Computer-Aided Evaluation of Screening Mammograms Based on Local Texture Models

Jiří Grim, Petr Somol, Michal Haindl, *Senior Member, IEEE*, and Jan Daneš

***Abstract*—We propose a new approach to diagnostic evaluation of screening mammograms based on local statistical texture models. The local evaluation tool has the form of a multivariate probability density of gray levels in a suitably chosen search window. First, the density function in the form of Gaussian mixture is estimated from data obtained by scanning of the mammogram with the search window. Then we evaluate the estimated mixture at each position and display the corresponding log-likelihood value as a gray level at the window center. The resulting log-likelihood image closely correlates with the structural details of the original mammogram and emphasizes unusual places. We assume that, in parallel use, the log-likelihood image may provide additional information to facilitate the identification of malignant lesions as untypical locations of high novelty.**

***Index Terms*—Gaussian mixture, local statistical model, log-likelihood image, screening mammography, texture information.**

I. Introduction

BREAST cancer is one of the most common forms of cancer in women and also one of the most frequent causes of cancer death. Approximately one out of ten women could develop breast cancer during her lifetime [1]. It appears that, at present, the only effective tool to decrease the high mortality rates is the early detection of malignant abnormalities. For this reason, numerous screening programs have been initiated in the last few years.

Evaluation of screening mammograms is known to be difficult. Mammograms are complex in appearance and the signs of early breast cancer are often small or subtle, embedded in a complex background and showing great diversity in shape, size, and location. The incidence of malignant abnormalities in screening mammograms is only about 0.1%–0.3% [1]. In the case of suspect abnormalities, the radiologists can recommend additional imaging or follow-up mammograms and about 5%–10% of the findings are proposed for surgical verification of suspect tissue by biopsy [3], [4]. Many biopsies (about 60%–80%) expectedly result in benign diagnoses as false positive [5]. On the other hand, retrospective analyses report relatively large false negative evaluation rates of about 10%–20% [5], [7].

The total number of screening mammograms evaluated worldwide in one year may be in the order of millions. Thus, even a small decrease in false negative evaluation rates may save many lives in the final effect. Similarly, the decrease in false positive biopsy rates may help to avoid unnecessary physical trauma and emotional stress to many women.

II. Computer-Aided Screening for Mammography

Mammographic screening programs produce large number of mammograms to be routinely evaluated under strong reliability requirements. Double reading may increase the sensitivity of radiologists by 5%–15% but as a standard of care, it appears to be too resource demanding [4]. In the last decade, there has been an increasing effort to help radiologists by introducing different computer-aided decision-supporting systems.

Unfortunately, in screening mammography, the diagnosis usually includes a substantial descriptive part concerning the location and form of the lesion, and, therefore, it cannot be treated simply as a standard problem of statistical classification into mutually exclusive classes [16]. According to the classification terminology in [2], the malignant abnormalities may be essentially of two categories: "calcifications" or "masses." The calcifications may be of different types (e.g., punctate, amorphous, pleomorphic, round and regular, vascular, coarse, dystrophic) and may differ in distribution (e.g., clustered, linear, segmented, regional, diffusely scattered). Masses may also have different shapes (e.g., round, oval, lobulated, irregular, tubular) and margins (e.g., circumscribed, microlobulated, obscured, ill-defined, spiculated), may differ in size, location, and orientation, and may have different backgrounds. In addition, there is a high natural variability of normal mammograms and a "positional" noise due to manual placement of the breast by the radiographer. For all these reasons, it would be nearly impossible to specify some suitable diagnostic classes which could be characterized by class-conditional probability distributions of a reasonable dimensionality.

To create a more standard classification problem the computer-aided detection (CADe) systems usually assume a small region of interest (ROI) as a subject of recognition. However, in view of the above-mentioned high diversity of malignant abnormalities and natural variability of normal mammograms, the application of the available pattern recognition methods remains to be a difficult task. For this reason, the CADe systems just only direct the attention of the radiologists toward potential abnormalities. Typical CADe systems use algorithms to mark potentially suspicious areas by prompts and then, with the information

Manuscript received February 04, 2008; revised October 20, 2008. First published February 18, 2009; current version published March 13, 2009. This work was supported in part by the projects GACR No. 102/07/1594 and 102/08/0593 of the Czech Grant Agency and in part by the MŠMT projects 2C06019 ZIMOLEZ and 1M0572 DAR. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Scott T. Acton.

J. Grim, P. Somol, and M. Haindl are with the Institute of Information Theory and Automation of the Czech Academy of Sciences, 18208 Prague 8, Czech Republic (e-mail: grim@utia.cas.cz; somol@utia.cas.cz; haindl@utia.cas.cz).

J. Daneš is with the First Faculty of Medicine of the Charles University, 18208 Prague 2, Czech Republic (e-mail: jan.danes@lf1.cuni.cz).

Digital Object Identifier 10.1109/TIP.2008.2011168

provided by the CADe system, the radiologist makes the final decision.

In literature, there are many CADe approaches based on different principles (see, e.g., [5]–[7] for extensive references) but, as it appears, the detection quality of the present CADe algorithms is not comparable with expert radiologists [9]. The latest CADe systems achieve sensitivity of about 80%–90% with respect to different diagnoses but the false positive rates of placing prompts are also high, usually in units per image [3], [8]. It should be noted that, because of the low incidence of malignant findings (0.1%–0.3%), standard CADe systems generate hundreds or even thousands of false positive prompts in order to correctly mark one malignant lesion. Nevertheless, according to different studies [3], [10]–[12], CADe can improve breast cancer detection by more than 20%.

In the last few years, there has been an increasing effort to utilize textural information for the detection of suspicious regions in mammograms. The textural approaches frequently refer to the renowned paper by Haralick [13], who introduced the concept of gray-level co-occurrence matrix (GCM). Formally, GCM can be viewed as a 2-D distribution of gray levels of two pixels in a specific mutual position defined by distance and direction. We recall that the underlying 2-D probability distribution describes the statistical relation between the two pixels in full generality. However, the size of GCM, which depends on the number of gray levels, may be very large and, therefore, unsuitable as a feature set. For this reason, different texture features like entropy, second moment, difference moment, inverse difference moment, etc. (cf. [14]), have been computed for differently chosen GCM in order to make the final classification feasible [8]. Unfortunately, in this way, the resulting features express only simple global properties of the respective GCM and the related statistical information is almost completely lost. A more general approach is to use statistical texture models as local evaluation tools. In view of the fact that most mammograms are pathology free, it has been proposed to perform CADe via novelty detection [15]–[17]. With this aim, a generative statistical model of mammographic appearance has been developed which produces synthetic mammograms of remarkable quality [17]–[19]. The ultimate goal is to model entire pathology-free mammograms in order to perform abnormality detection as a novelty (outlier) detection task.

In this paper, we propose preprocessing of screening mammograms by means of local texture models with the aim to emphasize diagnostically important details. First, we estimate local statistical texture model as a joint probability density of gray levels in a suitably chosen search window.[1] In the second phase, we compute the estimated density at each position of the window and display the corresponding log-likelihood value as a gray level at the central reference pixel of the window. The resulting "log-likelihood image" has the same resolution as the original mammogram, closely correlates with the underlying structural details and maps the "typicality" or "novelty" of different textural parts.

[1]Note that, within the scope of the search window, any GCM can be obtained as a 2-D marginal of the local statistical model.

We assume that, in parallel use, the log-likelihood image may provide additional information to facilitate the identification of malignant lesions as untypical locations of high novelty. In the present form, the method does not include any decision making, the diagnostic interpretation of the log-likelihood image is completely a domain of radiologists. However, as the approach is new, there is no diagnostic experience available. In order to clarify the basic features of the log-likelihood image, we explain the computational details of the main "detection" mechanisms in the following sections. The log-likelihood image is a purely statistical construct without specific relation to screening mammography. The appearance of the log-likelihood image is totally different from the original image and, for this reason, we plan to extend the set of illustrating examples on our website and in the near future we intend to make our software freely available.

III. Local Statistical Model

We have shown in a series of papers that gray-scale textures can be modeled locally by estimating the joint probability density of gray levels in a suitably chosen search window [20], [21]. Application of the estimated local statistical model to texture synthesis provides a unique visual evidence of the informativity of the model. In the case of a successful texture synthesis (cf. [22]), we may assume that the underlying mixture density locally describes all essential statistical properties of the texture. The main idea of our approach is to apply the local statistical model to the original image [23]. In this way, we can evaluate how probable (or typical) are the different locations of the image appearing in the search window.

A. Gaussian Mixture Model Estimation

We assume a digitized mammogram in the form of a matrix of discrete variables specifying the gray levels at the corresponding pixels. Formally, it is irrelevant if the digitized image has been obtained by scanning screen-film images or by digital mammography. We assume that the statistical properties of the image can be described locally by a joint probability density of gray levels of a suitably chosen search window. Denoting

$$\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$$

as the vector of gray levels of the window in a fixed pixel arrangement, we assume the joint probability density $P(\boldsymbol{x})$ in the form of a mixture of Gaussian components with diagonal covariance matrices [22], [23]

$$P(\boldsymbol{x}) = \sum_{m \in \mathcal{M}} w_m F(\boldsymbol{x}|\boldsymbol{\mu}_m, \boldsymbol{\sigma}_m), \boldsymbol{x} \in \mathcal{R}^N. \tag{1}$$

Here $\mathcal{M} = \{1, \ldots, M\}$ and $\mathcal{N} = \{1, \ldots, N\}$ denote the index sets of components and variables, respectively, and the mixture components are defined as products of univariate Gaussian densities

$$F(\boldsymbol{x}|\boldsymbol{\mu}_m, \boldsymbol{\sigma}_m) = \prod_{n \in \mathcal{N}} f_n(x_n|\mu_{mn}, \sigma_{mn}) \tag{2}$$

$$f_n(x_n|\mu_{mn}, \sigma_{mn}) = \frac{1}{\sqrt{2\pi}\sigma_{mn}} \exp\left\{-\frac{(x_n - \mu_{mn})^2}{2\sigma_{mn}^2}\right\}. \tag{3}$$

The standard way to compute maximum-likelihood estimates of the mixture parameters $w_m, \mu_{mn}, \sigma_{mn}$ is to use the EM algorithm [22]–[24]. The data set $\mathcal{S}$ is obtained by pixelwise scanning of the original mammogram with the search window

$$\mathcal{S} = \left\{ \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots \right\}, \quad \boldsymbol{x}^{(i)} \in \mathcal{R}^N. \tag{4}$$

The corresponding log-likelihood function

$$L = \frac{1}{|\mathcal{S}|} \sum_{x \in \mathcal{S}} \log \left[\sum_{m \in \mathcal{M}} w_m F(\boldsymbol{x} | \boldsymbol{\mu}_m, \boldsymbol{\sigma}_m) \right] \tag{5}$$

can be maximized by means of EM iteration equations.

We recall that the conditional independence model (1) does not imply the assumption of independence of variables. The diagonal covariance matrices in mixture components simplify EM computation essentially while the restrictive effect is less relevant. It can be seen that, with the increasing number of components, the product mixture (1) approaches the well known nonparametric Parzen estimates. In literature, the Gaussian product mixtures are often applied to normalized data—with zero means and unity variances. However, such a preprocessing of data is superfluous because Gaussian product mixture density is invariant with respect to arbitrary linear transform of variables (cf. Section IV and Appendix I).

B. Statistical Model of Screening Mammogram

Unlike supervised methods, evaluation of a screening mammogram based on a local statistical model need not be trained by using other mammograms. The local statistical model is estimated for every patient individually from a single mammogram, and, therefore, the method is not confronted with the natural variability of mammograms. This circumstance is relevant in view of the extreme diversity of malignant lesions.

To accumulate maximum information in the statistical model, we prefer as a source the full-field digital mammography images consisting of two medio-lateral views and two cranio-caudal views (cf. Fig. 1). In our experiments, we have used the full-field (four-view) mammograms from the DDSM database of the University of South Florida [2], which have been obtained by scanning screen-film images. For the sake of model estimation, the digitized mammograms have been sub-sampled to the pixel size of about 0.1 mm. Obviously, the method is applicable to the two-view digital mammograms (medio-lateral images only) without any change. To utilize the underlying symmetry, we apply a mirror transform to the right-hand-side images. In this way, the structural information from the symmetrical parts is compatible.

From the computational point of view, the choice of the search window is of great importance. Concerning the form, a square-shaped window is practical for texture synthesis [22], but in the case of evaluation of mammograms, the window should rather correspond to a circular neighborhood of the center pixel. The dimension of the estimated density should be as small as possible, but, on the other hand, the size of the window should be large enough to capture diagnostically important details. The choice of the window size is closely related to the image resolution. In the case of small pixel size, we should enlarge the window but the resulting increased dimension is computationally inconvenient. The global appearance of the log likelihood image is not essentially influenced by the window size. Generally, with a shrinking window, the image becomes more "discontinuous" but without loss of relevant information. On the other hand the increased window size tends to smooth out small details (cf. Fig. 3). After extensive experiments, we have chosen as an intuitive compromise a square window of 13×13 pixels with trimmed corners, assuming the pixel size of approximately 0.1 mm. The resulting dimension of data vectors $\boldsymbol{x} \in \mathcal{S}$ is $\mathrm{N} = 145 (= 169 - 4 \times 6)$. Another window size could appear to be more suitable to emphasize some specific diagnostic details but the proposed compromise solution seems to be generally acceptable.

Estimation of a Gaussian mixture in 145-dimensional space by means of EM algorithm is a specific task since a dimension of this order is usually considered to be prohibitive. In the case of

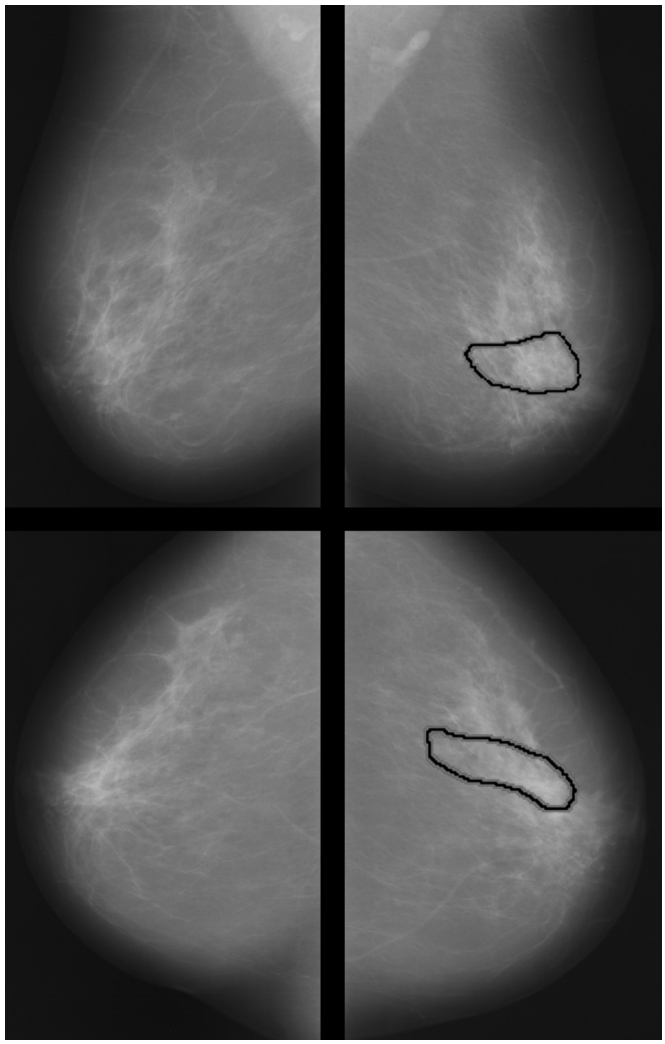

Fig. 1. Original image with a highlighted malignant lesion: calcification of pleomorphic type, segmentally distributed, assessment 4, subtlety 1–2 (modified full-field digital mammogram C-0002-1 from the DDSM database of the University of South Florida, cf. http://marathon.csee.usf.edu/Mammography/Database.html, [2]). The location of malignant lesion is derived from the corresponding overlay file.

our multidimensional local statistical model, the numerical solution is actually enabled by the assumption of diagonal covariance matrices of Gaussian components (2). The underlying conditional independence model simplifies computation and avoids difficulties with the potentially ill-conditioned covariance matrices. In this respect, it is advantageous that, estimating the window-space mixture density from a screening mammogram, we usually have a large data set ($|\mathcal{S}| \approx 10^5 - 10^6$) obtained by scanning the image with the search window. The traditional problem of a proper choice of the number of components and the initial parameter values is discussed extensively in literature (cf., e.g., [24]); however, it becomes less relevant in high-dimensional spaces and with increasing number of components. In all our experiments, we have a fixed number of components $M = 36$ and the mixture parameters have been initialized randomly.

IV. Log-Likelihood Image

A. Excluding Background Region

Normally, a large part of any mammogram represents dark (black) background which can be ignored in the model estimation process. Reliable solution to this problem is assumed to be an important prerequisite of a successful processing of mammograms [25]. In general, by excluding the background region, we save computing time and can also expect higher accuracy of mammogram analysis. The problem has become a subject of serious research in the last decade. Essentially, the proposed approaches are based on gray-level histogram thresholding combined with gradient analysis, morphological filtering and/or contour modeling (cf. [25]–[28]).

In this paper, the informative parts of the mammogram are chosen by computing a selection mask by means of a special mixture-based procedure. Actually, we use the local statistical model of Section III in a simplified version having only three components. We assume the mammogram to contain a relatively large and homogeneous dark background. In such a case, one of the components reliably fits the dark gray levels and can be used to identify the background region. The procedure is rather robust because each pixel is classified according to the window neighborhood. Obviously, defects of resulting selection mask are possible, but, in our experience, they have little influence on model accuracy.

B. Log-Likelihood Image Computation

Having estimated the parameters we evaluate the mixture density $P(\boldsymbol{x})$ at each position of the search window for the underlying window-patch vector $\boldsymbol{x}$ and display the corresponding log-likelihood value $\log P(\boldsymbol{x})$ as gray level at the central reference pixel of the window. The interpretation of the resulting log-likelihood image is straightforward. The light gray levels correspond to the "typical" highly probable parts of the image and the dark values reflect the less-probable, "untypical" or "unusual" locations. In this way, the resulting log-likelihood image should enable the identification of malignant abnormalities as locations of high novelty—exactly in the sense of the idea proposed by Rose and Taylor and others (cf. [15]–[19]). Let us remark, however, that the mammograms may contain many "normal" structural elements which are rare and, on the other hand, large or homogeneous malignant lesions could occur as "not-so-unusual" from the point of view of the local statistical model.

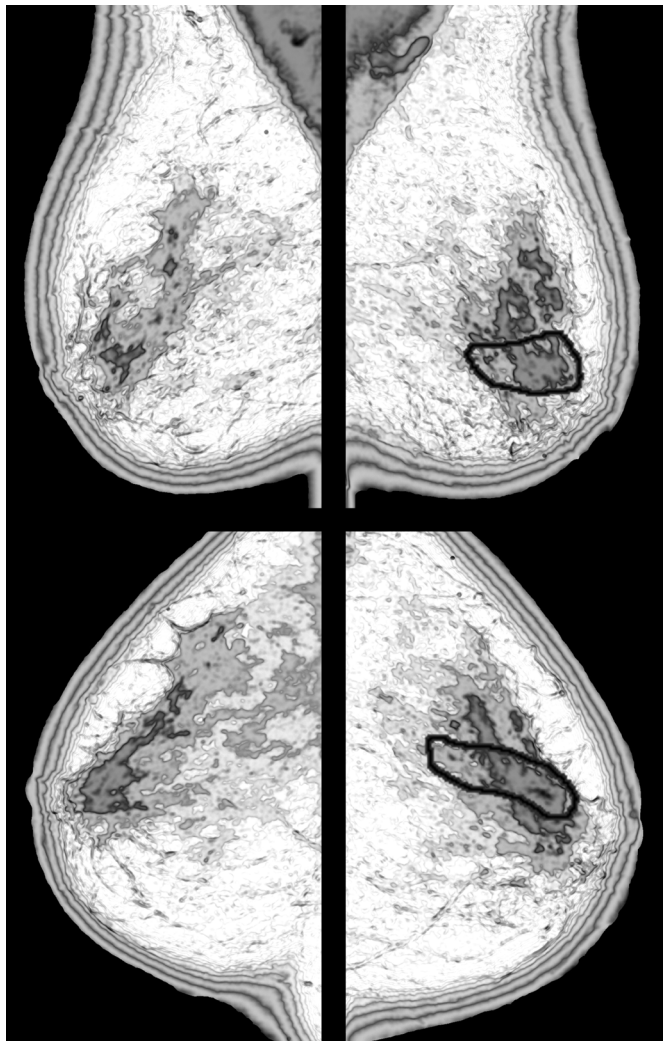

Fig. 2. Log-likelihood image for the full-field digital mammogram C-0002-1 from the DDSM database [2]. The log-likelihood image has the same resolution as the original mammogram. Each pixel value is defined by $\log P(\boldsymbol{x})$ where $\boldsymbol{x}$ is a 145-dimensional vector describing the pixel neighborhood. The malignant lesion is partly emphasized by contour lines. The log-likelihood image shows contralateral locations having similar textural properties.

It is obvious that the window patch is too small to enable a reliable "pixelwise" identification of malignant lesions. Nevertheless, by definition, the resulting log-likelihood image exactly correlates with the fine structural details of the original mammogram, has the same resolution and may provide additional information to facilitate the diagnostic evaluation of screening mammograms (cf. Figs. 1 and 2). Other examples of log-likelihood images in high resolution can be found at our web page http://www.utia.cas.cz/RO.

C. Detection of Microcalcifications

The presence of microcalcifications is one of the two most important features of clinically occult breast cancers. Automatic detection of microcalcifications is a frequent subject

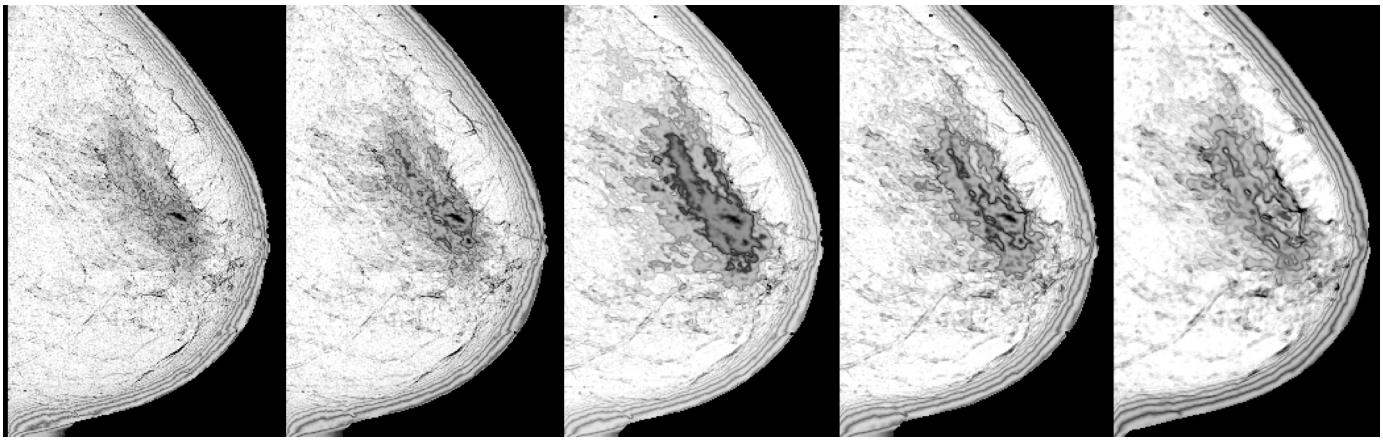

Fig. 3. Impact of window size on the log-likelihood image. Number of components: $M = 40 - 80$, window size in pixels from left to right: 9×9, 11×11, 13×13, 15×15, 17×17, pixel size $\phi = 100\ \mu$m (mammogram C-0002-1). With a shrinking window, the image becomes more "discontinuous."

of different CADe techniques usually based on specifically designed adaptive thresholding (cf., e.g., [7] for extensive references). In the case of log-likelihood image, the sensitivity to microcalcifications is closely related to the underlying Gaussian mixture model. The "detection" mechanism is very simple and illustrative.

Assume that a small micro-calcification causes a light isolated pixel in the mammogram. Then at all window positions containing the light pixel, we obtain essentially lower log-likelihood values [cf. (1)–(3)]. As a result, the log-likelihood image will contain a dark spot of window size centered on the light pixel. The minimum size of the spot is given by the search window and increases in the case of larger calcifications covering several pixels. The size and darkness of the spot continuously depend on the size and contrast of the underlying micro-calcification, respectively (cf. Fig. 4). A cluster of near micro-calcifications could appear as inhomogeneous dark spot of irregular form (cf. Figs. 1 and 2). In this sense each microcalcification is visualized in a way that continuously reflects the quantitative features of the original finding.

D. Interpretation of Contour Lines

Comparing the original mammogram and the corresponding log-likelihood image, we can see that even hardly visible details and structures closely correlate in both images but the boundaries of different regions are emphasized. The tendency of the local statistical model to create "contour lines" has a simple theoretical reason. Multidimensional spaces are "sparse," and, therefore, the mixture components tend to be almost nonoverlapping. For this reason, in the course of pixelwise evaluation of the mammogram, the log-likelihood values $\log P(\boldsymbol{x})$ are typically "dominated" by a single component of the mixture, which is most adequate to the underlying region. Thus, different regions are dominated by different components. A detailed numerical observation shows that the change of dominating component is accompanied by lower log-likelihood values. In this way, "switching" of mixture components is responsible for the arising dark contour lines at the boundaries of regions having different textural properties.

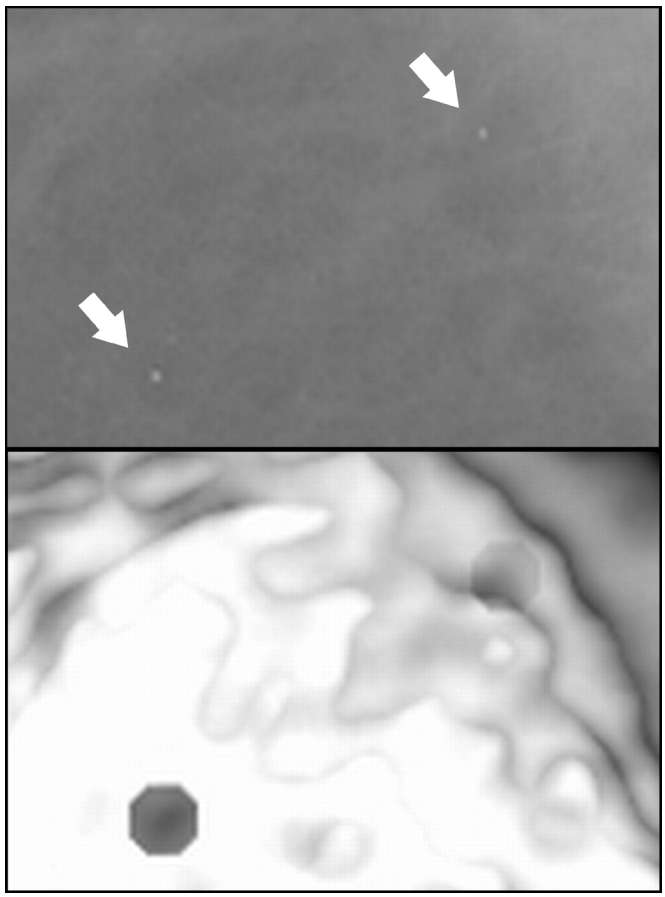

Fig. 4. Detection of microcalcifications. Top: Enlarged part of the original mammogram containing two microcalcifications. Bottom: Corresponding part of the log-likelihood image. Both microcalcifications are displayed as dark spots of window size.

The most apparent demonstration of this mechanism can be seen at the margins of the breast region which are characterized by continuously decreasing gray levels (cf. Figs. 1 and 2). The local statistical model describes the changing gray levels by means of several different dominating components and creates separate contour lines as a by-product of the component

switching process. We observe that, unlike iso-intensity contours (cf., e.g., [32]), the contour lines produced by mixture components should generally emphasize regions of similar textural properties.

The contour lines can be useful as a tool to distinguish mammogram regions of different densities (cf. Fig. 5). The identification of masses is especially of interest as one of the most important early signs of malignant lesions. Masses may be quite subtle, and may have smooth boundaries and different shapes. Detection and classification of masses is known to be more difficult than detection of microcalcifications as they often occur in dense glandular tissue (cf., e.g., [7]). The contour lines may help to solve this problem because they emphasize the shape of masses which may be of diagnostic meaning (cf. Fig. 5).

The log-likelihood image emphasizes locations having similar textural properties in all parts of the four-view mammogram and, in this way, any asymmetry becomes well visible. For this reason, the log-likelihood image can be helpful to evaluate possible contralateral findings (cf. [41]–[43]) or multifocal lesions because regions having similar properties are easily identified visually.

E. Model-Based Mammogram Segmentation

The CADe methods frequently make use of texture segmentation approaches to distinguish different kinds of breast tissue and facilitate the interpretation of mammogram (cf. [33]–[39]). In this respect, the local statistical model also provides a theoretically well-justified solution, which is similar to [38] in some respect. In particular, we can easily obtain an "oversegmented" image by simply classifying pixels according to the maximum conditional probabilities $q(m|\boldsymbol{x})$ [cf. Appendix I and (6)]. By using the over-segmented image as a starting point, we could obtain a more rough segmentation in a way proposed in [40]. However, the log-likelihood image appears to be more informative than any result of this kind. The log-likelihood image shows contour lines corresponding to segment boundaries but, in contrast to the segmented image, detailed textural information is visible simultaneously.

F. Invariance With Respect to Gray-Level Transformation

In the log-likelihood image, the gray scale can be fixed without information loss. In this paper, all log-likelihood images have been obtained in the same way by fixing the displayed gray-level interval to $\langle L - 3D_L, L + D_L\rangle$ where L is the final log-likelihood criterion and D_L is the standard deviation of the displayed values of $\log P(\boldsymbol{x})$. Thus, the value of the log-likelihood criterion (5) plays the role of a reference point only but it may be useful as a global characteristic of the mammogram.

In the literature, many different preprocessing approaches have been proposed to increase the contrast of mammograms by means of different gray-scale transforms (cf., e.g., [29]–[31]). In the case of a local statistical model (1), such a gray-level manipulation becomes partly superfluous because the log-likelihood image can be shown to be invariant with respect to arbitrary linear transform of the gray scale. In particular, if we assume a general linear transform of the gray scale of the original mammogram, then the transformed data and parameters can be shown to satisfy the EM iteration equations, and, therefore, the corresponding log-likelihood image of the transformed data is identical (for proof, see Appendix I).

V. Concluding Remarks

We propose a new approach to the evaluation of screening mammograms by means of local statistical texture models with the aim to facilitate the identification of the suspicious areas. The resulting log-likelihood image exactly correlates with the original mammogram, maps the typicality and novelty of the image and emphasizes regions of identical textural properties by contour lines. In parallel comparison the log-likelihood image may help to identify potential abnormalities in screening mammograms.

There are different possibilities to modify the log-likelihood image. The local data vector $\boldsymbol{x}$ could include more sophisticated features in order to describe explicitly the directional or morphological properties of the window patch. Also, instead of log-likelihood, we can use log-likelihood ratio with the product of unconditional marginals in the denominator (cf. [23]). In this form, the resulting image is more sensitive to structural irregularities and probably more suitable to some types of malignant lesions.

In the present form, the method does not include any decision making but the log-likelihood image is well applicable to generate additional explicit "black-box" decision information. Roughly speaking, any untypical location of high novelty could be marked by a prompt as a suspicious area. However, the log-likelihood image alone seems to be more informative in this respect since possible reasons of the generated prompt would be usually well visible in full complexity.

The log-likelihood image is a purely statistical construct based on the local statistical model without any specific relation to screening mammography. Formally, the method can be viewed as a new generally applicable principle of structure visualization in medical imaging. Using textural information the method may emphasize unusual locations and creates artificial contour lines around image regions characterized by different textural properties. In mammographic screening, it may be helpful to identify malignant abnormalities as locations of high novelty. Nevertheless, the diagnostic interpretation of the log-likelihood images represents a new problem for which the corresponding diagnostic experience can be accumulated only in cooperation with radiologists. For this reason, we plan to extend the set of illustrating high-resolution examples at our website (http://www.utia.cas.cz/RO) and in the near future we make our software freely available to interested professionals.

Appendix I
Proof of the Log-Likelihood Image Invariance

In view of standard calibration of mammograms, the invariance of the log-likelihood image with respect to arbitrary linear transform of gray levels is of importance. Logically, the invariance means that by using some linearly transformed gray level data in EM algorithm we would obtain the same log-likelihood image. However, we cannot prove the invariance property in this form because the parameter estimates obtained by

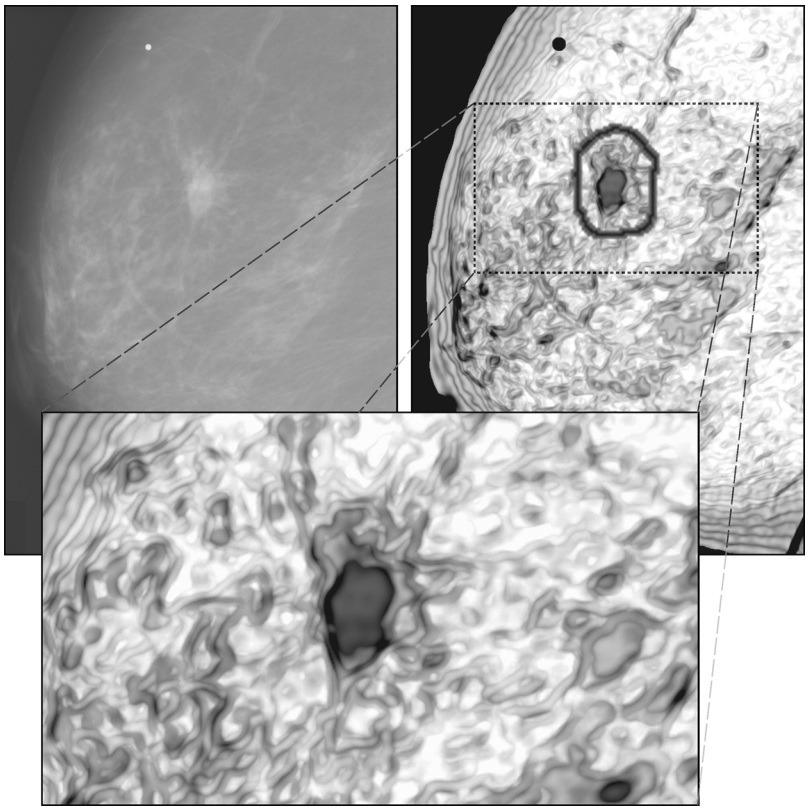

Fig. 5. Left upper part displays a detail of the mammogram C-0016-1 (DDSM database [2]) with malignant mass having irregular shape and spiculated margins. The right upper part contains the corresponding log-likelihood image combined with ground-true overlay file. The lower part of the image shows the malignant mass in more detail.

EM algorithm may be starting-point dependent and, therefore, not uniquely determined. We prove instead that if a linear transform is applied, both to the data set $\mathcal{S}$ and to some estimated mixture parameters, then the transformed parameters also satisfy the EM iteration equations.

In particular, let w_m, μ_{mn}, σ_{mn}, $m \in \mathcal{M}$, $n \in \mathcal{N}$ be the mixture model parameters obtained by EM algorithm for the data set $\mathcal{S}$, i.e., they satisfy the EM iteration equations

$$q(m|\boldsymbol{x}) = \frac{w_m F(\boldsymbol{x}|\boldsymbol{\mu}_m, \boldsymbol{\sigma}_m)}{\sum_{j\in\mathcal{M}} w_j F(\boldsymbol{x}|\boldsymbol{\mu}_j, \boldsymbol{\sigma}_j)} \tag{6}$$

$$w_m = \frac{1}{|\mathcal{S}|}\sum_{x\in\mathcal{S}} q(m|\boldsymbol{x}),$$

$$\mu_{mn} = \frac{1}{w_m|\mathcal{S}|}\sum_{x\in\mathcal{S}} x_n q(m|\boldsymbol{x}) \tag{7}$$

$$(\sigma_{mn})^2 = \frac{1}{w_m|\mathcal{S}|}\sum_{x\in\mathcal{S}} (x_n - \mu_{mn})^2 q(m|\boldsymbol{x}). \tag{8}$$

If we assume a general linear transform of the gray scale of the original mammogram, then it can be shown that the transformed data and parameters

$$\tilde{x}_n = a x_n + b, \quad \tilde{\boldsymbol{x}} = T(\boldsymbol{x}), \quad \tilde{\mathcal{S}} = \{T(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{S}\} \tag{9}$$

$$\tilde{\mu}_{mn} = a\mu_{mn} + b, \tilde{\sigma}_{mn} = a\sigma_{mn} \tag{10}$$

also satisfy the EM iteration (6)–(8) and the corresponding log-likelihood values differ only by a constant. Therefore, by fixing the mean gray level, we obtain an identical log-likelihood image of the transformed data.

By using (9) and (10), we can write

$$\tilde{P}(\tilde{\boldsymbol{x}}) = \sum_{m\in\mathcal{M}} w_m F(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{\mu}}_m, \tilde{\boldsymbol{\sigma}}_m) = \frac{1}{a^N} P(\boldsymbol{x}). \tag{11}$$

Further, we can show that the EM iteration (6)–(8) are satisfied for the transformed data and parameters

$$q(m|\tilde{\boldsymbol{x}}) = \frac{w_m F(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{\mu}}_m \tilde{\boldsymbol{\sigma}}_m)}{\tilde{P}(\tilde{\boldsymbol{x}})} = q(m|\boldsymbol{x}), \tag{12}$$

$$w_m = \frac{1}{|\mathcal{S}|}\sum_{x\in\mathcal{S}} q(m|\boldsymbol{x}) = \frac{1}{|\tilde{\mathcal{S}}|}\sum_{y\in\tilde{\mathcal{S}}} q(m|\tilde{\boldsymbol{x}}) = \tilde{w}_m \tag{13}$$

$$\tilde{\mu}_{mn} = a\mu_{mn} + b = \frac{1}{\tilde{w}_m|\mathcal{S}|}\sum_{\tilde{x}\in\tilde{\mathcal{S}}} \tilde{x}_n q(m|\tilde{\boldsymbol{x}})$$

$$(\tilde{\sigma}_{mn})^2 = a^2(\sigma_{mn})^2 = \frac{1}{\tilde{w}_m|\mathcal{S}|}\sum_{\tilde{x}\in\tilde{\mathcal{S}}} (\tilde{x}_n - \tilde{\mu}_{mn})^2 q(m|\tilde{\boldsymbol{x}}). \tag{14}$$

The proof is complete since, by (11), the values $\log\tilde{P}(\tilde{\boldsymbol{x}})$, $\log P(\boldsymbol{x})$ differ only by a constant which is eliminated by fixing the mean gray level of the log-likelihood image.

REFERENCES

[1] Eurostat, "Health statistics atlas on mortality in European Union," *Official J. Eur. Union*, 2002.

[2] M. Heath, K. W. Bowyer, and D. Kopans *et al.*, "Current state of the digital database for screening mammography," in *Digital Mammography*. Boston, MA: Kluwer, 1998, pp. 457–460.

[3] R. F. Brem, J. A. Rapelyea, G. Zisman, J. W. Hoffmeister, and M. P. DeSimio, "Evaluation of breast cancer with a computer-aided detection system by mammographic appearance and histopathology," *Cancer*, vol. 104, no. 5, pp. 931–935, 2005.

[4] D. B. Kopans, *Breast Imaging*. New York: Lippincott-Raven, 1998, p. 798.

[5] H. D. Cheng, X. P. Cai, X. W. Chen, L. M. Hu, and X. L. Lou, "Computer-aided detection and classification of microcalcifications in mammograms: A survey," *Pattern Recognit.*, vol. 36, no. 12, pp. 2967–2991, 2003.

[6] S. K. Yang, W. K. Moon, and N. Cho *et al.*, "Screening mammography-detected cancers: Sensitivity of a computer-aided detection system applied to full-field digital mammograms," *Radiology*, vol. 244, no. 1, pp. 104–111, 2007.

[7] H. D. Cheng, X. J. Shi, R. Min, L. M. Hu, X. P. Cai, and H. N. Du, "Approaches for automated detection and classification of masses in mammograms," *Pattern Recognit.*, vol. 39, pp. 646–668, 2006.

[8] N. R. Mudigonda, R. M. Rangayyan, and J. E. L. Desautels, "Detection of breast masses in mammograms by density slicing and texture flow-field analysis," *IEEE Trans. Med. Imag.*, vol. 20, pp. 1215–1227, 2001.

[9] N. Karssemeijer, "Use of prompt magnitude in computer-aided detection of masses in mammograms," in *Proc. 8th Int. Workshop on Digital Mammography IWDM*, S. M. Astley, Ed. *et al.*, Berlin-Heidelberg, Germany, 2006, vol. 4046, LNCS, pp. 54–60.

[10] R. F. Brem, J. K. Baum, S. Kaplan, S. Sounders, L. G. Naul, and J. Hoffmeister, "Improvement in sensitivity screening mammography with computer-aided detection (CAD): A multi-institutional trial," *Amer. J. Roentgenol.*, vol. 181, pp. 687–693, 2003.

[11] W. L. J. Burhenne, S. A. Wood, and C. J. D'Orsi *et al.*, "Potential contribution of computer-aided detection to the sensitivity of screening mammography," *Radiology*, vol. 215, pp. 554–562, 2000.

[12] T. W. Freer and M. J. Ulissey, "Screening mammography with computer-aided detection: Prospective study of 12,860 patients in a community breast center," *Radiology*, vol. 220, pp. 781–786, 2001.

[13] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–622, Nov. 1973.

[14] G. N. Lee, T. Hara, and H. Fujita, "Classifying masses as benign or malignant based on co-occurrence matrix textures: A comparison study of different grey-level quantizations," in *Proc. 8th Int. Workshop on Digital Mammography*, S. M. Astley, Ed. *et al.*, Berlin-Heidelberg, Germany, 2006, vol. 4046, LNCS, pp. 332–339.

[15] L. Tarassenko, P. Hayton, and N. Cerneaz *et al.*, "Novelty detection for the identification of masses in mammograms," in *Proc. 4th IEE Int. Conf. Artificial Neural Networks*, Cambridge, U.K., 1995, pp. 442–447.

[16] C. J. Rose and C. J. Taylor, "A statistical model of texture for medical image synthesis and analysis," *Med. Image Understand. Anal.*, pp. 1–4, 2003.

[17] C. J. Rose and C. J. Taylor, "A generative statistical model of mammographic appearance," *Med. Image Understand. Anal.*, pp. 89–92, 2004.

[18] P. R. Bakic, M. Albert, and D. Brzakovic *et al.*, "Mammogram synthesis using 3D simulation. I. Breast tissue model and image acquisition simulation," *Med. Phys.*, vol. 29, pp. 2131–2139, 2002.

[19] C. Spence, L. Parra, and P. Sajda, "Detection, synthesis and compression in mammographic image analysis with a hierarchical image probability model," in *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, L. Staib, Ed., 2001.

[20] J. Grim and M. Haindl, "Texture modelling by discrete distribution mixtures," *Comput. Statist. Data Anal.*, vol. 41, no. 3–4, pp. 603–615, 2003.

[21] J. Grim and M. Haindl, "A discrete mixtures colour texture model," in *Proc. 2nd Int. Workshop on Texture Analysis and Synthesis*, M. Chantler, Ed., Copenhagen, The Netherlands, 2002, pp. 59–62, Heriot-Watt Univ., Glasgow, U.K.

[22] J. Grim, M. Haindl, P. Somol, and P. Pudil, "A subspace approach to texture modelling by using Gaussian mixtures," in *Proc. 18th IAPR Int. Conf. Pattern Recognition*, B. Haralick and T. K. Ho, Eds., Los Alamitos, CA, 2006, pp. 235–238, IEEE Computer Soc.

[23] J. Grim, P. Somol, M. Haindl, and P. Pudil, "A statistical approach to local evaluation of a single texture image," in *Proc. 16th Annu. Symp. PRASA*, F. Nicolls, Ed., 2005, pp. 171–176 [Online]. Available: http://www.prasa.uct.ac.za/, Univ. Cape Town

[24] G. J. McLachlan and D. Peel, *Finite Mixture Models*. New York: Wiley, 2000.

[25] T. Ojala, J. Näppi, and O. Nevalainen, "Accurate segmentation of the breast region from digitized mammograms," *Comput. Med. Imag. Graph.*, vol. 25, no. 1, pp. 47–59, 2001.

[26] T.-K. Lau and W. F. Bischof, "Automated detection of breast tumors using the asymmetry approach," *Comput. Biomed. Res.*, vol. 24, no. 3, pp. 273–295, 1991.

[27] U. Bick, M. L. Giger, R. A. Schmidt, R. M. Nishikawa, D. E. Wolverton, and K. Doi, "Automated segmentation of digitized mammograms," *Acad. Radiol.*, vol. 2, no. 1, pp. 1–9, 1995.

[28] A. J. Méndez, P. G. Tahoces, M. J. Lado, M. Souto, J. L. Correa, and J. J. Vidal, "Automatic detection of breast border and nipple in digital mammograms," *Comput. Meth. Progr. Biomed.*, vol. 49, no. 3, pp. 253–262, 1996.

[29] N. Petrick, H.-P. Chan, D. Wei, B. Sahiner, M. A. Helvie, and D. D. Adler, "Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification," *Med. Phys.*, vol. 23, no. 10, pp. 1685–1696, 1996.

[30] E. D. Pisano, S. Zong, B. M. Hemminger, M. DeLuca, R. E. Johnston, K. Muller, M. P. Braeuning, and S. M. Pizer, "Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms," *J. Digit. Imag.*, vol. 11, no. 4, pp. 193–200, 1998.

[31] K. Wongsritong, K. Kittayaruasiriwat, F. Cheevasuvit, K. Dejhan, and A. Somboonkaew, "Contrast enhancement using multipeak histogram equalization with brightness preserving," in *Proc. IEEE Asia-Pacific Conf. Circuits and Systems*, 1998, pp. 455–458.

[32] D. Guliato, R. M. Rangayyan, J. A. Zuffo, and J. E. L. Desautels, "Detection of breast tumor boundaries using iso-intensity contours and dynamic thresholding," in *Proc. 4th Int. Workshop on Digital Mammography IWDM*, N. Karssemeijer, Ed. *et al.*, Nijmegen, The Netherlands, Jun. 1998, pp. 253–260.

[33] M. A. Kupinski and M. L. Giger, "Automated seeded lesion segmentation on digital mammograms," *IEEE Trans. Med. Imag.*, vol. 17, no. 4, pp. 510–517, Apr. 1998.

[34] G. M. Brake and N. Karssemeijer, "Segmentation of suspicious densities," *Med. Phys.*, vol. 28, no. 2, pp. 258–266, 2001.

[35] N. Petrick, H.-P. Chan, B. Sahiner, and M. A. Helvie, "Combined adaptive enhancement and region-growing segmentation of breast masses on digitized mammograms," *Med. Phys.*, vol. 26, no. 8, pp. 1642–1654, 1999.

[36] B. Sahiner, N. Petrick, H.-P. Chan, L. M. Hadjiiski, C. Paramagul, M. A. Helvie, and M. N. Gurcan, "Computer-aided characterization of mammographic masses: Accuracy of mass segmentation and its effects on characterization," *IEEE Trans. Med. Imag.*, vol. 20, no. 12, pp. 1275–1284, Dec. 2001.

[37] B. Zheng, Y. H. Chang, and D. Gur, "Computerized detection of masses in digitized mammograms using single-image segmentation and a multilayer topographic feature analysis," *Acad. Radiol.*, vol. 2, no. 11, pp. 959–966, 1995.

[38] R. Zwiggelaar, P. Planiol, and J. Marti *et al.*, "EM texture segmentation of mammographic images," in *Proc. 6th Int. Workshop Digital Mammography*, Bremen, Germany, 2002, pp. 223–227.

[39] A. Tsai, A. Yezzi, Jr., W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 137–154, Feb. 2003.

[40] J. Grim, P. Somol, M. Haindl, and P. Pudil, "Color texture segmentation by decomposition of Gaussian mixture model," in *Proc. Progress in Pattern Recognition, Image Analysis and Applications*, 2006, vol. 19, no. 4225, pp. 287–296.

[41] H. H. Storm and O. M. Jensen, "Risk of contralateral breast cancer in Denmark 1943–80," *Brit. J. Cancer*, vol. 54, no. 3, pp. 483–492, 1986.

[42] G. F. Robbins and J. W. Berg, "Bilateral primary breast cancers," *Cancer*, vol. 17, pp. 1501–1527, 1964.

[43] T. J. Rissanen, H. P. Makarainen, M. A. Apaja-Sarkkinen, and E.-L. Lindholm, "Mammography and ultrasound in the diagnosis of contralateral breast cancer," *Acta Radiol.*, vol. 36, no. 4, pp. 358–366, 1995.

[44] D. Brzakovic, X. M. Luo, and P. Brzakovic, "An approach to automated detection of tumors in mammograms," *IEEE Trans. Med. Imag.*, vol. 9, no. 3, pp. 233–241, Sep. 1990.

Jiří Grim graduated in physical electronics (M.S. level) from the Czech Technical University, Prague, in 1968, and received the Ph.D. degree in computer science from the Czech Academy of Sciences, Prague, in 1981.

Since 1968, he has been with the Institute of Information Theory and Automation, Czech Academy of Sciences. His current research interests include statistical pattern recognition, application of distribution mixtures to texture modeling, and the probabilistic approach to neural networks.

Petr Somol received the M.Sc. and Ph.D. degrees in computer science from the Faculty of Mathematics and Physics, Charles University, Prague.

He has been with the Department of Pattern Recognition at the Institute of Information Theory, Academy of Sciences of the Czech Republic, since his return from the Medical Informatics Unit, IPH, University of Cambridge, U.K. His current interests include statistical approach to pattern recognition, combinatorial algorithms, graphics, and modern programming.

Michal Haindl (SM'04) graduated in control engineering from the Czech Technical University (CTU), Prague, in 1979, and received the Ph.D. degree in technical cybernetics from the Czechoslovak Academy of Sciences (CAS) and the Sc.D. (Dr.Sc.) degree from CTU in 1983 and 2001, respectively.

From 1990 to 1995, he was with the University of Newcastle, Newcastle; the Rutherford Appleton Laboratory, Didcot; the Centre for Mathematics and Computer Science, Amsterdam; and the Institute National de Recherche en Informatique et en Automatique, Rocquencourt. In 1995, he rejoined the Institute of Information Theory and Automation, CAS, where he is Head of the Pattern Recognition Department. His current interests are random field applications in pattern recognition and image processing and automatic acquisition of virtual reality models.

Dr. Haindl is a fellow of the IAPR.

Jan Daneš is an Associate Professor of radiology, Chief of Radiology and the Polak Family Breast Diagnostic Center of the First Faculty of Medicine, Charles University, Prague, Czech Republic. He serves as co-Leader of the Czech National Breast Cancer Screening Program and Chairman of breast imaging of the Czech Radiological Society. His interests include statistical approach to recognition of ultrasonic and mammographic images and organization aspects and quality indicators of screening program and breast imaging.