Use of EdgeFlow based ALOE features In Joint Analysis of Multiple Mammographic Views

Zoltán Ludányi

***Abstract*—In screening X-ray mammography two different views are captured of both breasts. First the individual images are analysed independently looking for signs of cancer, but due to the high false positive hits good result can be achieved only by joint analysis of the images. One such approach is based upon the experience that masses and calcifications emerge on both views; so if no matching pair is found, the given object is a false positive hit. Since 3D correspondence is theoretically impossible, a "2.5D" reference system (based on the line parallel to the pectoral muscle and placed in the nipple) is evolved to find corresponding regions on the two images – matching a stripe on one view of a breast to any segment on another view of the same breast.**

Since masses have a distinctive texture, further examination is possible within this stripe. Using intensity, co-occurrence and GLD ("gray level differences") based texture features our algorithm performed 23% loss of false negative hits keeping 93% of true positive ones. These features are characteristic for usual masses but not for spiculated ("stellar") ones therefore adding such features are needed. Kegelmayer et al. suggested using the ALOE ("analysis of local oriented edges") texture feature that is based on the fact that linear structures have the same direction in normal tissues but significantly vary in speculated ones (due to their stellar shape) hence variance of the orientation histogram is distinctive for them. Since classical orientation estimation methods based on differential filtering (eg. Prewitt, Sobel or filtering with Gauss derivatives) result in noise for flat areas and do not provide information far from edges, the use of EdgeFlow for ALOE is more than a rational choice. Adding this feature the loss of false positive hits increased to 31% while still keeping 92% of true positive ones.

I. Introduction

BREAST cancer is one of the most frequent cancers and the leading cause of mortality for women, affecting almost one eighth of them and giving one third of cancers. Evidences show that X-ray mammography is the only reliable screening method giving nearly 95% chance of survival within 5 years due to early detection. [1] Due to the huge number of images captured per year and the high number of false positive diagnoses done by doctors (80–93%), development of mammographic decision support systems is under heavy research.

During X-ray screening two different views are captured of each breast: a CC (cranio-caudal) from above and a leaning lateral view, the MLO. The two most important symptoms are microcalcifications (small spots that have high intensity compared to their environment) and masses (big, high intensity blobs). Our results and other publications on this topic show that obtaining high hit rate while keeping the number of false positive detections low is extremely difficult – alike human evaluation. [2,3]

Microcalcification and mass detector algorithms developed in our department also have quite a good hit rate (nearly 95%) but the false positive hits per image are 3 per image in the case of microcalcifications and 6 in the case of masses. [2,3] However – this rate can be reduced if we manage to make use of an experience of radiologists: microcalcification clusters and masses should appear on both views of the same breast. Therefore finding a pair to a mass- or microcalcification-candidate should increase the probability that the hit is a true positive one.

Since in X-ray mammography perfect 3-D reconstruction is impossible due to breast deformation, we implemented a simple "2.5-D" positioning system between CC and MLO images for this joint analysis. [12] This means that we can assign a stripe on the MLO image to every mass-candidate on the CC image and vice versa. The stripe is based on the position of the nipple and the angle of the pectoral muscle. According to this reference system we could make a hypothesis: "the distances of a mass (measured from the tangent that is parallel to the pectoral muscle and placed in the nipple) in the CC and MLO pictures are equal". This hypothesis was tested by statistical analysis and was found to be true.

Due to the fact that masses have characteristic texture, the reference system can be improved by textural analysis. This is done through the following steps. First the image is segmented by EdgeFlow [4], and then texture features are calculated for each segment, k-means clustering is applied to these features resulting better segmentation. Once we have good quality segmentation, we recalculate features for the new segments. After these preliminary steps we establish the reference system and for each mass-candidate we try to find pairs within the corresponding stripe on the other view – based on texture features. (Note that this pairing may result no pair at all.)

Previously [12] we used texture features that were characteristic only for high intensity masses. However – in some cases masses have a "stellar" shape (sometimes even without a blob in the middle). Kegelmeyer et al [5] suggested using the ALOE (analysis of local oriented edges) texture

Manuscript received June 29, 2006.

Zoltán Ludányi is PhD student at the Budapest University of Technology and Economics, Hungary (email: zoltan.ludanyi@gmail.com).

feature for finding these signs on the image. Combining his idea with orientation information gained from EdgeFlow seems to give a more than 30% better performance with shorter running times than comparable methods.

II. EDGEFLOW

Image segmentation is beset with difficulties. In fact edge detection and segmentation are ill-posed problems, since it is usually undefined what we regard as an edge or one segment. EdgeFlow [4] is an algorithm that is based on well known techniques as differential filtering and orientation estimation but is also enhanced with a clever energy propagating trick that eliminates usual problems of noise and image-dependent thresholding. The only practical free parameter of it is the so called scale (σ) that focuses edge detection in a given scope. (The higher σ is, the larger edges it finds.)

Intensity edges: The EdgeFlow algorithm is based on a differential filtering if edges are defined as local gradient maxima. The differential filter used here is the derivative of the 2D Gaussian function: $G_\sigma(x,y)$.

$$E(s,\theta)=\left|\frac{\partial}{\partial\vec{n}}I_\sigma(x,y)\right|=\left|\frac{\partial}{\partial\vec{n}}\left[I(x,y)*G_\sigma(x,y)\right]\right|=\left|I(x,y)*\frac{\partial}{\partial\vec{n}}G_\sigma(x,y)\right|$$

where $s = (x,y)$ is a pixel, $I(x,y)$ is the image value at (x,y) pixel, and $\vec{n}$ represents the unit vector in the θ direction.

Edge flow vector: The traditional edge detection approach uses a threshold. If the edge energy falls above this value for a given pixel, this pixel is considered as a point of an edge. EdgeFlow also uses thresholding but only after an energy propagation trick with which edge energies are shifted and accumulated during iterative steps. The direction of this propagation is based on probabilities. $P(s,\theta)$ gives the probability of finding an edge in the direction θ within the distance $d = 4\sigma$:

$$P(s,\theta)=\frac{Error(s,\theta)}{Error(s,\theta)+Error(s,\theta+\pi)}$$

where

$$Error(s,\theta)=\left|I_\sigma(x,y)*DOOG_{\sigma,\theta}(x,y)\right|$$

is a kind of predictive coding error in direction θ at scale σ, and

$$DOOG_\sigma(x,y)=G_\sigma(x,y)-G_\sigma(x+d,y),$$
$$x'=x\cos\theta+y\sin\theta,\ y'=-x\sin\theta+y\cos\theta$$

The edge flow vector $\vec{F}(s)$ is derived from $E(s)$ and $\theta(s)$. Its magnitude is proportional to the edge energy, and its phase is directed towards the nearest edge on the given scale σ. (Summing up probabilities instead of finding the maximum reduces the effects of error in prediction.)

$$\vec{F}(s)=\sum_{\theta(s)\le\theta<\theta(s)+\pi}E(s,\theta)\cdot\exp(j\theta)$$

$$\Theta(s)=\arg\max_\theta\left\{\sum_{\theta\le\theta'<\theta+\pi}P(s,\theta')\right\}$$

Edge flow propagation and boundary detection: Edge flow vectors are propagated with an iterative algorithm:

1. $n = 0,\ \vec{F}_0(s)=\vec{F}(s)$
2. $\vec{F}_{n+1}(s)=0$
3. for each s' neighbour of s: if $\vec{F}_n(s)\cdot\vec{F}_n(s')>0$, then $\vec{F}_{n+1}(s')=\vec{F}_n(s')+\vec{F}_n(s)$

The iteration stops when no change occurs. Edge detection ends with the usual thresholding for the magnitudes of the edge flow vectors. According to [4,6] and our experiments a fixed value can usually be found for this purpose, there is no need for any kind of analysis of the resulting image.

Results of EdgeFlow are used in 3 cases: 1) the pectoral muscle is found by a high scale filtering, 2) segmentation is based on smaller scale filtering, 3) orientation information from a low scale filtering is used for ALOE features.

III. THE REFERENCE SYSTEM

For single view analysis three landmarks are named in publications [5]: the pectoral muscle, the nipple and the boundary of the breast. These landmarks segment the breast to its anatomical regions. Many complex algorithms tried to establish 3-D reconstruction [6,7] of breast segments. With 3-D reconstruction the shape of the mass, microcalcification distributions and matching objects could be determined, which could help to distinguish between malignant/benign cases and to reduce false positive cases. Because of the difficulties of 3-D reconstruction our main aim was only to build a simple "2.5-D" positioning system, which can find the approximate corresponding region to a region on the other view. CC and MLO are two-dimensional projections of the three dimensional object, therefore a stripe will correspond to a region on the other image. The system works similar in concept to the procedure a radiologist applies at comparing the two pictures: once he/she found a suspicious symptom, he/she starts looking for a similar one on the other view within a given stripe. The stripe should be parallel to the pectoral muscle and should have the same distance from the nipple as the original sign had on the first view. So, the reference system is to calculate the position of this stripe. The algorithm is founded on three simple hypotheses:

1. The position of the nipple can be estimated by laying a tangent on the breast border parallel with the pectoral muscle.

2. The pectoral muscle on a CC image is assumed to be the vertical axis.
3. The distance covered from the nipple perpendicular to the pectoral muscle on MLO approximately corresponds to the distance measured up on the horizontal axis from the nipple on CC.

The first step of the algorithm is to find the angle enclosed by the pectoral muscle and the horizontal axis on MLO views. With the angle a tangent is laid on the breast border marking the nipple. The distances of the observed region from the tangent (u and v) – are measured. The same distances are measured up on the perpendicular line to the tangent from the nipple of the other view. The two points and the angle of the tangent mark out the stripe. (See Fig. 1.) The correctness of the reference system was tested by a statistical analysis. Cases with 400μ/pixel resolution (600*400 pixels) from the DDSM database [8] were selected indiscriminately, where these contained only one pathological growth on each views according to the radiologists' assessments. Therefore it could be assumed, that those two masses or calcification clusters correspond to each other on the two views. The pixel corresponding to the centroid of the growth on the MLO was determined, and the deviation of the result from the centroid of the growth on the CC was measured in pixel. The results (See Fig. 2.) show that the assumption of the hypotheses was correct though there is some variance caused by the failures of the algorithm, wrong radiologist assessment or the flaw of the hypotheses (because of breast deformation) for a few cases. To compensate the effect of variance the width of the stripe can be increased by a constant or by a number relative to the width of the stripe to counteract the deviation of the algorithm.

Pectoral muscle is a roughly triangular region with high intensity and is located at the upper corner of the MLO mammogram. It has a higher intensity than the surrounding tissues therefore its border appears as a sharp intensity change, as an edge. Therefore boundary detection – Edgeflow – is the first step of finding the pectoral muscle, then the elimination of weak edges with cutting at a threshold. Finding the nipple and transformation in the reference system: The nipple is marked out by a tangent parallel to the pectoral muscle laid on the breast border. With the knowledge of the nipple position and the angle of pectoral muscle connection between the two views is provided by simple coordinate transformations.

The correctness of the reference system and our hypothesis were tested by a statistical analysis. Results showed that the assumption was correct though there is some variance caused by the failures of the algorithm, wrong radiologist assessment or the flaw of the hypothesis (because of breast deformation) for a few cases. To compensate these effects the width of the stripe can be increased by a constant or by a number relative to the width of the stripe to counteract the deviation of the algorithm.

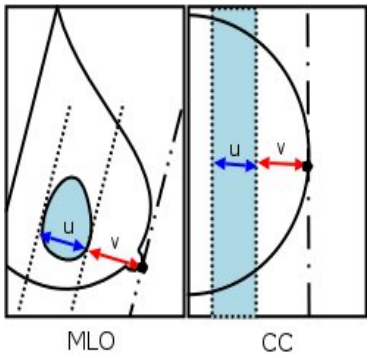

Fig. 1. The corresponding stripe on the CC of a selected region on the MLO

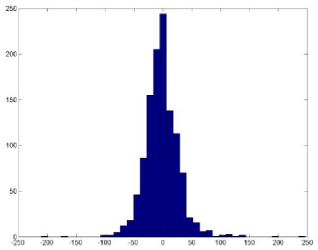

Fig. 2. Histogram of errors in reference system positioning (total: 1159 cases)

IV. Pairing of Masses

Making use of the reference system one could easily improve results of any mass- or calcification-detection algorithm simply by checking if any symptom-candidate on one view has any possible pairs on the other view – that is "if any candidate falls in the corresponding stripe". If a pair is found, the identifying probability of the given mass should be increased, otherwise decreased.

The expression identifying probability is used since finding pair to a mass merely says that there is something characteristic in the breast – because it can be seen from both views – but it might be either malignant or benign. Alike – if no pair is found, it says the mass supposed to be recognized on one of the images is only virtual, its appearance is the result of some overlaying tissues. Note also that this correspondence can solely be done for "clear" breasts. For dense ones even experienced radiologists can rarely find the mass on both images.

However – masses have a distinctive texture that makes it possible to do texture-based pairing within the stripe. Since the given mass detecting algorithms are fairly characteristic in size and shape of the identified mass, a good segmenting algorithm is also needed beforehand. We used the results of EdgeFlow for this purpose.

The first question arising when trying to apply EdgeFlow is the selection of the proper scale. After running it for a wide variety of mammographic images and range of scales,

scales 1, 2 and 3 are seemed to be significant in our case. Since the EdgeFlow algorithm itself only detects edges, some further steps are necessary to create a segmentation from its output: line segments should be linked creating continuous borders and closed segments. With some basic morphological operations (removing isolated pixels, dilation, removing disjoined line segments) one can get a practically good segmentation, but the result is sometimes too detailed, or may also contain unduly small segments. Computing some texture features and using clustering for the segments based on them can solve these problems. Note that in this case the number of clusters is not equal to the numbers of segments created after merging the members of each cluster, since these members may form more isolated groups on the image. With about maximum 100 isolated areas the resulting segmentation proved to be adequate for our aims. By binary search for the number of clusters needed this number can be approximated in 2-3 steps. (The number of segments on the original segmentation varies from about 80 up to even 300. Small regions are forced to merge even if this causes less than 100 segments.) Texture features used are as follows: mean of intensity, variance of intensity, mean and variance of co-occurrence values, mean and variance of grey-level-differences. (Co-occurence matrix and grey-level differences are image features used with great success for mass detection in the project, one can find descriptions in [9].) For clustering four methods have been tested: single linkage hierarchic, k-means, fuzzy k-means and subtractive clustering [9,10,11]. According to our experiments, k-means has been chosen for its simplicity and reliability.

Once a good segmenting algorithm and characteristic texture features are given, the accuracy of mass detecting can be increased by texture based matching on different views. Matching goes by the following steps: [12]

1. In the beginning results of a mass detection algorithm are given – usually a binary mask covering the mass-candidate area with the probability of that hit (Fig. 3b and d). (During the matching mass-candidates of one image called source are to be paired with mass-candidates of the other one called target.)
2. A mass-candidate – in our example the upper one on Fig. 3b – is chosen for investigation from the source image.
3. The reference system is established for both views. (This assigns a stripe to the mass chosen. See Fig. 3e)
4. We identify possible pairs of all the segments on the source image overlapped by the source mask (the mass-candidate). This pairing for a segment goes by identifying all segments that are close in texture feature (threshold is set dynamically depending on the minimal distance), and not significantly far in intensity, size or distance (assigned by the reference system) from the original segment. (See Fig 3f – segments overlapped by the stripe are highlighted; and Fig. 3g, where segments eliminated by the rule based system are not shown and intensity is proportional with similarity.)
5. Taking the hits on the source image one by one, we examine if its pairs overlap with any of the mass candidates on the target image. If so, the similarity of this pair is the mean of those nonzero elements on the non-binary mask (see Fig. 3h, where texture feature based thresholding is applied) that are covered by the given mask pair on the target image. (We found two pairs, one of them cover the mass candidate on the target image. This fact increases the probability that both of them are true positive hits.)
6. This pairing is done in reverse direction as well. Mass-candidates that are not paired are dropped.

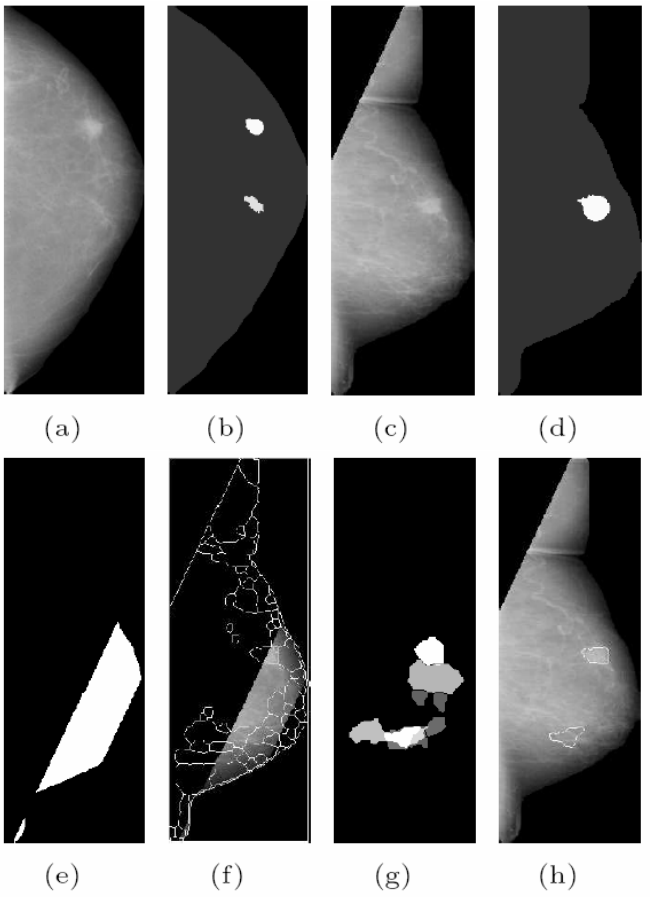

Fig. 3. Steps of pairing

We tested the above detailed algorithm on 363 pairs of the DDSM database. In this test set cancer is marked on both pictures for 97 pairs, marked only on one image (though it is visible on the other also) in 20 cases, and in 2 more cases the cancer is visible only on one image of the pair. The set consisted of 3618 mass candidates that means 5 candidates per image.

Using the pairing algorithm we could reach a performance of 92% TP / 23% FP (while keeping 92% of true positive hits we could eliminate 23% of false positive ones) or with stricter parameter setting 96% TP / 12 % FP.

There seemed to be two ways to gain better results: either by doing a cleverer pairing or by adding even more characteristic features. Since these features are characteristic only for high intensity masses, while other types also exist, the latter seemed to be more promising.

V. Using EdgeFlow Based ALOE As Texture Feature

Analysing our results we found that our texture features are not characteristic enough for distinguishing normal tissue from masses – especially in cases where the mass is not the usual high intensity homogenous blob. If the breast tissue also has high intensity or the mass is a form of "architectural distortion" (it consist only of radial stellar extensions without a high intensity core in the middle), our pairing algorithm pairs malignant segments with normal tissue with a higher probability. Therefore using such spiculated characteristic features is also needed.

Kegelmeyer et al. suggested using ALOE (analysis of local oriented edges) as such a feature. The basic idea of ALOE is as follows. Edges, linear structures in normal tissue of the breast tend to have the same direction. On the other hand: spiculated lesions usually have a radial form, or at least its extensions have a wider variety in direction. This variation can easily be analysed by variance of the pixelwise orientation histogram. (See Fig. 4 and 5. Note that orientation information is shown in 256 bins.)

The necessary orientation information can be gained from spatial differential filtering. Eg. using classical filters as Prewitt or Sobel resulting $D_v(x,y)$ vertical and $D_h(x,y)$ horizontal filter responses, the necessary pixelwise value is $\theta(x,y) = \arctan(D_v(x,y)/D_h(x,y))$; or using directional filters as rotated Gauss derivative functions, we may choose $\theta(x,y)$ to be the angle in which filtering results in maximum. Note that this feature is also not bad for usual masses since they have a strong outline, giving significant orientation information in all directions.

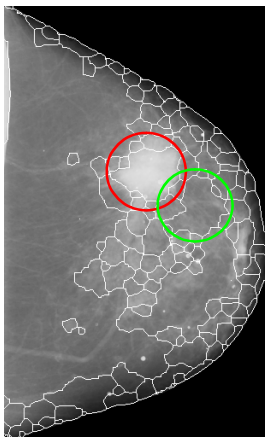

Fig. 4 segmentation of a breast

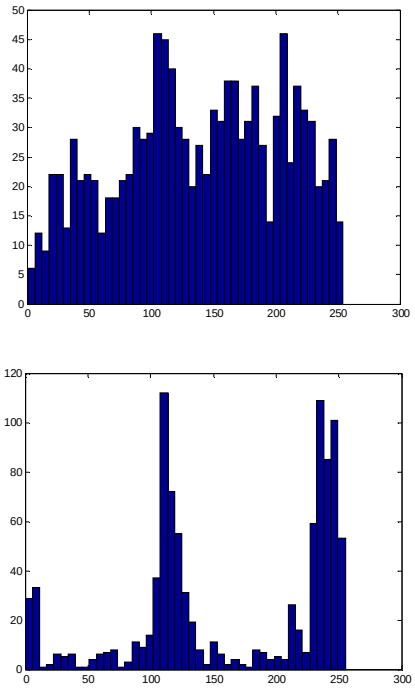

Fig. 5 orientation histogram of a spiculated mass segment and a segment of normal tissue respectively (see Fig. 4 also)

The aforementioned classical filters are almost unusable for our purposes since they result in noise for even the smallest flat parts of the image, and therefore flat orientation histogram – that should be characteristic for spiculated lesions. This problem can be solved by using better filters that gives no significant response for noisy images. Of course the more sophisticated methods we use for better performance, the more computation time is to be paid. However – in our case the results of EdgeFlow (and also results of filtering with Gauss derivatives) is already given, so it is only logical to try using them as a source of orientation information. The other advantage of using EdgeFlow is that orientation in a given pixel means "the direction in which the closest edge is to be found", so EdgeFlow not only eliminates noisy data for pixels of flat segments but also replaces them by meaningful information.

Putting this EdgeFlow ALOE texture feature in our pairing system increased reduce of false positive hits from 23% to 31% that means more than a 34% better result while still keeping 92% of true positive ones. (Note that the increase with Prewitt and Sobel operator was only a mere 1.5% and 2% respectively.)

For the sake of comparison we also tried using orientation gained from pure Gauss derivatives. The maximum performance achieved stopped at 28% FP for 91% TP but for this the resolution of orientation had to be risen from 8 to 32 – Gauss derivatives are rotated in 32 directions for filtering instead of only 8. This multiplies filtering times by 4 but since additional steps of EdgeFlow also eat lots of time, this only means that the EdgeFlow based version is about

20% faster than the best Gauss derivative based version – that is still worse in performance. (Note that with pure Gauss using 8 orientations the maximum achievable performance stopped at 89% TP.)

REFERENCES

[1] L. Tabár: Diagnosis and In-Depth Differential Diagnosis of Breast Diseases. Breast Imaging and Interventional Procedures, ESDIR, Turku, Finland, 1996.

[2] 2. G. Horváth, J. Valyon, Gy. Strausz, B. Pataki, L. Sragner, L. Lasztovicza, N. Székely: Intelligent Advisory System for Screening Mammography. Proc. of the IMTC 2004 - Instrumentation and Measurement Technology Conference, Como, Italy, 2004, Vol. pp.

[3] 3. N. Székely, N. Tóth, B. Pataki: A Hybrid System for Detecting Masses in Mammographic Images. Proc. of the IMTC 2004 - Instrumentation and Measurement Technology Conference, Como, Italy, 2004. Vol. pp.

[4] Wei-Ying Ma, B. S. Manjunath: EdgeFlow: A Technique for Boundary Detection and Image Segmentation. IEEE Trans. on Image Processing, Vol. 9, No 8, pp. 1375-1388, August 2000.

[5] W. P. Kegelmeyer Jr., J. M. Pruneda, P. D. Bourland, A. Hillis, M. W. Riggs, and M. L. Nipper, "Computer-aided mammographic screening for spiculated lesions," Radiology, vol. 191, pp. 331–336, May 1994.

[6] R.J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, R. A. Borges, A. F. Frre: Automatic Identification of the Pectoral Muscle in Mammograms. IEEE Trans. On Image Processing, Vol. 23, No 2, pp. 232-245, February 2004.

[7] M. Yam, M. Brady, R. Highnam, Ch. Behrenbruch, R. English and Y. Kita: Three-Dimensional Reconstruction of Microcalcification Clusters from Two Mammographic Views. IEEE Trans. on Image Processing, Vol. 20, No. 6, June 2001.

[8] M. Heath, K. Bowyer, D. Kopans, R. Moore, K. Chang, S. Munishkumaran and P. Kegelmeyer: Current Status of the Digital Database for Screening Mammography. In: Digital Mammography, N. Karssemeier, M. Thijssen, J. Hendriks and L. van Erning (eds.) Proc. of the 4th International Workshop on Digital Mammography, Nijmegen, The Netherlands, 1998. Kluwer Acamdemic, pp. 457-460.

[9] I. Pitas: Digital Image Processing and Algorithms and Applications. John Wiley & Sons, New York, 2000.

[10] D. E. Gustafson and W. C. Kessel: Fuzzy Clustering with a Fuzzy Covariance Matrix. Proc. IEEE-CDC, Vol. 2, pp. 761-766, 1979.

[11] R.O. Duda, P.E. Hart: Pattern Classification and Scene Analysis. John Wiley & Sons, 1973.

[12] Márta Altrichter, Zoltán Ludányi, Gábor Horváth: Joint Analysis of Multiple Mammographic Views in CAD Systems for Breast Cancer Detection. Lecture Notes in Computer Science, Springer-Verlag GmbH, Vol. 3540 / 2005: Image Analysis: 14th Scandinavian Conference, SCIA 2005, Joensuu, Finland, June 19–22, 2005. pp. 760–769. (ISBN: 3-540-26320-9)