Evaluation of Screening Mammograms by Local Structural Mixture Models

Jiří Grim & Gobert N. Lee

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic, Prague
and
Faculty of Science and Engineering
Flinders University, Adelaide, Australia

SPMS 2012, June 25-30, 2012, Sázavka, Czech Republic

Outline

Breast Cancer and Mammographic Screening

Statistical Data of Breast Cancer:

- breast cancer happens to about 8% of women during their lifetime
- malignant findings are rare: about 1 to 3 in 1000 screening mammograms
- 5 to 10% of findings is proposed for surgical verification by biopsy
- expectably: about 60 to 80% of biopsies result in benign diagnoses
- retrospective examinations: about 10 to 20% false negative results

Meaning of Mammographic Screening:

- even hardly palpable breast tumors can make metastases
- $\Rightarrow$ early detection of malignant lesions is extremely important
- **mammographic screening** is the only effective tool to decrease the breast cancer mortality rates
- $\Rightarrow$ screening programs: millions of mammograms in one year
- $\Rightarrow$ strong motivation for computer-aided evaluation

Local Evaluation of Screening Mammograms

Idea of the log-likelihood image: ▸ LITERATURE

to emphasize mammographic lesions as “untypical” locations of high “novelty” and facilitate diagnostic evaluation of screening mammograms

local properties: inside pixels of a square window with trimmed corners
$\mathbf{x} = (x_1, x_2, \ldots, x_N) \in \mathcal{X}, \quad x_n \approx$ grey-levels of the window inside

local statistical model: multivariate probability density $P(\mathbf{x})$
$\log P(\mathbf{x}) \approx$ **to measure how unusual is the window inside** $\mathbf{x}$

METHOD: approximation of the density $P(\mathbf{x})$ by Gaussian mixture

- **data set:** by scanning the mammogram with the search window
- **EM algorithm:** to estimate the Gaussian mixture $P(\mathbf{x})$
- **log-likelihood image:** $\log P(\mathbf{x})$ displayed as grey-levels at window center
- **interpretation:** dark grey-levels indicate “suspect” locations

UTIA

Estimation of Local Statistical Model

Gaussian mixture of product components:

$$P(\mathbf{x}) = \sum_{m=1}^{M} w_m F(\mathbf{x}|\boldsymbol{\mu}_m, \boldsymbol{\sigma}_m) = \sum_{m=1}^{M} w_m \prod_{n=1}^{N} \left[\frac{1}{\sqrt{2\pi}\sigma_{mn}} \exp\{-\frac{(x_n - \mu_{mn})^2}{2\sigma_{mn}^2}\} \right]$$

log-likelihood function: (data set $\mathcal{S}$ by scanning the image)

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

EM iteration equations: (initial parameters chosen randomly)

$$q(m|\mathbf{x}) = \frac{w_m F(\mathbf{x}|\boldsymbol{\mu}_m, \boldsymbol{\sigma}_m)}{\sum_{j \in \mathcal{M}} w_j F(\mathbf{x}|\boldsymbol{\mu}_j, \boldsymbol{\sigma}_j)}, \quad w_m^{'} = \frac{1}{|\mathcal{S}|} \sum_{x \in \mathcal{S}} q(m|\mathbf{x})$$

$$\mu_{mn}^{'} = \frac{1}{w_m^{'}|\mathcal{S}|} \sum_{x \in \mathcal{S}} x_n q(m|\mathbf{x}), \quad (\sigma_{mn}^{'})^2 = \frac{1}{w_m^{'}|\mathcal{S}|} \sum_{x \in \mathcal{S}} (x_n - \mu_{mn}^{'})^2 q(m|\mathbf{x})$$

NOTATION: $w_m^{'}, \mu_{mn}^{'}, \sigma_{mn}^{'}$ $\approx$ new parameter values

Computational Experiments

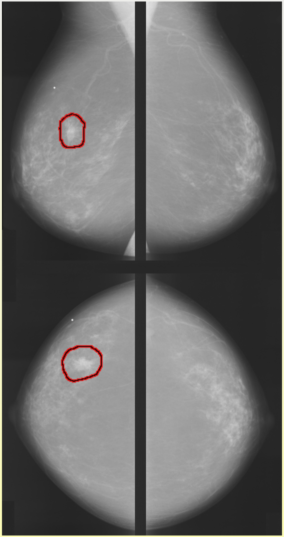
- **local statistical model:** estimated from a single mammogram
- each mammogram is evaluated individually
- mirror transform is applied to right-hand-part of images to utilize the underlying symmetry
- $\Rightarrow$ the method need not be trained by other images
- $\Rightarrow$ it is not confronted with high natural variability of mammograms

---

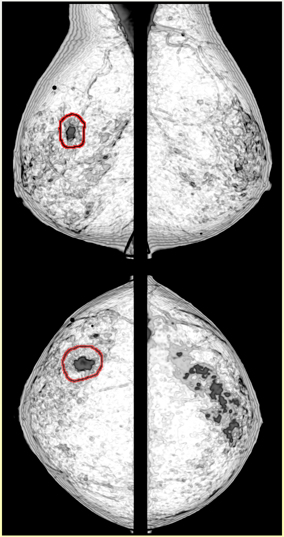
- **source database:** University of South Florida http://marathon.csee.usf.edu/Mammography/Database.html
- **chosen search window:** square window of 13 x 13 pixels with trimmed corners, dimension of $\mathbf{x}$ is $N = 145$ $(= 169 - 4 \times 6)$
- **model data set:** by scanning the four-view mammogram with the search window $(|\mathcal{S}| \approx 10^5 - 10^6)$
- **computing time:** cca 2 hours for 36 components but the computation can be parellelized

C-0016-1: segmentally distributed calcification

original image



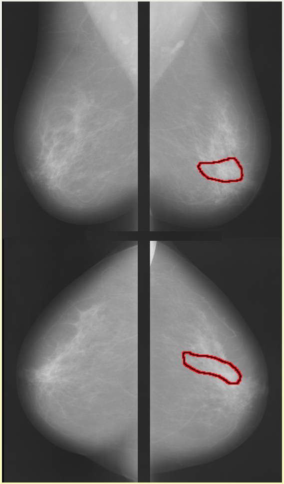
log-likelihood image



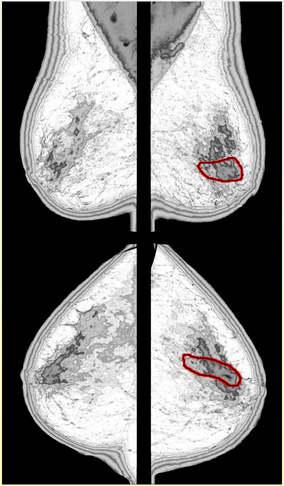
Remark: Pixel grey levels $\log P(\mathbf{x})$ defined by 145 neighboring pixels

C-0002-1: segmentally distributed calcification

original image



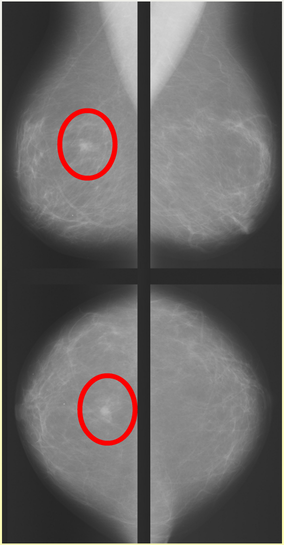
log-likelihood image



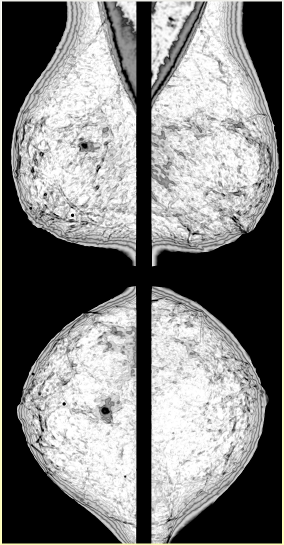
Remark: Large calcification emphasized by contour lines.

C-0143-1: mass, irregular shape, ill-defined margins

original image

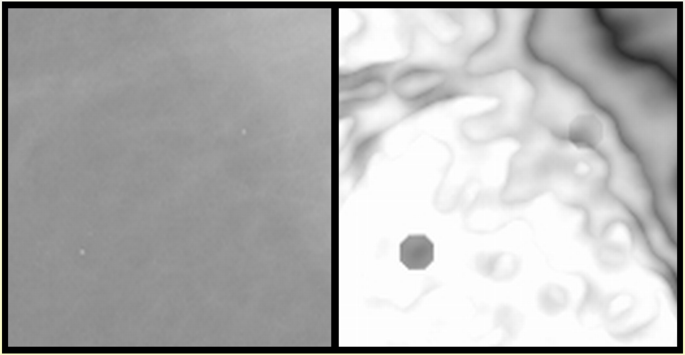


log-likelihood image



Identification of "Micro-Calcifications" by Spots

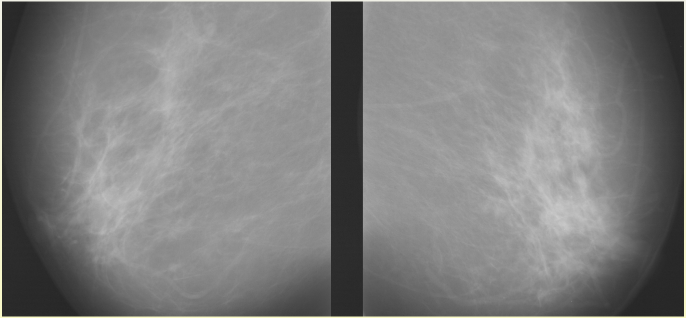
original micro-calcifications **corresponding dark spots**



Remark: window position containing a light pixel
$\Rightarrow$ implies decreased value of $\log P(\mathbf{x})$
$\Rightarrow$ light pixel is identified as a window-like dark spot

Identification of “Masses” by Contour-Lines

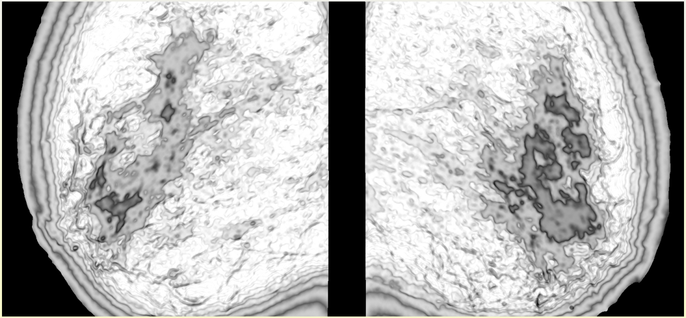
example of screening mammogram containing suspect masses



Remark: The masses may be quite subtle, may have smooth boundaries and different shapes. Detection and classification of masses is more difficult than detection of micro-calcifications.

Identification of “Masses” by Contour-Lines

contour lines around “masses” and at the mammogram boundaries



Remark: In high-dimensional spaces ($N \approx 10^2$) the log-likelihood values $\log P(\mathbf{x})$ are typically dominated by a single component of the mixture which is most adequate to the underlying region.
The “switching” of components at the boundaries of different regions is accompanied by decreased $\log P(\mathbf{x})$ values which produce contour lines.

Structural Mixture Model

IDEA: to exclude the less informative "noisy" variables in components

binary structural parameters: $\phi_m = (\phi_{m1}, \ldots, \phi_{mN}) \in \{0, 1\}^N$

$$F(\mathbf{x}|m) = \prod_{n \in \mathcal{N}} f_n(x_n|m)^{\phi_{mn}} f_n(x_n|0)^{1-\phi_{mn}}, \quad \sum_{m \in \mathcal{M}} \sum_{n \in \mathcal{N}} \phi_{mn} = s < MN$$

$\phi_{mn} = 0 \;\Rightarrow\;$ $f_n(x_n|m)$ replaced by common fixed density $f_n(x_n|0)$ $(\approx P_n(x_n))$

$$P(\mathbf{x}) = \sum_{m \in \mathcal{M}} F(\mathbf{x}|m) f(m) = F(\mathbf{x}|0) \sum_{m \in \mathcal{M}} G(\mathbf{x}|m, \phi_m) f(m)$$

$$G(\mathbf{x}|m, \phi_m) = \prod_{n \in \mathcal{N}} \left[\frac{f_n(x_n|m)}{f_n(x_n|0)} \right]^{\phi_{mn}}, \quad F(\mathbf{x}|0) = \prod_{n \in \mathcal{N}} f_n(x_n|0) \;\approx\; \textbf{background}$$

the fixed background density $F(\mathbf{x}|0)$ reduces in Bayes formula:

$$p(\omega|\mathbf{x}) = \frac{P(\mathbf{x}|\omega) p(\omega)}{P(\mathbf{x})} = \frac{\sum_{m \in \mathcal{M}_\omega} G(\mathbf{x}|m, \phi_m) f(m)}{\sum_{j \in \mathcal{M}} G(\mathbf{x}|j, \phi_j) f(j)} \approx \sum_{m \in \mathcal{M}_\omega} G(\mathbf{x}|m, \phi_m) f(m)$$

Remark: The model performs component specific feature selection

Structural Optimization - General EM Algorithm

structural optimization can be included into EM algorithm

$$L = \frac{1}{|\mathcal{S}|}\sum_{\mathbf{x}\in\mathcal{S}} \log P(\mathbf{x}) = \frac{1}{|\mathcal{S}|}\sum_{\mathbf{x}\in\mathcal{S}} \log \left[\sum_{m\in\mathcal{M}} F(\mathbf{x}|0)G(\mathbf{x}|m,\phi_m)f(m)\right]$$

E-step:

$$q(m|\mathbf{x}) = \frac{f(m)G(\mathbf{x}|m,\phi_m)}{\sum_{j=1}^{M} f(j)G(\mathbf{x}|j,\phi_j)}, \quad m = 1,2,\ldots,M$$

M-step:

$$f^{'}(m) = \frac{1}{|\mathcal{S}|}\sum_{\mathbf{x}\in\mathcal{S}} q(m|\mathbf{x}), \quad m = 1,2,\ldots,M,\ \mathbf{x}\in\mathcal{S}$$

$$\mathrm{G}^{'}(.|m,\phi_m^{'}) = \arg\max_{G(.|m,\phi_m)} \left\{\frac{1}{|\mathcal{S}|}\sum_{x\in\mathcal{S}} q(m|\mathbf{x}) \log F(\mathbf{x}|0)G(\mathbf{x}|m,\phi_m)\right\}$$

$$\Rightarrow \quad G^{'}(.|m,\phi_m^{'}) = \arg\max_{G(.|m,\phi_m)} \left\{\frac{1}{|\mathcal{S}|}\sum_{x\in\mathcal{S}} q(m|\mathbf{x}) \log G(\mathbf{x}|m,\phi_m)\right\}$$

Remark. EM algorithm transforms the difficult problem of maximization of the mixture log-likelihood function L to the repeated maximization of the weighted log-likelihood functions of mixture components.

Proof of the Monotonic Property of EM Algorithm

sequence of log-likelihood values $\{L^{(t)}\}_{t=0}^{\infty}$ is nondecreasing:

$$L^{(t+1)} - L^{(t)} \geq 0, \quad t = 0, 1, 2, \ldots$$

In view of Kullback-Leibler information divergence we can write

$$\frac{1}{|\mathcal{S}|}\sum_{\mathbf{x}\in\mathcal{S}} I(q(\cdot|\mathbf{x})||q^{'}(\cdot|\mathbf{x})) = \frac{1}{|\mathcal{S}|}\sum_{\mathbf{x}\in\mathcal{S}}\left[\sum_{m=1}^{M} q(m|\mathbf{x})\log\frac{q(m|\mathbf{x})}{q^{'}(m|\mathbf{x})}\right] \geq 0$$

making substitution for $q(m|\mathbf{x}), q^{'}(m|\mathbf{x})$ in the logarithm we obtain:

$$\frac{1}{|\mathcal{S}|}\sum_{\mathbf{x}\in\mathcal{S}}\sum_{m=1}^{M} q(m|\mathbf{x})\log\frac{P^{'}(\mathbf{x})}{P(\mathbf{x})} + \frac{1}{|\mathcal{S}|}\sum_{\mathbf{x}\in\mathcal{S}}\sum_{m=1}^{M} q(m|\mathbf{x})\log\left[\frac{f(m)G(\mathbf{x}|m,\phi_m)}{f^{'}(m)G^{'}(\mathbf{x}|m,\phi_m^{'})}\right] \geq 0$$

whereby the first sum corresponds to the increment of L:

$$\frac{1}{|\mathcal{S}|}\sum_{\mathbf{x}\in\mathcal{S}}\sum_{m=1}^{M} q(m|\mathbf{x})\log\frac{P^{'}(\mathbf{x})}{P(\mathbf{x})} = \frac{1}{|\mathcal{S}|}\sum_{\mathbf{x}\in\mathcal{S}}\log\frac{P^{'}(\mathbf{x})}{P(\mathbf{x})} = L^{'} - L.$$

Proof of the Monotonic Property of EM Algorithm

making substitution for the increment we obtain the inequality

$$L^{'} - L \geq \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} \sum_{m=1}^{M} q(m|\mathbf{x}) \log \left[\frac{f^{'}(m) G^{'}(\mathbf{x}|m, \phi_m^{'})}{f(m) G(\mathbf{x}|m, \phi_m)} \right]$$

which can be rewritten in the form

$$L^{'} - L \geq \sum_{m=1}^{M} \left[\frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(m|\mathbf{x}) \right] \log \frac{f^{'}(m)}{f(m)} + \sum_{m=1}^{M} \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(m|\mathbf{x}) \log \frac{G^{'}(\mathbf{x}|m, \phi_m^{'})}{G(\mathbf{x}|m, \phi_m)}$$

again, by using the substitution from the M-step, we can write

$$\sum_{m=1}^{M} \left[\frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(m|\mathbf{x}) \right] \log \frac{f^{'}(m)}{f(m)} = \sum_{m=1}^{M} f^{'}(m) \log \frac{f^{'}(m)}{f(m)} \geq 0.$$

since Kullback-Leibler information divergence is nonnegative

Proof of the Monotonic Property of EM Algorithm

in view of the M-step definition we can write the inequality

$$\sum_{m=1}^{M} \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x}\in\mathcal{S}} q(m|\mathbf{x}) \log G^{'}(\mathbf{x}|m, \phi^{'}_m) \geq \sum_{m=1}^{M} \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x}\in\mathcal{S}} q(m|\mathbf{x}) \log G(\mathbf{x}|m, \phi_m)$$

which can be rewritten in the form:

$$\sum_{m=1}^{M} \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x}\in\mathcal{S}} q(m|\mathbf{x}) \log \frac{G^{'}(\mathbf{x}|m, \phi^{'}_m)}{G(\mathbf{x}|m, \phi_m)} \geq 0$$

therefore the increment of the log-likelihood criterion is nonnegative:

$$L^{'} - L \geq \sum_{m=1}^{M} f^{'}(m) \log \frac{f^{'}(m)}{f(m)} + \sum_{m=1}^{M} \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x}\in\mathcal{S}} q(m|\mathbf{x}) \log \frac{G^{'}(\mathbf{x}|m, \phi^{'}_m)}{G(\mathbf{x}|m, \phi_m)} \geq 0$$

Remark. The proof is important for any new application of EM algorithm.

Structural Optimization Criterion

"structural" EM algorithm: explicit solution of the equation

$$G^{'}(.|m,\phi^{'}_m) = \arg \max_{G(.|m,\phi_m)} \left\{ \sum_{x \in \mathcal{S}} \frac{q(m|\mathbf{x})}{|\mathcal{S}|} \log G(\mathbf{x}|m,\phi_m) \right\}, \quad m \in \mathcal{M}$$

making substitution for $G(\mathbf{x}|m,\phi_m)$ we obtain:

$$\sum_{\mathbf{x} \in \mathcal{S}} \frac{q(m|\mathbf{x})}{|\mathcal{S}|} \log \prod_{n \in \mathcal{N}} \left[\frac{f_n(x_n|m)}{f_n(x_n|0)} \right]^{\phi_{mn}} = \sum_{n \in \mathcal{N}} \phi_{mn} \sum_{\mathbf{x} \in \mathcal{S}} \frac{q(m|\mathbf{x})}{|\mathcal{S}|} \log \left[\frac{f_n(x_n|m)}{f_n(x_n|0)} \right]$$

⇒ **the parameters $f^{'}_n(.|m)$ and $\phi^{'}_m$ can be optimized separately**

⇒ for any fixed structural parameters ϕ_{mn} we can write:

$$f^{'}_n(.|m) = \arg \max_{f_n(.|m)} \left\{ \sum_{\mathbf{x} \in \mathcal{S}} \frac{q(m|\mathbf{x})}{|\mathcal{S}|} \log f_n(x_n|m) \right\}$$

and given the new parameters of densities $f^{'}_n(x_n|m)$ we can write:

$$\phi^{'}_m = \arg \max_{\phi_m} \left\{ \phi_{mn} \gamma^{'}_{mn} \right\}; \qquad \gamma^{'}_{mn} = \sum_{\mathbf{x} \in \mathcal{S}} \frac{q(m|\mathbf{x})}{|\mathcal{S}|} \log \left[\frac{f^{'}_n(x_n|m)}{f_n(x_n|0)} \right]$$

Structural Optimization - Gaussian Mixture Model

Gaussian univariate densities: ▸ multivariate discrete mixture

$$f_n(x_n|m) = \frac{1}{\sqrt{2\pi}\sigma_{mn}} \exp\{-\frac{(x_n - \mu_{mn})^2}{2\sigma_{mn}^2}\}, \quad n \in \mathcal{N}, \ m = 0, 1, 2, \ldots, M$$

EM iteration equations: $(m \in \mathcal{M}, n \in \mathcal{N}, \mathbf{x} \in \mathcal{S})$

$$q(m|\mathbf{x}) = \frac{G(\mathbf{x}|m, \phi_m) f(m)}{\sum_{j \in \mathcal{M}} G(\mathbf{x}|j, \phi_j) f(j)}, \qquad w_m^{'} = \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(m|\mathbf{x})$$

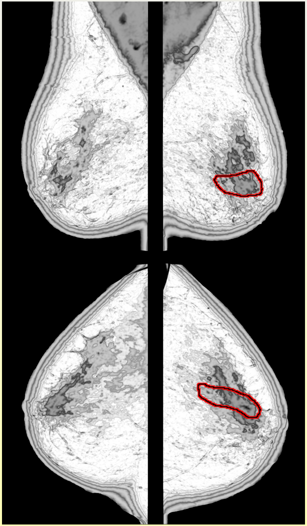
$$\mu_{mn}^{'} = \frac{1}{w_m^{'}|\mathcal{S}|} \sum_{x \in \mathcal{S}} x_n q(m|\mathbf{x}), \qquad (\sigma_{mn}^{'})^2 = \frac{1}{w_m^{'}|\mathcal{S}|} \sum_{x \in \mathcal{S}} (x_n - \mu_{mn}^{'})^2 q(m|\mathbf{x})$$

structural optimization: $\phi_{mn}^{'} = 1$ for the s highest values $\gamma_{mn}^{'}$

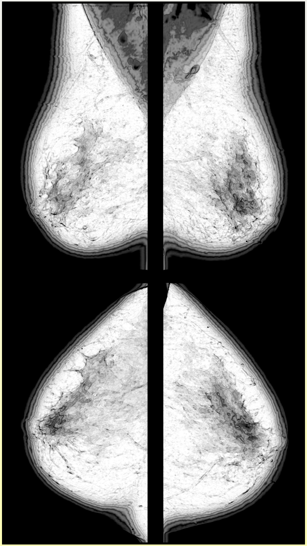
$$\gamma_{mn}^{'} = \frac{w_m^{'}}{2} \left[\frac{(\mu_n^{'(m)} - \mu_n^{(0)})^2 + (\sigma_n^{'(m)})^2}{(\sigma_n^{(0)})^2} - 1 - 2 \log \frac{\sigma_n^{'(m)}}{\sigma_n^{(0)}} \right]$$

C-0002-1: pleomorphic calcif., segmentally distributed

log-likelihood image

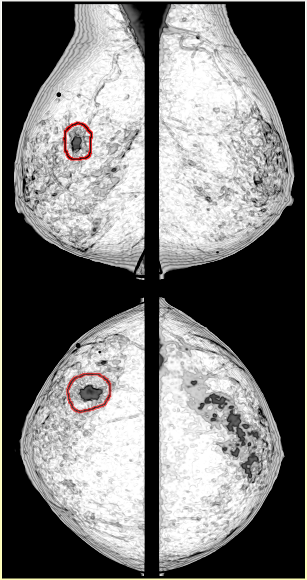


modified log-likelihood image

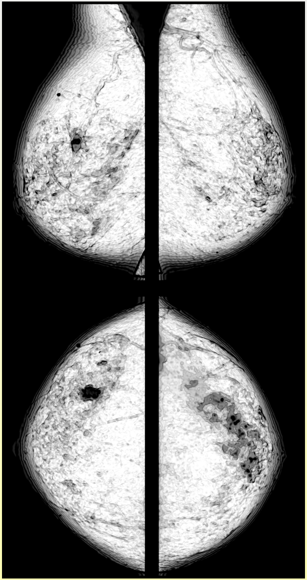


C-0016-1: segmentally distributed calcification

log-likelihood image

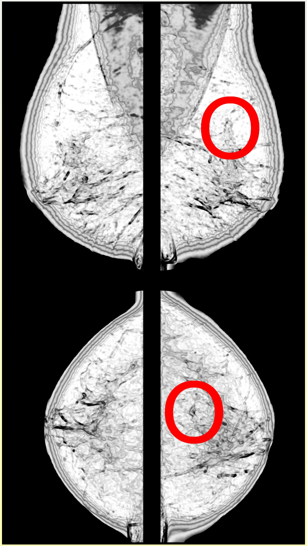


modified log-likelihood image

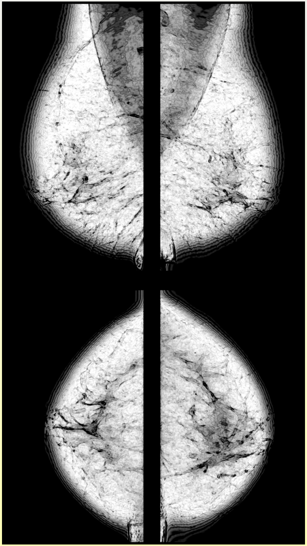


C-0188-1: malignant mass, oval margins

log-likelihood image

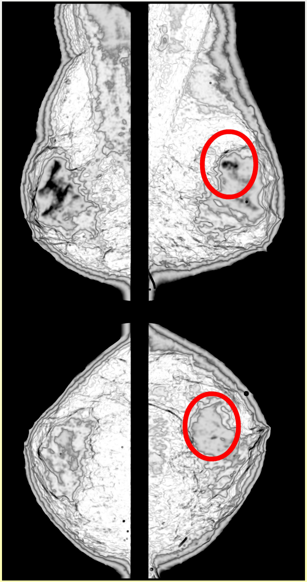


modified log-likelihood image

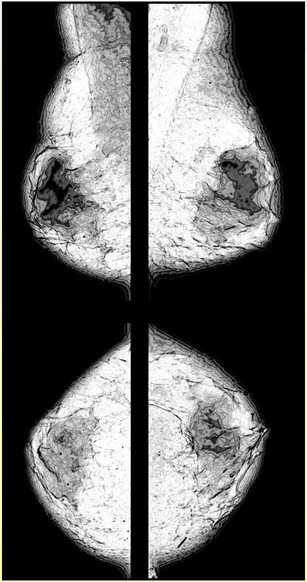


B-3056-1: mass, focal-asymmetric density, margins n/a

log-likelihood image

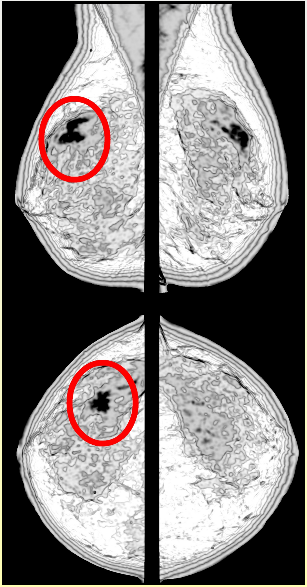


modified log-likelihood image

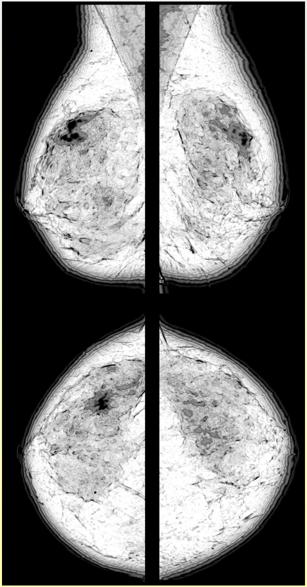


C-0001-1: : mass, irregular shape, spiculated margins

log-likelihood image

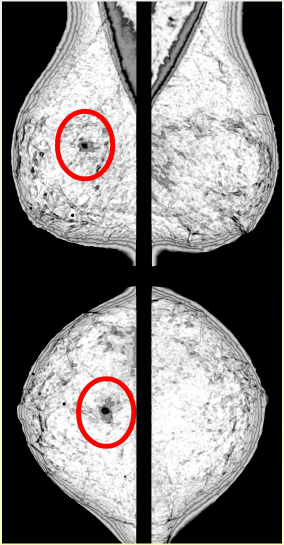


modified log-likelihood image

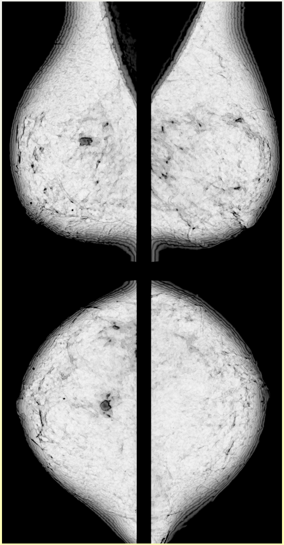


C-0143-1: mass, irregular shape, ill-defined margins

log-likelihood image

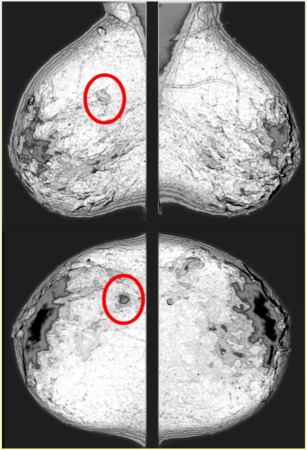


modified log-likelihood image

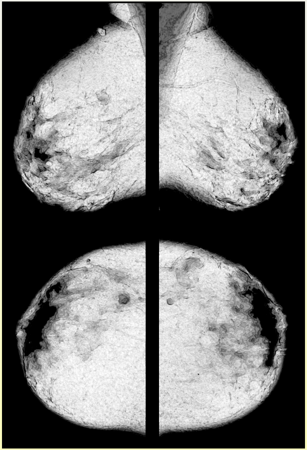


B-3020-1: mass, lobulated shape, ill defined margins

log-likelihood image

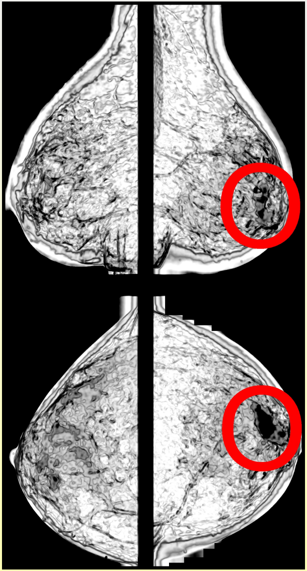


modified log-likelihood image

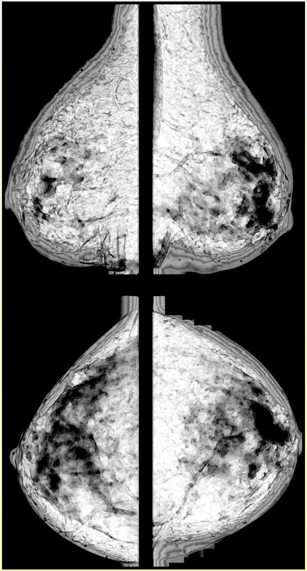


C-0206-1: malignant mass, lobulated margins

log-likelihood image

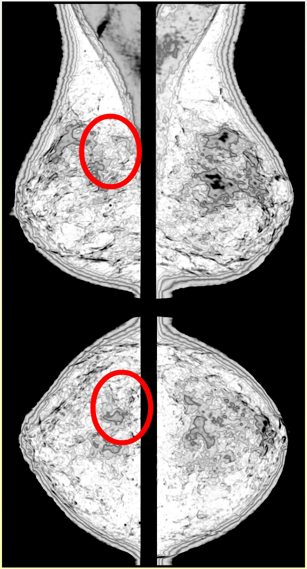


modified log-likelihood image

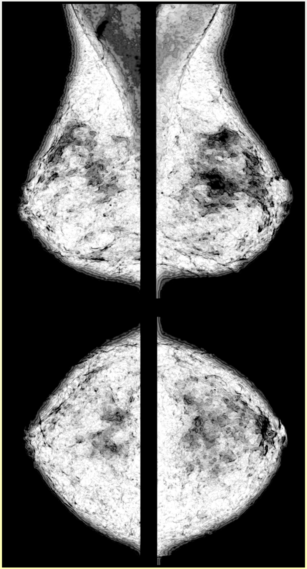


D-4163-1: punctate calcification, clustered distribution

log-likelihood image



modified log-likelihood image



Concluding Remarks

Log-Likelihood Image of Screening Mammogram:

- **purely statistical construct without any medical context**
- **aim:** to facilitate diagnostic evaluation
- **masses:** emphasized as dark regions with contour lines
- **micro-calcifications:** dark spots of the size and form of window
- **topological continuity:** image disintegrates for complex densities

APPLICATION in Monitoring Systems

- evaluation of multivariate time series by sliding window
- detection of unusual, suspect or unsafe states
- sequential optimization of mixture parameters (non-supervised training)
- clear statistical interpretation of the output
- invariant with respect to arbitrary linear transform of data Proof

Literatura 1/4

- Dempster A.P., Laird N.M. and Rubin D.B. (1977): Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc., B*, Vol. 39, pp. l-38.
- Grim J. (1982): On numerical evaluation of maximum - likelihood estimates for finite mixtures of distributions. *Kybernetika*, Vol.l8, No.3, pp. 173-190.
- Grim J., Haindl M. (2003): Texture Modelling by Discrete Distribution Mixtures. Computational Statistics and Data Analysis, 41, 3-4, pp. 603-615.
- Grim J., Somol P., Haindl M. and Daneš J. (2009): “Computer-Aided Evaluation of Screening Mammograms Based on Local Texture Models,” *IEEE Trans. on Image Processing*, Vol. 18, No. 4, pp. 765-773.

Literatura 2/4

- Grim, J. (2011): Preprocessing of Screening Mammograms Based on Local Statistical Models. In: Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies, ISABEL 2011, Barcelona, ACM, pp. 1-5,
- Grim J., Haindl M., Somol P., Pudil P. (2006): "A Subspace approach to texture modelling by using Gaussian mixtures." In: Proc. 18th IAPR International Conference on Pattern Recognition ICPR 2006. (Haralick B., Ho T. K. eds.), IEEE Computer Society, Los Alamitos 2006, pp. 235-238.
- Haindl M., Grim J., Somol P., Pudil P., Kudo M. (2004): A Gaussian mixture-based colour texture model. In: Proceedings of the 17th IAPR International Conference on Pattern Recognition. IEEE, Los Alamitos 2004, pp. 177-180.
- Haindl M., Grim J., Pudil P., Kudo M. (2005): A Hybrid BTF Model Based on Gaussian Mixtures. In: Proceedings of the 4th International Workshop on Texture Analysis and Synthesis "Texture 2005", Beijing, China, October 2005

Literatura 3/4

- Grim J., Somol P., Pudil P. (2010): "Digital Image Forgery Detection by Local Statistical Models", In: Proc. 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Los Alamitos, Calif., IEEE comp. soc., eds. Echizen I. et al., pp. 579-582.
- Heath M., Bowyer K. W., Kopans D. et al. (1998): "Current State of the Digital Database for Screening Mammography." In: *Digital Mammography*, Kluwer Academic Publishers, pp. 457–460. http://marathon.csee.usf.edu/Mammography/Database.html
- McLachlan G.J. and Peel D. (2000): *Finite Mixture Models*, John Wiley & Sons, New York, Toronto.
- Schlesinger, M.I. (1968): "Relation between learning and self-learning in pattern recognition." (in Russian), *Kibernetika*, (Kiev), No. 2, pp. 81-88.
- Titterington, D.M., Smith, A.F.M., & Makov, U.E. (1985): *Statistical analysis of finite mixture distributions.* New York: John Wiley & Sons.

Literatura 4/4

- L. Tarassenko, P. Hayton, and N. Cerneaz et al., (1995): "Novelty Detection for the Identification of Masses in Mammograms," in *4th IEEE International Conference on Artificial Neural Networks*, pp. 442–447. Cambridge, UK.
- C. J. Rose and C. J. Taylor, (2003): "A Statistical Model of Texture for Medical Image Synthesis and Analysis," in *Medical Image Understanding and Analysis*, pp. 1-4.
- C. J. Rose and C. J. Taylor, (2004): "A Generative Statistical Model of Mammographic Appearance," in *Medical Image Understanding and Analysis*, pp. 89-92.
- C. Spence, L. Parra, and P. Sajda, (2001): "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.

◂ Back

Invariance with Respect to Grey-Level Transform

Invariance Property:

the statistical model is invariant with respect to arbitrary linear transform of variables

the transformed data and transformed mixture parameters

$$y_n = ax_n + b, \quad \tilde{\mu}_{mn} = a\mu_{mn} + b, \quad \tilde{\sigma}_{mn} = a\sigma_{mn}, \quad \mathbf{y} = T(\mathbf{x}), \; \mathbf{x} \in \mathcal{S}$$

can be shown to satisfy the EM iteration equations

$$F(\mathbf{y}|\tilde{\boldsymbol{\mu}}_m, \tilde{\boldsymbol{\sigma}}_m) = \frac{1}{a^N} F(\mathbf{x}|\boldsymbol{\mu}_m, \boldsymbol{\sigma}_m), \quad \tilde{P}(\mathbf{y}) = \frac{1}{a^N} P(\mathbf{x})$$

$$q(m|\mathbf{y}) = q(m|\mathbf{x}), \quad \mathbf{x} \in \mathcal{S}, \quad \tilde{w}_m = w_m, \quad m \in \mathcal{M}$$

and therefore the corresponding log-likelihood values differ only by a constant

$$\log \tilde{P}(\mathbf{y}) = -N \log a + \log P(\mathbf{x}), \quad \mathbf{x} \in \mathcal{S}$$

which can be removed by norming the output log-likelihood values Back

Structural Optimization - Multivariate Discrete Mixtures

univariate discrete distributions: $f_n(x_n|m), n \in \mathcal{N}, \; m = 0, 1, \ldots, M$

EM iteration equations: $(m \in \mathcal{M}, n \in \mathcal{N}, \mathbf{x} \in \mathcal{S})$

$$q(m|\mathbf{x}) = \frac{G(\mathbf{x}|m, \phi_m) f(m)}{\sum_{j \in \mathcal{M}} G(\mathbf{x}|j, \phi_j) f(j)}, \qquad w_m^{'} = \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(m|\mathbf{x})$$

$$f_n^{'}(\xi|m) = \frac{1}{\sum_{\mathbf{x} \in \mathcal{S}} q(m|\mathbf{x})} \sum_{\mathbf{x} \in \mathcal{S}} \delta(\xi, x_n) q(m|\mathbf{x})$$

structural optimization: $\phi_{mn}^{'} = 1$ for the s highest values $\gamma_{mn}^{'}$

$$\gamma_{mn}^{'} = \sum_{\mathbf{x} \in \mathcal{S}} \frac{q(m|\mathbf{x})}{|\mathcal{S}|} \log \left[\frac{f_n^{'}(x_n|m)}{f_n(x_n|0)} \right] = f^{'}(m) \sum_{x_n \in \mathcal{X}_n} f_n^{'}(x_n|m) \log \frac{f_n^{'}(x_n|m)}{f_n(x_n|0)}$$

$\gamma_{mn}^{'} \approx$ Kullback-Leibler information divergence

Back