Machine Vision and Applications (2013) 24:695–705
DOI 10.1007/s00138-012-0422-8

ORIGINAL PAPER

Erosion band signatures for spatial extraction of features

Eduard Vazquez · Xiaoyun Yang · Greg Slabaugh

Received: 15 July 2011 / Revised: 9 February 2012 / Accepted: 28 February 2012 / Published online: 17 March 2012

Abstract We introduce *Erosion Band Signatures* (EBS), which are a codification of the spatial coherence of features extracted from a region. This coherence is often lost in traditional global and local feature extraction methods, thereby diminishing a feature's discriminative strength. The erosion band signature is generated through iterative erosions of the region of interest, forming what we call *erosion bands*. Features are then extracted from each band and accumulated in a specific order to form the EB signature, which preserves spatial information of the features. To demonstrate the versatility of EBS, we have implemented the method in two very different applications: polyp detection and region-based head tracking. In polyp detection, EBS provides an effective way to characterize spatial differences between the perimeter and core of a polyp candidate, and improves a state-of-the-art computer-aided detection method with an improved 27.6 % reduction of false positives. We also apply EBS analysis to region-based tracking yielding a very clear improvement in both robustness and accuracy.

Keywords CT colonography · CAD · False positive reduction · Tagged stool · Tracking · Spatially-aware feature extraction

1 Introduction

Feature extraction is a fundamental component of many computer vision applications including object recognition, classification, retrieval and tracking [1]. Often, in both local and global feature extraction approaches, the spatial coherence (i.e., the relative spatial position) of features is lost in the feature extraction process. As a result, the discriminative strength of the features may be diminished. For instance, features describing the eyes of a face are expected to appear near the center of the face. Consequently any object with eye-like features on its perimeter has less likelihood to be a face. The same lack of spatial coherence is present in computer-aided detection (CAD) methods applied to CT Colonography (CTC) images [2–4], as well as histogram [5] or spatiogram [6]-based methods used in tracking.

In CTC, features based on intensity, gradient orientation, shape index, and curvedness [7] are computed from a 3D polyp candidate without reference to the spatial location within the region the features are computed. Patients are given oral contrast agents (tagging agents) that result in higher intensities for liquid and solid remains in the colon, in order to distinguish residual waste from tissue. The relative variations between the core and perimeter intensity values of a polyp candidate provide a very useful characteristic for false positive detection. Specifically, tagged stool [8] is one of the main sources of false positives in computer-aided detection systems [9], and is characterized by a higher intensities in its core region compared to polyps. This has motivated recent work on false positive detection and visualization through translucency rendering [10,11]. This method finds differences in opacity between the perimeter and the core of a polyp candidate using volume rendering [12]. The main shortcoming of this technique is that a good rendering requires robust computation of an axis perpendicular to the colon wall, which is a challenging task given the varying morphology of polyps and stool. The reduction of false positives resulting from tagged stool and the shortcomings derived from translucency rendering form the main motivations of the present work.

E. Vazquez (✉) · X. Yang · G. Slabaugh
Medicsight PLC, Kensington Centre, 66 Hammersmith Road, London W14 8UD, UK
e-mail: eduard.vazquez1@gmail.com

In this paper, we introduce *Erosion Band Signatures* (EBS) for spatially aware feature extraction. EBS is applied to reduce tagged stool false positives in CTC CAD. Furthermore, we also analyze its potential as a general method for feature extraction. For a given region, we perform iterative erosions from its border to its core, generating a series of bands. For each band, we extract features to form a single 1D signature that implicitly codifies the spatial information. This signature is used to differentiate false positives from polyps. We point out that EBS is not limited to a 1D representation and can be extended to form an n-dimensional signature. This spatial codification is the main strength of the method compared with purely global or local methods, as results support.

Morphology is also used in other CTC CAD approaches. Commonly, it is used for polyp candidate detection [2,13]. Once polyp candidates have been found, these approaches perform a feature extraction process which does not include morphology or spatial information as in EBS. In [14], the authors use morphology as a feature for polyp detection. Nevertheless, it is just focused on detecting elliptical shapes which a polyp is expected to have in several cases. Spatial information is not included in this approach.

This lack of spatial information is also present in histogram-based approaches. Spatiograms, introduced in [6], propose adding some spatial information to the basic chromatic histogram, by adding the mean vector and the covariance matrix of the image pixels. Another approach focused on including spatial information to histograms consists on partitioning the image or object using a specific pattern and computing a histogram for each part [15]. Some other approaches have been proposed to add spatial information in object tracking, as template matching [16] consisting on tracking an object by extracting an example image of the object. The initial proposal did not consider that the object changes during the sequence, what requires updating the template [17]. In [18], a probabilistic framework is presented to track an object based on its appearance. Feature values and feature location are considered random variables.

In this paper, we analyze the effect of applying erosion bands to histograms and spatiograms. In both cases clear improvements are achieved in both accuracy and robustness. These results point out that the treatment of spatial coherence in spatiograms is purely local and the spatial coherence of the whole region is not included.

The rest of the article is organized as follows. In Sect. 2, we introduce erosion bands (EB). In Sect. 3, we apply EB signature for false positive reduction in CTC images. Afterwards, in Sect. 4, we apply EB signatures to region-based tracking. In Sect. 5, we show results obtained. Subsequently, in Sects. 6 and 7, we present a discussion of EB signatures and conclusions.

2 Erosion bands

For the explanation of EB we use polyp detection as example. Indeed, it has been the main motivation of our work.

Tagging material can be present in the perimeter of both polyp and stool (covering them) but stool commonly partially absorbs the tagging material. Erosion bands reflect these differences independently of the shape of the candidate.

Figure 1a shows an example of a 2D view of a 3D polyp candidate and the original mask, the latter in yellow. The mask is determined automatically by the CAD system. For tagged data we apply a threshold of $HU < -64$ to include polyps and tagging material, as suggested in [10] (Hounsfield units (HU) are the intensity values of the CT image). Figure 1b depicts in gray color the mask after the thresholding and in white and red lines the discarded part ($HU < -64$) of the original mask. Afterwards, we apply iterative 3D erosions to generate a set of Erosion bands (B_l), as depicted in Fig. 1c, where each of the three erosion bands is showed with a different color. Let M_0 be the initial mask of a given candidate and E a spherical structuring element. Firstly, we

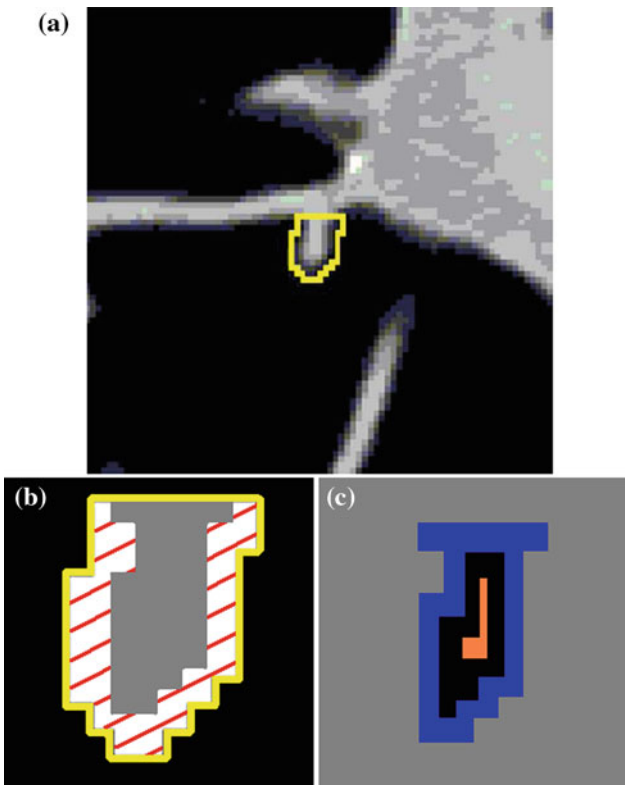


Fig. 1 Erosion bands generation. **a** Polyp candidate and its associated mask. **b** *In grey* Mask after thresholding at −64 HU. **c** We perform a set of erosions to the mask obtained in **b**, from which we find three erosion bands (B_l) for $l = \{1, 2, 3\}$ here labeled with three colors (*blue*, *black*, *orange*) (color figure online)

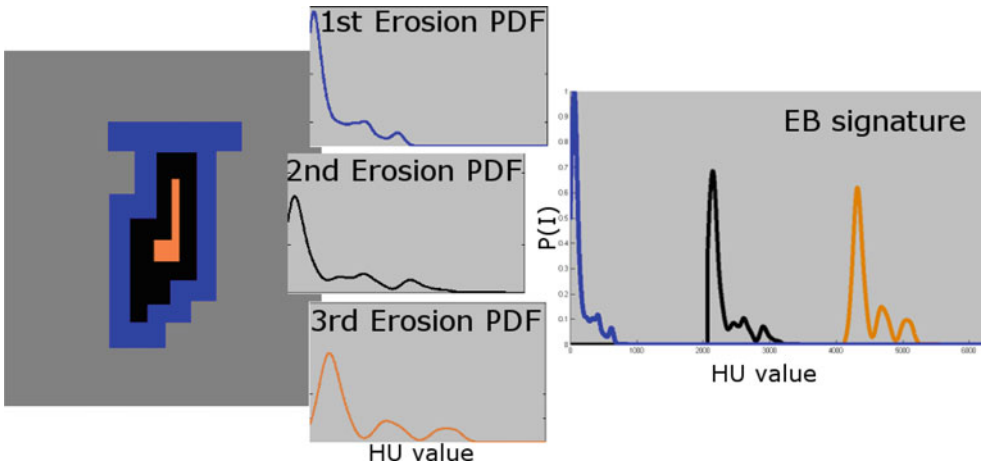


Fig. 2 EB-signature: a PDF computed from the intensity values of each band is calculated. Each PDF is placed one after the other thereby preserving the order of the bands from which they have been computed. It forms a single signature as a representative of the candidate (the EB signature)

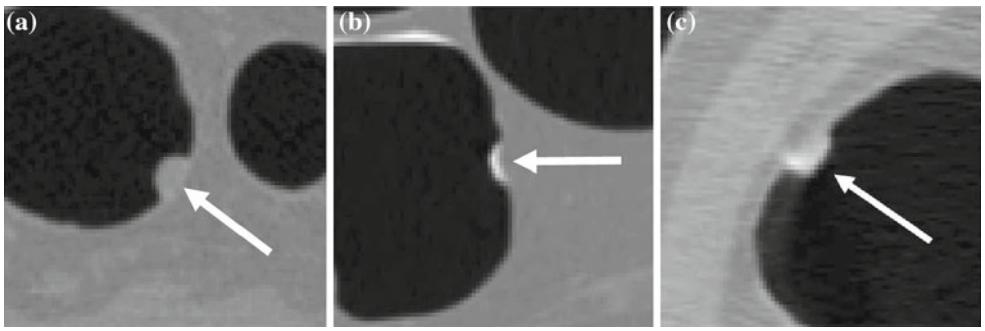


Fig. 3 An example of a **a** polyp, **b** tagged stool and **c** tagged polyp

compute the mask at the lth erosion step M_l as:

$$M_l = M_{l-1} \ominus E, \tag{1}$$

where $\ominus$ stands for the morphological erosion operator. The lth erosion Band B_l is then defined as the XOR operation (denoted by $\dot{\vee}$) between M_l and M_{l-1}:

$$B_l = M_l \dot{\vee} M_{l-1}, \tag{2}$$

where l ranges from $[1, \ldots, L]$ and L is the maximum number of erosions until $\emptyset$ ($L = 3$ in Fig. 1). We refer to the *size* of a candidate as its number of erosion bands. Each erosion band provides a mask to extract features. This approach does not require finding any particular plane or projection as the translucency approaches [19] although differences between the perimeter and the core of the candidate are represented.

2.1 EBS: single signature from B_l

The information given by each B_l has to be presented in such a way that the spatial coherence is preserved. We are also interested in having a single signature for a candidate to facilitate the task of classification, as we will show in the next section.

We place, preserving the order from perimeter towards core, the information given by each B_l, one after the other, thereby generating a single signature for a candidate. The information extracted from each band can be intensity-based, chromatic, shape-based (e.g. shape index [3], gradient or histogram of gradients [20]), statistical (mean, maximum), among others. Figure 2 depicts an example of our approach for false positive reduction. In this case, we compute a PDF from the intensity values of each B_l. Finally, from the three PDFs we generate a single PDF which will be the EB signature of the polyp candidate shown in Fig. 1a.

False positive detection in CTC CAD is based on distance metric learning and classification as described in Sect. 3. As an alternative application of EBS, we also present an approach based on the direct comparison of the EB signature, i.e., without projecting them, in the framework of region-based tracking. This second application is detailed in Sect. 4.

3 EBS for false positive reduction in CTC CAD

An example of a polyp and a false positive resulting from tagged stool is shown in Fig. 3a, b. Whereas the difference

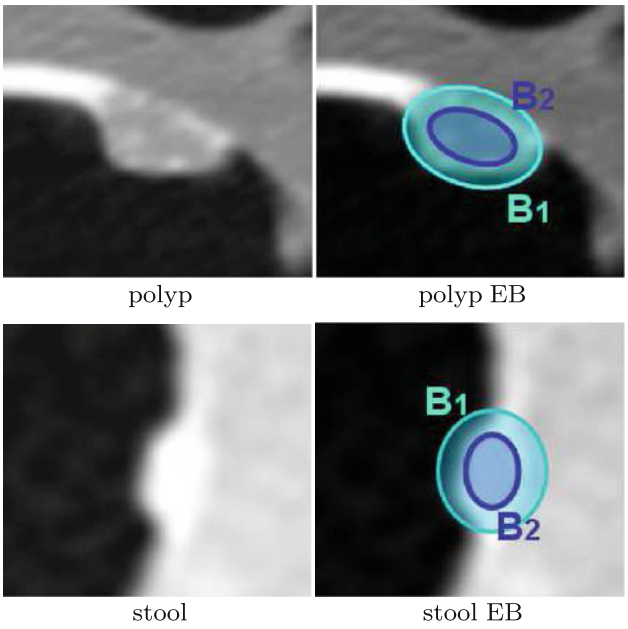


Fig. 4 EBS captures the spatial distribution of features along the candidate. The first band (B_1) in the polyp contains a fairly large amount of bright tagging and might be classified as false negative. However, the second band contains low presence of tagging in the core of the candidate. The second band of tagged stool will contain values not corresponding to tissue

between polyp and tagged stool may be straightforward, there is much more confusion when a polyp is covered by tagging (Fig. 3c). For detecting tagged stool, we use the intensity information of the polyp candidate as obtained from each B_l. Although intensity is already used in computer-aided detection systems, we argue that the introduction of spatial information provided by the EB improves false positive detection. Using EB, we have a signature for each polyp candidate which codifies the spatial differences in the region intensity from its perimeter towards its core. These differences are useful to distinguish polyps from some other sources of false positives as tagged stool. Figure 4 shows an example. The first band (B_1) of the polyp contains tagging agent. Nonetheless, the second band contains a small amount of tagging agent. In the case of tagged stool, the second band also contains values not corresponding to tissue. Furthermore, another advantage of having a single signature is that it facilitates the classification task.

From each B_l we compute a PDF and build the candidate's signature as previously explained. We then project each EB signature into a lower dimensional space where a polyp candidate is represented by a single point in the lower-dimensional space. Then, we can run a classifier to detect false positives, which have similar EB signatures. Similarity is expressed in terms of proximity in the lower-dimensional space.

3.1 PDF generation

A graphical example of the PDFs obtained is depicted in Fig. 2. The three PDFs associated with these three B_l are displayed at the right of the mask using corresponding colors. The abscissa axis depicts the CT values of the candidate. The ordinate axis displays the estimated density. The PDF has been computed using the off-the-shelf non-parametric *kernel-smoothing density estimation* method as described in [21]. This method estimates a PDF of a set of independent and uniformly distributed random variables $(x_1, x_2, \ldots, x_n)$. The density estimation is obtained by summing a set of kernels distributed among these variables:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \tag{3}$$

where K is the kernel used and h is the bandwidth that controls the smoothing. In our case the kernel used is a standard normal distribution:

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \tag{4}$$

Note that $\sigma = 1$, a common choice for the kernel [22]. The bandwidth has been selected using the method proposed in [23], which is a completely data-driven method that yields a good estimation even in multimodal densities.

These PDFs are therefore representative of the B_l from the border of the candidate and towards its core. Candidates in the same class (i.e., polyp or tagged stool) are expected to have similar EB signatures.

3.2 EBS projection

Dimensionality reduction can be applied to the EB signature to simplify classification [24]. In general, if we have two points x_i and x_j in a $\mathbb{R}^n$ space, we must find a proper distance measure $d(x_i, x_j)$ to preserve the relationship between the points. Formally:

$$\begin{aligned} d(x_i, x_j) &= (x_i - x_j)^t A (x_i - x_j) \\ &= \left(A^{\frac{1}{2}} x_i - A^{\frac{1}{2}} x_j\right)^t \left(A^{\frac{1}{2}} x_i - A^{\frac{1}{2}} y_j\right), \end{aligned} \tag{5}$$

where the projective matrix $A^{\frac{1}{2}}$ is to be learned.

In the present work, we have tested different techniques to find the best distance metric for the characteristics of our EB signature, namely Principal Component Analysis (PCA), ISOMAP [25] and Laplacian Eigenmaps (LE) [26]. A comprehensive survey of these techniques can be found in [24,27].

To test these dimensionality reduction methods, we have interpolated all EB signatures into the same number of bins independently of the candidate's size. Additionally, we have

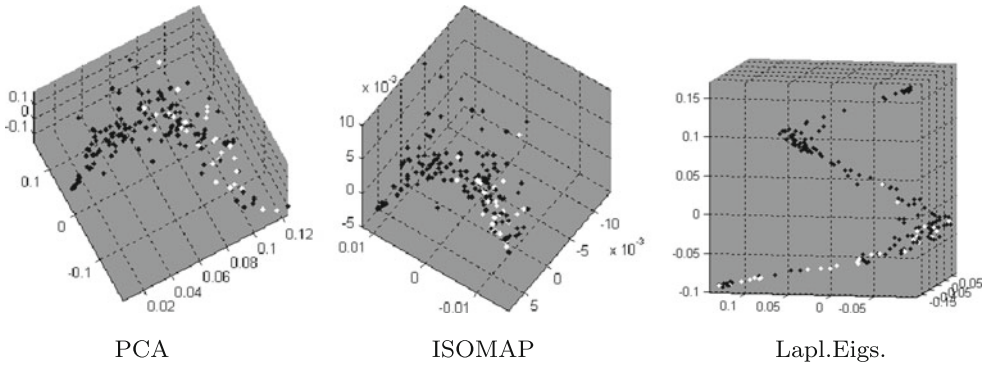


Fig. 5 Projections in 3D space of polyp candidates of the training datset using PCA, ISOMAP, and Laplacian Eigenmaps. These methods show a good separability between false positives (*black points*) and polyps (*white*)

zero-padded signatures corresponding to candidates of different sizes. Figure 6 depicts an example. It shows a EB signature of a candidate of size 4 and of a candidate of size 2. These signatures ought to have the same number of bins to be compared. Interpolated signature and zero-padded signature are shown in Fig. 6, second row. In both cases the number of bins of the signature is the same as that of the candidate of size 4. Each approach follows a different idea. Interpolating the signature represents a common way to deal with the different number of bins. Furthermore, we also used an adaptive size of the kernel depending on the candidate's size in order to obtain signatures with the same number of bins. Results obtained were similar to those obtained with interpolation. It can be explained because the singularities in the level of absorption of tagging agent for a big candidate are different of those of a small candidate. A higher precision for each band is able to explain these differences, whereas a big kernel cannot do so. For this reason we have also included the zero-padded signatures, which are able to preserve this information, therefore, yielding a higher discriminative strength. Results obtained are detailed in Sect. 5.

For the purpose of visualization, each EB signature is projected into a 3D space using the projective matrix $A^{\frac{1}{2}}$ as learned by each technique. Figure 5 shows results for the training dataset. As can be seen PCA, ISOMAP and Laplacian eigenmaps yield a good separability between polyps (white) and false positives (black), suggesting their use in a classification approach, which will be described in Sect. 5.

4 EBS in region-based tracking

To demonstrate the versatility of EBS, we evaluate the technique in a second application of region-based tracking. We compare the performance of color histograms as proposed in [5] and such of the color histograms extracted from each B_l and forming the EB signature. The same is performed for the spatiograms introduced in [6].

4.1 Region-based tracking with colour histograms

The first feature we compute for region-based tracking is the histogram of the region as proposed in [5]. Then, histogram intersection is computed between the histogram of a model h_I and a new frame $h_{I'}$:

$$D(h_I, h_{I'}) = \frac{\sum_{b=1}^{\beta} \min(h_I(b), h_{I'}(b))}{\sum_{b=1}^{\beta} h_{I'}(b)}, \tag{6}$$

where $h_I(b), h_{I'}(b)$ are the number of pixels in the bth bin and β the number of bins.

The color space proposed in [5] is $B-G, G-R$ and $R+G+B$ with 8 bins for the first (chromatic) channels and 4 bins for the intensity channel $R+G+B$. In our experiments we used the Opponent color space [28], which is normalized and represents the data in a better way. The three channels of the Opponent space are: (Red − Green) $= (R-G)/2$, (Blue − Yellow) $= (R+G-2B)/4$ and Intensity $= (R+G+B)/3$. Furthermore, the number of bins used was 32 instead of 8. With these two modifications, the tracker yields acceptable results that can be used as a baseline.

4.2 Region-based tracking with spatiograms

Spatiograms are an extension of histograms where local spatial information is added. As in [6], we use the second-order spatiograms for region-based tracking, defined as:

$$h_I^2(b) = \{n_b, \mu_b, \Sigma_b\}, \tag{7}$$

where n_b is the number of pixels in the bth bin, μ_b and Σ_b are the mean vector and the covariance matrix and $b = \{1, \ldots, \beta\}$ being β the number of bins.

The distance measure proposed in [6] has some disadvantages as detailed in [29]. Its main problem is that comparing a spatiogram with itself does not give the closest distance. We use the distance measure proposed in [29], based on the Bhattacharyya distance between two spatiograms. First, the second-order spatiogram is transformed into a histogram by

adding an extra spatial dimension, forming the transformed spatiogram n. Then, the distance measure proposed between two spatiograms n and n' is:

$$\rho(n, n') = \sum_{b=1}^{B} \sqrt{n_b n'_b} \left[8\pi |\Sigma_b \Sigma'_b|^{1/4} N(\mu_b; \mu'_b, 2(\Sigma_b + \Sigma'_b)) \right] \quad (8)$$

with $N(x; \mu, \Sigma)$ being a normalized Gaussian evaluated at point x.

5 Experimental results

In this section we show results obtained using EBS for false positive reduction in CTC CAD as well as in region-based tracking. EBS, when applied to CTC CAD, improves state-of-the-art results. In a second experiment we show that EBS can be applied to improve another computer vision application as region-based tracking.

5.1 False positive reduction in CTC

In this section, we present results obtained in false positive reduction in CTC CAD. Firstly, we describe the data used. Although the EBS method was originally designed for reducing tagged stool false positives, EBS is also tested on regions that are untagged.

5.1.1 CAD system

The starting point of our study is the output given by the computer-aided detection system, detailed in [2]. The set of features used in this system is derived from intensity, differential geometric features (shape index and curvedness [7]), gradient concentration, texture, volume and other shape and binary features. This system gives a performance of 90.1 % sensitivity (percentage of polyps detected) with 4.02 false positives per scan, the latter is the number that we seek to reduce. The CAD system as well as our new approach for false positive reduction is robust to different sources of data and patient preparation. Tested data comes from eight different institutions, with different imaging software (Siemens, GE, Philips and Toshiba) and different scanning protocols. The output of the system after being applied to these eight datasets is formed by a total of 412 scans and 2,048 candidates, from which 392 are polyps and 1,656 are false positive. There is no further information about the specific cause of the false positives.

A different dataset consisting in 1,368 candidates is used for testing. In this dataset there are 168 polyps and 1,200 cases of false positives. Thus, in these experiments there is a total of 3,416 polyp candidates, of which 60 % form the training set and 40 % is used for independent testing.

5.1.2 Tagging and non-tagging data

The separation between the tagging data and the rest has been done using a threshold of 250 HU. Thus, if a candidate has a single voxel with a CT value $\geq$250, it is considered as tagging data. In the training dataset of 2,048 candidates, there are 645 candidates (139 polyps and 506 false positives) that belong to this group. Our aim is to distinguish between polyps and false positives in the presence of tagging. A threshold in CT image values of -64 is applied to this dataset for the classification as suggested in [10] to remove air from the region, so what remains include soft tissue, tagging and bone.

As stated before, we also test EBS in non-tagging data. The rest of the training data, i.e., 1,403 candidates (253 polyps and 1,150 false positives) form the non-tagging data. For these candidates, since we want to distinguish between cases of polyps and false positives, which can result from a variety of sources including fat tissue and untagged stool. We apply a threshold of -439 HU [10] to just exclude the air.

Following the same procedure, the test dataset is divided in 636 cases of tagging (42 polyps and 594 false positives) and 732 cases of non-tagging data (126 polyps and 606 false positives).

5.1.3 Results obtained

As depicted in Fig. 5, good separability is achieved with PCA, ISOMAP and Laplacian Eigenmaps. In our experiments we apply PCA for two main reasons. First, PCA is faster than ISOMAP and Laplacian Eigenmaps. Second, the generalization for new data with Laplacian Eigenmaps is not as straightforward, whereas the generalization with PCA is trivial to perform.

To generate the PDF we selected 1,064 bins which cover the values from -64 HU to tagging agent values, which are around 1,000 HU. The dimensionality and priors used for projecting the data with the PCA have been found by a cross-validation along the first 100 dimensions in the training data. The number of dimensions for tagging and non-tagging data is 12 and 15, respectively. Note that more dimensions are required by non-tagging data. Afterwards, we classify the projected data to identify polyps from false positives. In this case we applied Naive Bayes classifier for tagging data and QDA [30] for non-tagging data. Other classifiers have been tested as linear discriminant analysis [31] or support vector machines [32], with nearly identical results.

Finally, we point out that treating separately tagging and non-tagging data yields a more reliable classification. Tagging material causes undesired effects in CTC intensity val-

Table 1 Results obtained using EB signature for false positive reduction for the testing data after projection with PCA and classifying with Naive Bayes classifier for tagging data and QDA for non-tagging data

	No. of false positives (%)	Sensitivity (%)
Tagging-data (12D)	18.3	97.7
Non-tagging (15D)	9.3	98.81
Total (%)	27.6	96.4

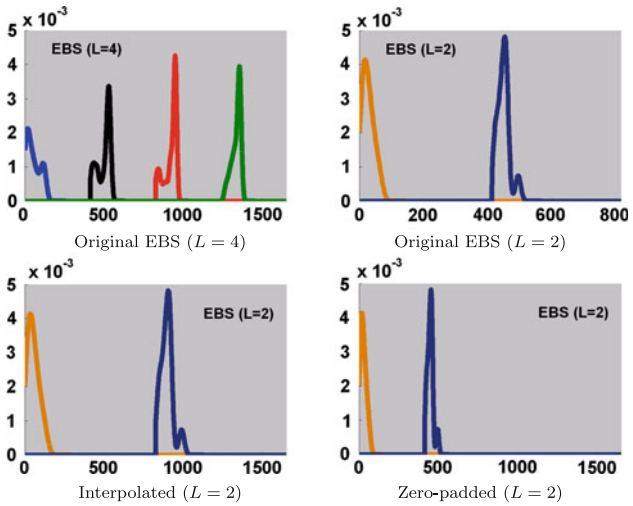


Fig. 6 EB signatures for candidates of size 4 and 2 (*top*), 2 interpolated (*bottom left*) and 2 zero-padded (*bottom right*)

ues. Therefore, mixing tagging with non-tagging material may result in poorer discrimination by the classifier.

Table 1 shows results obtained. As we can see, EBS-based classification using intensity values obtains 27.6 % false positive reduction at 96.4 % of sensitivity of the original CAD system. These results have been obtained by zero-padding the EBS corresponding to candidates of different sizes. Results obtained by interpolation (see Fig. 6) are 8 % lower in false positive reduction.

Finally, we point out that applying PCA directly to a PDF obtained from the whole candidate, i.e., without performing the iterative erosions, yields less than 12 % of FP reduction, showing the improvement achieved by including EBS. Furthermore, summarizing the information of each band with statistical measures as standard deviation, maximum, minimum, median or mean obtains meaningfully worse results. All experiments performed in this direction, either by combining these features in a single classifier or using a cascade of classifiers, obtained less than 10 % of FP reduction. In addition, such approach implies the selection of multiple thresholds, whereas EBS just requires a single linear classifier.

Clinically, this 27.6 % false positive reduction obtained from a state-of-the-art CAD system implies one false positive less per volume, which leads to a considerable improvement to radiologists review of CAD findings.

5.2 Region-based tracking

The aim of this experiment is to show how EBS improves a feature's discriminative strength (as happened with intensity in CTC images), rather than proposing a state-of-the-art region-based tracking approach. To this aim, we used the dataset presented in [5], which has been also used in [6]. The dataset consists of 15 sequences. One of the sequences was discarded since the face appearing initially was replaced by other faces along the video until appearing again. In these experiments we want to evaluate the accuracy and robustness when tracking the same face/head. Among the rest of the other 14 sequences, two of them have a ground-truth, consisting in the coordinates of a manually selected center of the head. These sequences contain 360° rotations of the head (then the face disappearing) as well as changes in scale.

For the EBS version of both histograms and spatiograms, we placed the histograms/spatiograms corresponding to each B_l one after the other to form the EB signature of the region. The initial mask was manually placed. The search on a new frame was done as in [5,6] with displacements in images coordinates of ± 6 and in scale of ± 1. No gradient information was included in order to demonstrate the potential advantages of using EBS in a simple manner. Furthermore, the addition of EBS yields a successful tracking in most of the sequences without gradient information.

5.2.1 Comparison with histograms

Region-based tracking using histograms was applied to the 14 sequences as detailed in Sect. 4 and further detailed in [5]. The same procedure was applied with the incorporation of EB signatures (EB histogram). For efficiency, the number of bands for each region was fixed at $L = 5$. We manually analyzed results to detect two objective error measures: *number of times lost* (head lost, mask somewhere else), and *number of frames lost* (the amount of frames the head has been lost). Table 2 summarizes results obtained. Histogram-based tracking loses the head a total of 22 times along the 14 sequences summing up a total of 381 frames. EB Histogram approach loses the head 6 times (54 % better) along 22 frames (94 % improvement).

Table 2 Comparison between Histogram tracking and EB histogram tracking

	Histogram	EB histogram
No. of frames lost	381	22
No. of times lost	13	6

5.2.2 Comparison with histograms and spatiograms

In this experiment we used the ground truth facilitated with two sequences called '*sb*' and '*mb*' of 500 frames each (which represents the 52 % of the total frames in the dataset). For each frame, an error is calculated with the Euclidean distance from the center of the mask (as found by each method) and the ground truth. Methods compared in this experiment were histograms [5], EB Histograms, spatiograms [6] and EB Spatiograms. Results are graphically depicted in Fig. 7 and Table 3.

As explained in previous section, EB-histogram outperforms histograms. Quantitatively, histogram-based tracking has an average error of 32.5 pixels in *sb* sequence against 23.84 pixels for EB-histogram and 35.83 pixels in *mb* sequence against 12.12 pixels for EB-histogram.

Regarding spatiograms, unless that the curve depicted in Fig. 7 draws a low error, spatiograms do actually not succeed in these two sequences. An example of the problem is shown in Figs. 8 and 9 second row. The mask of the head tends to grow until it covers almost the entire image, common error when no gradient information is used. EB-spatiogram avoids this error, even without gradient information. Since the head tends to be placed nearby the center of the image, the error is apparently low. In order to give the error for spatiograms shown in Table 3, we have subjectively selected a frame from which we consider that tracking is incorrect. The error is computed until those frames (see Fig. 8). Actually, even when averaging the error along all the frames for spatiogram, the error for EB-spatiogram is lower (even the error for EB-histogram in sequence *mb* is lower).

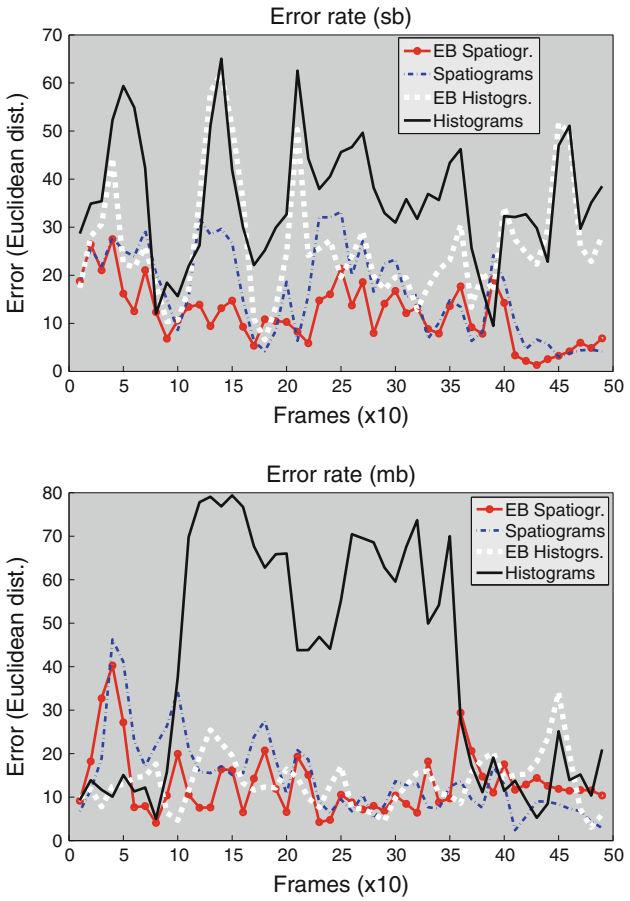


Fig. 7 Error rates obtained for sequences *'sb'* and *'mb'* for all region-based methods tested. Histogram clearly offers the worse performance. EB versions of histograms and specially of spatiograms provide the best performance

Table 3 Mean error along all sequence

	Sequence *sb*	Sequence *mb*
Hist.	32.57	35.83
EB hist.	23.84	12.12
Spatiog.	21.08[a]	21.57[a]
EB spat.	**10.71**	**11.94**

Error computed as the Euclidean distance between the benchmark and the center of the mask as found by each method

Bold values indicate the lowest errors

[a] Until last correct frame

Finally, in Fig. 9, we show some frames of sequence *mb*. In the first row, we show the frame 97, where both histogram and spatiogram approaches lose the head. In the second row, we show frame 400 where spatiogram, although having an apparent low error, has actually lost the head.

These results point out that by adding erosion bands, the performance of region-based tracking is improved in robustness and accuracy.

6 Discussion

In this article we have presented EB signature for feature extraction. This method, originally developed for polyp detection in CTC images, codifies the spatial coherence (relations) of the features extracted in a single signature. Its advantages are clear when compared to global measures such as histograms. The importance of erosion bands is graphically depicted in Fig. 10. First row shows the frame 120 from sequence *mb* and, in this order, presents the histogram for the blue-yellow channel, histogram from intensity channel and EB-Histograms for blue-yellow and intensity as well. The second row in Fig. 10 shows the same but for a different mask. It is clear that the portion of the image contained in the first mask is different than the information contained in the second mask. This difference is not captured in the histograms, which are visually identical. However, these differences are well represented by EB signatures.

Fig. 8 Mask computed using spatiograms tends to grow until the tracking fails, as the tracker starts to focus more on the background. Adding EB to spatiograms significantly improves the result

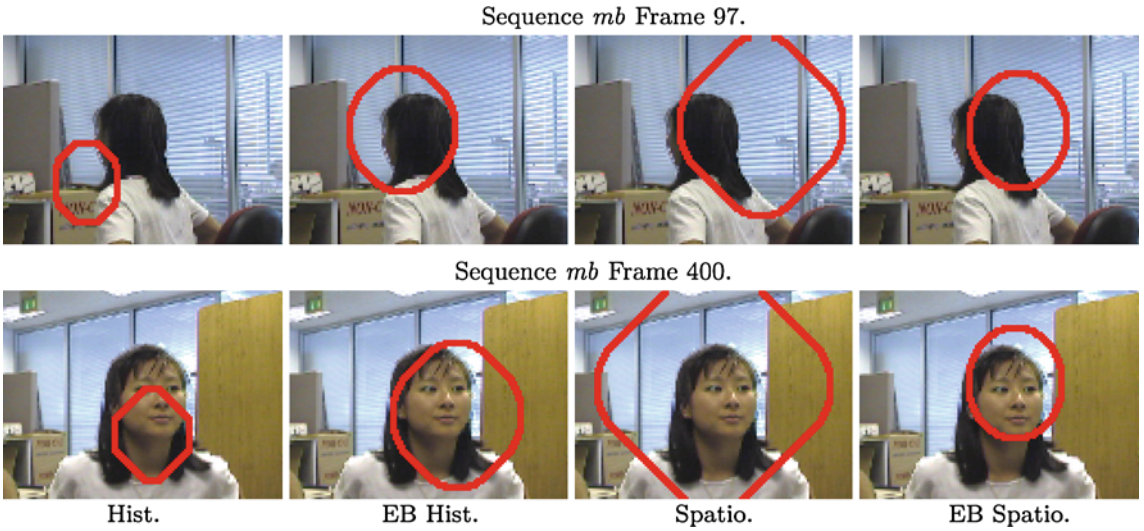


Fig. 9 *First row* Frame 97 (when histogram approach loses the head). *Second row* Frame 400. Apparently all methods provide a similar error rate for this frame. Nonetheless, spatiogram tracking covers almost the whole image (head lost). EB spatiogram avoids this problem

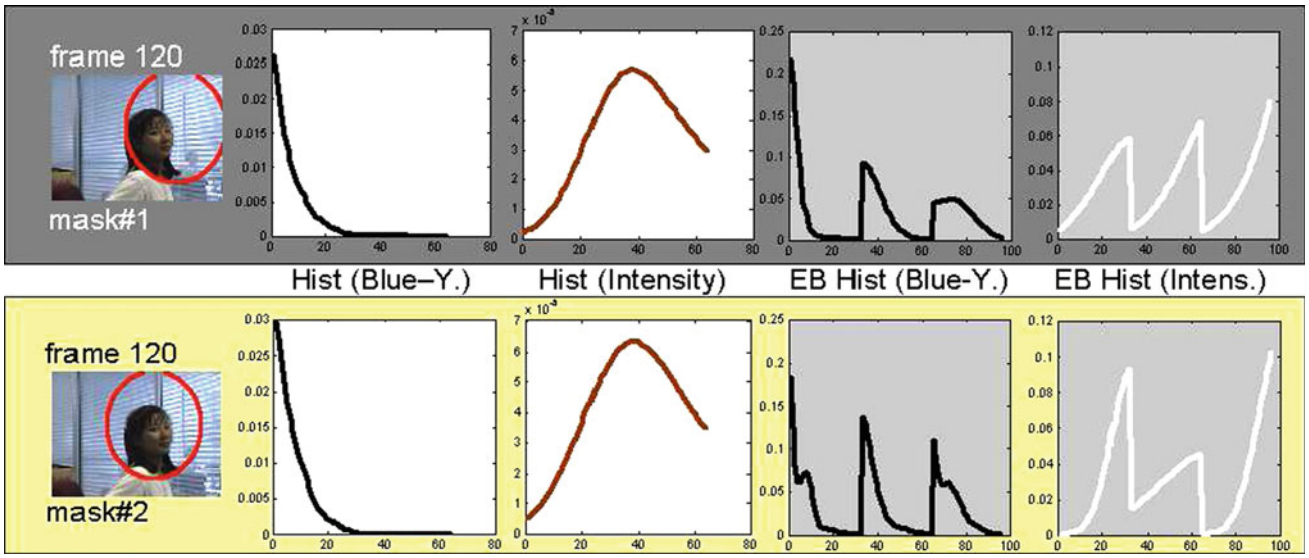


Fig. 10 Including spatial coherence: visually clear differences in the portion of images contained by these two masks (depicted in *red*) are not reflected in the blue-yellow and intensity histograms. EB signatures differ as the masks (and therefore the portion of image inside the mask) also differ (color figure online)

Interesting results are achieved when comparing to spatiograms which, contrary to histograms, do include spatial information. By adding EB analysis to spatiograms, tracking does not lose the head, whereas spatiograms fails. The main advantage is that the spatial codification of the EB signature is neither pure global nor strictly local, but somehow in between. Hence, EB can be a method to add spatial codification of the features, as in the case of polyp detection or tracking with histograms, or complement other methods as spatiograms.

Another point to consider is that the shape of the bands can be adapted for a given application. For instance, in head tracking, extracting a band for the hair, another for the chin and neck and two for the face might help to further improve the performance.

Finally, note that the size of the band in each application is different. For the case of CTC images, the size of the structure element (SE) should be as small as possible, since the objects to detect are indeed small. In tracking, we found small differences using a smaller SE. In this case is enough to include enough bands as to distinguish between face, hair, and surrounding, arguably the important transitions when looking for a head. Smaller SEs would capture the same transitions, but not a bigger one, which, in the limit, would result in the global histogram. Summarizing, the size of the SE have to be chosen in such a way that relevant parts of the object to be detected are not included in the same band, as far as possible.

7 Conclusions

We have shown how EBS improves the discriminative strength of features used for object classification. It has been tested in two different frameworks, namely, false positive reduction in CTC CAD, as well as region-based tracking. The implicit spatial codification of the EB signature better describes an object (polyp, stool and fat tissue or a head in the second experiment) than global or local computation of features. The false positive reduction from a state-of-the-art system which already used intensity is 27.6 % with a minimal loss in sensitivity. This false positive reduction implies one false positive less per volume, which leads to a considerable improvement to radiologists review of CAD findings. We point out that the data used in CTC CAD come from a state-of-the-art CAD system. Consequently, our input is formed by a very challenging data.

Clear increasing in robustness and accuracy of region-based tracking are also clearly better when adding EB analysis of the region. Showing again how EBS can be successfully used to include spatial information, improving detection performance. It occurs also with spatiograms, proposed as a way to include spatial information to histograms.

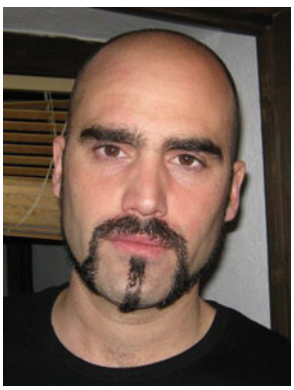
For both frameworks (CTC and tracking) more features could be tested, as shape index in CTC or gradient information for tracking. Nonetheless, results clearly demonstrate the improvement achieved by EBS-based analysis.

References

1. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Trans. Pattern Anal. Mach. Intell. **27**(10), 1615–1630 (2005)
2. Slabaugh, G., Yang, X., Ye, X., Boyes, R., Beddoe, G.: A robust and fast system for ctc computer-aided detection of colorectal lesions. Algorithms **3**(1), 21–43 (2010)
3. Nappi, J., Yoshida, H.: Feature-guided analysis for reduction of false positives in CAD of polyps for computed tomographic colonography. Med. Phys. **30**(7), 1592–1601 (2003)
4. Wang, Z., Liang, Z., Li, L., Li, X., Li, B., Anderson, J., Harrington, D.: Reduction of false positives by internal features for polyp detection in CT-based virtual colonoscopy. Med. Phys. **32**(12), 3602–3616 (2005)
5. Birchfield, S.: Elliptical head tracking using intensity gradients and color histograms, pp. 232–237. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'98) (1998)
6. Birchfield, S.T., Rangarajan, S.: Spatiograms versus histograms for region-based tracking, vol. 2, pp. 1158–1163. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPRć605) (2005)
7. Yoshida, H., Nappi, J.: Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. IEEE Trans. Med. Imaging **20**(12), 1261–1274 (2002)
8. Lefere, P., Gryspeerdt, S., Marrannes, J., Baekelandt, M., Van Holsbeeck, B.: CT colonography after fecal tagging with a reduced cathartic cleansing and a reduced volume of barium. Am. J. Roentgenol. **184**(6), 1836–1842 (2005)
9. Lawrence, E.M., Pickhardt, P.J., Kim, D.H., Robbins, J.B.: Colorectal polyps: stand-alone performance of computer-aided detection in a large asymptomatic screening population. Radiology **256**(3), 791–798 (2010)
10. Zhu, H., Liang, Z., Pickhardt, P.J., Barish, M.A., You, J., Fan, Y., Lu, H., Posniak, E.J., Richards, R.J., Cohen, H.L.: Increasing computer-aided detection specificity by projection features for CT colonography. Med. Phys. **37**(4), 1468–1480 (2010)
11. Pickhardt, P.J.: Translucency rendering in 3D endoluminal CT colonography: a useful tool for increasing polyp specificity and decreasing interpretation time. Am. J. Roentgenol. **183**(2), 429–436 (2004)
12. Levoy, M.: Rendering of Surfaces from Volumetric Data. Technical Report 87-016, Computer Science Department, University of North Carolina (1987)
13. Shahbazi, M., Sattari, M., Ghazic, M.: Automatic polyp detection from CT colonography using mathematical morphology, pp. 823–828. In: Proceedings of the 21st ISPRS Congress (2008)
14. Wang, Z., Li, L., Anderson, J., Harrington, D., Liang, Z.: Colonic polyp characterization and detection based on both morphological and texture features, pp. 23–26. In: Proceedings of the 18th International Congress and Exhibition on Computer Assisted Radiology and Surgery (2004)
15. Sarkar, A., Singh, V., Ghosh, P., Manjunath, B.S., Singh, A.: Efficient and robust detection of duplicate videos in a large database. IEEE Trans. Circuits Syst. Video Technol. **20**(6), 870–885 (2010)
16. Lucas, B., Kanade, T.: An iterative image registration technique with an application to stereo vision, pp. 674–679. In: Proceedings of the International Joint Conference on Artificial Intelligence (1981)

17. Matthews, L., Ishikawa, T., Baker, S.: The template update problem. IEEE Trans. Pattern Anal. Mach. Intell. **26**(6), 810–815 (2004)
18. Elgammal, A., Duraiswami, R., Davis, L.S.: Probabilistic tracking in joint feature-spatial spaces, vol. 1, pp. 773–781. In: IEEE Computer Society Conference on Proceedings of the Computer Vision and Pattern Recognition (2003)
19. Hong, W., Qiu, F., Kaufman, A.: A pipeline for computer aided polyp detection. IEEE Trans. Vis. Comput. Graphics **12**(5), 861–868 (2006)
20. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection, vol. 1, pp. 886–893. In: IEEE Computer Society Conference on Proceedings of CVPR05 (2005)
21. Bowman, A.W., Azzalini, A.: Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-plus Illustrations. Oxford University Press, USA (1997)
22. Sheather, S.J.: Density estimation. Stat. Sci. **19**(4), 588–597 (2004)
23. Botev, Z.I., Grotowski, J.F., Kroese, D.P.: Kernel density estimation via diffusion. Ann. Stat. **38**(5), 2916–2957 (2010)
24. De Silva, V., Tenenbaum, J.B.: Global versus local methods in nonlinear dimensionality reduction. Proc. NIPS **1256**, 721–728 (2003)
25. Tenenbaum, J.B., Silva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. Science **290**(5500), 2319–2323 (2000)
26. Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput. **15**(6), 1373–1396 (2003)
27. Yang, L.: Distance Metric Learning: A Comprehensive Survey. Technical report. Michigan State University, East Lansing (2006)
28. Sangwine, S.J., Horne, R.E.N.: The Colour Image Processing Handbook. CRC Press, Boca Raton (1998)
29. Conaire, C.O., O'Connor, N.E., Smeaton, A.F.: An improved spatiogram similarity measure for robust object localisation, pp. 1069–1072. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2007) (1998)
30. Jain, A.K., Duin, R.P.W., Mao, J.: Statistical pattern recognition: a review. IEEE Trans. Pattern Anal. Mach. Intell. **22**(1), 4–37 (2000)
31. Fisher, R.A.: The use of multiple measurements in taxonomic problems. Ann. Eugen. **7**(1), 179–188 (1936)
32. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)

Author Biographies



Eduard Vazquez received his MSc in 2007 and his PhD in 2011 in the University Autonoma de Barcelona, where he was a teaching assistant in artificial intelligence for five years. Eduard was a research engineer in Medicsight PLC, London. He is currently a scientific research engineer at Cortexica (Imperial College), London. His research interests are color, segmentation, saliency, perception, classification, machine learning, medical imaging, video processing, and object recognition.

Xiaoyun Yang is a Senior Statistician Engineer at Telemetry from January 2012. His previous job is a Senior Researcher of R&D at Medicsight, where he has worked for the last seven years. He earned a PhD in computing from the University of Ulster in Northern Ireland, UK. Xiaoyun has roughly 15 publications (5 journals), and 2 patents pending in the field of medical image processing. Xiaoyun's research interests include Machine Learning, Convex Optimization, Partial Differential Equations, Time Series Analysis, Linear Algebra, and Computer Vision. He is a member of IEEE and has served as reviewer on the IEEE conferences (ICIP) and MICCAI (2011). He is the principal investigator on a grant recently awarded by the UK Technology Strategy Board (TSB) on endoscopic image segmentation.

Greg Slabaugh is a Senior Lecturer (Associate Professor) in the Department of Computing in the School of Informatics at City University, London. Greg has a broad background in computer vision, medical image processing, computational geometry, and video games. He earned a PhD (highest honours) in Electrical Engineering from Georgia Institute of Technology in Atlanta, GA. Greg has over 10 years' experience working in industry, holding research positions at small and large companies, including Medicsight, Siemens, Hewlett-Packard, and Friendly Software. Greg has over 70 publications, 18 granted patents, and over 40 patents pending. He is a Senior Member of IEEE and served as an Associate Editor of IEEE Signal Processing Magazine from 2007–2012.