IMPROVING COMPUTER AIDED BREAST CANCER DETECTION USING JOINT ANALYSIS OF MULTIPLE MAMMOGRAPHIC VIEWS

Márta Altrichter*, Zoltán Ludányi* and Gábor Horváth*

* Department of Measurement and Information Systems, Budapest University of Technology and Economics, Budapest, Hungary

grimma@sch.bme.hu, quad@dpg.hu, horvath@mit.bme.hu

Abstract: During X-ray mammography two different views are captured and analysed of both breasts. First radiologists and CAD systems look for special signs of cancer in the four individual images independently. But as radiologists' methodology shows that good result can be achieved only if joint analysis of the images is also done, a method for joint analysis of the breast's two views in Computer Aided Diagnostics is needed to be developed. The procedure is based upon the experiences of radiologists: masses and calcifications should emerge on both views, so if no matching is found, the given object is a false positive hit. First a reference system is evolved for positioning on the two views. Calcification clusters obtained in individual images are matched in "2.5 D" provided by the reference system. Masses detected in individual images are further examined with texture segmentation. The proposed approach can significantly reduce the number of false positive hits both in calcification and in mass detection. To analyse the "2.5 D" reference system's sensitivity and validity statistical methods are applied.[1]

Introduction

Mammography is currently the best method for breast cancer screening. Several experiments conducted to analyse the effect of large scale screening with mammography demonstrated reduction in mortality (HIP experiment in 1963). But the nationwide screening projects result in enormous amount of mammograms to be analysed in each year. To ease and assist the work of overburdened radiologists computer aided diagnostic (CAD) systems are developed.

A CAD system has to analyse two views captured from the two breasts in each case: a CC (cranio-caudal) from above and a leaning lateral view, the MLO. The two most important symptoms are microcalcifications and masses. Microcalcification clusters have a higher X-ray attenuation than the normal breast tissue and appear as a group of small localized granular bright spots in the mammograms. Masses appear as areas of increased density on mammograms.

In the past many algorithms were developed to detect these abnormalities on a single mammogram. The main problem of these algorithms is that the cost of sensitivity needed to detect most of the positive cases – a hit rate of 90-95% – is the too high false positive hit rate. [2], [3], [4], [5]

Many attempts at using information from two mammographic views have been explored in literature. [6] Most of them tried to reconstruct masses or calcification clusters in 3D to help their assessments as benign or malignant signs of cancer. [7], [8] We try to use the information laying in the comparison of two views to decrease the number of false positive hits, in a way that the true positive ones are not or barely decreased. The method sets off from the fact that the images of calcifications and masses have to appear on both views (MLO and CC). To be more precise they must be on positions of the two views that correspond to each other. Therefore we need to construct a reference system with which we can position within the breast on the 2-D images. Instead of a full 3-D reconstruction of breast we suggest a simpler procedure which we call "2.5-D" correspondence.

The two pathological abnormalities have different texture characteristics, therefore the joint analysis of them is slightly different. As tissues with calcifications and normal breast tissues are rather similar in texture for the regions of calcification clusters only the 2.5-D matching method is used. However, for masses texture segmentation could be developed to further refine their assignment.

The construction of a reference system is presented in Sect. 2. The use of the reference system on calcification clusters is examined in Sect. 3. A texture segmentation is described in Sect. 4, and Sect. 5. describes how the reference system and the texture analysis can combine. The performance of the joint analysis is evaluated in Sect. 6. The sensitivity of the reference system for the joint analysis is examined in Sect. 7. and finally conclusions are drawn in Sect. 8.

Construction of the reference system

For the joint analysis of two mammograms a reference system is needed. Many complex algorithms can be found in literature, which try to establish complete 3-D reconstruction of breast segments. [7],[8] With 3-D reconstruction the shape of the mass, microcalci-

[1]This work was supported by National Office for Research and Technology under contract IKTA 102/2001.

ISSN: 1727-1983

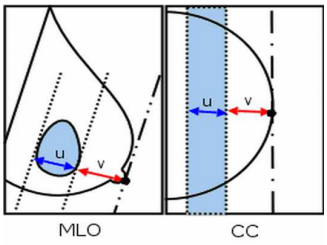


Figure 1: The corresponding stripe on the CC of a selected region on the MLO

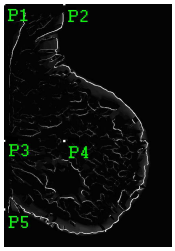


Figure 2: ROI selection

fication distributions and the matching objects can be determined, which help to distinguish between malignant/benign cases. 3-D reconstruction would attempt to build compressed 3-D breast model from two differently compressed breast views (perfect reconstruction is impossible due to different deformation). Thus the success would also mean the decrease of error in the reference system by the deformation. Unfortunately these complex algorithms usually need extra information about the images, how they were taken, the thickness of the compressed breast, and so on.

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, thus it is able to help joint analysis of the CC and MLO views. The system works similar in concept to the procedure a radiologist applies at comparing the two pictures.

As CC and MLO views are 2-D projections of the 3-D object a stripe will correspond to a region on the other image. The reference system will be able to determine the position of the corresponding 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) on Fig. 1. – 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.)

A, Finding the angle of the pectoral muscle on the MLO view

Pectoral muscle is one of the 3 main landmarks – besides nipple position and boundary of the breast – which segment the breast to its anatomical regions. [9]

Pectoral muscle is always located at the upper corner of the breast. It is roughly triangular with high intensity compared to the surrounding tissues. Most of the algorithms try to find the intensity change at the boundary of the pectoral muscle.

i, Muscle search with the use of adaptive iterative threshold

The results of adaptive iterative threshold method described in literature [10] showed that problems emerge mainly in case of extreme images, where our initial assumptions – like the sharp intensity change – don't stand. To provide a more accurate angle detection of pectoral muscle a special edge detection algorithm the EdgeFlow was examined.

ii, Muscle search based on EdgeFlow

EdgeFlow is an edge detection method based on differential filtering with Gabor wavelets, Derivatives of 2D Gaussian functions and Difference of Offset Gaussians. [11] Traditional boundary detection algorithms determine the location of the edge on the filtered images by thresholding it. In contrast to this EdgeFlow uses first a special edge energy accumulation process, ensuring more accurate results. Due to the accumulation, and the filter scalability the refinement of the segmentation can be adjusted with a so-called scale parameter (σ).

EdgeFlow's result can be used not only at the search for the pectoral muscle, but at the texture segmentation needed for mass matching as well, as it can accumulate edge energy of intensity and texture change as well.

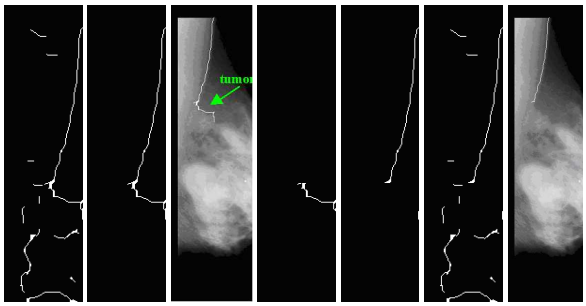
The first step of muscle search is boundary detection with EdgeFlow, then the elimination of weak edges with cutting at a threshold. Secondly, a region of interest (ROI) containing the pectoral muscle is obtained according to the commendation of paper [9]. Five control points are used. P1: top-left corner pixel, P2: top-right pixel of the breast boundary, P5: lowest pixel on the left of breast boundary, P3: 2/3 between P1 and P5, P4: completes a rectangle with P1, P2 and P3 and forms the ROI. (See Fig. 2.)

As the whole line of the pectoral muscle is not needed for the matching, just an approximation of the angle enclosed by the pectoral muscle and the horizontal axis, the iteration processing the lines diverges from paper [9]. The deletion of disturbing line segments will be allowed.

The pseudo code of the algorithm is:

(1) $n = 0$, $BW_0 = ROI$
(2) $L_n =$ longest object on BW_0
(3) L_n is divided to parts with uniform length along the vertical axis
(4) $L\_bad_n =$ objects which enclose $< 40°$ or $> 90°$ angles with the horizontal axis
(5) $L\_good_n = L_n - L\_bad_n$, $BW_{n+1} = BW_n - L\_bad_n$

ISSN: 1727-1983



(a) (b) (c) (d) (e) (f) (g)

Figure 3: (a) BW_0, (b) L_0, (c) $L_0 + Picture$, (d) L_bad_0, (e) L_good_1, (f) BW_1, (g) $L_1 + Picture$

(6)
- if $BW_{n+1} == BW_n$
- then iteration stops, the pectoral muscle is the object L_good_n
- else $n = n+1$ and go to Step 2.

The steps 3, 4 and 5 are used to increase the robustness of the algorithm for cases, where a mass or blood vessel deflects the edge of the pectoral muscle. (See Fig. 3.) With the deletion of segments that cannot be part of the pectoral muscle the angle is better approached.

B, 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. (See. Fig 4.)

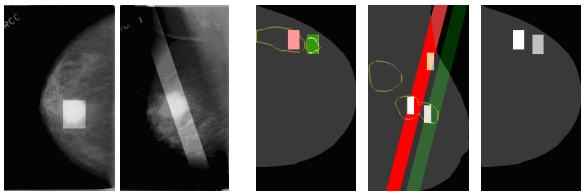
Application of 2.5-D reconstruction on microcalcification clusters

The probability accompanied to a calcification cluster is modified with the area ratio of the stripe corresponding to it and of other calcification clusters found on the other view. Fig. 5. represents such a matching. 5(a) shows the result of the calcification detection algorithm on a CC view. Two ROIs were found where the brightness (intensity) values of these regions are proportional to the "probability" of beeing a calcification cluster. The Fig. 5(b). image shows the corresponding stripes on the MLO view, while Fig. 5(c). shows the result: the "probabilities" are changed – increased in this case – according to the matching rations.

Texture segmentation

For masses a more sophisticated texture-based joint analysis is possible for they have distinctive texture. Since the given mass detecting algorithms are fairly characteristic in size and shape of the identified mass, a good segmenting algorithm is needed.

The first question arising when trying to apply EdgeFlow is the selection of the proper scale. After running it



(a) (b) (a) (b) (c)

Figure 4: (a) Original image, (b) Strip on MLO

Figure 5: Figure showing the process of matching

for a wide variety of mammographic images and range of scales, scales 1, 2 and 3 seem to be useful at 400μ resolution. 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.

However — 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 the number of clusters is not equal to the numbers of segments created after clusterization, since one cluster may contain more isolated groups on the image. (The number of segments on the original segmentation varies from about 80 up to even 300.) With binary search for about 100 isolated areas on the resulting image, an adequate segmentation can be achieved in 2-3 steps. (Small regions are forced to be merged even if we have less than 100 segments.)

The 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-occurrence matrix and grey-level differences are image features used with great success for mass detection in the project. Reviewing these features is beyond the scope of this article, but one can find descriptions in [12].) For clustering four methods have been tested: single linkage hierarchic, k-means, fuzzy k-means and subtractive clustering [12], [13], [14]. Reasoned by our experiments the first two ones have been chosen for their simplicity and reliability.

Matching of masses

Once a good segmenting algorithm and characteristic texture features are given, the accuracy of mass detection can be increased matching their results on the different views. If pairs can be found on both pictures of the same side, the identifying probability of that mass should be increased, finding no pair reduces this probability.

The expression *identifying probability* is used since finding a counterpart to a mass merely says that there is something characteristic *in the breast* – because it can be

ISSN: 1727-1983

seen from both views – but it might be either malignant or benign. Alike — if no counterpart 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 a mass on both images.

Pairing goes by the following steps:

(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 (see Fig. 6(b)., 6(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 this case the upper one – is chosen from the source image.
(3) Segments overlapped by a mass-candidate are identified, and for each segment the distance – a measure of similarity – is computed from segments on the target image creating a non-binary mask where nonzero elements cover the paired segments. For each segment the covering values are the reciprocal of the above detailed distance. Thus on Fig. 6(e). intensity is proportional to similarity. (Areas fairly overlapping with the source mask or having radically different size or intensity are omitted.) Since one mass-candidate area may overlap more segments, these masks should be added. In fact as EdgeFlow may be run for more scales and for more clustering algorithms, one may also sum along these dimensions as well. (In our experiments 3 scales and k-means segmentation are used.) At the end this non-binary mask is thresholded. The resulting pairs can be seen on Fig. 6(f).
(4) 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 that are covered by the given mask pair on the target image. It seems that the values of similarity are not representative enough to make further thresholding at resulting pairing, so they are treated as binary values.
(5) This pairing is done in reverse direction as well. Mass-candidates that are not paired are dropped.

The algorithm above can be combined with the reference system used in the case of calcifications by computing distances only for those segments which overlap the corresponding stripe.

Performance

The calcification matching was analysed over 188 cases (376 pairs of mammographic images). 66 of these cases contained malignant calcifications. The original calcification searching algorithm solved 63 of these cases but only 1 (in 122) of the negative ones. With the combined matching 61 malignant cases and 15 normal cases are solved.

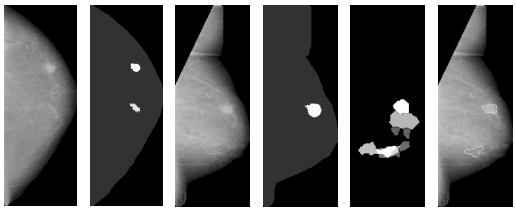


Figure 6: Steps of pairing for masses based on area and texture matching

- 12.4% of false positives hits were dropped
- 96.8% of true positive hits were kept
- 3.2% of true positive hits were dropped

The main reasons for the loss of positive markers are summarised in the conclusion.

In the case of mass detection the above detailed algorithm has been tested on 363 pairs of mammographic images – with 256 masses – from the DDSM database. Surprisingly using only texture based pairing gave best performance compared to only reference system based and combined methods. (In the case of combined method errors of the reference system and pairing seemed to accumulate, while the merely texture based pairing could also exclude major of the breast area fulfilling its task with less error.)

- 94% of true positive hits were kept
- 20% of false positives were dropped

There are 4 main reasons for dropping true positive hits: (i) a mass can be seen on both views but only one of them is marked as mass-candidate (in 20% of pairs), (ii) the mass can be seen only on one of the images (only in the case of a few pairs), (iii) the reference system is not accurate enough (in 7 cases of the 18 false negative pairs when combined method is used), (iv) miscellaneous errors of the pairing algorithm. The results of mass-pairing are worse than that of calcification-pairing since the number of mass-candidates per image is higher than the number of calcification-candidates per image with worse hit probabilities.

Sensitivity analysis

Results showed a decrease of False Positive cases, but the loosing of true cases made the further analysis of the matching necessary.

Statistical analysis of correspondence error

The correctness of the reference system was tested by a statistical analysis. Cases with 400μ/pixel resolution (around 600*400 pixels/image) from the DDSM database

ISSN: 1727-1983

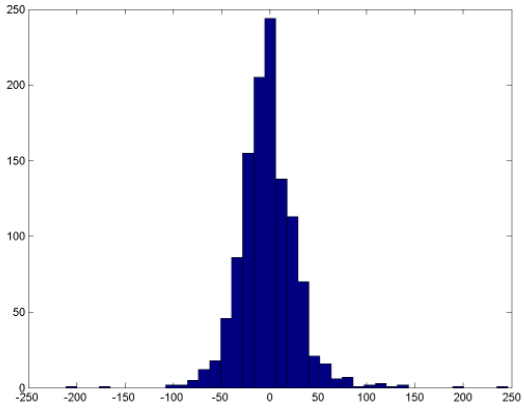


Figure 7: Histogram of pixel errors, number of cases 1159

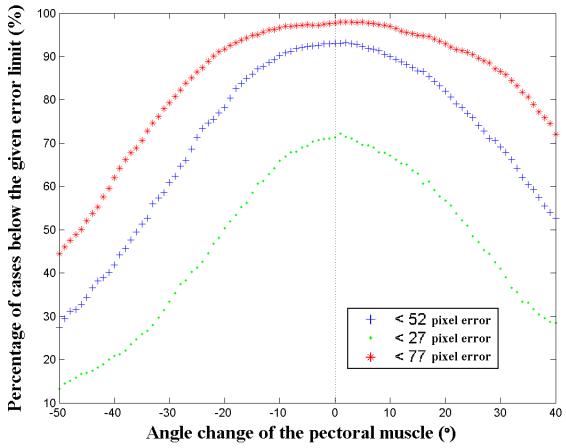
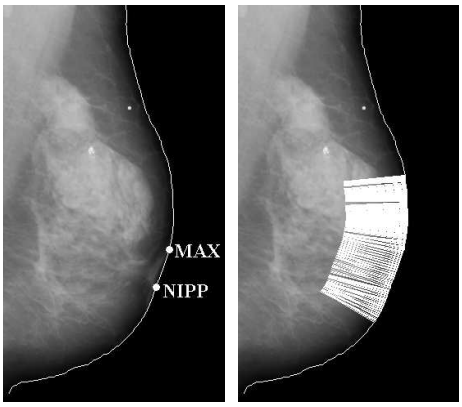


Figure 8: The corresponding stripe on the CC of a selected region on the MLO



(a) MLO with *MAX* and *NIPP* (b) Normals laid on the border in search of *NIPP*

Figure 9: Example of search for *NIPP*

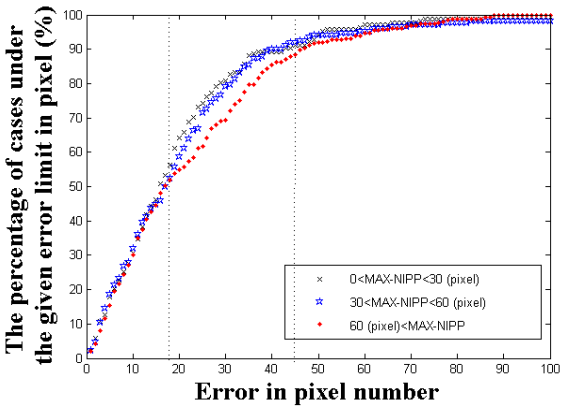


Figure 10: Percentage of cases having smaller error in reference system along the limits of x-axis

were selected, so that the cases contained only one pathological growth on each views based on the radiologists' assessments. Therefore it has a high probability, 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. 7.) 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. As 48 pixel is the average diameter of the regions marked by the calcification algorithm, in 93% of the cases the stripe and the corresponding region will cut each other, even without the increase of the stripe width.

A, Search for constant error

The above mentioned statistical analysis was made with specially modified reference systems, to conclude that the variables (angle, nipple position) calculated so far does not contain any constant error. The angle was modified from $k = -50 \ldots 40$, and a new reference system was established with it. Fig. 8. shows that there is no constant correction factor for the pectoral muscle's angle.

B, New parameter collection

To deal with the correspondence error we tried to find new features, parameters of a case, which would show and separate cases with inclination toward having a big error at matching and those with a good reference system. The following paragraphs will show an attempt of finding such a feature.

During the construction we assumed the nipple to be marked out by the tangent parallel to the pectoral muscle. This assumption is often not correct. C. Olse'n and Georgeson examination showed, that the nipple is located with 99% probability within $MAX \pm 100 pixel$ along the breast border, where *MAX* is the so far estimated nipple. [15] To find the real *NIPP* nipple position the algorithm searches through these breast border pixels. The algorithm assumes, that the intensity drop as we enter into breast tissue deeper is the largest at the location of the nipple.

Thus the algorithm first determines the average intensity gradient G_i along the normal – to the tangent laid

ISSN: 1727-1983

on the border at that pixel – of each pixel running on $MAX \pm 100$ border. The pixel having the maximum G_i is the new nipple position: *NIPP*. (See Fig. 9(a).)

On Fig. 10. we can see that those cases which had a distance between *MAX* and *NIPP* more than 60 pixels had a slightly larger variance in the statistics of reference system error. The difference is around 10% in between 18-45 pixel error.

The 10% difference is promising, though not yet enough to distinguish accurately between good and bad reference systems. Thus we are still searching for a good feature to do so. If such a parameter could be found with the above mentioned statistical analysis, it might also enable the incorporation of the parameter in the construction of the reference system to make the result system more accurate. Currently our main search field for the feature is by modelling the compression.

Conclusions

The paper proposed a relatively simple way of combining the results of mass and microcalcification detection algorithms applied for individual X-ray breast images. The joint analysis follows the way applied by skilled radiologists: if a suspicious area can be found in one view, usually its corresponding pair should be detected in the other view of the same breast. The first results – based on a few hundred of cases - show that using this approach the number of false positive detections can be reduced significantly while the decrease of true positive hits is relatively small. The loss of a few true positive cases comes from three problems: (i) the variance of the corresponding distances (Fig. 7), (ii) the lack of detected microcalcification cluster or mass in one of the corresponding views, (iii) the lack of microcalcification cluster or mass in one of the corresponding views. The variance can be decreased if the different deformation caused by breast compression is taken into consideration. The reason of the second problem is that although there are signs of tumor in both views, the primal algorithms can detect them only in one of the views, in these cases the primal algorithms should be improved. The third problem cannot be solved as in these cases the signs cannot be seen in one of the images even by a skilled radiologist. This means that in such cases other modality like ultrasound should also be used. The proposed joint analysis system is under testing: the whole 2.600 cases of the DDSM data base will be analysed in the near future.

References

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

[2] HORVÁTH G., VALYON J., STRAUSZ GY., PATAKI B., SRAGNER L., LASZTOVICZA L., SZÉKELY N. (2004): 'Intelligent Advisory System for Screening Mammography', Proc. of the IMTC 2004 - Instrumentation and Measurement Technology Conference, Como, Italy

[3] SZÉKELY N., TÓTH N., PATAKI B. (2004): 'A Hybrid System for Detecting Masses in Mammographic Images', Proc. of the IMTC 2004 - Instrumentation and Measurement Technology Conference, Como, Italy

[4] SONGYANG YU, LING GUAN (2000): 'A CAD System for the Automatic Detection of Clustered Microcalcifications in Digitized Mammogram Films', *IEEE Trans. on Medical Imaging*, Vol. 19, No. 2

[5] VERMA B. and ZAKOS J. (2001): A Computer-Aided Diagnosis System for Digital Mammograms Based on Fuzzy-Neural Feature Extraction Techniques. *IEEE Trans. on Information Technology in Biomedicine*, vol. 5. No. 1. pp. 46-54.

[6] S. VAN ENGELAND, KARSSEMEIJER N., HENDRIKS J. (2002): 'Using Information from Two Mammographic Views to Improve Computer-Aided Detection of Mass Lesions', Digital Mammography, IWDM

[7] LING SHAO, 'Three-Dimensional Mass Reconstruction in Mammography', Thesis for Master of Science

[8] YAM M., BRADY M., HIGHNAM R., BEHRENBRUCH CH., ENGLISH R., KITA Y. (2001): 'Three-Dimensional Reconstruction of Microcalcification Clusters from Two Mammographic Views', *IEEE Trans. on Image Processing*, Vol. 20, No. 6

[9] FERRARI R.J., RANGAYYAN R. M., DESAUTELS J. E. L., BORGES R. A., FRRE A. F.(2004): 'Automatic Identification of the Pectoral Muscle in Mammograms', *IEEE Trans. on Image Processing*, Vol. 23, No 2, pp. 232-245

[10] SZE MAN KWOK, CHANDRASEKHAR R., ATTIKIOUZEL Y., MARY T. R. (2004): 'Automatic Pectoral Muscle Segmentation on Mediolateral Oblique View Mammograms', *IEEE Trans. on Medical Imaging*, Vol. 23, No. 9

[11] WEI-YING MA, MANJUNATH B. S. (2000): 'EdgeFlow: A Technique for Boundary Detection and Image Segmentation', *IEEE Trans. on Image Processing*, Vol. 9, No 8, pp. 1375-1388

[12] PITAS I. (2000): 'Digital Image Processing and Algorithms and Applications', (John Wiley & Sons), New York

[13] GUSTAFSON D. E., KESSEL W. C. (1979): 'Fuzzy Clustering with a Fuzzy Covariance Matrix', Proc. IEEE-CDC, Vol. 2, pp. 761-766

[14] DUDA R.O., HART P.E. (1973): 'Pattern Classification and Scene Analysis', (John Wiley & Sons)

[15] OLSÉN C., GEORGSSON F. (2005) 'Problem Related to Automatic Nipple Extraction', 14th Scandinavian Conference on Image Analysis, Finland

ISSN: 1727-1983