

Adaptation of multifractal analysis to segmentation of microcalcifications in digital mammograms

Tomislav Stojić^{a,*}, Irini Reljin^b, Branimir Reljin^b

^a*Faculty of Mechanical Engineering, University of Belgrade, 11000 Belgrade, Kraljice Marije 16, Serbia and Montenegro*

^b*Faculty of Electrical Engineering, University of Belgrade, 11000 Belgrade, Bulevar Kralja Aleksandra 73, Serbia and Montenegro*

Received 10 October 2005; received in revised form 10 November 2005

Available online 21 December 2005

Abstract

A method for detecting microcalcifications in digital mammograms is proposed. After recognizing basic features of microcalcifications we introduced several modifications in multifractal analysis, obtaining an efficient method adapted to enhance only small light parts not belonging to surrounding tissue, possibly microcalcifications. Started with a mammogram image, a method creates corresponding multifractal image from which a radiologist has the freedom to change the level of segmentation in an interactive manner and to find suspicious regions, which may contain microcalcifications. Additional postprocessing, based on mathematical morphology, refines the procedure by selecting and outlining regions that contain clusters with microcalcifications. The proposed method was tested through referent mammograms from MiniMIAS database, which is available at public domain. The proposed method successfully extracted microcalcifications in all (clinically approved) cases belonging to this database.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Mammography; Microcalcifications; Image processing; Multifractal analysis; Image segmentation

1. Introduction

Breast cancer is a leading cause of mortality in women population in developed countries [1]. Clinical studies have been demonstrated that survival is greatly improved if the breast anomalies are detected at early stages [2]. One of the significant signs of possible cancerous changes is an existence of small mineral deposits in the breast tissue, usually referred as microcalcifications [3]. Currently, a noninvasive method, known as mammography, is the most effective means for detecting possible breast anomalies, including microcalcifications. Conventional mammography comprises a film-screen X-ray imaging under specific conditions [3,4]. After screening, X-ray films are analyzed by physicians. However, the microcalcifications, having the small size (typically up to a few of millimeters) and low contrast, are very difficult for detection. Thus, diagnosis may be doubtful even for skilled radiologists [3]. Certainly, by digitizing radiology films and applying digital image processing algorithms, significant improvements of image analysis are possible. Moreover, the newest

*Corresponding author. Tel.: +381 64 61 41 088; fax +381 11 33 70 364.

E-mail addresses: tstojic@mas.bg.ac.yu (T. Stojić), ireljin@ptt.yu (I. Reljin), reljinb@etf.bg.ac.yu (B. Reljin).

technique, known as digital mammography, allows direct digitalization: by replacing the X-ray film cassette with appropriate solid-state device, the radiology image is converted directly to digital image. This filmless technology drastically improves the overall screening procedure enabling even lower radiation doses at the same (or better) image quality. However, the detection of microcalcifications still remains a hard job. Conventional contrast enhancement algorithms [5] are not quite appropriate methods since they are globally oriented—they affect all image changes not only small light details like microcalcifications. Recently, the use of mathematical morphology is suggested for deriving an efficient method for local contrast enhancement, being applicable for extracting only small light details without affecting other image details [6].

In this paper, a method for detecting small-sized light details in mammograms, based on multifractal (MF) approach, is proposed. The main premise in our investigation was related to the fact that human tissue is characterized by high degree of *self-similarity* [7]. A broad class of irregularly shaped self-similar objects may be evaluated, described and classified from the fractal point of view [7–9]. Moreover, the anomalies, if any, are then considered as structural “defects”, i.e., as deviations from global regularity of the background. In a MF approach, conversely from the classical image analysis [5], edges are not considered as points where sharp variations of the signal still exist after smoothing, but rather as points whose regularity differs from a background. More general, this approach is capable to describe image features both from local and global points of view. For instance, MF analysis permits the sharp distinction between edge points and isolated points. Although both types of points differ from background, edge points are locally connected while isolated ones are not [10]. Reader may find more examples in [11–14].

The paper is organized as follows. In Section 2 basic principles of multifractal image analysis is briefly described and more attention is devoted to a calculation of MF quantities: the Hölder exponent, α , and the multifractal spectrum, $f(\alpha)$ [8–10]. In Section 3, after analyzing the main characteristics of microcalcifications, we suggest several modifications in calculation of multifractal parameters, leading to the MF method adapted to extraction of isolated light objects in digital mammograms (possible microcalcifications). The efficiency of proposed method is tested through referent mammograms from Mammographic Image Analysis Society (MiniMIAS) database [15] and obtained results are presented in Section 4. Section 5 consists of concluding remarks.

2. Multifractal image analysis

2.1. Multifractal analysis basics

Many natural objects and phenomena exhibit self-similar or fractal property: a structure is assumed as made of parts similar to the whole, exactly or statistically. The parameter characterizing such feature is called the fractal dimension, D_f , a noninteger number, in general, describing how the irregular structure of objects and/or phenomena is replicated in iterative way from small to large scales [7]. The most popular way for calculating the fractal dimension is a box-counting method. The observed structure S is covered by a grid of boxes of size ε and one can count the number of nonempty boxes, $N(\varepsilon)$. The limiting value of $N(\varepsilon)$, when ε tends to zero, follows the power law, $N(\varepsilon) \sim \varepsilon^{-D_f}$, i.e., the fractal dimension is estimated as

$$D_f = - \lim_{\varepsilon \rightarrow 0} \frac{\ln(N(\varepsilon))}{\ln(\varepsilon)}. \quad (1)$$

Fractal structures may be artificially generated by applying an exact rule. Such structures are known as *deterministic* (or, *mathematical*) *fractals* and they are characterized by the same fractal dimension in all scales: a structure is composed of objects whose smaller scales replicate *exactly* their larger ones, up to infinity. Consequently, these structures are referred as *exact self-similar*, or *monofractals*. There are a lot of such structures, for instance, Cantor sets, Koch’s curves, Sierpinski gasket and carpet, etc. [7,8]. Instead, a variety of natural objects, structures, and phenomena, may also exhibit self-similarity but in a statistical way, since the reproduced detail is not an exact copy of the previous. These structures are known as *random fractals*. Furthermore, natural fractals are not self-similar over all scales. There are both upper and lower size limits, beyond which a structure is no longer fractal. Structures characterized by fractal dimension which varies with

the observed scale are known as *multifractals*, and for describing them more sophisticated mathematical quantities are necessary [8–10].

The quantitative description of multifractal property can be derived in several ways. Usually, the procedure starts with finding the noninteger exponent α , known as the *Hölder* exponent, describing the pointwise singularity of the object, and then deriving the distribution of this quantity, known as the *multifractal spectrum*, $f(\alpha)$, as will be briefly described.

Let the structure S be divided into nonoverlapping boxes S_i of size ε such that $S = \cup_i S_i$. Each box S_i is characterized by some amount of measure, $\mu(S_i)$, and boxes may be assumed as measure domains. An equivalent parameter suggested to the MF analysis is defined by

$$\alpha_i = \frac{\ln(\mu(S_i))}{\ln(\varepsilon)}, \quad (2)$$

which is denoted as the *coarse Hölder exponent* of the subset S_i . If ε tends to zero the coarse Hölder exponent approaches to limiting value α at observed point

$$\alpha = \lim_{\varepsilon \rightarrow 0}(\alpha_i). \quad (3)$$

Parameter α depends on the actual position on the fractal and describes *local regularity* of the structure. In the whole structure there usually are many boxes (or points) with the same parameter α . Consequently, the next step is to find the distribution of this quantity, i.e., to find the function $f(\alpha)$, known as the MF spectrum, over subsets characterized by α . This function describes the *global regularity* of observed structure [8–11]. The MF spectrum can be assumed as the *fractal dimension* over the subsets characterized by α

$$f_\varepsilon(\alpha_i) = -\frac{\ln(N_\varepsilon(\alpha_i))}{\ln(\varepsilon)}, \quad (4)$$

where $N_\varepsilon(\alpha_i)$ is the number of boxes S_i containing particular value of α_i . From (4) one can obtain the limiting value

$$f(\alpha) = \lim_{\varepsilon \rightarrow 0}(f_\varepsilon(\alpha)). \quad (5)$$

The MF spectrum $f(\alpha)$ calculated as above is also known as the *Hausdorff dimension* of the distribution of α . Note again that the box-counting is only one of the several possible methods for estimation of the MF spectrum, but due to its simplicity and fast computing procedure this method is the most frequently used [10,11]. Irrespective of particular technique for deriving MF quantities α and $f(\alpha)$, they describe both local and global regularities of the process under investigation. Consequently, MF analysis may be used in a broad class of signal processing problems, as a robust method for describing and/or extracting some features probably hidden in large amount of data [10–14].

2.2. Multifractal analysis in digital image processing

Our investigation is addressed to multifractal-based digital image processing. For this purpose it is necessary to derive some kind of two-dimensional multifractal transform—a procedure enabling bi-directional mapping of pixel values from input image (original domain) to corresponding values of α and $f(\alpha)$ (transform domain). By applying such a procedure it is possible to extract pixels (details) belonging to particular image regions from the multifractal point of view [10].

The classical procedures for finding α and $f(\alpha)$ described by (2)–(5) relate to arbitrary continuous-space continuous-magnitude functions. Our research is addressed to digital gray-scale images which are described by two-dimensional real and nonnegative functions of gray, $g(x,y)$, with discrete spatial coordinates $x = m \cdot \Delta x$ and $y = n \cdot \Delta y$, $m = 1, 2, \dots, M$, $n = 1, 2, \dots, N$, known as pixels, each characterized by discrete K -level magnitude, $g(x,y) = g(m,n) = k \cdot \Delta g$, $k = 0, 1, 2, \dots, (K-1)$. Discrete space introduces several limitations and difficulties in determining of the exponent α and, consequently, its distribution $f(\alpha)$. The basic limitation lies in a discrete nature of box sizes covering the image space—possible size is an integer multiple of the pixel size.

Instead of (2), which holds in a continuous space, in a discrete space each pixel may be characterized by a discrete set of coarse Hölder exponents as

$$\alpha_i(m, n) = \frac{\ln(\mu_i(m, n))}{\ln(i)}, i = 1, 2, 3, \dots, \tag{6}$$

where $\mu_i(m, n)$ is the amount of a measure within observed box size $\varepsilon = i$. It follows that limiting procedure given by (3) and/or (5), meaning $\varepsilon \rightarrow 0$, is not possible, since $\varepsilon_{\min} = 1$. Moreover, α and $f(\alpha)$ are even undefined if $\varepsilon = \varepsilon_{\min} = 1$. Instead, we may use an indirect method for estimating these quantities.

For estimating Hölder exponents we can calculate natural logarithms of measure value, $\ln(\mu_i(m, n))$, and of the box size, $\ln(i)$, and plot corresponding points in bi-logarithmic diagram $\ln(\mu_i(m, n))$ vs. $\ln(i)$. Then, the limiting value of $\alpha(m, n)$ is estimated as a slope of linear regression line.

After finding values of α we may create an ‘ α -image’—a matrix of the same dimension $M \times N$ as initial image but filled by values of $\alpha(m, n)$ with one-by-one correspondence to image pixels. From this matrix the MF spectrum $f(\alpha)$, also in a matrix form, $f(m, n) = f(\alpha(m, n))$, $m = 1, 2, \dots, M$, $n = 1, 2, \dots, N$, is estimated as follows. First, continuous Hölder exponents are discretized into R values of α_r :

$$\alpha_r = \alpha_{\min} + (r - 1)\Delta\alpha_r, \quad r = 1, 2, \dots, R. \tag{7a}$$

We used the uniform division with

$$\Delta\alpha_r = \Delta\alpha = (\alpha_{\max} - \alpha_{\min})/R. \tag{7b}$$

If the actual value of α lies within the subrange r , i.e., if $\alpha_r \leq \alpha < (\alpha_r + \Delta\alpha)$, it is replaced by α_r . Such α -image is covered by a regular grid of boxes with integer box sizes $j = 1, 2, \dots$. The boxes containing at least one value of α_r are counted giving the number $N_j(\alpha_r)$. Boxes of different sizes are recursively taken into account and corresponding Hausdorff measures are calculated for each image pixel according to (4) as¹

$$f_j(\alpha_r) = -\frac{\ln N_j(\alpha_r)}{\ln(j)}, \quad j = 1, 2, \dots \tag{8}$$

Finally, from a set of discrete points in bi-logarithmic diagram of $\ln N_j(\alpha_r)$ vs. $-\ln(j)$, the MF spectrum $f(\alpha)$ is estimated from linear regression, in similar manner as in case of estimation of α . Procedure is repeated for whole α -matrix thus obtaining ‘ $f(\alpha)$ -image’—a matrix of dimension $M \times N$ filled by pixel-wise values of $f(\alpha)$ with one-by-one correspondence with initial image pixels.

Different measures $\mu_i(m, n)$, may be used for estimating α . Some of the most frequently used measures, known as *capacity measures*, are those introduced by Lévy Véhel and Pascal Mignot [10]:

$$\text{Maximum : } \mu_i(m, n) = \max_{(k, l) \in \Omega} g(k, l), \tag{9a}$$

$$\text{Minimum : } \mu_i(m, n) = \min_{(k, l) \in \Omega^*} g(k, l), \tag{9b}$$

$$\text{Sum : } \mu_i(m, n) = \sum_{(k, l) \in \Omega} g(k, l), \tag{9c}$$

$$\text{Iso measure : } \mu_i(m, n) = \text{card}\{(k, l) | g(m, n) \equiv g(k, l), (k, l) \in \Omega\}, \tag{9d}$$

as well as measures referred in [12]

$$\text{Absolute difference : } \mu_i(m, n) = \max_{(p, q), (k, l) \in \Omega} |g(p, q) - g(k, l)|, \tag{9e}$$

$$\text{Central absolute difference : } \mu_i(m, n) = \max_{(k, l) \in \Omega} |g(m, n) - g(k, l)|, \tag{9f}$$

¹The dimension of boxes in (8) is denoted as j , instead of i used in (6), indicating to different domains: i relates to pixel neighborhood in image domain, while j relates to size of regular grid in α -matrix domain.

$$\text{Maximum sum difference : } \mu_i(m, n) = \sum_{(k,l) \in \Omega} |g(m, n) - g(k, l)|, \tag{9g}$$

$$\text{Modified iso measure : } \mu_i(m, n) = \text{card}\{(k, l) \mid |g(m, n) - g(k, l)| < T, (k, l) \in \Omega\}, \tag{9h}$$

where i is a size of a measure domain around observed pixel (m, n) , Ω is a set of all pixels (k, l) within a measure domain, Ω^* is a set of all nonzero pixels within a measure domain, $g(k, l)$ is a gray-scale intensity at point (k, l) , T is a threshold. Certainly, different measures lead to different Hölder exponents and may be used to produce different effects on the processed image. According to our research, we found measures “Maximum” and “Minimum” to be more appropriate for extracting small details, light or dark, respectively. So, these two measures will be considered further.

2.3. Illustrative examples

We will illustrate the described procedure through the concrete example. Consider a gray-scale image $g(m, n)$, $m = 1, 2, \dots, M$, $n = 1, 2, \dots, N$, with $M = N = 128$ (128×128 pixels) and 8 bits per pixel (bpp), as in Fig. 1(a). Given image has actual gray levels ranged from 102 to 200. Let us observe an arbitrary pixel, for instance at position $m = 97$, $n = 78$, labeled by a black cross in Fig. 1(a). In Fig. 1(b) the intensity levels of this pixel (central pixel, gray shaded) and its nearest neighborhood up to $i = 5$ are given.

Since the multifractal analysis operates with values in the range $[0, 1]$, each image pixel is divided by its maximum value (the value of 200 in this case). The normalized pixel values in observed domain are presented in Fig. 2. The measure “Maximum” given by (9a) is considered. The procedure is illustrated in Fig. 2, upper row. The smallest domain, $i = 1$, contains only central pixel, labeled as dark gray in Fig. 2, left column. The measure value in the observed domain is $\mu_1(97, 78) = 0.79$, and its natural logarithm is $\ln \mu_1(97, 78) = \ln(0.79) = -0.2357$. Corresponding point in log-log diagram in Fig. 3(a) is labeled as ‘1’. Next measure domain, $i = 3$, consists of light-gray shaded pixels in Fig. 2, central column. Maximal level in this domain is $\mu_3 = 0.91$ (dark-shaded pixel in upper row), corresponding natural logarithm is $\ln \mu_3(97, 78) = -0.0943$, point ‘3’ in Fig. 3(a). The third measure domain ($i = 5$), Fig. 2, right column, contains maximum intensity of 0.98 (dark-shaded pixel), hence $\ln \mu_5(97, 78) = -0.0202$, point ‘5’. From those three points the value of Hölder exponent at given image pixel is estimated as $\alpha(97, 78) = 0.133$, from the slope of linear regression line.

The use of a measure “Minimum” is illustrated in Fig. 2, lower row, and Fig. 3(b). The characteristic measure values now are: $\ln \mu_1(97, 78) = \ln(0.79) = -0.2357$, $\ln \mu_3(97, 78) = \ln(0.75) = -0.2877$, $\ln \mu_5(97, 78) = \ln(0.74) = -0.3011$, and corresponding points 1, 3 and 5 in log-log diagram are depicted in Fig. 3(b). Since a linear regression line is descending one, the Hölder exponents will be estimated as the absolute value of the slope, and amounts $\alpha(97, 78) = 0.0417$.

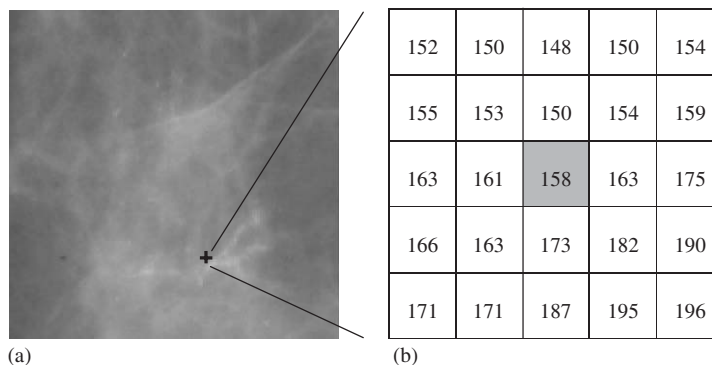


Fig. 1. (a) Gray-scale image of 128×128 pixels and 8 bpp, and (b) pixel intensities corresponding to cross-labeled pixel $g(97, 78)$ and its closest neighborhood.

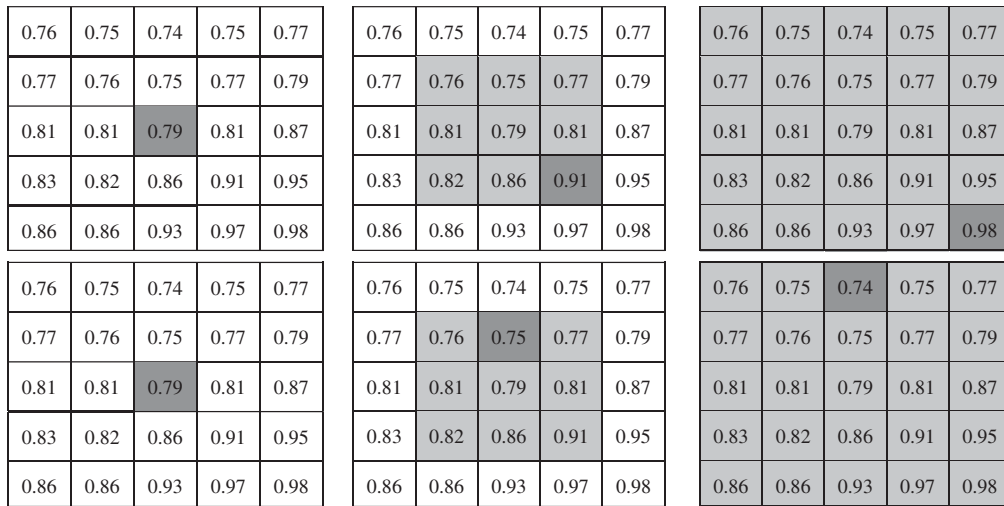


Fig. 2. Illustration of the use of measures “Maximum” (upper row) and “Minimum” (lower row) for different measure domains $i = 1$, $i = 3$, and $i = 5$, light-gray shaded, from left to right. Pixels having maximum/minimum intensity in observed domain are dark-gray shaded.

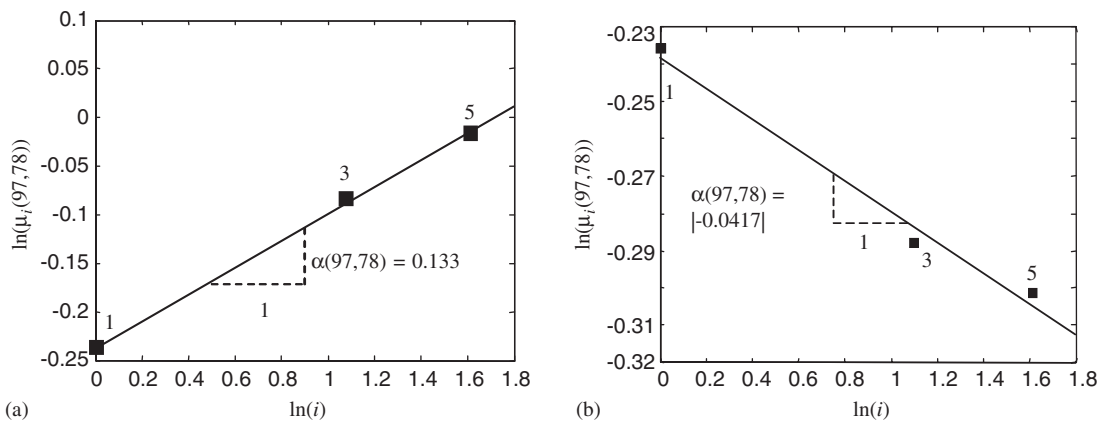


Fig. 3. Illustration of estimation of Hölder exponent α , at given pixel (97,78) from Fig. 1(a), by using measures: (a) “Maximum” and (b) “Minimum”.

After applying the described procedure for a whole image one can obtain Hölder exponents α having one-by-one correspondence with image pixels. For a given example, values of α are ranged from 0 to 0.1572, by applying a measure “Maximum”, and from 0 to 0.1457, for a measure “Minimum”. These α -images are depicted in Fig. 4. For better visualization, α -values are normalized to a full scale [0,1]. Note that the α -images derived in these ways enhance just local contrast, irrespectively of the actual level of gray.

Estimation of MF spectrum $f(\alpha)$ is derived according to (7)–(8). The entire range of actual α values is divided into $R = 100$ equally sized sub-ranges, and then Hausdorff dimension is calculated for each α sub-range, for different boxes sized from $j = 1$ (one ‘ α -pixel’) to $j = 16$. As an example, consider the α -image as in Fig. 4(a), obtained by using a measure “Maximum”, and let us observe its central sub-range α_{50} . ‘Pixels’ from α -image having the actual values $\alpha_{50} \leq \alpha < \alpha_{51}$, i.e., belonging to an interval [0.0770,0.0786) in this case, are replaced by $\alpha_{50} = 0.0770$, and displayed as black dots on white surrounding, as in Fig. 5(a). For given box-dimension j , we count the number of boxes, $N_j(\alpha_{50})$, covering at least one value of α_{50} . For instance, if $j = 16$, i.e., if the whole α -image of 128×128 pixels is covered with regular grid of 64 boxes, one can obtain $N_{16}(\alpha_{50}) = 20$ boxes

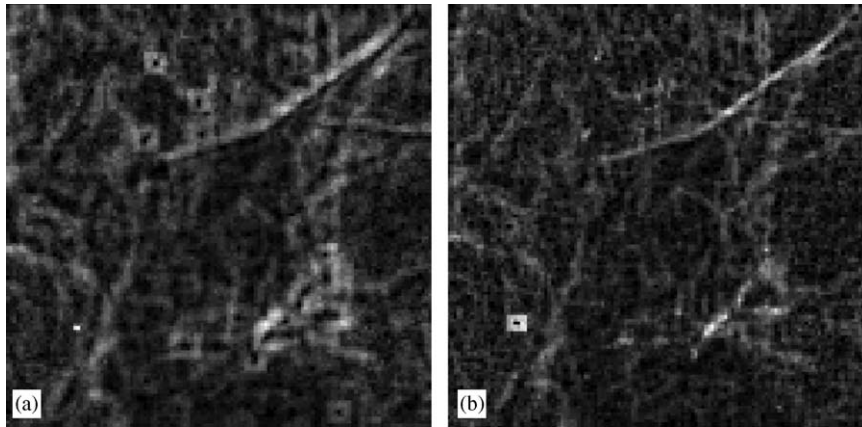


Fig. 4. (a) α -images derived from original image as in Fig. 1(a), by using measures: (a) “Maximum” and (b) “Minimum”. Values of α are displayed in the full magnitude range [0,1], for better visualization.

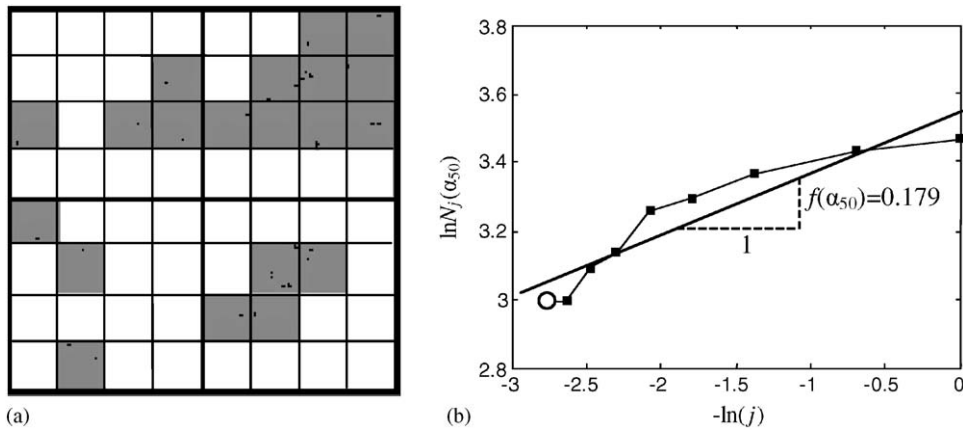


Fig. 5. Estimation of $f(\alpha)$ from α -image in Fig. 4(a). (a) Part of α -image, sized 128×128 , containing ‘pixels’ from α_{50} sub-range (black dots) is covered by 64 regular boxes sized $j = 16$. Boxes containing α_{50} values are gray shaded. (b) Bi-logarithmic diagram for estimating $f(\alpha_{50})$. Point corresponding to Fig. 5(a): $\ln(N_{16}(\alpha_{50}))$ and $-\ln(16)$ is marked by a circle. Black dots correspond to other box sizes ($j = 1, 2, 4, 6, 8, 12$, and 14) in α_{50} -image domain. The slope of linear regression line estimates the value of $f(\alpha_{50}) = 0.179$.

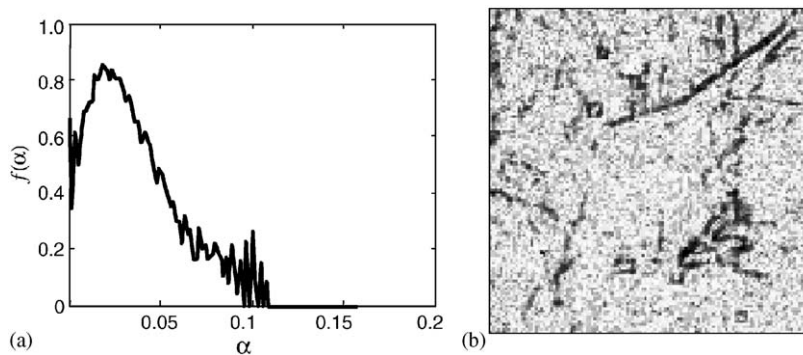


Fig. 6. (a) The multifractal spectrum $f(\alpha)$ of an image from Fig. 1(a), derived from α -image in Fig. 4(a) (measure “Maximum”), assuming $R = 100$ sub-ranges of parameter α . (b) Corresponding ‘ $f(\alpha)$ image’. Actual $f(\alpha)$ values are in range $[0, 0.859]$, but for better visualization image is displayed in the full magnitude range $[0, 1]$.

shaded as gray in Fig. 5(a). The corresponding point with coordinates $[-\ln(16), \ln(20)]$ is marked by a circle in bi-logarithmic plot depicted in Fig. 5(b). By repeating this procedure for all different sized boxes ($j = 1, 2, 4, 6, 8, 10, 12, 14$ and 16 α -pixels) a set of points is obtained as in Fig. 5(b). The box-dimension of α_{50} sub-range, i.e., the value of $f(\alpha_{50})$, is then estimated as a slope of linear regression line and amounts $f(\alpha_{50}) = 0.179$. After applying this procedure for all α_r sub-ranges, $r = 1, 2, \dots, R$, the MF spectrum $f(\alpha)$ has been obtained. Moreover, each image pixel has been described by a pair of $(\alpha, f(\alpha))$ values. For a given example, the MF spectrum $f(\alpha)$ and corresponding $f(\alpha)$ image are depicted in Figs. 6(a)–(b).

3. Adaptation of multifractal analysis to segmentation of microcalcifications

Multifractal approach is already used in digital image processing, mainly for texture classification and segmentation [10–14]. This approach exploits both local regularity of a given measure, described by the pointwise Hölder exponent α , and the global distribution of the regularity in a whole scene, described by $f(\alpha)$. By appropriate choice of a pair α and $f(\alpha)$, different features may be recognized, extracted and even classified, both in geometric and probabilistic sense. However, it must be pointed out that there is not an unique way for describing an image through a pair $(\alpha, f(\alpha))$ since the value of α strongly depends on the measure used—see expressions (9a)–(9h) and examples in previous section. For particular problem, the appropriate measure have to be choosen. Our goal is to extract microcalcifications in digital mammogram and to investigate possible use of MF analysis for this purpose. In radiology image the microcalcifications are seen as small (of size 0.1–1 mm) bright and relative smooth surfaces not belonging to surrounding, with small gray-level variation. Sharp change in gray-level arises just around the edge of microcalcification [16,17]. The presence of this local contrast is de facto necessary condition for the detection of microcalcifications. The size and shape of microcalcifications can significantly vary—12 different types of microcalcifications are reported [3]. Moreover, microcalcifications usually are clustered, with five or more parts each less than 5 mm apart. Recognizing microcalcifications is not easy in many cases. If microcalcifications are located in radiology “dense” tissue, which strongly reflects X-rays, the contrast between them and surrounding tissue is very low, making their detection difficult even for skilled radiologists. In addition, the parts of blood vessels or connective tissue can produce even higher contrast then microcalcifications.

Let us summarize the main features of microcalcifications. They are: (i) small bright parts, (ii) not belonging to background tissue (rare events), with (iii) relatively high local contrast, having (iv) different size and shape, and (v) usually clustered.

We have used features (i)–(v), listed above, as guidelines for adaptation of multifractal analysis targeted to detection and extraction of microcalcifications in radiology images. First, if we had created a MF image with one-by-one pixel-wise correspondence with an original image, we should be able to select original image pixels having high α (high local contrast) and low $f(\alpha)$ (rare events). Consequently, MF image may be used for detecting and extracting microcalcifications, having in mind their features (ii) and (iii). The next question is: which measure is appropriate?

From feature (i), the measure “Maximum” seems to be adequate for extracting bright details. However, determination of Hölder exponents depends on the logarithm of normalized measure, μ . Having in mind the plot of logarithmic curve $\ln(\mu)$, it is evident that local contrast expressed by $\Delta \ln(\mu)/\Delta \mu$ is very high in the dark-level domain (low μ) and very low in the light-level domain (high μ), which is just the opposite to our

Table 1
Comparative values of measures “Maximum” and “Minimum” and their logarithms, for pixel (97,78) of original image as in Fig. 1(a) and its negative

i	1	3	5	α estimated
$\ln(i)$	0	1.0986	1.6094	
$\mu_i = \text{Max}(g_i(97, 78))$	0.79	0.91	0.98	
$\ln \mu_i(97, 78)$	-0.2357	-0.094	-0.0202	0.133
$\mu_i = \text{min}(g_i(97, 78))$	$1-0.79 = 0.21$	$1-0.91 = 0.09$	$1-0.98 = 0.02$	
$\ln \mu_i(97, 78)$	-1.5606	-2.4079	-3.912	1.351

need to enhance small light details. Instead, we suggest to use an inverted (negative) image, described as a complement of an original image: pixels in negative image have normalized values $gn(m, n) = 1 - g(m, n)$. In this way bright anomalies (possible microcalcifications) migrate to dark region being thus under the strong influence of the “logarithmic amplifier”. Consequently, the capacity measure “Minimum”, given by (9b), becomes well suited to describe local image regularity in this domain. This conclusion is easy to approve from comparative values of measures “Maximum” and “Minimum” and their logarithms, for the same example as in Fig. 1(a) and its negative, respectively, given in Table 1. From the negative image and a measure “Minimum”, the estimated value of α at observed point (97,78) amounts $\alpha(97,78) = 1.351$ which is more than 10 times greater than corresponding value derived from the measure “Maximum” applied to the same original image. In general, the effect of “logarithmic amplifier” strongly enhances just small gray-level variations in the dark zone of inverted image ($gn \ll 0.5$). This procedure does not reduce the sensitivity within regions in the middle gray ($gn \cong 0.5$), while within radiology ‘sparse’ tissue a contrast between microcalcifications and surrounding tissue is naturally high and it still remains high enough even after the inverting of the image.

By choosing the shape and size of the measure domain used in MF analysis, particular singularity may be extracted from observed image. In examples described in Section 2.3, square measure domains of different sizes are used. In our research, considering feature (iv), we found that disk and/or diamond-shaped domains are well suited to most of microcalcifications.

Finally, as will be presented in Section 4, by additional morphology postprocessing applied to already MF processed image, only clustered light parts may be extracted satisfying feature (v).

The multifractal spectrum is determined by box-counting method according to (8) as a slope of linear regression line from a bi-logarithmic plot $\ln(N_j(\alpha))$ vs. $-\ln(j)$. Fitting procedure is more accurate if more points are used. The box sizes may vary from $j = 1$ (one α pixel) to $j = \max(M, N)$ (a whole $M \times N$ image). Since our goal is to favor small singularities, i.e., high-frequency components in α distribution, we used small boxes sized $j = 1$ to 16 α -pixels. Namely, by the increasing of the box size the number of the nonempty boxes decreases. However, if the box size is large enough the number of nonempty boxes may remain unchanged through the box size increases. Thus, the saturation effect is obtained: points on bi-logarithmic plot stay on horizontal line. These points significantly reduce the resolution of calculated $f(\alpha)$ values—a slope of linear regression line tends to zero. On the contrary, by using smaller boxes, the number of nonempty boxes significantly varies with the box dimension change, preventing saturation and enabling high resolution of estimated $f(\alpha)$ values.

The number of α subranges, denoted as R , also influences the accuracy of multifractal spectrum. Small number of subranges has an effect of low-frequency filtering, yielding to smooth spectrum but with small resolution and small “sharpness”. Conversely, too much subranges produce saw-toothed (“erratic”) spectrum, with more details. In our research, a compromise solution with $R = 100$ is adopted.

4. Application of multifractal algorithm adapted to segmentation of microcalcifications in digital mammograms

Our research was concentrated to segmentation and visualization of microcalcifications in digital mammograms, as already noted. The efficiency of the method proposed is verified through referent mammograms from public domain MiniMIAS database [15]. Mammograms are sized 1024×1204 pixels with 8 bpp pixel depth, i.e., with gray level from 0 (black) to 255 (white), and with $200 \mu\text{m}$ spatial resolution. The mammograms have been reviewed by a consultant radiologist and abnormalities have been identified and marked. Each individual image contains a set of information: a breast type, type and location of abnormality, and a radius of circle enclosing the abnormality. We have selected cases with calcifications (25 cases in this database), from which several interesting examples will be presented here. For better visualization, only a part (sized 128×128 or 256×256 pixels) of whole mammograms, containing regions with approved microcalcifications, will be examined. For better visualization, all images will be displayed with maximally enhanced contrast—full normalized level range from 0 (black) to 1 (white). Multifractal quantities, α and $f(\alpha)$, are calculated from the procedure described in Section 3. Original images were inverted, first, and capacity measure “Minimum” was used over a diamond-shaped measure domains of sizes 1, 3, and 5 image pixels. The number of α subranges was $R = 100$, and the covering box sizes were $j = 1, 2, 4, 6, 8, 10, 12, 14$ and 16 α -pixels.

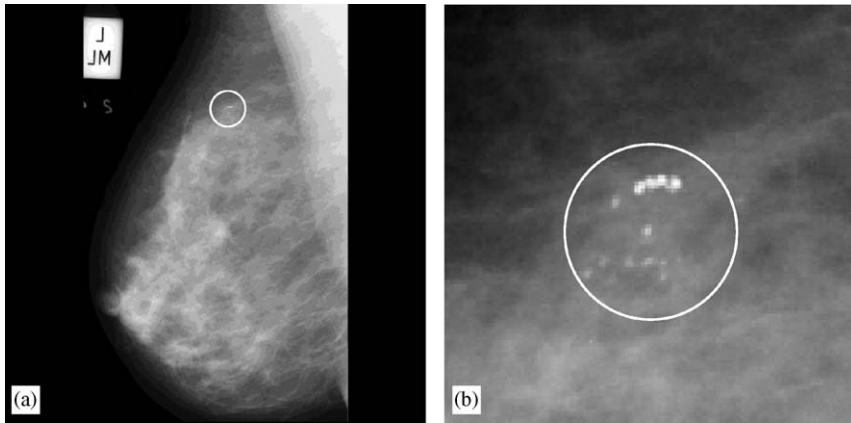


Fig. 7. (a) Original mammogram (1024×1024 pixels) mdb219.pgm from MiniMIAS database, displaying Medio Lateral view of the left breast, and (b) its selected part (128×128 pixels) around clinically approved microcalcifications.

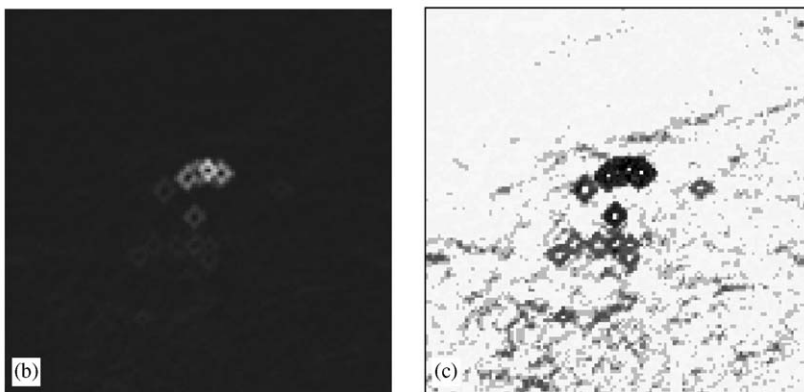
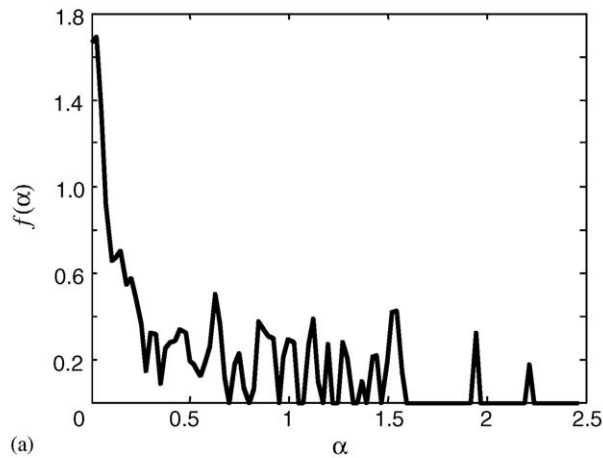


Fig. 8. (a) A MF spectrum of an image from Fig. 7(b) and (b) α image, and (c) corresponding $f(\alpha)$ image. White pixels correspond to high levels of α and $f(\alpha)$.

4.1. Radiology easy case

Fig. 7(a) represents full-sized mammogram (1024×1024 pixels) mdb219.pgm from MiniMIAS database and its part (128×128 pixels) around clinically approved microcalcifications (somewhere under the white

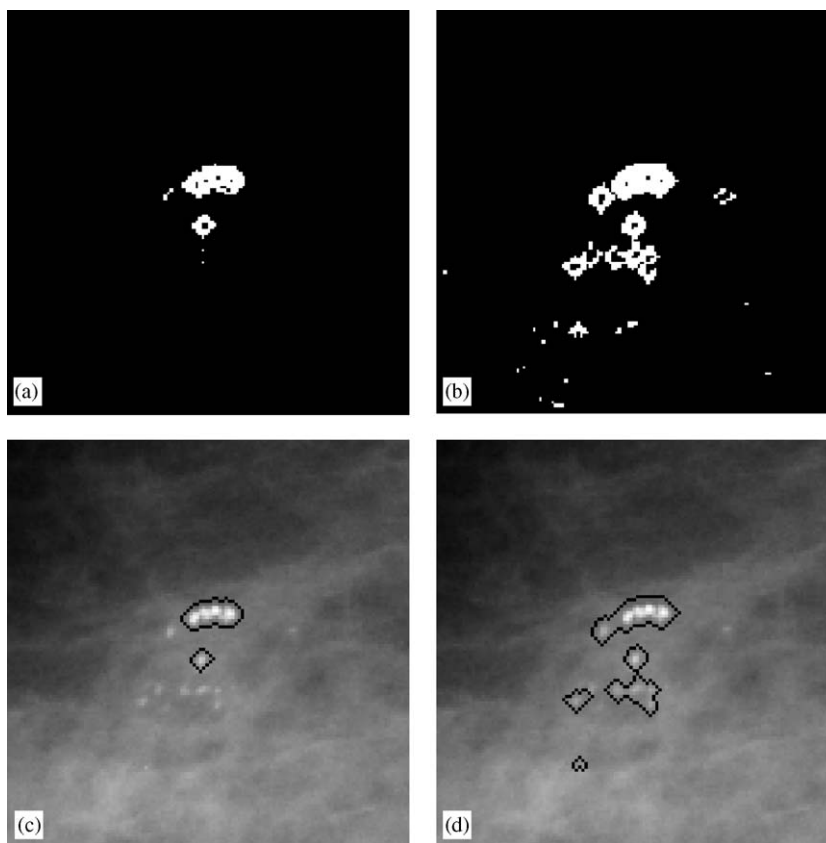


Fig. 9. Illustration of the use of proposed method. (a) and (b): Image details containing possible microcalcifications obtained from $f(\alpha)$ image by selecting pixels with $0 \leq f(\alpha) < 0.4$, and $0 \leq f(\alpha) < 0.6$, respectively. (c) and (d): Segmentation of regions containing microcalcifications after morphological postprocessing applied to Figs. 9(a) and (b).

circle) in Fig. 7(b). This case belongs to an “easy” one, since the tissue is radiology sparse and microcalcifications are visible even for not so skilled radiologist. The MF spectrum $f(\alpha)$ is depicted in Fig. 8(a), while α and $f(\alpha)$ images are presented in Figs. 8(b) and (c), respectively.

Figs. 8(b) and (c) clearly confirm basic features of microcalcifications: they are small light local tissue anomalies—parts in a mammogram not belonging to background tissue. In geometrical interpretation, they are seen as singular sets of points. From the multifractal standpoint they are characterized by both high α and low $f(\alpha)$ values, because, they represent sharp local changes and rare events. In Fig. 8 the edges of microcalcifications are represented as white pixels (high values) in an α -image and as black ones (low values) in $f(\alpha)$ image. Note that in α -image the background is not pure black but, due to nonlinear characteristic of human visual system, we are not able to distinguish small variations in this dark domain.

Once obtaining α - and $f(\alpha)$ images we can select desired parts from an original image by extracting pixels having particular values of α and/or $f(\alpha)$. In order to extract microcalcifications we have to choose image pixels characterized by high values of α and low values of $f(\alpha)$. Referring to Fig. 8(a), if we select image pixels from desired range of $f(\alpha)$ values, and present them as white points on black background, the MF segmentation of desired region is possible, as depicted in Figs. 9(a) and (b), for cases $0 \leq f(\alpha) < 0.4$ and $0 \leq f(\alpha) < 0.6$, respectively. Segmented regions cover possible microcalcifications. Since microcalcifications usually are clustered, we may refine the segmentation by using some postprocessing method. We applied successive morphological closing and opening procedures [5] to images as in Fig. 9(a)–(b), with a small diamond-shaped structuring element. By morphological closing (dilation followed by erosion), the holes within an object are filled, and unlinked contours are connected. Then, by applying opening (erosion followed by dilation), small

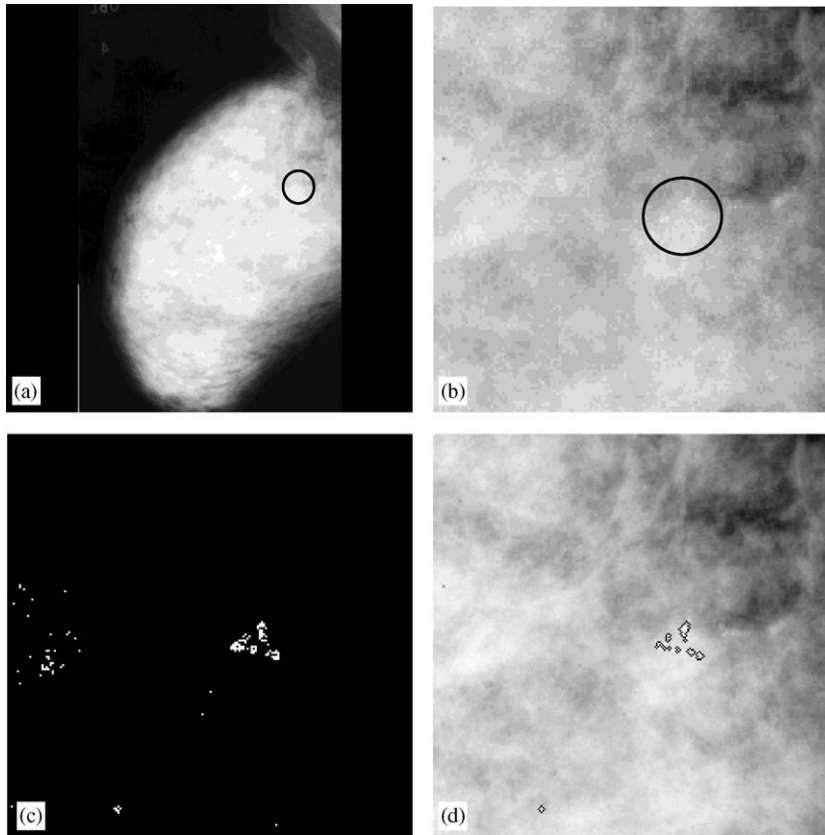


Fig. 10. (a) Original mammogram (1024×1024 pixels) mdb253.pgm from MiniMIAS database; (b) its selected part (256×256 pixels) around clinically approved microcalcifications; (c) segmented details from image in Fig. 10(b) having $0 \leq f(x) < 0.3$, and (d) original image with superimposed contour lines around detected microcalcifications after morphology postprocessing.

objects that cannot hold the structuring element (for instance, on the left and bottom from the central part of an image in Fig. 9(a)), are removed. Finally, the contours of such created objects are superimposed to original image, as displayed in Figs. 9(c) and (d), strongly indicating to a cluster of microcalcifications. The procedure described may be performed in an interactive way enabling the physician to find the best result for further diagnosis.

4.2. Radiology hard case

A mammogram mdb253.pgm from MiniMIAS database is shown in Fig. 10(a). Clinically approved microcalcifications are located within a small black circle. A part of this mammogram (256×256 pixels) around approved microcalcifications is presented in Fig. 10(b). The breast tissue is very dense causing very poor contrast between anomalies and surrounding tissue. Thus, the visual detection of microcalcifications is extremely difficult even for skilled radiologists. Image pixels characterized by $0 \leq f(x) < 0.3$ are presented as white in Fig. 10(c). From this image, after morphology postprocessing, regions containing possible microcalcifications are marked by contour lines as in Fig. 10(d). As we can conclude, clinically approved microcalcifications are highly visible.

4.3. Mammogram with high masking effect caused by background tissue

As a third example, we will observe a mammogram mdb223.pgm from MiniMIAS database as shown in Fig. 11(a). Clinically approved microcalcifications are located somewhere within small black circles. A part of

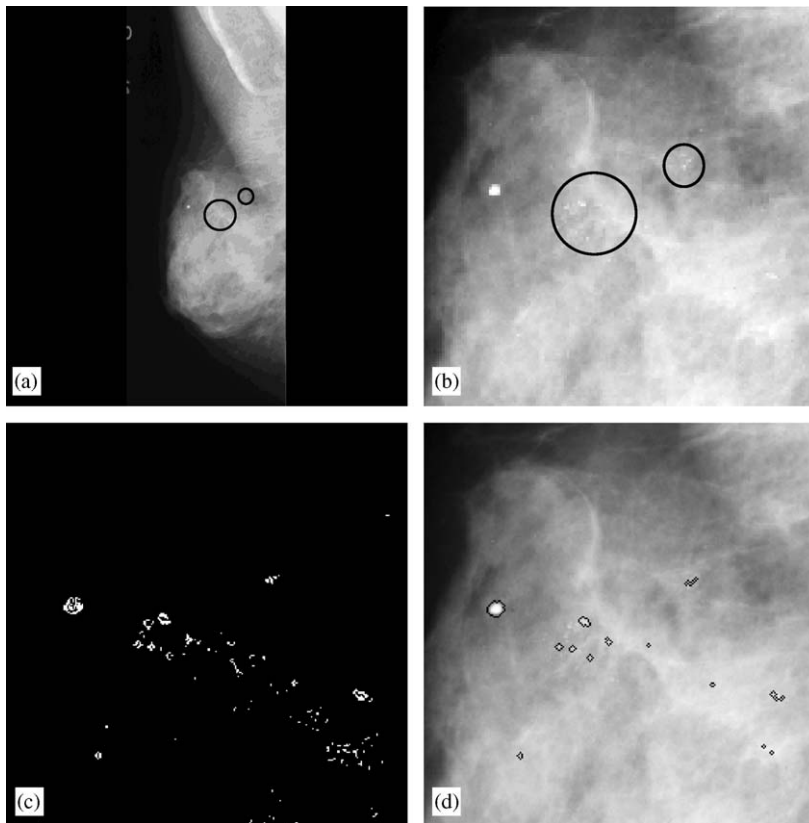


Fig. 11. (a) Original mammogram (1024×1024 pixels) mdb223.pgm from MiniMIAS database; (b) its selected part (256×256 pixels) around clinically approved microcalcifications; (c) segmented details from image in Fig. 11(b) having $0 \leq f(x) < 0.4$, and (d) original image with superimposed contour lines around possible microcalcifications after morphology postprocessing.

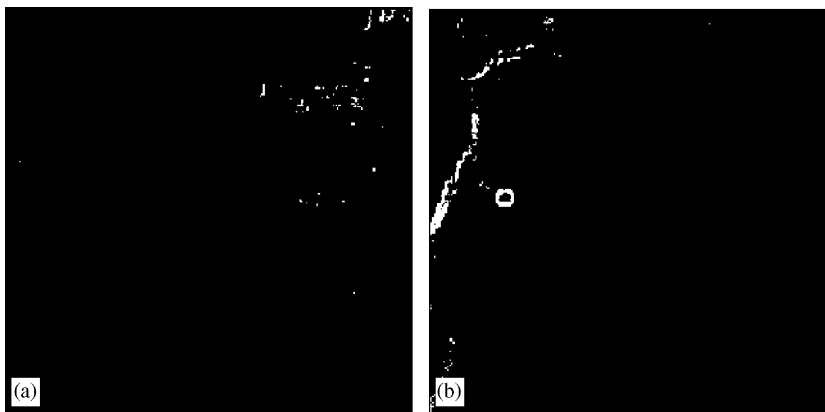


Fig. 12. Multifractal segmentation obtained by using software FracLab and the measure “Maximum”: (a) image as in Fig. 10(b) and $f(x) < 0.3$; (b) image as in Fig. 11(b) and $f(x) < 0.4$.

this mammogram (256×256 pixels) around approved microcalcifications is presented in Fig. 11(b). Microcalcifications are visible for radiologists, but automated selection is not so simple since background tissue has similar local contrast producing masking effect. However, our algorithm may extract very efficiently regions containing microcalcifications without affecting background tissue, as shown in Figs. 11(c) and (d).

4.4. Comparison with ‘classical’ multifractal analysis

The efficiency of the method proposed may be compared to other known multifractal methods. For instance, let us observe the use of public available software package *FracLab* [18] applied to the same examples. The capacity measure “Maximum”² on square-shaped measure domains of sizes 1, 3 and 5 pixels was applied. For an easy case, as in Fig. 7, the package *FracLab* gives quite good result in extracting microcalcifications. But, for radiology harder cases, this software was not so efficient. For instance, let us observe cases as in Figs. 10(b) and 11(b) and apply the *FracLab* software. From $f(x)$ images, pixels belonging to output ranges $f(x) < 0.3$ and $f(x) < 0.4$, respectively, are displayed as white in Figs. 12(a) and 12(b). Note that in both cases none of the approved microcalcifications is selected—compare these two images with those already presented in Figs. 10(c) and 11(c). Instead, a package *FracLab* extracts parts having high contrast such as a region on the upper right-hand side in Fig. 10(b) and a borderline between breast and the background, as depicted on the left-hand side of Fig. 11(b).

5. Conclusions

Our research was addressed to extraction of microcalcifications in digital mammograms. In original image domain, calcifications are represented by small bright spots not belonging to background tissue, usually in the form of clusters, and characterized by sharp change of local contrast just at their edges. In multifractal terminology these features are described both by high values of Hölder exponent α (high local changes) and low values of its distribution $f(x)$ (rare events in global sense). We introduced several modifications in MF approach targeted to improve the segmentation of microcalcifications. First, instead of original image we used negative image and measure “Minimum”. In this way, microcalcifications migrate from bright to dark domain where the effect of ‘logarithmic amplifier’ highly enhances them from surrounding. Furthermore, we determined parameter α and its distribution $f(x)$ from box-counting procedure as slopes of linear regression lines fitted through discrete points in corresponding bi-logarithmic diagrams. For estimating parameter α we used near neighborhood around observed image pixel as a measure domain corresponding to expected dimension of calcifications. More precisely, since our analysis was tested through mammograms from MiniMIAS database (characterized by 200 μm spatial resolution), we used box sized $i = 1, 3, \text{ and } 5$ pixels (i.e., 0.2, 0.6 and 1.0 mm). The shape of a measure domain is adjusted to more evident microcalcification types: disk and/or diamond ones. For estimating $f(x)$ from α -image we used small boxes sized from $j = 1$ to 16 ‘pixels’, since local changes (or high-frequency components) of this parameter are relevant for microcalcifications extraction. From these modifications, a multifractal analysis is adapted to enhance small contrast changes permitting very good detection of calcifications even in radiology dense tissue when classical visual detection and/or image processing algorithms fall down. Objects extracted in this way are not only microcalcifications but also other small bright details including a noise, film failure and/or other artifacts. By introducing morphological post-processing we may both remove details not belonging to the tissue anomalies (for example, film emulsion artifacts) and select regions containing clusters with possible microcalcifications.

Note that the calculation of multifractal spectrum is slightly time consuming, particularly for large images. However, after obtaining multifractal image a radiologist has the freedom to change the level of segmentation in interactive way, by setting the range of $f(x)$ values, and to find regions which may contain microcalcifications. The efficiency of proposed method was tested through public domain mammograms from MiniMIAS [15] database. In all cases the method successfully found declared microcalcifications even in radiology “dense” images where they are highly invisible within the background tissue.

Acknowledgements

This work was partially supported by Serbian Ministry of Science and Environmental Protection, under the Project No. 2105 of Basic Researches.

²In available *FracLab* software [18] the measure “Minimum” does not work properly.

References

- [1] C. Borling, T. Squires, T. Tong, *Cancer Statist.* 42 (19) (1992).
- [2] L. Tabar, S. Duffy, L. Burhenne, New Swedish breast cancer detection results for women aged 40–49, *Cancer* 72 (suppl.) (1993) 1437.
- [3] J. Suri, S. Setarehdan, S. Singh (Eds.), *Advanced Algorithmic Approaches to Medical Image Segmentation*, Springer, New York, 2002.
- [4] J. Bronzino (Ed.), *Biomedical Engineering Handbook*, CRC Press, Boca Raton, FL, 1995.
- [5] R. Gonzales, R. Woods, *Digital Image Processing*, second ed, Prentice-Hall, Englewoods, NJ, 2002.
- [6] T. Stojic, I. Reljin, B. Reljin, Local contrast enhancement in digital mammography by using mathematical morphology, in *Proc. Int. Symp. on Signals, Circuits and Systems—ISSCS 2005*, Vol. 2, Iasi, Romania, pp. 609–612, July 14–15, 2005.
- [7] B.B. Mandelbrot, *The Fractal Geometry of Nature*, WH Freeman, Oxford, 1983.
- [8] C. Evertsz, B. Mandelbrot, “Multifractal Measures”, Appendix B, in: H. Peitgen, H. Jurgens, P. Andrews (Eds.), *Andrews Chaos and Fractals*, Springer, Amsterdam, 1992.
- [9] P. Iannaccone, M. Khokha (Eds.), *Fractal Geometry in Biological Systems*, CRC Press, Boca Raton, FL, 1996.
- [10] J. Lévy Véhel, P. Mignot, Multifractal segmentation of images, *Fractals* 2 (3) (1994) 379–382.
- [11] J. Lévy Véhel, Introduction to the multifractal analysis of images, Technical Report INRIA, 1996.
- [12] M.J. Turner, J.M. Blackledge, P.R. Andrews, *Fractal Geometry in Digital Imaging*, Academic Press, New York, 1998.
- [13] I. Reljin, B. Reljin, Fractal geometry and multifractals in analyzing and processing medical data and images, *Arch. Oncol.* 10 (4) (2002) 283–293.
- [14] B. Pesquet-Popescu, J. Lévy Véhel, Stochastic fractal models for image processing, *IEEE Signal Process. Mag.* 19 (5) (2002) 48–62.
- [15] Suckling J, The MiniMIAS database, Mammographic Image Analysis Society—MIAS, www.wiau.man.ac.uk/services/MIAS/MIAScom.html.
- [16] P.R. Bakic, M. Albert, D. Brzakovic, A.D.A. Maidment, Mammogram synthesis using a 3D simulation, *Med. Phys.* 29 (9) (2002) 2140–2151.
- [17] A.-K. Carton, H. Bosmans, C. Van Ongeval, G. Souverijns, F. Rogge, A. Van Steen, G. Marchal, Development and validation of a simulation procedure to study the visibility of micro calcifications in digital mammograms, *Med. Phys.* 30 (8) (2003) 2234–2240.
- [18] *FracLab v1.1.1*, <http://www-rocq.inria.fr/fractales/Software/FRACLAB>, 2003.