boundaries between regions should be characterized by an abrupt change of probability values. Thus, each $h_{ik} = p(k|f_i, \theta)$ field should be a piecewise constant function across the image and this result can be achieved [26] by a system of k anisotropic diffusion equations

$$\frac{\partial h_{ik}(t)}{\partial t} = \nabla \cdot (g(\nabla h_{ik}) \nabla h_{ik}(t)) \quad (12)$$

one for each label probability plane; $g(\cdot)$ is a suitable conductance function, monotonically decreasing. Hence small differences of h_{ik} among pixels close to each other are smoothed out, since diffusion is allowed, whereas large variations are preserved. As in the isotropic case, anisotropic diffusion is proved to increase the spatial entropy H_s [26]. The DEM algorithm obtains the maximization of $\log \mathcal{L}$ by iteratively

computing $p(k|f, \theta)$, $p(F|k, \theta)$, $P(k)$ along the expectation and maximization steps via Eqs. (6), (7) and interleaving this computation with diffusion on $p(k|f, \theta)$, through Eq. (12) which in practice regularizes each k labeling field by propagating anisotropically such labels. Eventually, the segmentation is performed by using the estimated parameters k , θ_k .

2.2. Extension to multi-valued images

A vector-valued (multi-valued) image can be considered a set of N unlabeled samples $\mathbf{F} = \{f_1, f_2, \dots, f_N\}$, on a 2D discrete support $\Omega \subseteq Z^2$, with $N = |\Omega|$, where each sample f_i is a D -dimensional vector $[f_i^1, f_i^2, \dots, f_i^D]^T$; in other terms, \mathbf{F} is a set of single-valued images $\{F^1, F^2, \dots, F^D\}$, sharing the same domain. A color image can be considered a vector-valued image, each component corresponding to one of the three color channels; for instance an RGB image, can be represented as $\{F^R, F^G, F^B\}$.

A color space is a geometrical and mathematical representation of color; there is a variety of such representations either derived from hardware considerations (e.g., RGB, YCrCb, NTSC, YIQ, CMYK, etc.), or colorimetry issues (e.g., XYZ, UCS, CIELAB, CIELUV), or visual perception motivations (Opponent colors, IHS, HSV, etc.). A survey can be found in [27,28]. We will use the YCrCb color space; this choice is motivated by the fact that YCrCb de-correlates the original tristimulus color components, thus granting independence of the color channels. The mapping $\{F^R, F^G, F^B\} \mapsto \{F^Y, F^{Cr}, F^{Cb}\}$ from the RGB space to YCrCb space is accomplished through linear transformation [28].

2.3. Multiresolution representation of images via pyramids

It has been observed previously that the DEM approach takes into account the spatial interactions among pixels in the image; it should be noted, however, that different types of regions are, in general, characterized by different interactions lengths, e.g. points in almost uniform regions have longer interaction range than those in weak texture, and, moreover such range depends on the scale on which the analysis is performed [29]. Then the problem arises on which is the optimal scale at which segmentation must be carried out: a low resolution might result in a too coarse segmentation, with different regions merged together, whereas a high resolution, on the contrary could result in many fragmented images and involves a heavy computational load.

In particular, when dealing with very complex images, such as vector-valued images, clearly one should be concerned in obtaining reliable segmentation results, while keeping acceptable computational costs. Both issues can be accounted for by resorting to a multiresolution representation of the original image, e.g., a pyramidal representation [8]. On the one hand, propagation of information gained from coarse resolution levels makes significant objects/regions in the image more relevant respect to weak textures and noise. On the other hand, pyramids provide an efficient tool to reduce the computational load both as regards the iterations necessary to maximize the likelihood—by initializing the parameter set on the basis of the

parameters estimated at the coarser level—and for what concerns the diffusion step, which, at each iteration needs only to work upon a sub-sampled version the probability maps.

A multiresolution representation can be derived from the original color image, by means of a Gaussian pyramid [8] obtained by performing a low-pass filtering via convolution with a gaussian $G(\sigma_l)$ [8], l being the level of resolution, followed by a sub-sampling of the smoothed scalar field, here the color channel F^d , where the index d denotes the color component [30]:

$$f_i^{d,(l+1)} = S \downarrow G(\sigma_l) * f_i^{d,(l)}, \quad (13)$$

where $S \downarrow$ is the down-sampling operator. Note that the transition from scale l to scale $l+1$ produces a coarser representation of the image, and hence $f_i^{d,(0)} = f_i^d$, is the highest resolution level and corresponds to the original image.

We will denote by $\mathcal{P}\{F^d\}$ the Gaussian pyramid build on the scalar field, F^d , while $\mathcal{P}\{\mathbf{F}\} = \{\mathcal{P}\{F^1\}, \dots, \mathcal{P}\{F^D\}\}$, is the Gaussian pyramid representation of the vector-valued image. The level l of the pyramid $\mathcal{P}\{\mathbf{F}\}$ will be indexed by $\mathcal{P}^{(l)}\{\mathbf{F}\}$, or, for notational simplicity, by $\mathbf{F}^{(l)}$.

Fig. 1 provides an example of the pyramidal decomposition of a computed tomography (CT) scan image rendered as a gray-level image. The pyramidal representation of a vector-valued image is shown in Fig. 2 by using a skin lesion image; in particular, for visualization purposes, the “color” pyramid of the same image is shown on the left (Fig. 2(a)), actually obtained through the decomposition $\mathcal{P}\{\mathbf{F}\} = \{\mathcal{P}\{F^Y\}, \mathcal{P}\{F^{Cr}\}, \mathcal{P}\{F^{Cb}\}\}$ on the Y , Cr , Cb channels respectively, presented in the right picture Fig. 2(b).

In our implementation, pyramid depth is automatically computed given the size of the image at the coarsest level.

2.4. The MDEM algorithm

As previously remarked, a key property of pyramids is that long range interactions can be captured by short paths in the coarse levels, as the paths are through points representing blocks of pixels instead of pixels. Such multiresolution representation is used as follows. At a certain level l of the pyramid, different from the lowest resolution level, maximization of $\log \mathcal{L}$ is obtained by iteratively computing $p(k^{(l)}|\mathbf{F}^{(l)}, \theta^{(l)})$, $p(\mathbf{F}^{(l)}|k^{(l)}, \theta^{(l)})$, $P(k^{(l)})$ while diffusing on $p(k^{(l)}|\mathbf{F}^{(l)}, \theta^{(l)})$ over sites i :

$$\frac{\partial p(k^{(l)}|\mathbf{f}_i^{(l)}, \theta^{(l)})}{\partial t} = \nabla \cdot (g(\nabla p(k^{(l)}|\mathbf{f}_i^{(l)}, \theta^{(l)})) \nabla p(k^{(l)}|\mathbf{f}_i^{(l)}, \theta^{(l)})); \quad (14)$$

Eq. (14), defines a system of D diffusion equations for each labeling plane k .

As previously discussed, this step can be used to incorporate spatial coherence into segmentation techniques. Using anisotropic diffusion on the posterior probabilities $p(k^{(l)}|\mathbf{F}^{(l)}, \theta^{(l)})$ to capture local spatial constraints is motivated by the intuition that posteriors with piecewise uniform regions result in segmentations with piecewise uniform regions. Application of anisotropic smoothing on the posterior

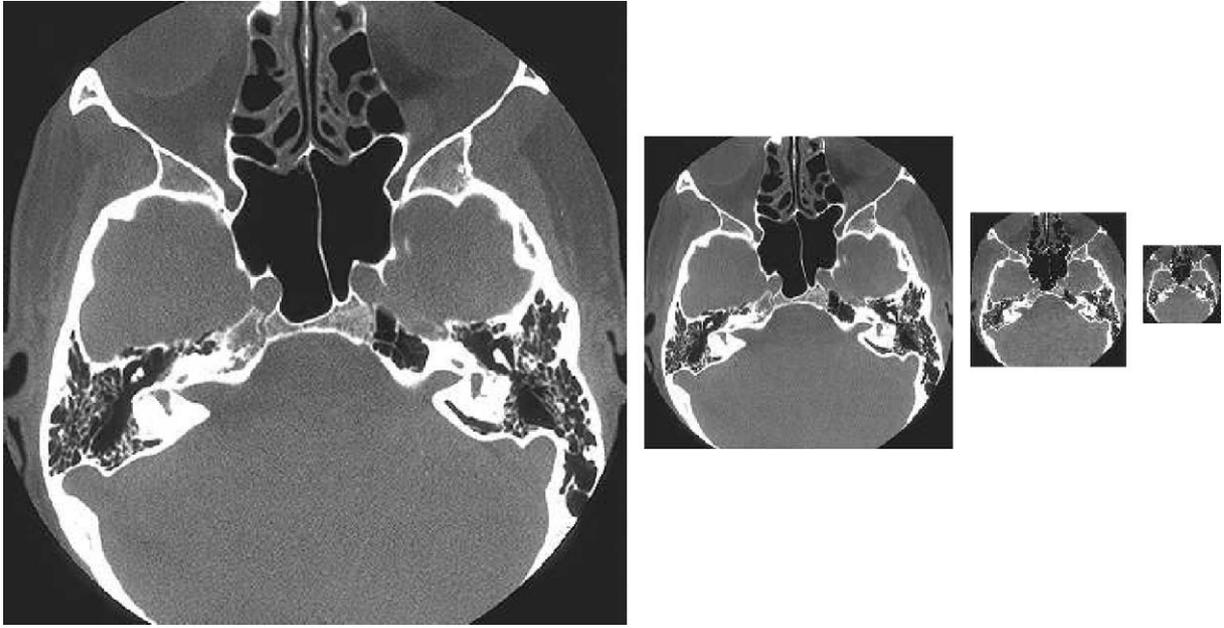


Fig. 1. A pyramidal representation of a CT scan image using four resolution levels of decreasing resolution, from left to right, $f^{d,(0)}$ (the original image), $f^{d,(1)}$, $f^{d,(2)}$, $f^{d,(3)}$.

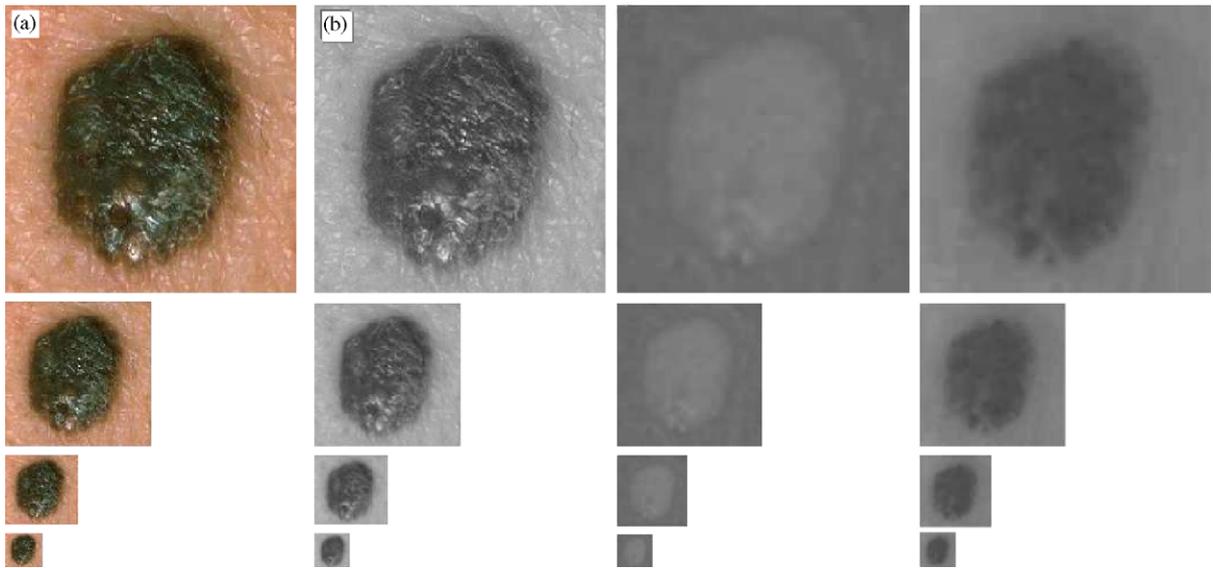


Fig. 2. A pyramidal representation of a color image (skin lesion); (a) the “color” pyramid of the original image; (b) from left to right the scalar pyramids $\mathcal{P}\{F^Y\}$, $\mathcal{P}\{F^{Cr}\}$, $\mathcal{P}\{F^{Cb}\}$ of the Y , Cr , Cb channels, respectively.

probabilities is possible even in case classes are described by general probability distribution functions [31].

At such level, the labeling planes $p(k^{(l)}|F^{(l)}, \theta^{(l)})$ are initialized by up-sampling the probability maps that have been previously derived at the coarser level $l + 1$:

$$p(k^{(l)}|F^{(l)}) = S \uparrow p(k^{(l+1)}|F^{(l+1)}). \quad (15)$$

This way the algorithm uses the probabilistic labeling proposed at coarser levels, while reducing the iteration steps necessary to achieve convergence of the expectation–diffusion–maximization cycle at that level.

The probabilistic model is assumed to be a mixture of multivariate gaussians

$$p(F|k, \mu_k, \Sigma_k) = \frac{\exp(-(1/2)(F - \mu_k)^T \Sigma_k^{-1} (F - \mu_k))}{(2\pi)^{D/2} |\Sigma_k|^{1/2}}, \quad (16)$$

$\theta_k = \{\mu_k, \Sigma_k\}$ being the unknown mean vectors and covariance matrices, respectively, weighted by mixing proportions $\alpha_k = P(k)$. Note that, we can consider the covariance matrices being diagonal because of the choice of the $YCrCb$ color space, and, furthermore we assume K fixed, in that we are not concerned here with the problem of model selection.

In summary, the MDEM algorithm works by performing two steps: in an unsupervised learning stage, parameters of the mixture and mixing coefficients are derived; in the classification step, the learned mixture is used to segment the image.

The learning step is articulated as follows:

1. Transform the RGB image (F^R, F^G, F^B) to (F^Y, F^{Cr}, F^{Cb}) .
2. Compute the pyramidal representation of the image $\mathcal{P}\{F^Y\}, \mathcal{P}\{F^{Cr}\}, \mathcal{P}\{F^{Cb}\}$ via Eq. (13).
3. Initialize $h_{L,ik}, \alpha_{L,k}, \mu_{L,k}, \Sigma_{L,k}$ at the coarsest level L of the pyramid.
4. **for** $(l = L - 1, \dots, 0)$ **do**.
5. Propagate to the upper level probabilities $h_{l,ik} \leftarrow h_{l+1,ik}$, by up-sampling, according to Eq. (15), and parameters $\alpha_{l,k} \leftarrow \alpha_{l+1,k}, \mu_{l,k} \leftarrow \mu_{l+1,k}, \Sigma_{l,k} \leftarrow \Sigma_{l+1,k}$
6. $t \leftarrow 1$
7. **repeat**
8. {E-step: given θ_l , obtain the distribution of the hidden variables}
9. **for** $(i = 1, \dots, N)$ **do**
10. **for** $(k = 1, \dots, K)$ **do**
- 11.

$$h_{l,ik}^{(t)} \leftarrow \frac{\alpha_{l,k}^{(t)} P(f_i^{(l)} | k^{(l)}, \mu_{l,k}^{(t)}, \Sigma_{l,k}^{(t)})}{\sum_k \alpha_{l,k}^{(t)} P(f_i^{(l)} | k^{(l)}, \mu_{l,k}^{(t)}, \Sigma_{l,k}^{(t)})} \quad (17)$$

12. **end for**
13. **end for**
14. {D-step: propagate $h_{l,ik}$ by $T(l)$ iterations of the discrete form of anisotropic diffusion (14)}
15. **for** $(k = 1, \dots, K)$ **do**
16. **for** $\tau = 0 \dots T(l) - 1$ **do**
17. **for** $(i = 1, \dots, N)$ **do**

$$h_{l,ik}^{(t+\tau+1)} \leftarrow h_{l,ik}^{(t+\tau)} + \lambda \nabla \cdot (g(\nabla h_{l,ik}^{(t+\tau)}) \nabla h_{l,ik}^{(t+\tau)}) \quad (18)$$

18. **end for**
19. **end for**
20. $\tilde{h}_{ik}^{(t)} \leftarrow h_{l,ik}^{(t+T(l))}, \forall i$
21. **end for**
22. {M-step: with $\tilde{h}_{l,ik}^{(t)}$ fixed, calculate the parameters that maximise $\log \mathcal{L}$ }
23. **for** $(k = 1, \dots, K)$ **do**

$$\alpha_{l,k}^{(t+1)} \leftarrow \frac{1}{N} \sum_i \tilde{h}_{l,ik}^{(t)} \quad (19)$$

$$\mu_k^{(t+1)} \leftarrow \frac{\sum_i \tilde{h}_{l,ik}^{(t)} f_i^{(l)}}{\sum_i \tilde{h}_{l,ik}^{(t)}} \quad (20)$$

$$\Sigma_{l,k}^{(t+1)} \leftarrow \frac{\sum_i \tilde{h}_{l,ik}^{(t)} [f_i^{(l)} - \mu_{l,k}^{(t+1)}]^2}{\sum_i \tilde{h}_{l,ik}^{(t)}}, \quad (21)$$

26. **end for**
27. $t \leftarrow t + 1$
28. **until** $|\log \mathcal{L}^{(t+1)} - \log \mathcal{L}^{(t)}| < \varepsilon$
29. **end for**

The initialization of the MDEM algorithm at the coarsest level is performed, by running the E-, D-, M-steps but with parameters $\alpha_{L,k}, \mu_{L,k}, \Sigma_{L,k}$ initialized as reported in the experimental section.

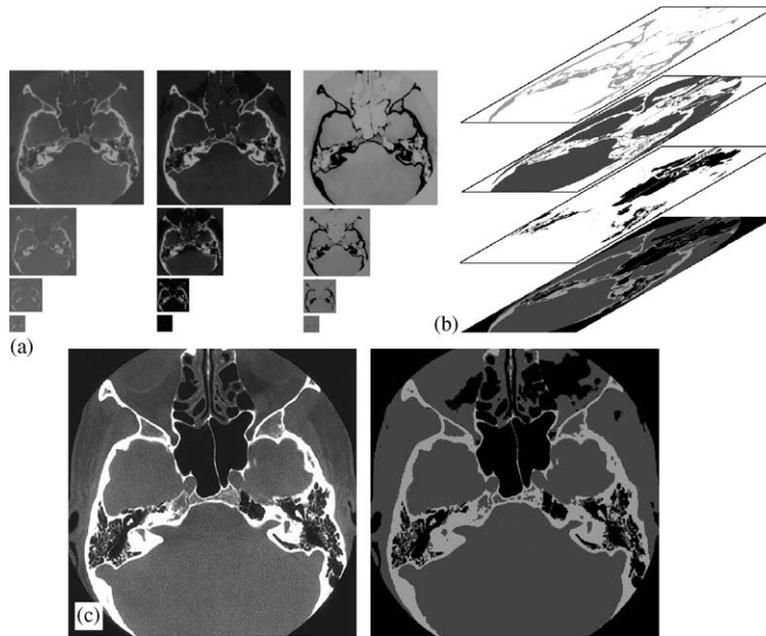


Fig. 3. Intermediate representations and final output of the method obtained on the CT image: (a) the labeling planes $\{p(k^{(l)}|F^{(l)})\}_{k=1,2,3}^{l=0,1,2,3}$; (b) a layered representation of the three region classes (top three planes) and the segmented image (bottom plane); (c) the final segmentation result of the MDEM algorithm (right image) compared with the original (left), by using μ_k parameter as a gray-level to graphically represent the regions of class k .

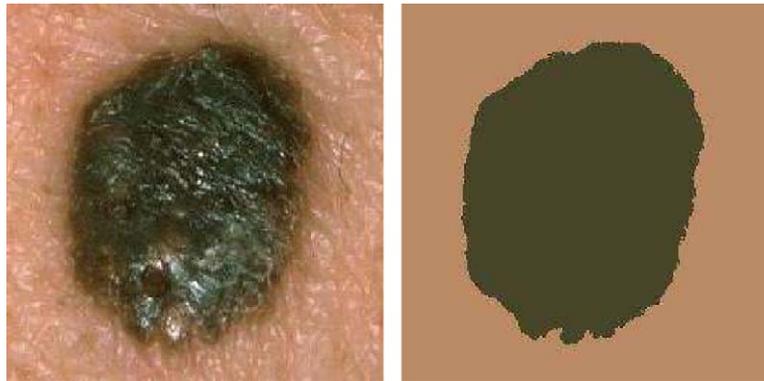


Fig. 4. Skin lesion segmentation (right) using the MDEM algorithm compared with the original image (left).

After that the parameter estimation stage has been completed, segmentation is achieved for each pixel $i \in \Omega$ by assigning to i , the label k for which $\max_k \{p(f_i^{(0)}|k^{(0)}, \mu_{0,k}^{(t)}, \Sigma_{0,k}^{(t)})\}$ holds. Eventually, the segmented YCrCb image is back-transformed to RGB space.

The algorithm has been outlined for color images ($D = 3$). Note that the algorithm applies to gray-level images as a special case ($D = 1$); also, the case of noncolor vector-valued images can be handled. The final output and intermediate representations produced by the proposed method are summarized at a glance in Fig. 3; for graphical simplicity we will use the scalar CT image shown in Fig. 1. Assuming a number of classes equal to 3, i.e. $k = 1, 2, 3$, in Fig. 3(a), the labeling planes $\{p(k^{(l)}|F^{(l)})\}_{k=1,2,3}^{l=0,1,2,3}$ are shown for each resolution level $l = 0, 1, 2, 3$, as gray-level images, brighter levels denoting higher probability; note how, for each of

the three classes the probability maps are progressively refined when going from the coarsest level $l = 3$ to the finest level $l = 0$. At the end of the learning stage, segmentation of the original image is obtained which can be conceived as a layered representation of the original image, each layer grouping regions/objects of the same class (Fig. 3(b)). By collapsing the three layers on a single plane, the segmentation result is obtained and compared with the original (Fig. 3(c)). Fig. 4 presents an example of skin lesion classification performed on the color image of Fig. 2. Two classes have been used, skin and lesion. It can be noted that, beyond spatial accuracy, by using the vector mean μ_k as the color to represent the regions of class k , the segmented result is chromatically coherent with the original image.

The performance of the MDEM algorithm with respect to both EM and DEM can be evaluated in terms of maximization rate of the log-likelihood. In Fig. 5 $\log \mathcal{L}$ is plotted versus the

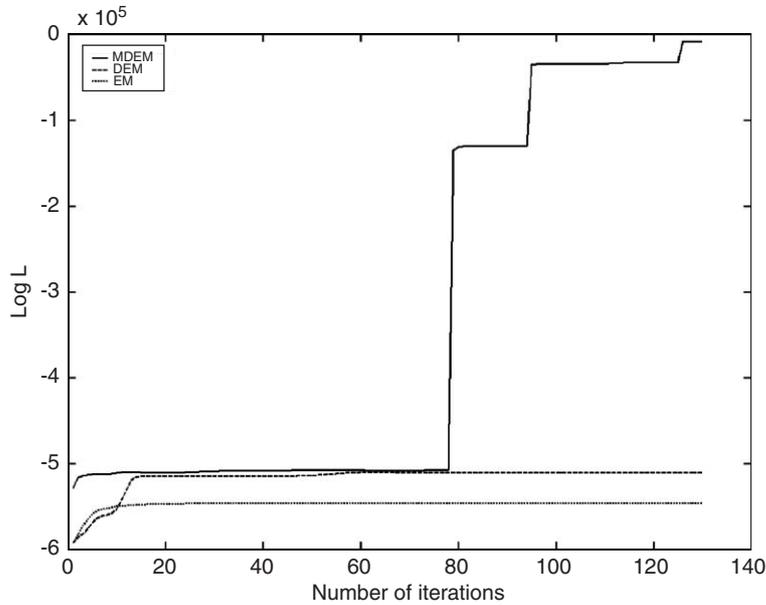


Fig. 5. Plot of $\log \mathcal{L}$ vs. number of iterations: MDEM (solid line), DEM (dash line), EM (dotted line).

number of iterations; it is apparent that the log-likelihood is much higher in the MDEM case which already at the coarsest level yields a value of $\log \mathcal{L}$ which is larger than those of the other methods. The steps in the continuous line correspond to the transitions from a coarser to a finer scale, and clearly demonstrate the advantage of a multiscale representation. The number of iterations in the case of EM and DEM algorithms, was equal to the total iterations of MDEM (sum of iterations carried out at each level), and the same parameter initialization has been adopted.

3. Procedures and results

A first set of experiments aimed at comparing the proposed method with other methods either feature or spatial based. Among feature based technique we have chosen the EM algorithm and its DEM variant; also, for scalar, gray-level images we implemented a version of entropy-based thresholding which can iteratively account for more than two classes, based on a modified version of well known Kapur, Sahoo and Wong algorithm [15]. The algorithm works as follows. Consider the histogram of the gray-level f , and define $p(f') = n(f')/N$, where $n(f')$ is the number of pixels with gray-level f' and N the total number of pixels. Next define $\Phi(f) = \sum_{f'=f_{\min}}^f p(f')$ and form distributions $F_1 = p(f')/\Phi(f)$, $f' \in [f_{\min}, f]$ and $F_2 = p(f'')/(1 - \Phi(f))$, $f'' \in [f + 1, f_{\max}]$ For each level f , the Shannon entropies of distributions F_1 , and F_2 can then be computed. They are

$$\begin{aligned} H_S(F_1) &= - \sum_{f'=f_{\min}}^f \frac{p(f')}{\Phi(f)} \log \frac{p(f')}{\Phi(f)}, H_S(F_2) \\ &= - \sum_{f''=f+1}^{f_{\max}-1} \frac{\Phi(f'')}{1 - \Phi(f)} \log \frac{\Phi(f'')}{1 - \Phi(f)}. \end{aligned}$$

These entropies are then added to give the total entropy for f : $H_{\text{tot}}(f) = H_S(F_1) + H_S(F_2)$. The optimal threshold t for distinguishing between two adjacent regions of different brightness is the level f which maximizes the uncertainty measured by H_{tot} : $t = \text{argmax}\{H_{\text{tot}}(f)\}$, in that this value has the same probability of belonging to either region. In order to achieve the multilevel thresholding for more than two classes, the procedure is iteratively applied to compute a sequence of thresholds t_1, t_2, \dots, t_{K-1} , for segmenting K classes. This algorithm is used only on scalar images, since its generalization to vector-valued ones is not straightforward.

A split-and-merge algorithm was chosen among spatial based techniques and preferred to region growing since not needing initialization of seed points. The algorithm used is the classic recursive algorithm where regions are split and merged until they have a variance σ below a specified threshold σ_{th} . In the case of scalar images the variance is simply the gray-level variance of the considered region, while for vector-valued images it is set to the joint total variance $\sum_{d=1}^D \sigma_d$ of the D channels.

Since the method proposed here is a general-purpose segmentation technique, we have experimented on different kinds of medical images. The data set included 90 images, namely CT scans, radiographs, dermatologic and eye fundus images. CT and radiographic images are useful to assess the performance of MDEM on scalar images, while the latter allow evaluation with respect to vector-valued ones. Examples from this set of experiments are presented using the test images shown in the following.

For showing results on gray-level images, we use an image of a CT scan of the head at the level of the orbits, middle ear, and paranasal sinuses, Fig. 6(a) and an image of a radiography of the hand, Fig. 6(b). In the case of the CT image, where gray-level values represent average X-ray absorption distorted by noise and artifacts, the initially perceived complexity of the CT can be reduced by first identifying the major anatomical



Fig. 6. Scalar images: (a) CT scan of the head at level of the orbits, middle ear, and paranasal sinuses; (b) Radiograph of a hand.

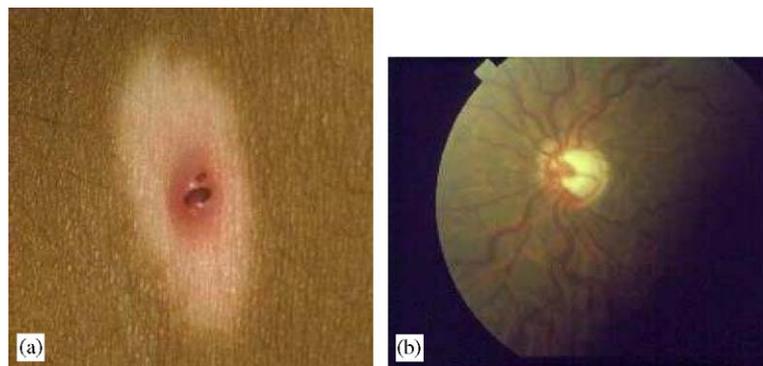


Fig. 7. Vector-valued images: (a) Skin lesion; (b) Glaucomatous eye.

subregions. In the case of radiographic image, it is of interest to detect the radiopaque parts of a radiograph. Results from vector-valued images are illustrated using one image of skin benign cancer, Fig. 7(a), and a image digitized from a photograph of the optic disc of a glaucomatous eye, Fig. 7(b). Skin lesions are of particular interest to test a vector-valued image segmentation algorithm, since they come in a variety of colors, and changes of colors from a lesion to its background can be effectively used for segmentation, although color variations may exist within the lesion or in the background [32].

A second set of experiments aimed at validating segmentation results of our method with a ground-truth. To this end we compared MDEM segmentation of images of skin lesions with segmentation of the same images provided by experts in the field of skin cancer.

The experimental setting was the following. For all the experiments, the depth of the pyramid, in the MDEM algorithm, is automatically computed by fixing the minimum length/width

to the size of 20 pixels. This value has been found to be a reasonable trade-off between an accurate segmentation, which requires the initialization of a number of classes/regions undistinguishable at the coarsest resolutions (e.g., 10, 5 and so on), and an efficient bootstrapping of the EM procedure.

Then, nonuniform initial estimates were chosen for $\alpha_{L,k}^{(0)}$, $\mu_{L,k}^{(0)}$, $\sigma_{L,k}^{(0)}$ parameters at such level. The values of the components of the mean vector $\{\mu_{L,k}^{(0)}\}$ ranged from the minimum to the maximum value of f_i , elements $\{\sigma_{L,k}^{(0)}\}$ of the covariance matrix took values in the range from 1 to $\max\{f_i\}$; finally $\{\alpha_{L,k}^{(0)}\}$ were set from $\max\{f_i\}$ to 1 and then normalized so that $\sum_k \alpha_{L,k}^{(0)} = 1$. In all cases the variations of the parameters were constant. The conductance function g used in the diffusion step of the algorithm can have a quite general form, but must be such that label boundaries are preserved, and numerical stability guaranteed. In order that diffused quantities remain probabilities, that is, are positive and add to one,

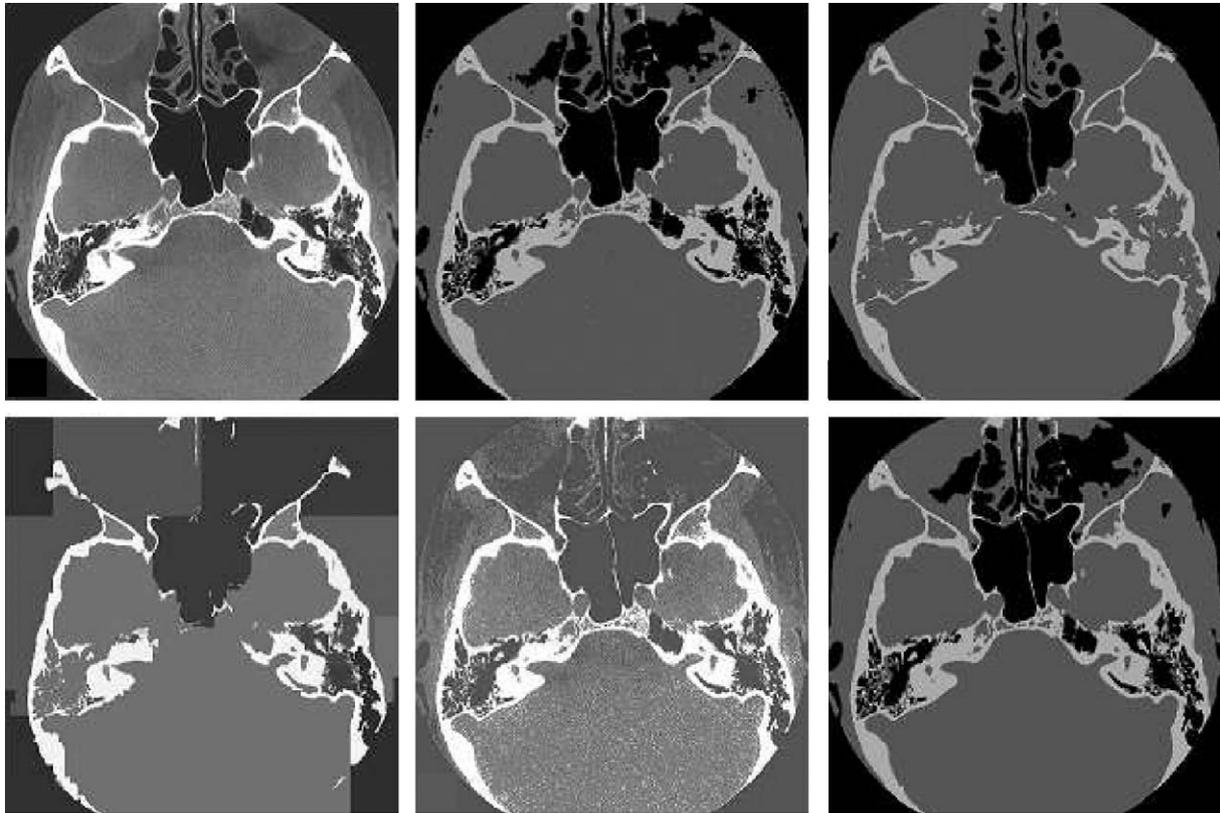


Fig. 8. CT image segmentation results. From left to right, the original image and segmentation results of DEM, EM, split-and-merge, iterative entropy thresholding, MDEM methods. Regions are coded with their average gray-value.



Fig. 9. Hand radiograph segmentation results. From left to right, the original image and segmentation results of DEM, EM, split-and-merge, iterative entropy thresholding, MDEM methods.

a suitable normalization must be provided after each iteration step; to this end, the functions $h_{ik}^{(l)}$ are renormalized so that their sum is one after each iteration [31]. In our experiments we set $g(\nabla h_{ik}) = |\nabla h_{ik}|^{-9/5}$, while the values of λ , the number of $T(l)$ iterations of (18) performed at level l is automatically set as $T(l) = 3 \times (1 + L - l)$, L being the maximum depth of the pyramid. The number of classes, K , was chosen to account for the classes to be determined in each experiment, with respect to different types of images, while convergence of the algorithm was controlled by $\varepsilon = 0.1$. For all the experiments, the same setting was used for EM, DEM and MDEM algorithms. The number K of classes automatically determines the number of thresholds $N_l = K - 1$ to be iteratively computed for entropy thresholding (e.g., 2 thresholds t_1 and t_2 are needed to discriminate among 3 classes of regions).

Fig. 8 shows the results obtained on the CT image using DEM, EM, split-and-merge, iterative entropy thresholding and MDEM by considering $K = 3$ classes; regularization parameter of anisotropic diffusion was set as $\lambda = 0.05$. For the split-and-merge method, the threshold value $\sigma_{th} = 1000$ was experimentally chosen because providing the best performance.

It can be noted that DEM and MDEM achieve the most reliable segmentations, but the latter is more precise in discarding small spurious regions, due to the multiresolution analysis. Worst performance is shown by the standard split-and-merge method, which is also outperformed by the thresholding method, though the latter exhibits higher sensitivity to noise.

Results obtained for the hand radiography are summarized in Fig. 9. In this case $K = 6$ classes were taken into account;



Fig. 10. Skin lesion segmentation. From left to right, the original image and segmentation results of DEM, EM, split-and-merge, MDEM methods. Regions are coded with their average color value.

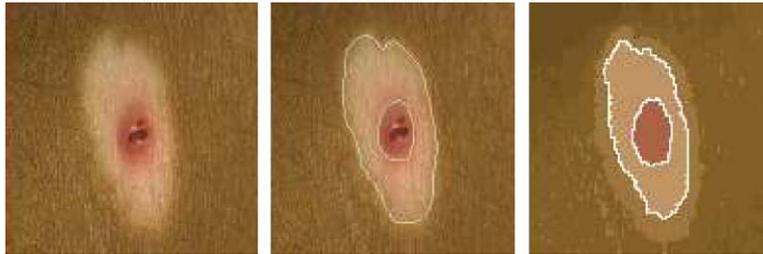


Fig. 11. Skin lesion contour: from left to right, original image, contours traced by a specialist, contours extracted after MDEM segmentation.

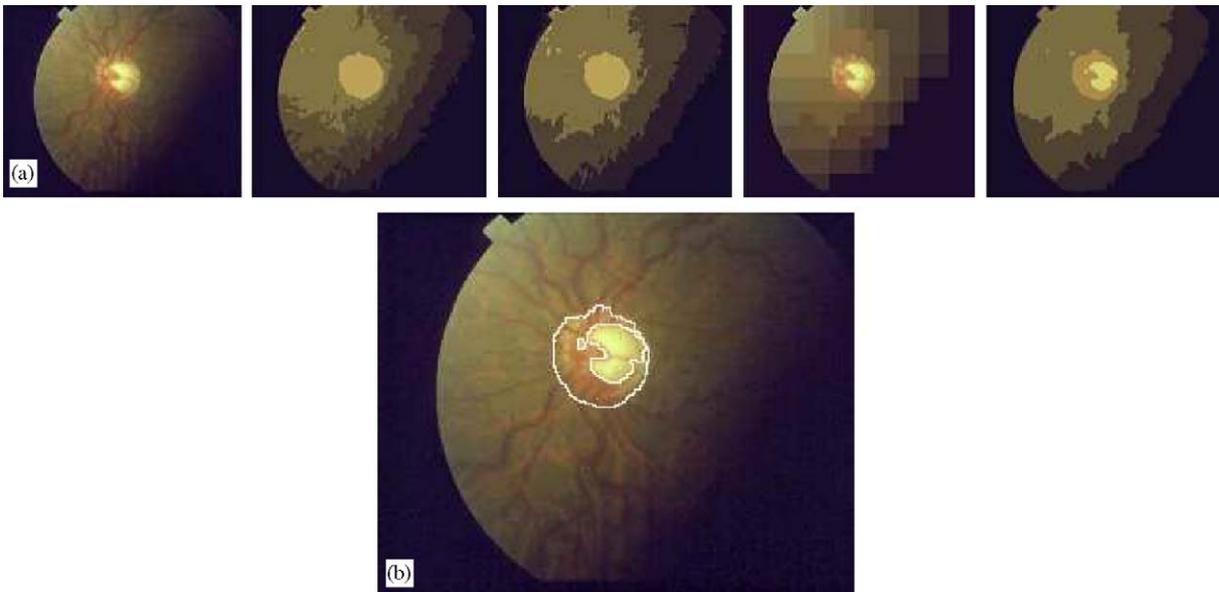


Fig. 12. Glaucoma image: (a) from left to right, original image, DEM, EM, Split and Merge, MDEM; (b) Contour of the Cup and the Disk of the Glaucoma.

anisotropic diffusion parameter was set as $\lambda = 0.01$. The split-and-merge method, was run with $\sigma_{th} = 1000$. Results achieved show the same ranking among the considered methods as in the previous experiment, in terms of quality of segmentation.

In the skin lesion case (Fig. 10) we have compared the methods considering $K = 4$ classes (three main region classes plus one class for outliers), $\lambda = 0.1$ and the split-and-merge method was run with $\sigma_{th} = 300$.

Note that for this specific case the classical EM performs better than DEM in terms of spatial precision, because of over-smoothing in the diffusion step. However, the MDEM algorithm is able to constrain smoothing due to multiresolution analysis. Interestingly enough, split-and-merge performance on vector-valued images is higher than its performance on

gray-scale since variance is jointly taken into account over the different channels.

To better appreciate the result obtained by MDEM, in Fig. 11, the ground truth (marked contours) found by an expert in the field of skin cancer, where a lesion boundary potentially exists, is compared with the contour automatically traced via edge detection performed on the segmented region.

The last example of this set of experiments (Fig. 12) shows results in the case of an image of the eye fundus. In this case a glaucoma is present. We used $K = 6$ classes, $\lambda = 0.1$ and the split-and-merge method was run with $\sigma_{th} = 200$. Segmentation results are illustrated in Fig. 12(a). Note that glaucoma is a neuropathy for which the evaluation of optic disk morphometry can be of cardinal importance. For instance, the measure of the

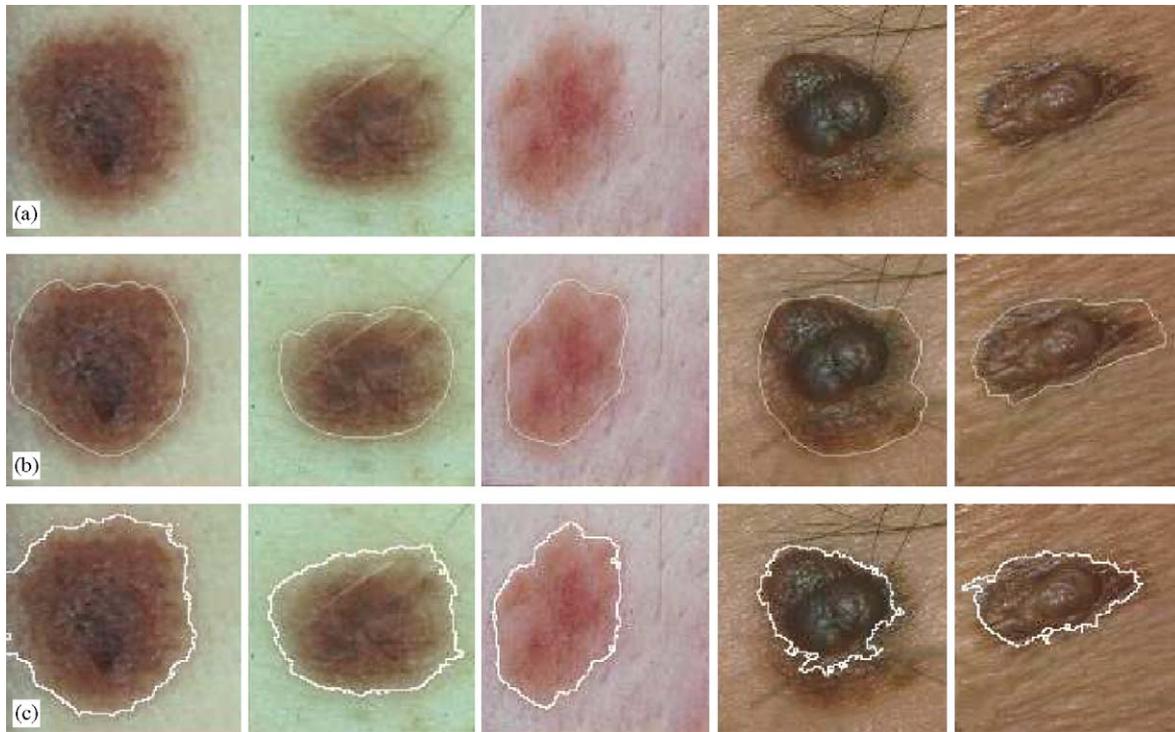


Fig. 13. Skin lesion segmentation results: (a) Original: from left to right, images 1–3 represent atypical lesions, 4 and 5, malignant lesions; (b) Ground truth (lesion boundaries manually traced); (c) boundaries automatically detected after MDEM processing and overlaid with the original.

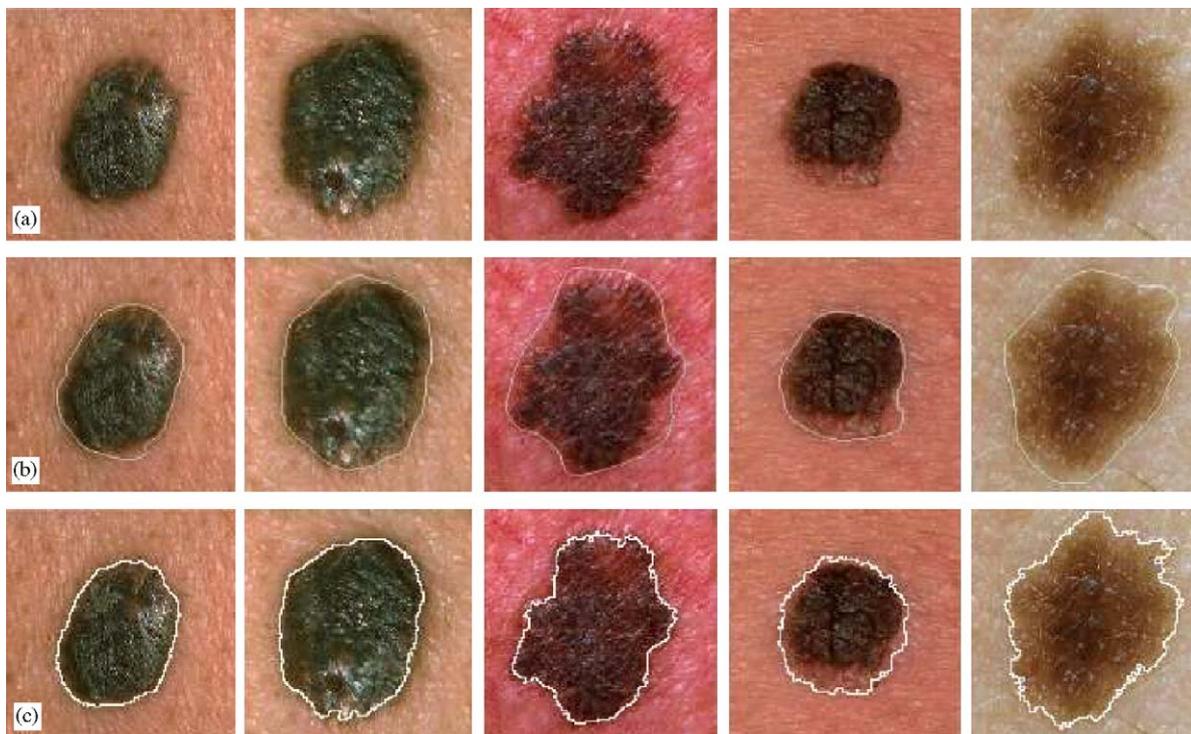


Fig. 14. Skin lesion segmentation results: (a) Original: from left to right, the images represent benign lesions; (b) Ground truth (lesion boundaries manually traced); (c) boundaries automatically detected after MDEM processing and overlaid with the original.

cup to disk ratio is a widely accepted figure of merit. Thus, to better appreciate segmentation results, in Fig. 12(b) boundaries of the cup and the disk have been extracted from the segmented image and overlaid on the original, to show how this preliminary

step may allow easy computation of such ratio, when followed by simple circle fitting.

Eventually, we present some examples from a second set of experiments performed on images of pigmented lesions

(either benign, atypical, malignant). To analyze such lesions it is necessary to accurately locate and isolate the lesions. Thus, results obtained by the MDEM algorithm have been compared with ground truths traced by experts in the field of skin cancer (Figs. 13 and 14). For ease of comparison, the contours of the segmented regions have been extracted via edge detection and overlapped onto the original image. Parameter settings were the same as in the case presented in Fig. 11.

4. Final remarks

MDEM is a novel scheme that results in a simple but effective segmentation algorithm for color images that: (1) retains the appealing characteristics of a feature clustering based approach; (2) takes into account spatial constraints while avoiding complex schemes such as MRFs; (3) it operates within a multiresolution framework, in order to reliably define regions of interest and efficiently perform required computations.

Note that the MDEM algorithm is different from previously proposed related methods. For instance, different approaches have tried to incorporate within the EM algorithm a prior term in order to maximize a log-posterior probability instead of log-likelihood, thus leading to quite complex EM steps [23,24]. On the other hand, Haker et al. [31], have suggested to compute an initial posterior probability map, through some kind of preliminary classification (e.g., clustering), followed by anisotropic diffusion performed on such initial map in order to diffuse spatial constraints among probability sites; clearly, in this way final results strictly depend upon the goodness of the initial labeling. Here, we follow a different approach: we operate on the maximization of the log-likelihood function, and spatial context is implicitly accounted for along maximization via diffusion. Further, in order to take advantage of the structure of the image as represented at different scales, this method has been carried out in a multiscale framework, thus yielding an accurate segmentation, without unduly increasing the computational load.

Here it has been assumed K fixed, in that we are not concerned with the problem of model selection. In general, this problem could be tackled by resorting to BIC or Akaike's information criteria [20]. However, this may not be necessary for biomedical images, since, depending on the application, the value of K is often assumed to be provided by prior knowledge of the anatomy being considered.

As a result we obtain a simple iterative segmentation algorithm which can be easily interpreted in terms of a multiresolution competition/cooperation scheme: at each resolution level of the pyramid, the E- and M-steps can be seen as an individual site competition between the k different label planes, while the interleaved diffusion step can be considered as a cooperation step within sites on the same plane.

With respect to the computational load, the whole algorithm is slightly slower than the EM procedure. Currently, it takes 30 s for a 256×256 2D image, using an Intel PIV 3.4 GHz processor, equipped with 2 GHz RAM, under Windows XP operating system.

Finally, simulations show that the method performs quite well on a variety of medical images either with respect to more

standard methods or also techniques specifically designed for an application (for instance, our skin lesion segmentation results can be compared with examples provided by Xu et al. [32] and available at www.cs.wright.edu/people/faculty/agoshtas/paper_fig.htm), and it is flexible enough to be used in a wide range of applications.

5. Summary

In the work presented here image segmentation of vector-valued images is carried out within a multiresolution framework. A multiscale representation is provided by Gaussian pyramids defined in the YCrCb color space, one for each color channel, and segmentation is performed in a hierarchical fashion, starting from the coarsest representation, so that results at a scale provide initial condition to segmentation in the next lower (finer) scale; this propagation of information from coarse resolution levels makes significant objects/regions in the image more relevant respect to weak textures and noise. Also, pyramids provide an efficient tool to reduce the iterations necessary to maximize the likelihood while optimizing the diffusion step, which, at each iteration acts upon a sub-sampled version of the probability maps. At each scale segmentation is obtained via an expectation–diffusion–maximization loop in which standard expectation–maximization is coupled with an anisotropic diffusion of the posterior probabilities of label assignment; this way MDEM takes into account spatial correlations in the image. The new approach has been validated on a variety of medical images and through comparisons with more standard methods.

References

- [1] L. Lucchese, S.K. Mitra, Color image segmentation: A state-of-the-art survey, Proceedings of the Indian National Science Academy (INSA-A), 67, New Delhi, India, March 2001, pp. 207–221.
- [2] J.S. Duncan, N. Ayache, Medical image analysis: progress over two decades and the challenges ahead, IEEE Trans. Pattern Anal. Mach. Intell. (PAMI) 22 (2000) 85–106.
- [3] D.L. Pham, C. Xu, J.L. Prince, Current methods in medical image segmentation, Annu. Rev. Biomed. Eng. 2 (2000) 315–337.
- [4] S. Pohlman, K.A. Powell, N.A. Obuchowski, W.A. Chilcote, S.G. Broniatowski, Quantitative classification of breast tumors in digitized mammograms, Med. Phys. 23 (1996) 1337–1345.
- [5] I.N. Manousakas, P.E. Undrill, G.G. Cameron, T.W. Redpath, Split-and-merge segmentation of magnetic resonance medical images: performance evaluation and extension to three dimensions, Comput. Biomed. Res. 31 (1998) 393–412.
- [6] J.J. Koenderink, The structure of images, Biol. Cybern. 50 (1984) 363–370.
- [7] S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, New York, 1998.
- [8] P.J. Burt, E.H. Adelson, The Laplacian pyramid as a compact image code, IEEE Trans. Commun. 9 (1983) 532–540.
- [9] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE Trans. Pattern Anal. Mach. Intell. (PAMI) 6 (1984) 721–741.
- [10] K. Held, E.R. Kops, B.J. Krause, W.M. Wells, R. Kikinis, H.W. Muller-Gartner, Markov random field segmentation of brain MR images, IEEE Trans. Med. Imaging 6 (1997) 878–886.
- [11] T. McInerney, D. Terzopoulos, Deformable models in medical image analysis: a survey, Med. Image Anal. 1 (1996) 91–108.
- [12] P. Perona, J. Malik, Scale-space and Edge Detection Using Anisotropic Diffusion, IEEE Trans. Pattern Anal. Mach. Intell. 12 (1990) 629–639.

- [13] D. Geiger, A. Yuille, A common framework for image segmentation, *Int. J. Comput. Vision* 6 (1991) 401–412.
- [14] D. Mumford, Bayesian rationale for the variational formulation, in: B.M. ter Haar Romeny (Ed.), *Geometry-Driven Diffusion in Computer Vision*, Kluwer Academic Publishers, Dordrecht, 1994, pp. 135–147.
- [15] P.K. Sahoo, S. Soltani, K.C. Wong, A Survey of Thresholding Techniques, *Comput. Vision, Graphics, Image Process.* 41 (1988) 233–260.
- [16] C. Lee, S. Hun, T.A. Ketter, M. Unser, Unsupervised connectivity-based thresholding segmentation of midsagittal brain MR images, *Comput. Biol. Med.* 28 (1998) 30–38.
- [17] G. Boccignone, A. Chianese, A. Picariello, Multiresolution spot detection via entropy thresholding, *J. Opt. Soc. Am. A* 17 (7) (2000) 1160–1171.
- [18] G. Boccignone, A. Chianese, A. Picariello, Computer Aided Detection of Microcalcifications in Digital Mammograms, *Comput. Biol. Med.* 30 (2000) 267–286.
- [19] R.O. Duda, P.E. Hart, D.G. Stork, *Pattern Classification*, Wiley and Sons, New York, 2001.
- [20] D.J.C. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, UK, 2003.
- [21] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum likelihood from incomplete data, *J. R. Stat. Soc.* 39 (1977) 1–38.
- [22] T.J. Hebert, Fast iterative segmentation of high resolution medical images, *IEEE Trans. Nucl. Sci.* 44 (1997) 1362–1367.
- [23] S. Sanjay-Gopal, T.J. Hebert, Bayesian pixel classification using spatially variant finite mixtures and the generalized EM algorithm, *IEEE Trans. Image Process.* 7 (1998) 1014–1028.
- [24] Y. Zhang, M. Brady, S. Smith, Segmentation of brain MR images through a hidden Markov random field model and the expectation–maximization algorithm, *IEEE Trans. Med. Imaging* 20 (1) (2001) 45–57.
- [25] G. Boccignone, M. Ferraro, P. Napoletano, Diffused expectation maximisation for image segmentation, *Electron Lett.* 40 (2004) 1107–1108.
- [26] J. Weickert, Applications of nonlinear diffusion in image processing and computer vision, *Acta Math. Univ. Comenianae* 70 (2001) 33–50.
- [27] G. Wyszecki, W.S. Stiles, *Color Science-Concepts and Methods, Quantitative Data and Formulas*, Wiley, New York, 1982.
- [28] I. Pitas, *Digital Image Processing Algorithms*, Prentice-Hall International, London, U.K., 1993.
- [29] G. Boccignone, M. Ferraro, An information-theoretic approach to interactions in images, *Spatial Vision* 12 (1999) 345–362.
- [30] D.A. Forsyth, J. Ponce, *Computer Vision: A Modern Approach*, Prentice-Hall International, London, UK, 2002.
- [31] S. Haker, G. Sapiro, A. Tannenbaum, Knowledge-based segmentation of SAR data with learned priors, *IEEE Trans. Image Process.* 9 (2000) 299–301.
- [32] L. Xu, M. Jackowski, A. Goshtasby, D. Roseman, S. Bines, C. Yu, A. Dhawan, A. Huntley, Segmentation of skin cancer images, *Image Vision Comput.* 17 (1999) 65–74.

Giuseppe Boccignone received the Laurea degree in theoretical physics from the University of Torino, Torino, Italy, in 1985. In 1986, he joined Olivetti Corporate Research, Ivrea, Italy. From 1990 to 1992, he served as a Chief Researcher of the Computer Vision Lab at CRIAI, Naples, Italy. From 1992 to 1994, he held a Research Consultant position at Research Labs of Bull HN, Milan, Italy, leading projects on biomedical imaging. In 1994, he joined the Dipartimento di Ingegneria dell'Informazione e Ingegneria Elettrica, University of Salerno, Salerno, Italy, where he is currently an Associate Professor of Computer Science. He has been active in the field of computer vision, image processing, and pattern recognition, working on general techniques for image analysis and description, medical imaging, and object-oriented models for image processing. His current research interests lie in active vision, theoretical models for computational vision, medical imaging, image and video databases. He is a Member of the IEEE Computer Society.

Paolo Napoletano received the Laurea degree in telecommunication engineering from the University of Naples Federico II, Naples, Italy, in 2003. He is currently working toward the Ph.D. degree in information engineering at the University of Salerno, Salerno, Italy. His current research interests lie in active vision, theoretical models for computational vision, medical imaging and image processing. He is a Student Member of the IEEE Computer Society.

Vittorio Caggiano received the Laurea degree in electronics engineering from the University of Salerno, Italy, in 2004. He is currently working toward the Ph.D. degree in computer science and engineering at the University of Naples Federico II, Naples, Italy. His current research interests are in the field of active vision, theoretical models for computational vision, medical imaging, image and video databases. He is a Student Member of the IEEE Computer Society.

Mario Ferraro received the Laurea degree in theoretical physics from the University of Turin (Italy) in 1973. He has worked in Universities in England, Canada, Germany and United States, carrying on research on fuzzy sets theory, human vision, invariant pattern recognition and computational vision. Presently he is associate professor of Physics at the University of Turin. His research interests include image and shape analysis, cellular biophysics and the theory of selforganising systems.