Mammographic CAD: Correlation of regions in ipsilateral views – a pilot study

Jayanethie Padayachee, *PhD*
Michael J Alport, *PhD*
Department of Physics, University of KwaZulu-Natal, Durban

William I D Rae, *MB BCh, PhD*
Department of Medical Physics, University of the Free State, Bloemfontein

Abstract

Background. Radiologists analyse both standard mammographic views of a breast to confirm the presence of abnormalities and reduce false-positives. However, at present, no computer-aided diagnosis system uses ipsilateral mammograms to confirm the presence of suspicious features.

Aim. The aim of this study was to develop image-processing algorithms that can be used to match a suspicious feature from one mammographic view to the same feature in another mammographic view of the same breast. This algorithm can be incorporated into a computer-aided diagnosis package to confirm the presence of suspicious features.

Method. The algorithms were applied to 68 matched pairs of craniocaudal and mediolateral-oblique mammograms. The results of this pilot study take the form of maps of similarity. A novel method of evaluating the similarity maps is presented, using the area under the receiver operating characteristic curve (AUC) and the contrast (C) between the area of the matched region and the background of the similarity map.

Results and conclusions. The first matching algorithm (using texture measures extracted from a grey-level co-occurrence matrix (GLCM) and a Euclidean distance similarity metric) achieved an average AUC of 0.80±0.17 with an average C of 0.46±0.26. The second algorithm (using GLCMs and a mutual information similarity metric) achieved an average AUC of 0.77±0.25 with an average C of 0.50±0.42. The latter algorithm also performed remarkably well with the matching of malignant masses and achieved an average AUC of 0.96±0.05 with an average C of 0.90±0.21. In conclusion, texture analysis methods used with suitable similarity metrics allow a suspicious feature from one mammographic view to be matched with the same suspicious feature in other mammographic views of the same breast.

Introduction

According to the Cancer Association of South Africa, breast cancer is currently the most common cancer among women worldwide, and is second to cervical cancer among South African women. While sometimes fatal, breast cancer can be successfully treated, provided that it is detected early. The most common method of detecting breast cancer in its early stages is mammography. Unfortunately, mammography suffers from the problem that radiologists, in their visual interpretation of the resulting mammograms, can sometimes miss the subtle signs of breast cancer.[1]

Computer-aided diagnosis (CAD) was developed to consistently draw the radiologist's attention to suspicious regions in a mammogram that could be missed. Commercial CAD systems are designed to be consulted **after** the radiologist has made an initial assessment of the mammogram, and it has been shown that prompting by a CAD system improves radiologists' detection performance.[2]

The basic algorithm in a CAD system is (*i*) the detection of suspicious features, (*ii*) the reduction of false-positives, and (*iii*) the classification of suspicious features as malignant or benign. While current CAD methods can achieve sensitivities of up to 100% in identifying microcalcification clusters, masses are detected with a lower sensitivity because of their variable appearance and similarity to normal tissue. Mass detection CAD algorithms also have high false-positive rates, which are not very practical because the radiologist would waste time examining the false-positive marks.

The minimum requirements of a CAD system should be to completely emulate the actions of a radiologist, who uses many methods to analyse a set of mammograms. These methods are summarised in Fig. 1. Most methods generally used by a radiologist in
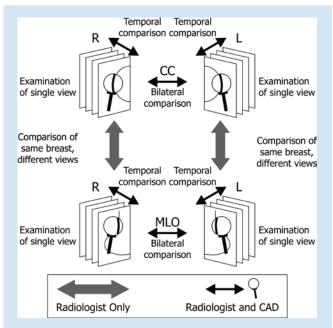

Fig. 1. Methods used by a radiologist to analyse a set of mammograms. Most methods have been implemented in a CAD system, but the examination of ipsilateral views has not.
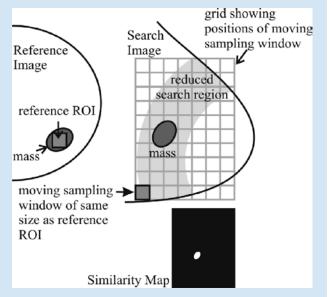
Fig. 2. Schematic (not to scale) of the matching algorithm. The location of the reference ROI was used to reduce the search region in the search image. Textural characteristics of the reference ROI are compared with textural characteristics of equivalently sized sub-images in the search image. The comparison process results in a similarity map. The brighter the regions on the similarity map, the greater the similarity. The similarity map is generally smaller than the search image because the sampling windows are >1 pixel and the windows are stepped in increments >1 pixel.
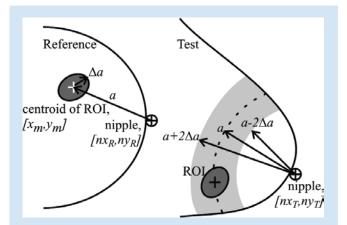

Fig. 3. Geometry of the arc method used to reduce the search region in the search image with the CC view as the reference view and the MLO view as the search view. The position of the nipple $[nx_R, ny_R]$ and the position of the centroid of the ROI $[x_m, y_m]$ in the reference view were used to define the arc distance a. The maximum extent Δa of the ROI border from the centroid in the reference view was also determined. Then the nipple in the search view was used as an origin to draw two arcs (of radii $a \pm 2\Delta a$) that were bounded by the breast border. The region enclosed between the arcs and the breast border defined the reduced search region in the search view. The method is independent of which view was used as a reference. The positions of the nipple in both standard mammographic views and the centroid of the selected ROI in one view were used to extract that portion of the breast in the other view where the ROI could possibly lie. The value of Δa was based on the size of the ROI in the reference view, which meant that the area of the annular region depended on the size of the selected ROI.

interpreting mammograms (e.g. examination of single views, bilateral comparison, temporal comparison) have been implemented in CAD systems, mostly for the detection of abnormalities. In practice,

radiologists also consider the distance from the nipple to the centroid of a suspicious feature in one mammographic view, and then search an annular region in another mammographic view at about the same radial distance from the nipple corresponding to the suspicious feature. This technique is known as the arc method and is used under certain circumstances to confirm a true-positive feature or to eliminate a false-positive feature.

While there have been a few studies indicating the usefulness of using two standard mammographic views for false-positive reduction,[3-6] these algorithms have not been incorporated into CAD systems. Most importantly in these dual-view algorithms, the suspicious features are identified in both standard mammographic views and information is correlated between pairs of suspicious features to identify matches and thereby reduce the number of false-positives. All these methods also rely on some form of training (e.g. linear discriminant analysis, artificial neural networks) and therefore only perform as well as the data set that was used for the training. Any algorithms based on training generally perform very poorly when applied to situations outside the scope of the training data.

This paper presents an algorithm that finds a suspicious feature in one standard mammographic view and then uses the position and characteristics of the feature to find it in another standard mammographic view. Two image-processing matching algorithms were developed that perform an exhaustive search of a reduced breast tissue region to match a suspicious feature identified in one mammographic view with the same feature in another mammographic view of the same breast. The algorithms developed have the advantage of not requiring any training and can be slotted into existing CAD systems as a method of providing further information to reduce false-positives.

Materials and methods

The general matching algorithm was based on three assumptions:[4] (*i*) at least two mammographic views of the breast are available, (*ii*) a mass is visible in at least two mammographic views, and (*iii*) a mass has similar image textural characteristics in all mammographic views.

A schematic of the image-processing matching algorithm is shown in Fig. 2. The steps in the matching algorithm were:

1. **Identification:** A radiologist identified a region of interest (ROI) in the reference image, as there was no access to a CAD system. The radiologist manually drew borders around suspicious features in both standard mammographic views.

2. **Pre-processing:** The mammogram background[7] and pectoral muscle[8] were removed to reduce the area of the search image that was analysed. The arc method was used to further reduce the area analysed and is detailed in Fig. 3. Variations on the arc method have been used by Paquerault *et al.*[4], Zheng *et al.*[5] and van Engeland *et al.*[6] to reduce the search region in their CAD algorithms.

3. **Quantification of image texture:** Textural characteristics of the ROI were quantified using Haralick's texture measures[9] and grey-level co-occurrence matrices (GLCMs). Texture measures extracted from GLCMs have been applied to the texture analysis of mammograms on numerous occasions.[3,10]

4. **Similarity:** The similarity between the ROI and the reduced search image was quantified with a Euclidean distance metric and mutual information. This comparison process resulted in a similarity map where maxima corresponded to positions of greatest similarity between the ROI and the search image.

GLCMs and texture measures extracted from the GLCMs form the basis of the textural analysis in this study. GLCMs have the advantage of including information about the distributions of the relative locations of pixels and their grey levels.[9] The GLCMs were computed at four angles ($\theta = [0°, 45°, 90°, 135°]$) and then averaged to remove any directional effects that may be introduced by the change in orientation of the breast tissue between mammographic views. Texture measures were calculated from the averaged GLCM.[9]

Similarity metrics were used to quantify how similar the reference ROI was to the search ROI. The Euclidean distance and mutual information similarity metrics were used in this study.

- **Euclidean distance metric,** D_E: The Euclidean distance metric is the most commonly used metric to calculate distance. For two points and $y = [y_0, y_1, \ldots, y_{n-1}]$ in n-dimensions, D_E is defined as:

$$D_E = \sqrt{\sum_{i=0}^{n-1}(x_i - y_i)^2} \qquad (1)$$

- **Mutual information,** *MI*: Mutual information has been shown to be a robust similarity metric in image registration problems,[11] but has also been applied to template matching,[12] feature selection and segmentation problems. Mutual information can be interpreted as a measure of the information that two quantities have in common and is defined as:[13]

$$MI(X \cdot Y) = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} p(x_i \cdot y_j) \log \frac{p(x_i \cdot y_j)}{p(x_i)p(y_j)} \qquad (2)$$

Mutual information is an acceptable similarity metric because *MI>0* unless the two quantities are completely independent, then *MI=0*. Also, *MI* increases as the dependency between both quantities increases and is independent of the actual value of the probability.[13]

In the first matching algorithm, referred to as **texture measure matching** or TM-matching, the image texture was quantified by 13 texture measures calculated from averaged GLCMs: maximum probability, entropy, energy, inertia, inverse difference moment, correlation, sum average, sum entropy, difference entropy, sum variance, difference average, difference variance, information measure of correlation. A 13-dimensional texture measure vector was calculated for each position of the sampling window and was then compared with the texture measures of the reference ROI using Euclidean distance as a measure of similarity. The result was a 2-dimensional map of distance values.

In the second matching algorithm, referred to as **mutual information matching** or MI-matching, GLCMs were used to quantify the image texture. Mutual information was used to quantify the

similarity between the reference and search GLCMs. This study uses the full GLCM as an estimate of the probability density function that incorporates spatial information. The only application of the full GLCM that used mutual information was for image registration,[14] but the full GLCM has not been applied to template-matching problems or to any problem in mammographic CAD. For the calculation of mutual information, individual GLCMs of the reference and search images as well as a joint GLCM between the reference and search images were calculated.

The results of the TM- and MI-matching algorithms are maps of similarity, defined to have the optimal match at maximum similarity map values. The accuracy of each matching algorithm was evaluated by comparing the similarity maps to ground-truth maps. Ground-truth maps were generated from the ROIs manually marked by a radiologist. Matching accuracy is defined as a combination of two quantities: the area under the receiver operating characteristic curve and contrast between the matched region and its background.

Receiver operating characteristic (ROC) analysis is a standard method of evaluating and ranking medical diagnostic tests.[15] To perform any evaluation, the 'truth' must be known so that it can be compared with the output of the test. The evaluation and ranking of CAD algorithms is analogous to that of a standard medical diagnostic test and is therefore perfectly suited to the use of ROC analysis. The basis of ROC analysis is the ROC curve, which is a plot of the true-positive fraction (TPF) v. the false-positive fraction (FPF).

For the evaluation of the results from the matching algorithms, the similarity maps were thresholded at grey-level values between 0 and 255 (8 bits of information). The thresholded maps (Fig 4g - i) were compared with the ground truth data (Fig. 4c) to compute values of the TPF and the FPF at each threshold. The TPF and FPF values were used to generate the ROC curve. The area under the ROC curve, AUC, (computed using the trapezoidal rule) is used as an indication of what proportion of the matched region was actually matched and $0 \leq AUC \leq 1$.

Contrast is used as a measure of how well the matched ROI stands out from its surroundings in the similarity map. Contrast refers to a local change in brightness and is defined as the ratio of the average brightness of an object to the average brightness of the background.[16]

$$c = \frac{f - b}{f + b} \qquad (3)$$

f is the average grey level of the foreground, *b* is the average grey level of the background and $-1 \leq C \leq 1$. Negative values arise when the foreground is darker than the background, and positive values arise when the foreground is brighter than the background. A contrast of 0 means that the object cannot be distinguished from its background. The ground-truth data were used to define regions in the similarity map used to calculate *f* and *b*. Ideally, the matched regions should be the brightest objects in the similarity map.

By definition, the best match should have the highest matching accuracy. The selection of the best combination of AUC and C values is facilitated by the novel use of a **combined AUC-C value** referred to as **matching accuracy**, κ, and calculated as follows:

$$\kappa = \begin{cases} -2(AUC-0.5)C & \text{if } AUC \leq 0.5 \text{ and } C \leq 0 \\ 2(AUC-0.5)C & \text{otherwise} \end{cases} \quad (4)$$

$-1 \leq \kappa \leq 1$ because $0 \leq AUC \leq 1$ and $-1 \leq C \leq 1$.

Acceptable matches should have AUC>0.5 as this indicates that the matching is better than random, and C should be positive as this indicates that the matched region is brighter than the background in the similarity map.

A PC with an AMD Athlon XP 2.4GHz processor and 512 Mb of RAM, running Microsoft Windows 2000, was used for the software development. All algorithms were implemented in IDL 6.1, a programming environment providing mathematical functionality with a graphic interface.

The algorithms were applied to 34 pairs of cranial-caudal (CC) and medio-lateral oblique (MLO) mammograms. The mammograms were arbitrarily selected from the patient archives at the Inkosi Albert Luthuli Central Hospital, Durban, to represent a range of breast densities, mass sizes and patient ages. The images were acquired on a Siemens Mammomat 3000 Nova mammography unit, with a focal spot size of 0.3 mm, a molybdenum anode and a 30 μm molybdenum filter. The image reader was a Digiscan M (Fuji Photo Film Co. Ltd). The computed radiography images were exported in the digital imaging and communications in medicine (DICOM) format from the hospital data archives, at a bit-depth of 10 bits and 0.05 mm per pixel. For processing, images were resampled to 0.254 mm per pixel (100 dpi).

Since the matching results are independent of which view is used as a reference, each CC and MLO view was used separately as a reference image and a search image, resulting in the matching algorithms being applied to 68 pairs of mammograms. The 68 individual mammograms were divided into four categories based on the pathology of the suspicious ROI or overall diagnosis of the mammogram: 28 benign, 18 malignant, 10 normal and 12 indeterminate. Most patients are referred to the Inkosi Albert Luthuli Central Hospital for diagnostic tests, so not all the mammograms had a full pathological history, since not all referring physicians recommended a biopsy. In these cases, the radiologist's report was used as a basis for the diagnosis of the mammogram.

The 'benign', 'indeterminate' and 'malignant' diagnoses refer to masses, while the 'normal' diagnosis refers to a suspicious-looking region in a normal mammogram. Masses were categorised as 'indeterminate' if the biopsy was inconclusive or the radiologist was unable to render a diagnosis based on the mammographic appearance. A radiologist marked the borders of the suspicious ROIs in MagicView, the software interface used to view standard digital imaging and communications in medicine (DICOM) medical images. The borders were saved as DICOM images and were automatically extracted in IDL, for use as ground truth data, which eliminated the need to register the ground truth data with the mammograms. The areas of the regions enclosed by the radiologist-drawn borders were automatically computed in MagicView.

The average area of the ROIs together with the average visibility of the ROIs is shown in Table I. Visibility was automatically determined from the original mammograms (at 0.254 mm per pixel)

Table I. Average area and visibility of ROIs for the 68 mammograms used			
Type	Number	Average area (cm^2)	Average visibility
Benign	28	1.54±2.50	0.18±0.08
Indeterminate	12	2.23±2.05	0.23±0.01
Malignant	18	11.0±12.8	0.40±0.19
Normal	10	1.46±2.26	0.14±0.04

before pre-processing. Visibility was defined to be the contrast of the ROI compared with the surrounding tissue and was computed from Eq. 3. Visibility ranges between 0 for a very subtle ROI and 1 for a very obvious ROI. There is a wide range of ROI sizes, and some of these are very subtle while others are more visible.

Results

Examples of similarity maps for TM-matching and MI-matching are shown in Figs 4e and 4f respectively. While both maps show the ROI as the brightest feature, both maps also have other features that have been matched. These false-positive detections lower the AUC value and C value, and therefore lower the overall matching accuracy. For example, the TM similarity map (Fig. 4e) has AUC=0.94, C=0.44 and κ=0.39, while the MI similarity map (Fig. 4f) has AUC=0.92, C=0.36 and κ=0.30.

The map for TM-matching has a bright band around the interface between the breast and the segmented background. However, the bright region does not lower the overall accuracy of the match, because the evaluation focuses on the feature of interest and its immediate surroundings. The MI map does not display any artefacts at the interface between the breast and the segmented background.

The averages of the best matching results for TM-matching are listed in Table II. Results show that the malignant masses were matched with the highest matching accuracy compared with the other mammogram classifications. Results for the benign and indeterminate masses and the normal mammograms were scattered across a range of AUC and C values, and there were 5 ROIs (2 benign, 1 indeterminate, 2 normal) that were not matched with TM-matching.

The average of the best matching accuracies for MI-matching are listed in Table III. The malignant masses were matched with the highest accuracy. The matching accuracy for the malignant masses was statistically different ($p<0.002$) from the matching accuracies for the other mammogram categories. All malignant masses were well matched, with the results generally clustered around AUC=1 and C=1.

Mammograms classified as non-malignant were poorly matched in both methods. Possible factors that contributed to the poor matching results were the ROI area and the ROI visibility. Most of the ROIs on mammograms classified as benign, indeterminate and normal were either very small (<0.5 cm^2) or had a low visibility, while the ROIs on mammograms classified as malignant were generally larger and more visible. Matching accuracy was generally well

Table II. Average of the best matching results for TM-matching

Diagnosis	Average AUC	Average C	Average κ
Benign	0.75±0.17	0.37±0.21	0.22±0.16
Indeterminate	0.80±0.15	0.50±0.26	0.34±0.27
Malignant	0.93±0.05	0.67±0.21	0.58±0.21
Normal	0.68±0.19	0.29±0.26	0.18±0.19
All	0.80±0.17	0.46±0.26	0.33±0.25

Table III. Average of the best matching results for MI-matching

Diagnosis	Average AUC	Average C	Average κ
Benign	0.68±0.26	0.29±0.36	0.21±0.30
Indeterminate	0.75±0.28	0.59±0.40	0.44±0.34
Malignant	0.96±0.05	0.90±0.21	0.84±0.23
Normal	0.71±0.24	0.26±0.29	0.17±0.23
All	0.77±0.25	0.50±0.42	0.41±0.39

correlated with ROI visibility (71% correlation), and the ROIs with low visibilities had very poor matching accuracies.

Matching accuracy is spread over a wide range for the small, low-visibility ROIs, while the large, high-visibility ROIs generally have high matching accuracies. Regarding TM-matching, matching accuracy was correlated with ROI visibility (62% correlation). There were 12 image pairs (8 benign, 2 indeterminate, 2 normal) that were not matched with MI-matching. Areas of the non-matched ROIs ranged from 0.16 cm^2 to 1.21 cm^2 and all had low visibilities.

Discussion

Fig. 5 shows the best matching results for each of the 68 pairs of mammograms, for each matching algorithm. The AUC and C values are generally quite scattered for both. Some results have AUC<0.5 and C<0, indicating that the match was unsuccessful. Matching accuracies are, however, generally quite high for both methods; this is confirmed by examining the average of the best matching accuracies (Tables II and III).

The results of performing a paired t-test analysis on the distribution of the best κ values for TM-matching and MI-matching yielded a t value of -1.40 and a p value of 0.16. For a significance level of 0.05, the average values of κ for TM-matching compared with MI-matching are not statistically different. Overall, both TM-matching and MI-matching show potential as matching schemes, as both methods yielded quite accurate matches on the small data set.

Results at a significance level of 0.05 show that the average results for matching the benign and indeterminate masses and normal ROIs are similar for each method (p=0.78, p=0.46, p=0.97, respectively), but that the results of matching malignant masses are
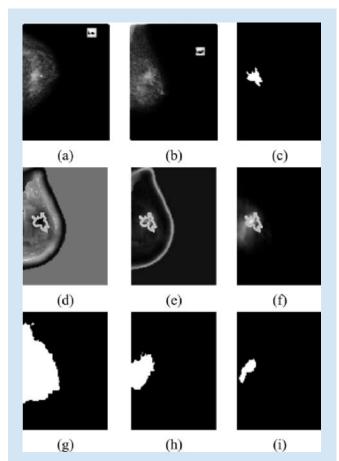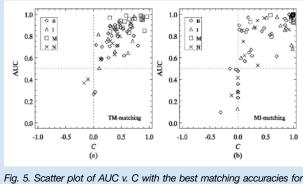

Fig. 4. Sample images: (a) reference image, (b) search image, (c) ground-truth map, (d) distance map, (e) TM similarity map, (f) MI similarity map (g - i) MI similarity map thresholded at grey levels of 1, 64 and 128. The thresholded maps are compared with the ground-truth map in (c) to determine the TPF- and FPF-values that were used to generate an ROC curve.

Fig. 5. Scatter plot of AUC v. C with the best matching accuracies for each mammogram pair for (a) TM-matching and (b) MI-matching.

statistically different (p<0.01), with the MI-matching results better than the TM-matching results.

MI-matching required less computational time and was more accurate at matching malignant masses, but there were fewer unmatched pairs of mammograms for TM-matching (5 out of 68 compared with 12 out of 68 for MI-matching). Therefore, a hybrid-

matching scheme using the results of both methods could yield better matching results.

Both matching algorithms showed great potential for use in a CAD scheme. The average best matching accuracy for MI-matching was κ=0.41±0.39, which corresponds with average best AUC and C values of 0.77±0.25 and 0.50±0.42, respectively. The average best matching accuracy for TM-matching was κ=0.33±0.25, which corresponds with average best AUC and C values of 0.80±0.17 and 0.46±0.26, respectively. The average best results for these two methods were not statistically different (p>0.05)

MI-matching showed the best matching accuracy for matching malignant masses (κ=0.84±0.23 corresponding to AUC=0.96±0.05 and C=0.90±0.21), while the results for the other types of ROI (benign, indeterminate, normal) were similar for both methods.

The TM- and MI-matching algorithms show potential for providing more information for use in a false-positive reduction scheme in a CAD system. The ideal solution would be to incorporate mutual information ideas into the texture measure method. If the suspicious object is present in both mammographic views, only one view needs to be analysed to detect the object, while the second view is analysed with information extracted from the object in the first view, for confirmation of a true object.

One advantage of using a distance similarity metric and mutual information for matching is that no training is required, which is quite important for a mammographic CAD system since breast tissue varies considerably from patient to patient. The TM- and MI-matching algorithms can be applied to any image-matching problem. Unfortunately, the current algorithms are very time-consuming and will have to be optimised for implementation in a CAD system.

Two shortcomings of this study are the quality of the ground truth data and the small data set. Only one radiologist marked the borders of the ROIs in each mammogram, and there was no method of confirming the accuracy of the identified borders. Also, results will be strengthened if the algorithms are tested on a larger database of mammograms.

Conclusion

Texture analysis methods used with suitable similarity metrics allow a suspicious feature from one mammographic view to be matched with the same suspicious feature in other mammographic views of the same breast. The matching algorithms (using grey-level co-occurrence matrices, distance similarity metrics and mutual information) perform especially well in matching malignant masses. This dual-view analysis method can most probably be used to provide complementary information to a false-positive reduction scheme in a mammographic CAD system.

1. Huynh P, Jarolimek A, Daye S. The false-negative mammogram. *Radiographics* 1998; 18: 1137-1154.
2. Freer T, Ulissey M. Screening mammography with computer-aided detection: prospective study of 12 860 patients in a community breast center. *Radiol* 2001; 220: 781-786.
3. Chang Y, Good W, Sumkin J, Do B, Gur D. Computerized localization of breast lesions from two views: an experimental comparison of two methods. *Invest Radiol* 1999; 34: 585-588.
4. Paquerault S, Petrick N, Chan H, Sahiner B, Helvie M. Improvement of computerized mass detection on mammograms: Fusion of two-view information. *Med Phys* 2002; 29: 238-247.
5. Zheng B, Leader J, Abrams G, *et al.* Multiview-based computer-aided detection scheme for breast masses. *Med Phys* 2006; 9: 3135-3143.
6. van Engeland S, Timp S, Karssemeijer N. Finding corresponding regions of interest in mediolateral oblique and craniocaudal mammographic views. *Med Phys* 2006; 9: 3202-3212.
7. Padayachee J, Rae W, Alport M. Identification of the breast edge using areas enclosed by iso-intensity contours. *Comput Med Imaging Graph* 2007; 31: 390-400.
8. Karssemeijer N. Automated classification of parenchymal patterns in mammograms. *Phys Med Biol* 1998; 43: 365-378.
9. Haralick R, Shanmugam K, Dinstein I. Texture features for image classification. *IEEE T Syst Man Cyb* 1973; 3: 610-621.
10. Sahiner B, Chan H, Petrick N, Helvie M, Hadjiiski L. Improvement of mammographic mass characterization using spiculation measures and morphological features. *Med Phys* 2001; 28: 1455-1465.
11. Pluim JW, Maintz JA, Viergever M. Mutual-information-based registration of medical images: A survey. *IEEE Trans Med Imaging* 2003; 22: 986-1004.
12. Tourassi G, Vargas-Voracek R, Catarious D, Floyd CE. Computer-assisted detection of mammographic masses: a template matching scheme based on mutual information. *Med Phys* 2003; 30: 2123-2130.
13. Abramson N. *Information Theory and Coding.* New York: McGraw-Hill, 1963.
14. Hseu H, Bhalerao A, Wilson R. *Image Matching Based on the Co-occurrence Matrix. Tech. Rep. CS-RR-358.* Department of Computer Science, University of Warwick, Coventry, UK, 1999. http://citeseer.ist.psu.edu/hseu99image.html (accessed 2003).
15. van Erkel A, Pattynama PT. Receiver operating characteristic (ROC) analysis: Basic principles and applications in radiology. *Eur J Radiol* 1998; 27: 88-94.
16. Sonka M, Hlavac V, Boyle R. *Image Processing, Analysis and Machine Vision.* Florence, Kentucky, USA: Thomson-Engineering, 1999.