Int. J. Appl. Math. Comput. Sci., 2010, Vol. 20, No. 3, 545–553
DOI: 10.2478/v10006-010-0040-4

amcs

SURROGATE DATA: A NOVEL APPROACH TO OBJECT DETECTION

ZBISŁAW TABOR

Institute of Applied Informatics
Cracow University of Technology, Al. Jana Pawła II 37, 31–864 Cracow, Poland
e-mail: ztabor@pk.edu.pl

In the present study a novel method is introduced to detect meaningful regions of a gray-level noisy images of binary structures. The method consists in generating surrogate data for an analyzed image. A surrogate image has the same (or almost the same) power spectrum and histogram of gray-level values as the original one but is random otherwise. Then minmax paths are generated in the original image, each characterized by its length, minmax intensity and the intensity of the starting point. If the probability of the existence of a path with the same characteristics but within surrogate images is lower than some user-specified threshold, it is concluded that the path in the original image passes through a meaningful object. The performance of the method is tested on images corrupted by noise with varying intensity.

Keywords: surrogate data, optimal paths, fuzzy connectedness.

1. Introduction

Recognition, description, classification, and grouping of patterns are tasks frequently arising in a variety of disciplines such as biology, medicine and computer science. Given a pattern, its recognition can be formulated in one of the following ways (Watanabe, 1985):

(i) supervised classification, in which the input pattern is identified as a member of a predefined class,

(ii) unsupervised classification, in which the pattern is assigned to an *a priori* unknown class.

In the former case the classes are defined by the system designer, while in the latter the classes are learned during the operation of the recognizing system based on similar patterns. The four best known approaches to pattern recognition include:

(i) template matching (Brunelli, 2009),

(ii) statistical classification (Fukunaga, 1990),

(iii) syntactic matching (Fu, 1982),

(iv) neural networks (Ripley, 2008).

Each method of pattern recognition has its advantages and drawbacks and, to date, no single method delivers the best result for an arbitrary input image. In fact, what means "the best" depends even on the data that have to be extracted from the analyzed image. No matter which of the above listed methods is chosen, it must be trained using an appropriate training set before being applied to classify unknown patterns. The performance of a classifier strongly depends on both the number of training samples and their diversity. A proper choice of the training set is always a very delicate issue which influences the future performance of the classifier. Poor performance and a poor generalization ability of a classifier can result from both too small or too large the size of the training set. In the first case the number of features can be too large in relation to the number of training samples. In the second case the classifier may become overtrained, that is, too intensively optimized on the training set.

In this study a novel approach is proposed to detect patterns in images. The features extracted from an analyzed image are tested against surrogate images, which constitute the set of training samples. In contrast to other pattern recognition methods, the choice of this training set is a well-defined procedure. Each surrogate image is a randomization of the original one but shares some features (like the gray-level histogram and the power spectrum) with it. The generic flow-chart of the proposed pattern-recognition system in depicted in Fig. 1. The user specifies features that will be computed for subsets of each surrogate image as well as classes of the subsets that would be considered. Then probability density functions for the

features values are estimated from surrogate data in the course of training. In the testing step, feature values are computed for subsets of the original image. Subsets are marked as meaningful if feature values are computed, that are sufficiently rare in surrogate data.

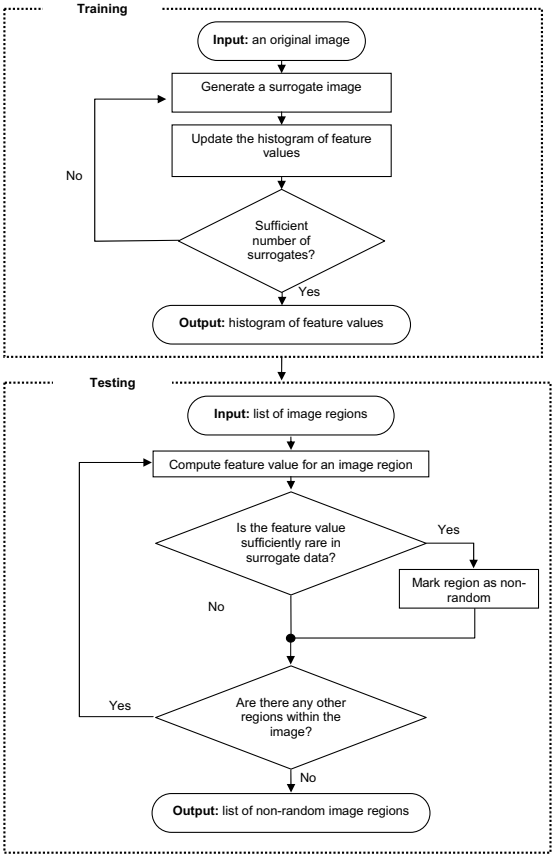

Fig. 1. Flow-chart of a generic pattern-detection procedure, utilizing surrogate data.

The presented approach is motivated by the surrogate data method (Theiler *et al.*, 1992; Schreiber and Schmitz, 2000) used in the analysis of time series. This method is reviewed in the next section. To test the presence of non-random patterns, fuzzy connectedness (Udupa and Saha, 2003) was selected in this study to quantify the strength of connectedness between image pixels. Existing fuzzy connectedness-based approaches to segmentation are, however, interactive—seeds must be specified to launch segmentation. In contrast, the presented approach is a non-interactive one. Fuzzy-connected components are extracted from an analyzed image by specifying a single parameter, related to the probability of the occurrence of some random patterns in surrogate images. In that way the detection of an object becomes similar to human perception—the content of an image is compared with some model of randomness without referring to a separate set of training images. Fuzzy connectedness of objects so defined must be high enough to be exceptionally rare in randomized images. Adopting the presented approach, one avoids to some extent the difficult problem of image denoising before object recognition (Buades *et al.*, 2005). The concept of fuzzy connectedness is described in Section 3. The application of surrogate data method to object detection is described in Sections 4 and 5 and some experimental results are presented in Section 6.

2. Background on surrogate data methods

The surrogate data method was developed to test the presence of nonlinearities in time series data (Schreiber and Schmitz, 2000; Theiler *et al.*, 1992). Nonlinearities in observed signals can exist because the system generating the signal may include nonlinear components. Therefore imprints of this nonlinearity would persist in the signal in the form of non-trivial patterns, characterized by some features. Before the existence of patterns is concluded, the least interesting explanation for the observed "nonlinearities" should be, however, excluded. Namely, it can be possible that the observed patterns come from a distortion of an originally linear stochastic signal by a nonlinear measurement procedure. Surrogate data testing attempts to find a basic explanation that cannot be ruled out based on the data.

In order to systematically exclude simple explanations for the observed patterns, formal statistical tests were developed. If a task involves the analysis of time series, the null hypothesis may, for example, state that the time series was generated by a Gaussian linear stochastic process. In an attempt to reject the null hypothesis, the probability is estimated that some parameter taken on the observed data is equal with some precision to the value actually measured. The probability density function of the measured parameter is known only in exceptional cases, but it can be estimated by a Monte Carlo resampling technique, a method known in the nonlinear science literature as the method of surrogate data.

As an example consider the simplest null hypothesis that the data consist of independent draws from a fixed probability distribution. Although in the case of time series analysis this hypothesis is not a very interesting one, gray-level intensities of image pixels are commonly assumed to be independent identically distributed random variables. Surrogate data for such a null hypothesis can be simply obtained by randomly shuffling the measured data. If significantly different correlations are found in the original and in the shuffled data, then the null hypothesis can be rejected.

The method of surrogate data tries to impose the desired structures onto randomized data. The weak point of the method is that the set of null hypotheses is limited because imposing arbitrary constraints onto otherwise randomized data can be intractable in a general case. The-

re are a few classes of null hypotheses described in the surrogate data literature, but here the most interesting and still computationally tractable one is considered—the hypothesis of the rescaled Gaussian linear process.

The rescaled Gaussian linear process null hypothesis states that deviations from the normal distribution are due to the action of an invertible, static measurement function. An appropriate method of generating surrogate data in this case—the method of iteratively refined surrogates—works as follows. Let $|S_k|^2$ be Fourier amplitudes of the original data $s_0, s_1, \ldots, s_{N-1}$, defined as

$$|S_k|^2 = \left| \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} s_n e^{i2\pi kn/N} \right|^2 . \tag{1}$$

Let c_k be a copy of the original data sorted by magnitude. At each iteration step (j) one has a sequence $\{\overline{r}_n^{(j)}\}$ that has the correct distribution of values as c_k and a sequence $\{\overline{s}_n^{(j)}\}$ that has the correct Fourier amplitudes $\{|S_k|^2\}$. One can start iterations with $\{\overline{r}_n^{(0)}\}$ being a random shuffle of $\{c_k\}$. The step $\overline{r}_n^{(j)} \rightarrow \overline{s}_n^{(j)}$ is a kind of filter in the Fourier domain. First, one takes the Fourier transform of $\{\overline{r}_n^{(j)}\}$:

$$\overline{R}_k^{(j)} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \overline{r}_n^{(j)} e^{i2\pi kn/N}. \tag{2}$$

Before transforming back, the actual amplitudes $|\overline{R}_k^{(j)}|$ are replaced by $|S_k|^2$, but the phases $e^{i\psi_k^{(j)}} = \overline{R}_k^{(j)}/|\overline{R}_k^{(j)}|$ are kept:

$$\overline{s}_n^{(j)} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{i\psi_k^{(j)}} |S_k| e^{-i2\pi kn/N}. \tag{3}$$

The step $\overline{s}_n^{(j)} \rightarrow \overline{r}_n^{(j+1)}$ proceeds by rank ordering:

$$\overline{r}_n^{(j+1)} = c_{\mathrm{rank}\left(\overline{s}_n^{(j)}\right)}. \tag{4}$$

It can be proven that the iteration is attracted to a fixed point $\overline{r}_n^{(j+1)} = \overline{r}_n^{(j)}$ and for finite N the fixed point is reached after a finite number of iterations (Schreiber and Schmitz, 2000).

3. Optimal paths

The 4-neighbourhood of a pixel p contains by definition pixels which share an edge with p. A path $P(p,q)$ between p and q is a sequence of $n > 1$ pixels $(p = p_1, \ldots, p_n = q)$ such that any two successive pixels of the sequence are adjacent, according to the assumed neighborhood relation.

Fuzzy connectedness of pixels p and q is defined as in the work of Rosenfeld (1983). For two selected pixels p and q there are multiple possible paths connecting them. A maximum gray-level value along some path $P(p,q)$ is denoted by $C_P(p,q)$. Then the fuzzy connectedness $C(p,q)$ is equal to the minimum of $C_P(p,q)$ over all possible paths connecting p and q, that is (Rosenfeld, 1983)

$$C(p,q) = \max_{P(p,q)} \left(C_{P(p,q)} \right). \tag{5}$$

A path $P(p,q)$ such that gray-level intensities at all its pixels are not lower than $C(p,q)$ is called an optimal path. According to that definition of optimal paths, pixels contained within objects of interest are connected by relatively dark (low gray-level intensities) paths.

The procedure evaluating $C(p,q)$ is based on Dijkstra's algorithm (Cormen *et al.*, 1990). The definition of an optimal path can be eventually further elaborated if necessary. An image can be transformed to a graph with pixels being graph nodes and graph edges present between every pair of neighboring pixels. Then the weights, computed according to some user specified rules, can also be tied to the edges of the graph, in which case the definition of a path should be appropriately modified. Then, the computation of an edge weight can be, for example, based on the notion of affinity, reviewed by Udupa and Saha (2003).

4. Surrogate data for image segmentation

The method of generating surrogate images was directly based on the iteratively refined surrogate method for time series. The original power spectrum was eventually filtered before using it in the surrogate data generating procedure. Butterworth and Gaussian filters were applied to filter the power spectrum and the width of the filters was a user-adjusted parameter. Varying the width of the filter between the limiting values 0 and infinity is a means of interpolating between pure iteratively refined surrogates and pure random shuffling, respectively. At the output of the surrogate generating procedure an image was obtained with the histogram of gray-level intensities identical to that of the original image and the power spectrum equal to the filtered power spectrum of the original image. The pseudocode of the surrogate generating procedure is presented in Fig. 2 for 2D implementation.

In surrogate images the spanning trees of optimal paths were generated for every root pixel such that its gray-level intensity was in the range from I_{LOW} to I_{HIGH}. Optimal path construction was abandoned whenever the length of the path became larger than L. Generally, I_{LOW} was equal to the minimal gray-level intensity within an image and I_{HIGH} was equal to the maximal gray-level intensity. In all experiments we have, however, chosen I_{HIGH} smaller than the maximal gray-level intensity within an image but higher than the maximal gray-level intensity expected for the objects of interest (assumed, according to the adopted definition of an optimal path, to

Input:
1. 1D array ***data*** containing N_x*N_y gray-level values of the original image sorted in ascending order;
2. 2D N_x x N_y real array ***amplitudes***, containing a filtered power spectrum of the original image;
3. 2D N_x x N_y complex array ***r***, with zero imaginary entries and real entries taken from ***data*** in random order.
4. error tolerance ε

Auxiliary variables:
1. 2D N_x x N_y real array ***angles***;
2. 2D N_x x N_y complex array ***s***;
3. 2D N_x*N_y x 3 real array ***dum***;
4. Real number ***err***.

Main loop:
While (err > ε)
- **Calculate FFT** of ***r***
- **For each** complex entry ***r***[i][j] of ***r*** compute:
  1. angle ***ϕ*** of ***r***[i][j], set ***angles***[i][j]=***ϕ***
  2. amplitude ***a*** of ***r***[i][j].
  3. Increase ***err*** by |***a***-***amplitudes***[i][j]|
- **For each** complex entry ***s***[i][j] of ***s*** set:
  1. Re(***s***[i][j])=***amplitudes***[i][j]*cos(***angles***[i][j])
  2. Im(***s***[i][j])=***amplitudes***[i][j]*sin(***angles***[i][j])
- **Calculate inverse FFT** of ***s*** (then, for all i,j, Im(***s***[i][j])=0) and:
  1. Insert Re(s[i][j]), i j to ***dum***[N*i+j][0], ***dum***[N*i+j][1] and ***dum***[N*i+j][2], respectively, in random order;
  2. Sort ***dum*** in an ascending order using dum[*][0] as the key.
- **For each** i,j set:
  1. Im(***r***[i][j])=0;
  2. Re(***r***[***dum***[N*i+j][1]][***dum***[N*i+j][2]])=***data***[N*i+j][0].

Output :
Surrogate image = r

Fig. 2. Pseudocode of the surrogate data generating procedure. Input and output images are rectangular matrices with N_x rows and N_y columns.

be dark on a bright background). Setting I_{HIGH} in this way is not a limitation of the model because I_{HIGH} needs not be carefully tuned. Choosing I_{HIGH} smaller than the maximal gray-level intensity within an image is primarily a way of saving computational time. L can be set to the length of the longest path which can be generated within an image, which is simply equal to the number of image pixels. In all experiments we have used smaller values of L, of the order of a few dozens, that is, a typical size of the objects of interest. The pseudocode of the procedure generating optimal paths is presented in Fig. 3.

Based on the constructed optimal paths, an extended histogram $H(I_{START}, D, I_{END})$ was computed. The value of $H(I_{START}, D, I_C)$ was equal to the number of pixels connected by an optimal path of length D and fuzzy connectedness equal to I_C to some seed pixel with gray-level intensity equal to I_{START}. Each histogram entry $H(I_{START}, D, I_C)$ defines the conditional probability $P_{I_{START},D}(I_C)$ that a path of length D, starting from a pixel having the gray-level intensity equal to I_{START} and such that the gray-level intensities at all path pixels do not exceed I_C exists in surrogate data.

A sample image and cumulative probability values $F_{I_{START},D}(I_C)$, calculated from the histogram $H(I_{START}, D, I_C)$ corresponding to different values of I_{START}, D and I_C, are presented in Fig. 4. The gray-level values of the image span the range from 100 to 230. The gray-level values of the dark elongated structures to be detected are lower than 160, but a lot of noise artifacts can also be found within this range of intensities. The size of the image is 128×128 pixels and the length of the structures is larger than about 20 pixels. It follows from the figure that the probability of the existence of an optimal path of length 20 and fuzzy connectedness equal to 155, starting from a pixel with gray-level intensity 125, is very low (less than 0.001) and approximately 50 % of optimal paths of length 20, starting from seed pixels with gray-level intensity 125, have fuzzy connectedness lower than 180. This means that if a path of length larger than 20 pixels can be found in the original image such that gray-level intensities of all pixels of the path are below 155, the path is almost certainly a non-random object.

Input:
1. 2D N_x x N_y integer array ***image***, containing a surrogate data image;
2. 3D L x G x G integer array ***histogram***, with all entries set to 0.

Auxiliary variables:
1. 2D N_x x N_y integer array d
2. 2D N_x x N_y integer array m
3. 2D N_x x N_y integer array t
4. 1D integer array ns of G elements
5. 1D integer array ne of G elements
6. 3D G x * x 2 integer array Q
7. integer M
8. integer ***N***, is, js, in, jn, $0 \le i \le N_x$, $0 \le j \le N_y$

For every $I_{LOW} \le$ ***image***[i][j] $\le I_{HIGH}$ **do**:
- **Set**:
  1. all entries of ***d***, ***m***, ***ns***, ***ne***, ***Q*** to zero
  2. ***N***=1
  3. ***M*** =***image***[i][j]
  4. ***Q*** [***M***] [0] [0]=i
  5. ***Q*** [***M***] [0] [1]=j
  6. ***m*** [i] [j]=1
  7. ***t*** [i] [j] = ***image*** [i] [j]
  8. ***ne*** [***M***]=1

- **While** ***N*** > 0 **do**:
  1. is = ***Q*** [***M***] [***ns***[***M***]] [0]
  2. js = ***Q*** [***M***] [***ns***[***M***]] [1]
  3. ***N***--, ***ns*** [***M***]++
  4. ***histogram*** [***d***[is][js]] [***image***[is][js]] [***t***[is][js]]++
  5. For every neighbor (in,jn) of (is,js) such that ***m*** [in] [jn] = 0
     a. ***N***++
     b. ***m*** [in] [jn] = 1
     c. ***d*** [in] [jn] = ***d*** [is] [js] + 1
     d. **if** (***image***[in][jn]>***t***[is][js]) ***t***[in][jn]=***image***[in][jn] **else** ***t***[in][jn]=***t***[is][js]
     e. ***ne***[***image***[in][jn]]++
     f. ***Q*** [***image***[in][jn]] [***ne***[***image***[in][jn]]-1] [0]=in ***Q*** [***image***[in][jn]] [***ne***[***image***[in][jn]]-1] [1]=jn
     g. ***if*** (***image***[in][jn] < ***M***) ***M***= ***image***[in][jn]
  6. **While** (***ne***[***M***] - ***ns***[***M***] ≤ 0) ***M***++

Output:
integer array ***histogram***

Fig. 3. Pseudocode of the optimal paths generating procedure. The gray-level intensities of an input image range from 0 to G. The maximal length of the analyzed optimal paths is equal to L. In the 3D matrix ***p***, queues of pixels with identical gray-level intensities are stored. The length of each queue, corresponding to any given gray-level intensity, can be statically bounded from above by simply analyzing the entries of the gray-level histogram of the original image. Each queue is managed according to the FIFO policy.

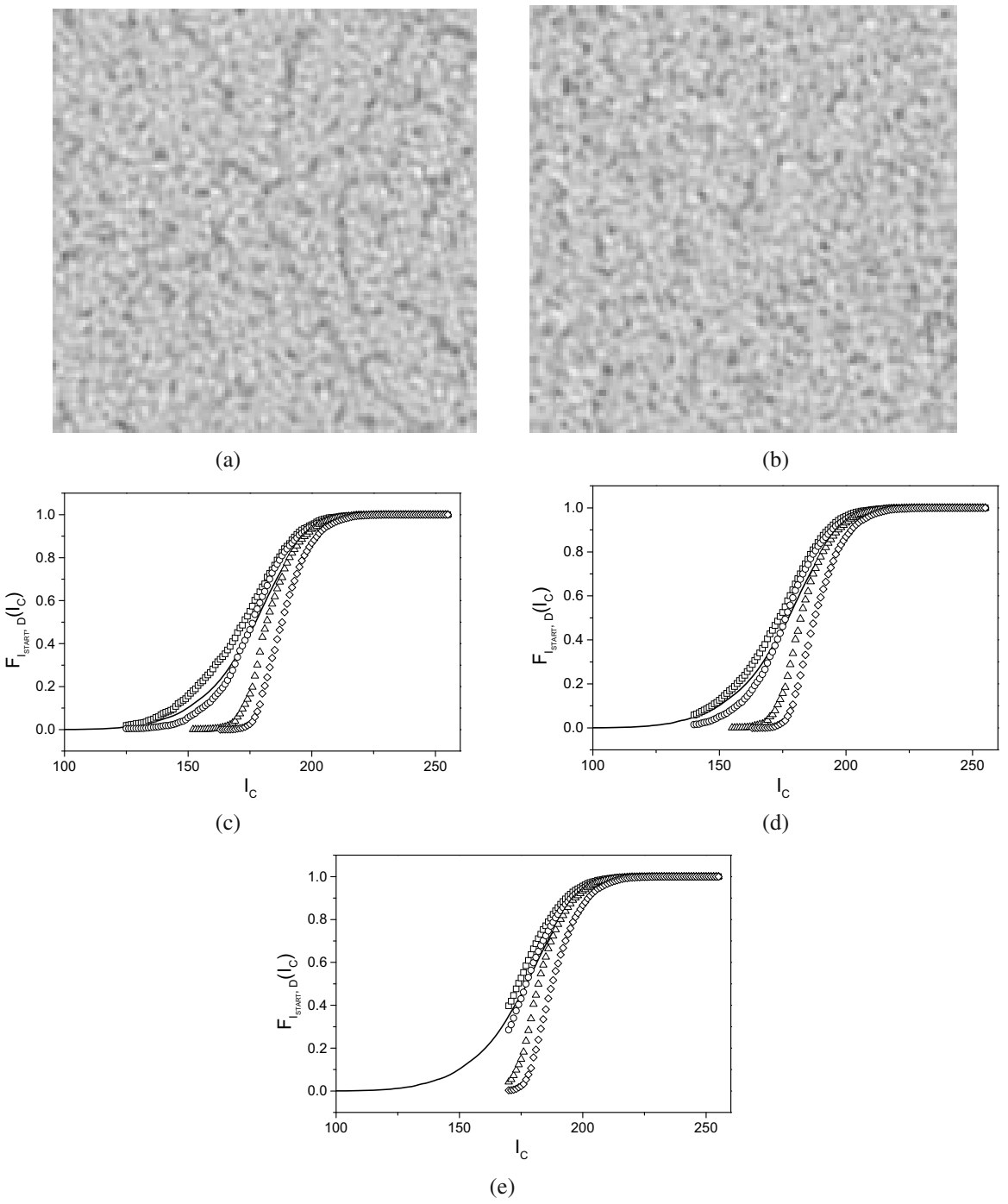

Fig. 4. Sample image (a), surrogate image generated from (a) (b), cumulative probability density functions for paths starting from pixels with gray-level intensity equal to 125 (c), 140 (d), 170 (e), length equal to 3 (squares), 5 (circles), 20 (triangles), 40 (diamonds), and fuzzy connectedness equal to I_C. The cumulative probability density function of the gray-levels distribution is shown for comparison (solid line).

5. Fuzzy object definition

The method of surrogate data and, in particular, the values of $F_{I_{START},D}(I_C)$ for different I_{START}, D and I_C are used to define significant regions of interest. These non-random regions which we are going to detect are composed of pixels connected by optimal paths with lengths, fuzzy connectedness values and gray-levels of terminating points which are sufficiently rare in surrogate data. For example, if in Fig. 4(a) an optimal path which starts from a pixel with gray-level intensity 125, has the length of 20 and fuzzy connectedness of 135 can be found, then we can conclude that the path is contained within some non-random region (possibly consisting of this path only), because the probability of the existence of such a path in surrogate data is extremely low.

To detect such "non-random" regions of an image the notion of an $\alpha - R$ subset of an image is introduced in the present study. For a selected significance level α, the $\alpha - R$ subset of an image I is a union of paths defined as follows:

$$\alpha - R = \bigcup \Big\{ P(p,q) : \min \big(F_{I(p),D(P(p,q))}(C(p,q)), F_{I(q),D(P(q,p))}(C(q,p)) \big) \le \alpha \Big\}. \quad (6)$$

In the definition of $\alpha - R$, $I(p)$ denotes gray-level intensity at pixel p, $D(P(p,q))$ is the length of the path $P(p,q)$, $C(p,q)$ is the fuzzy connectedness defined in Eqn. (5) and F is the cumulative density distribution function, computed from the extended histogram H. The objects of interest are then defined as the connected components of $\alpha - R$. The flow-chart of the procedure detecting the $\alpha - R$ subset of an image I is presented in Fig. 5. This procedure is to a large extent based on the optimal paths generating procedure. Because, however, this procedure creates in fact a spanning tree of optimal paths, in order to recover significant paths from this tree it is necessary to store some additional data. In particular, if some pixel $p((i_n, j_n)$-th entry of ***image*** matrix in Fig. 3) is added to queue ***Q*** after scanning the neighborhood of some pixel $q((i_s, j_s)$-th entry of ***image***), then a pointer to q is stored in the (i_n, j_n)-th entry of an auxiliary matrix ***parent***. Then, after detecting an optimal path in the spanning tree, these data can be used to trace back the path up to the root. In Fig. 5 the simplest implementation of the $a - R$ region detecting procedure is given. In fact, significant paths can be detected and traced back already during the stage of spanning tree construction. Also, it is not always necessary to trace a path back to the root. Because of the tree topology of the set of optimal paths, it is in fact sufficient to trace a path up to the nearest already marked (that is, belonging to $a - R$ region) pixel.

Clearly, for each $\alpha \le \alpha'$ one has $\alpha - R \subseteq \alpha' - R$, that is, the $\alpha - R$ region detecting method belongs to the class of region growing algorithms. In contrast to other fuzzy-connectedness algorithms, the proposed method needs no seed points to be launched. The novelty of the presented approach results from the fact that each time a region grows, a whole path of pixels is added. The growth is governed by a single parameter—a float number, which has a direct interpretation on statistical grounds.

Input:
1. 2D N_x x N_y integer array ***image***, containing original gray-level intensities;
2. 3D L x G x G float array ***histogram*** containing calculated probabilities F;
3. 2D N_x x N_y integer array ***α-Region*** with zero entries;
4. a float number α - significance level.

Auxiliary variables:
1. used is the optimal path generating procedure;
2. 3D N_x x N_y x 2 integer array parent.

For every pixel p such that its gray-level intensity I(p) fulfills the condition: $I_{LOW} \le I(p) \le I_{HIGH}$ **do**:
1. Generate spanning tree T of optimal paths with lengths not larger than L
2. **For** every pixel $q \in$ T **do**:
   if ($F_{I(p),D(P(p,q))}(C(p,q)) \le \alpha$)
   i. use ***parent*** to trace path back up to p
   ii. mark each entry of ***α-Region*** corresponding to pixels along the path

Output:
integer array ***α-Region***

Fig. 5. Pseudocode of the meaningful regions detecting procedure.

By construction, the $\alpha - R$ region detecting method is not sensitive to manipulations like contrast or brightness changes or even nonlinear transforms of gray-level values as long as such transforms are reversible. Indeed, the surrogate data method has been devised exactly to test if an observed signal is in fact random but distorted by a reversible but otherwise arbitrary nonlinear measurement procedure (Schreiber and Schmitz, 2000; Theiler *et al.*, 1992). In the next section it will be examined how the proposed definition conforms to real objects of interest, seen in different classes of analyzed images.

6. Experiments

The proposed method was applied to the analysis of gray-level images of binary structures. The performance of the method was compared with the Markov Random Field (MRF) classifier, implemented in the ITK library (http://www.itk.org) and available as the `ScalarImageMarkovRandomField1.exe` executable file. This implementation of the Markov Random field classifier requires at input an image containing initial classes, the number of iterations, smoothing factor, the number of classes and mean gray-level intensities for the classes.

The first example is presented in Fig. 6. The parameters used for generating surrogate data were as follows:

$I_{LOW} = 40$, $I_{HIGH} = 130$ (this range of gray-level intensities contained 60% of pixels, which is much larger than the area of the structure), $L = 40$. In the figure, the $\alpha - R$ region for $\alpha = 0.02$ is presented. Initial classes for the MRF classifier were obtained by simple thresholding of the original image, using a threshold equal to 125. Two classes were suggested to the MRF classifier, with means equal to 95 and 130. The number of iterations of the MRF classifier was not larger than 20 and the smoothing factor was in the range from 0.5 to 1.5. For the tested values of input parameters the proposed method was superior to the MRF classifier as it better detected thin and bright struts, present in the network.

The second example of an image of a binary structure—retinal vasculature—is presented in Fig. 7. The parameters used for generating surrogate data were as follows: $I_{LOW} = 110$, $I_{HIGH} = 130$ (this range of gray-level intensities contained 29% of pixels), $L = 40$. In the figure, the $\alpha - R$ region for $a = 0.002$ is presented. Initial classes for the MRF classifier were obtained by thresholding the original image, using a threshold equal to 135. Two classes were suggested to the MRF classifier, with means equal to 128 and 140, and the smoothing factor was equal to 0.5. For the presented low noise images (Figs. 6 and 7) it was sometimes advantageous to filter the power spectrum, before generating surrogates. After the filtering, surrogates were generated at smaller computational time (the convergence of iterative refining was faster). Generally, there was no difference between results obtained using either a Gaussian or a Butterworth filter.

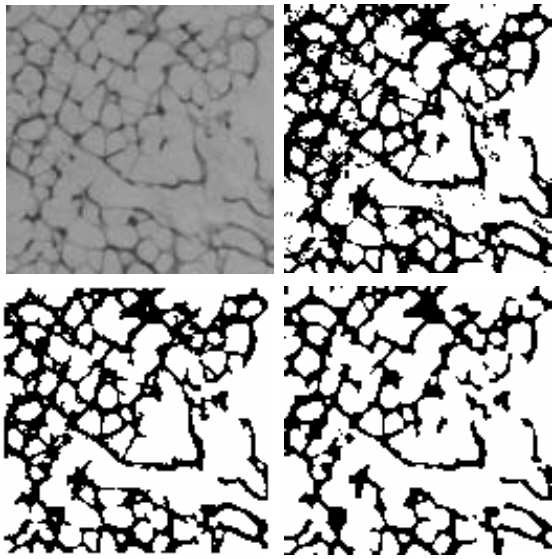

Fig. 6. High-resolution image of the trabecular bone network (top-left), initial classes for the MRF classifier (top-right), the detection of meaningful regions performed, using the developed procedure (bottom-left) and the MRF classifier (bottom-right).

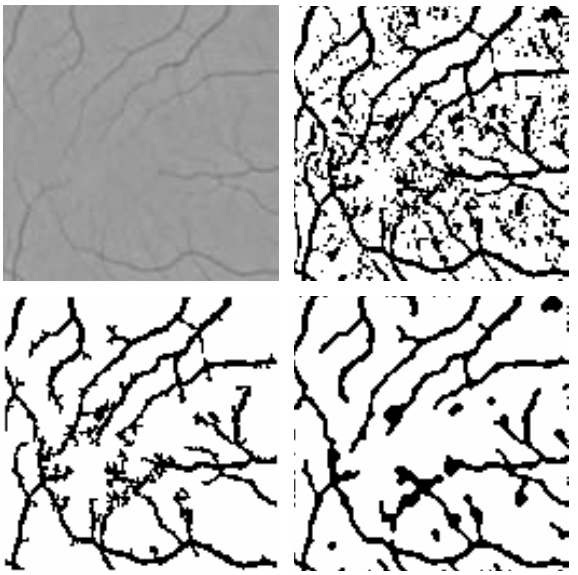

Fig. 7. Image of retinal vasculature (top-left), initial classes for the MRF classifier (top-right), the detection of meaningful regions performed using the developed procedure (bottom-left) and the MRF classifier (bottom-right).

The next example, a computed tomography image of a trabecular bone network, is presented in Fig. 8. The parameters used for generating surrogate data were as follows: $I_{LOW} = 90$, $I_{HIGH} = 170$ (this range of gray-level intensities contained 35% of pixels), $L = 40$. In the figure, the $a - R$ region for $a = 0.005$ is presented. Initial classes for the MRF classifier were obtained by thresholding the original image, using a threshold equal to 170. Two classes were suggested to the MRF classifier, with means equal to 150 and 185. The number of iterations of the MRF classifier was not larger than 20 and the smoothing factor was in the range from 0.5 to 1.5. The next example, a micrograph of a DNA filament, is presented in Fig. 9. The parameters used for generating surrogate data were as follows: $I_{LOW} = 50$, $I_{HIGH} = 130$ (this range of gray-level intensities contained 63% of pixels), $L = 40$. In the figure, the $a - R$ region for $a = 0.0001$ is presented. Initial classes for the MRF classifier were obtained by thresholding the original image, using a threshold equal to 107. Two classes were suggested to the MRF classifier, with means equal to 90 and 130. The other parameters were set as in the previous examples. In both cases the proposed method yields clearly better results than the MRF classifier.

7. Discussion

In the present study a novel method of pattern analysis was proposed. Although there can be various explanations for what a pattern is, generally it is viewed as the opposite of

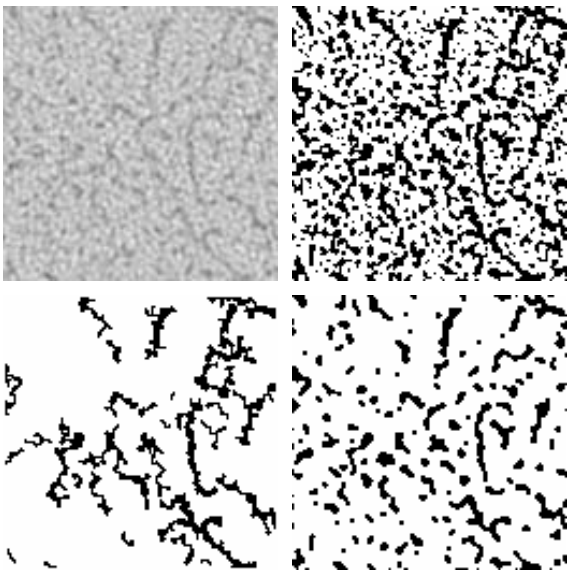

Fig. 8. Low-resolution computed tomography image of the trabecular bone network (top-left), initial classes for the MRF classifier (top-right), the detection of meaningful regions performed using the developed procedure (bottom-left) and the MRF classifier (bottom-right).

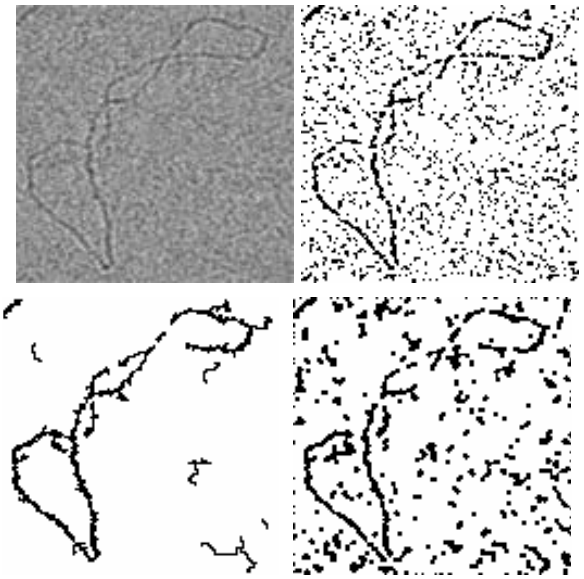

Fig. 9. Micrograph of a DNA filament (top-left), initial classes for the MRF classifier (top-right), the detection of meaningful regions performed using the developed procedure (bottom-left) and the MRF classifier (bottom-right).

randomness (Watanabe, 1985). The present study is motivated by this definition. If a pattern has to be detected in a scene, the first question can be: what will the scene look like if it is random? The surrogate data generating procedure attempts to provide an answer. Then, given a set of surrogate images, one may try to find features which differentiate random images from images containing patterns to be detected.

In this study, gray-level noisy images of binary structures were analyzed and fuzzy connectedness was selected as the feature indicating the presence of objects. The method was not devised for some specified type of noise. The adopted randomization procedure assures that an original image is compared with a random image with the same noise characteristics. It follows from experiments that the real strength of the proposed method lies in detecting meaningful regions in images of such structures, especially the ones strongly corrupted by correlated noise. The choice of the computed feature follows from the apparent importance of the concept of fuzzy connectedness for human perception. The generic fuzzy connectedness procedure is, however, interactive (the selection of some seed points is required) and outputs fuzzy connectedness values, which are relative numbers depending, e.g., on brightness. In contrast, the developed procedure requires no seed points and operates on absolute values—probabilities.

Finally, the method presented in this study is also motivated by the arguments originating from the theory of percolation (Stauffer and Aharony, 1994). Surrogate images are randomized versions of an original image. If there are patterns like gray-level runs of some specific value and length, they would almost certainly disappear in randomized images. If, nevertheless, such gray-level runs are observed in an original image, they are with high probability non-random, possibly meaningful objects.

Acknowledgment

The study was supported by the Polish government grant no. NN518423536.

References

Brunelli, R. (2009). *Template Matching Techniques in Computer Vision: Theory and Practice*, Wiley, New York, NY.

Buades, A., Coll, B. and Morel, J.M. (2005). A review of image denoising algorithms, with a new one, *Multiscale Modeling and Simulation* **4** (2): 490–530.

Cormen, T.H., Leiserson, C.E. and Rivest, R.L. (1990). *Introduction to Algorithms*, MIT Press, Cambridge, MA.

Fukunaga, K. (1990). *Introduction to Statistical Pattern Recognition*, 2nd Edn., Academic Press, New York, NY.

Fu, K.S. (1982). *Syntactic Pattern Recognition and Applications*, Prentice-Hall, Englewood Cliffs, NJ.

Ripley, B.D. (2008). *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge.

Rosenfeld, A. (1983). On connectivity properties of grayscale pictures, *Pattern Recognition* **16**(1): 47–50.

Schreiber, T. and Schmitz, A. (2000). Surrogate time series, *Physica D* **142** (3–4): 346–382.

Stauffer, D. and Aharony, A. (1994). *Introduction to Percolation Theory*, 2nd Edn., Taylor & Francis, Philadelphia, PA.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. and Farmer, J.D. (1992). Testing for nonlinearity in time series: The method of surrogate data, *Physica D* **58**(1–4): 77–94.

Udupa, J.K. and Saha, P.K. (2003). Fuzzy connectedness and image segmentation, *Proceedings of the IEEE* **91**(10): 1649–1669.

Watanabe, S. (1985). *Pattern Recognition: Human and Mechanical*, Wiley, New York, NY.

Zbisław Tabor was born in Czeladź, Poland, in 1970. He received an M.Sc. degree in physics from Jagiellonian University in Cracow, Poland, in 1994, and a Ph.D. degree from the same university in 1999. In 2000, he became an associate professor of biophysics at Jagiellonian University. Since 2006, he has been with the Cracow University of Technology, where he is an associate professor at the Department of Multimedia and Information Systems of the Institute of Applied Informatics. His research interests focus on biomedical engineering, biomedical image analysis, and simulations of biological processes.

Received: 29 October 2009
Revised: 15 March 2010