

Establishing correspondence in mammograms and tomosynthesis projections

Predrag R. Bakic^{*,1}, Nenad Vujovic^{1,2}, Andrew D.A. Maidment¹

¹Department of Radiology, University of Pennsylvania, 3400 Spruce St., Philadelphia, PA 19104

²Opex Corporation, 305 Commerce Dr., Moorestown, NJ 08057

ABSTRACT

We are developing a computer based aid for automated analysis of digital mammograms and digital breast tomosynthesis (DBT) images. The ultimate goal is to establish correspondence between regions in images obtained using different modalities. This paper is focused on establishing point correspondences between mammograms and DBT projections. Correspondence has been established utilizing two similarity criteria, one based on topology of prominent elongated anatomical structures, and another based on texture near the potential correspondence points. We evaluated robustness of the described technique with respect to variations in x-ray tube angle. The evaluation included 72 image pairs: 9 DBT projections from both left and right breasts of 4 women and corresponding mammograms. The evaluation was performed by a consensus between two trained observers. Two images with highlighted pairs of automatically established point correspondences were presented. The observers were asked to manually identify the correct correspondences and measure the displacement error. Points for which correspondences could not be manually identified were excluded. The topology method automatically generated an average of 12.2 correspondences per image pair, for which the average measured displacement error was 1.33 mm (N = 10.5). The texture method generated 18.6 correspondences with an average measured error of 1.80 mm (N = 14.7). The algorithms were found to be robust; the number of correspondences and the average displacement did not significantly change with variations in tube angle.

Keywords: Methods: Pre-processing, Feature extraction; Modalities: Mammography, Digital breast tomosynthesis; Diagnostic Task: Detection..

1. INTRODUCTION

Both clinical and automated analyses of breast images involving temporal image sequences and/or different modalities are complex tasks due to the multitude of factors that can alter image appearance; these factors include breast positioning and compression, acquisition parameters, measured tissue properties, and breast composition. Digital breast tomosynthesis (DBT) is a novel radiographic method in which 3D images of the breast are reconstructed from a small number of projections (source images), acquired at various x-ray focus positions, with total dose comparable to that of conventional mammography. The 3D nature of DBT allows for more accurate visualization of breast anatomy and pathological changes, thus making DBT a strong candidate to replace mammography as the screening modality of choice.^{1,2}

It is also possible to perform automated analysis of temporal sequences of mammograms. Sallam and Bowyer^{3,4} have proposed a technique based on unwarping for compensating differences in breast appearances, followed by image subtraction. Timp *et al.*^{5,6} developed a regional registration, which consists of a global alignment, identification of the lesion search area, and selection of the best location using a registration measure. Vujovic and Brzakovic⁷ have proposed the use of control points identified at intersections of elongated tissue structures, for establishing correspondence between temporal image pairs. Richard *et al.*⁸ utilized synthetic images generated by a computer model of breast tissue to evaluate methods for mammogram registration.

With the recent introduction of DBT in clinical trials, attempts have been made to detect breast cancer using computer-aided analysis of both source projection images and reconstructed 3D images.⁹⁻¹¹ We seek an automated method to compare temporal sequences of DBT images. If DBT replaces mammography for routine breast cancer screening, it will

* Predrag.Bakic@uphs.upenn.edu; Phone: 215 746 8758; Fax: 215 746 8764

be necessary to compare large number of mammograms acquired in previous years with current DBT images, during the transitional period. Thus, there is a need to develop a method for establishing correspondence between mammograms and DBT images. Here we present such a method, applied to a set of mammograms and the DBT projections of the same breast taken the same day.

2. METHODOLOGY AND MATERIALS

To date we have acquired DBT images of 52 women as a part of a clinical evaluation of multimodality breast imaging at the University of Pennsylvania.¹² Each woman underwent a mammography and DBT exam on the same day by the same technician. The DBT projections were acquired on a GE Senographe 2000D (General Electric Medical Systems, Milwaukee, WI), modified to allow the x-ray tube to be positioned at 9 locations, each 6.25 degrees apart. The breast was lightly compressed in an MLO position for the duration of the exam; the amount of compression used in DBT is significantly lower (5-7 dN) compared with the amount of compression typically used in conventional mammography (11-20 dN). The source images were acquired with a spatial resolution of 100 $\mu\text{m}/\text{pixel}$. The images were down-sampled to 200 $\mu\text{m}/\text{pixel}$ before analysis to reduce the effects of noise.

Our study establishes correspondences between two sets of potential correspondence points in a mammogram-DBT projection image pair. Potential correspondence points (PCPs) are defined as intersections of prominent anatomical structures. We presented the methodology employed for the PCP detection in more detail previously.⁷ Once detected, PCPs are compared and correspondences are established according to similarity of signatures or texture in the respective neighborhoods. Signatures are topological descriptors of prominent elongated structures converging to a PCP.⁷ Texture descriptors are computed using Laws convolution matrices.¹³ Automatically computed correspondence points were visually evaluated by two trained observers. The observers were asked to manually determine corresponding points in the image pair. A displacement error was calculated between the manually and automatically determined points.

2.1 Summary of the algorithm for establishing point correspondences

Establishing point correspondences is a multi-step process. In the first step PCPs are determined independently in two images being compared. A PCP is defined as the intersection of prominent elongated anatomical structures such as ducts, blood vessels, and parenchymal structures. The algorithm is focused on prominent elongated structures to reduce sensitivity to noise, thus improving reliability and robustness of the method.

In the second step, a pair of PCPs is selected according to the distance from corresponding reference points in the two images, typically the tip of the nipple. Next, we measure topological and texture properties in the vicinity of the pair of PCPs. A similarity criterion based on these properties is introduced to test whether the pair of PCPs constitutes a possible match. The match established this way is called the first order match. All other PCPs are then tested in a similar fashion, using the first order match pair as a reference; each match established this way is called the second order match. Note that in this step the first order match pair is used as a reference instead of the nipple; because matching based on the similarity criterion has been found to be more reliable than approximate matching of the nipple based on the topology alone.⁷

Combined, the first and second order matching processes result in a large number of attempts. To keep track of all matches established during these attempts, we utilize an accumulator matrix. In the matrix, each row corresponds to a PCP in the first image and each column to a PCP in the second image. Every time a match between a pair of points is established the appropriate accumulator matrix entry is incremented. The accumulator matrix reinforces true matches. In the final step the accumulator matrix is analyzed to identify the most reliable matches; these matches constitute the final correspondences.

Extraction of PCPs, detection of the reference point, similarity criterion, and establishing final correspondences are detailed in Sections 2.2 through 2.5, respectively.

2.2 Extraction of PCPs

PCPs are found at the intersections of prominent elongated anatomical structures, which are extracted from blurred versions of the clinical images. Blurring is utilized to reduce high spatial frequency variations in the images and to focus the analysis only on the prominent structures.

Image blurring is achieved through Gaussian filtering.¹⁴ The filter is approximated by a 5×5 kernel and is applied in three passes. This process transforms prominent elongated structures into smooth ridges with flat peaks. Thus, the problem of extracting prominent elongated structures is reduced to a problem of detecting flat peaks in the blurred images.

We utilized a technique that extracts ridges in the horizontal and vertical directions independently and then combined the two results to obtain ridges in arbitrary directions. Detection of flat peaks is accomplished using a modified monotonicity (or monotony[†]) operator.¹⁵ Modified monotonicity operators consider two neighborhoods and the pixel intensity relationship between them. For instance, a direction sensitive modified monotonicity operator that extracts flat peaks from vertical ridges could be constructed as follows:

- (a) Small neighborhood $a = \{ (k,l) \mid k=1, -p \leq l \leq p \}$.
- (b) Large neighborhood $A = \{ (k,l) \mid k=1, -q \leq l \leq q \}$.
- (c) An Intensity Criterion is the number of pixels in the large neighborhood A , that have values that are larger than the largest pixel value in the small neighborhood a .

The pixels in the original image are replaced by values computed according to the Intensity Criterion (c). The resulting image is called the modified monotonicity image. The flat ridge peaks are extracted from the modified monotonicity image by applying an appropriately selected threshold T .

Selection of the values of p , q , and T is dictated by spatial resolution and acceptable levels of noise. For images utilized in this work we have used $p = 1$, $q = 5$, and $T = 8$. The selection of these values is guided by the fact that after Gaussian smoothing the prominent elongated structures are wider than $2q + 1 = 11$ pixels with flat peaks around four pixels wide.

Extraction of horizontal ridges is accomplished similarly using the modified monotonicity operator rotated 90 degrees. Vertical and horizontal ridge images are extracted independently. Then, a logical *OR* operation is applied to these images

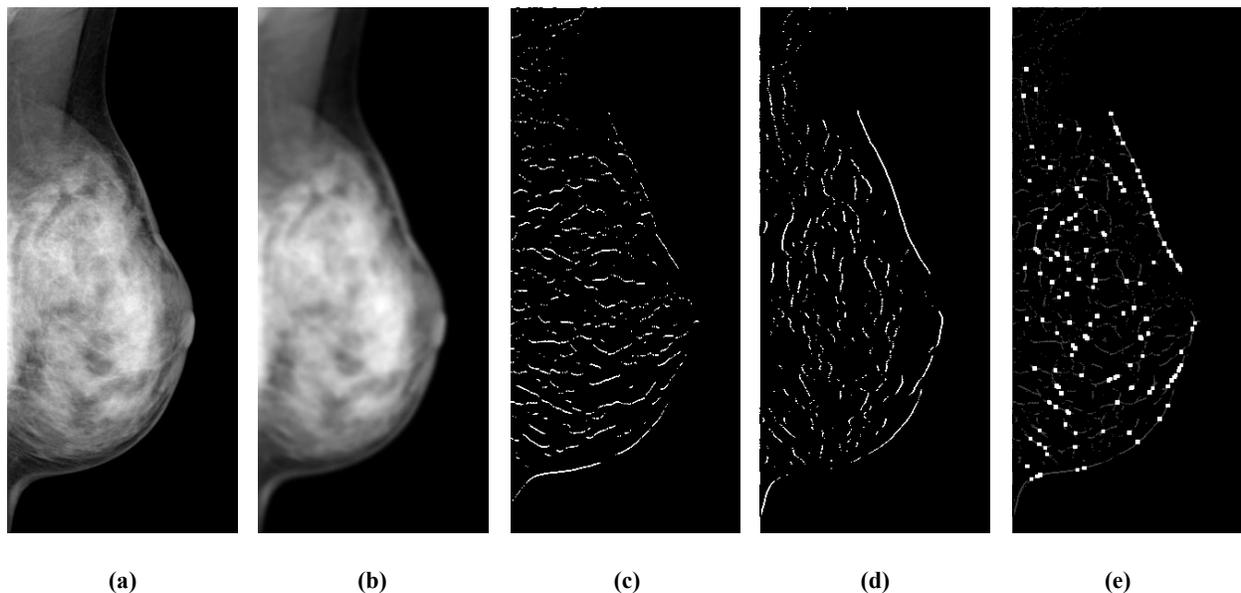


Figure 1: Extraction of potential correspondence points (PCPs). A Gaussian filter is applied to the original mammogram (a) to obtain a blurred version (b). Horizontal (c) and vertical (d) ridges are extracted from the blurred image using orthogonal modified monotonicity operators. PCPs are extracted as intersections of horizontal and vertical ridges (e). Correspondence between PCPs from a mammogram pair is established based on ridge topology or texture properties in the vicinity of the PCPs.

[†] In 1986 paper Kories and Zimmerman first introduced the term “monotony” operators; in the literature, these operators have been referred to both as “monotony” or “monotonicity” operators. In this paper, we use the term “monotonicity” operators.

to obtain ridges in arbitrary direction and a logical *AND* operation to obtain PCPs (as centroids of isolated blobs from *AND* images). Figure 1 summarizes the PCP extraction process.

2.3 Finding a reference point

A pair of reference points must be identified; one point in each of the two images being compared. The reference points are utilized to identify candidates for the first order matches. We selected the tip of the nipple to be the reference point because of its prominence in both mammographic and DBT projection images. A reference point could alternatively be selected at any other point in the image that could be consistently identified within the breast or on the breast surface (e.g., a marker). In this study we determined the reference point automatically using the least mean squares approximation of the breast outline with the quadratic function and then found the extremum point of the approximated outline. In a few cases, the reference point was determined manually when the nipple point was not visible in one of the images (typically on a large angle DBT projection due to partial beam obstruction by the collimator).

Once the reference pair of points had been established, a set of candidate PCPs for the first order matches is determined in both images by considering a 30×30 neighborhood from the reference point.

2.4 Similarity criterion

A pair of PCPs is matched if a similarity criterion is satisfied. Two similarity criteria were considered in this study. The first one is based on the topology of elongated structures while the other is based on texture characteristics near the PCPs.

- *Similarity criterion based on topology:* A pair of PCPs is considered a possible match if the prominent elongated structures that originate from the PCPs have similar arrangements (or topology). The arrangement is captured by signature vectors whose elements are equal to the number of non-zero pixels in a rectangle of size 20×13 pixels, originating in the PCP and rotating in steps of 10 degrees. To simplify the comparison of signature vectors we first determine the reference angle in the mammogram image as the angle of the largest element of the signature vector. Then the similarity criterion is approximated by considering only the first element of the signature vector in the mammogram and comparing it with the first, second, and last element in the DBT projection image to account for small structural differences in two images. Similarity comparison is made considering a specific threshold value that depends on resolution and level of detail to be detected. In this study, elongated structures were on average four pixels wide. Hence, the rotating 20×13 window oriented in the direction of the structure should ideally capture 20×4 = 80 structure pixels. Assuming that the elongated structure is not perfect, the threshold was selected to be 90% of the ideal value, i.e., 72 pixels. It should be pointed out that in our earlier study we experimented with various signature parameters including length and width of the rotating rectangle, threshold value, and the angle step; it has been shown that the signature formation and matching is robust to all of these parameters.⁷
- *Similarity criterion based on texture:* A pair of PCPs is considered a possible match if the texture characteristics in the respective neighborhoods are similar. We measured texture characteristics using Laws convolution *R5R5* matrix;¹⁶ the same matrix has been frequently used in mammogram texture analysis by other researchers.^{17, 18}

2.5 Establishing final correspondences

The accumulator matrix stores all possible matches between any two points in the mammogram and DBT projections, based on the similarity criterion. Since a single point in one image could be matched to multiple points in the other image, some possible matches are incorrect. The likelihood that the match is correct is related to the value of the entry of the accumulator matrix corresponding to the point pair; the higher the value, the more likely that the match is correct.

The following thresholding algorithm is utilized to identify final matches:

1. Determine the maximum value in each row and set all other values in the row to zero.
2. Determine the maximum value in each column and set all other values in the column to zero.
3. Threshold non-zero entries at a percent of the maximum value in the accumulator matrix to select the most robust matches (final correspondences).

Typical values for automatically established correspondences per image pair for images considered in this study are between 15 and 20, while the total number of PCPs in two images is around 300. If an insufficient number of matches

had been established the algorithm could be applied iteratively using successively more relaxed parameters, potentially at the expense of accuracy.

2.6 Evaluation of the proposed methodology

Evaluation was performed by a consensus between two observers with expertise in mammographic image processing. Two images with highlighted pairs of automatically established point correspondences were presented. The observers were asked to manually identify the correct correspondences and measure the displacement error between the automatically and manually found points. Points for which the correct correspondences could not be manually identified or if a consensus between the observers could not be reached were excluded.

We have analyzed the establishing correspondences between the mammogram and all nine DBT projection images of both breasts from four women. This analysis included computing the average number of correspondences established by the consensus of two observers and the average displacement error.

The displacement error is defined as:

$$e_D = \sqrt{(x_A - x_M)^2 + (y_A - y_M)^2}, \quad (1)$$

where (x_A, y_A) are the coordinates of an automatically established correspondence point, and (x_M, y_M) are the coordinates of manually identified position of the same correspondence point. We computed the displacement error on mammograms, using the automatically established correspondences in DBT projection images as reference. We also computed the ratio of the number of manually and automatically established correspondences and used this ratio as a surrogate for the reliability of the computed displacement error.

3. RESULTS

Figure 2 shows an example of established correspondences using the proposed methodology. The image pair consists of a mammogram and a DBT projection image (left and right image in the pair, respectively).

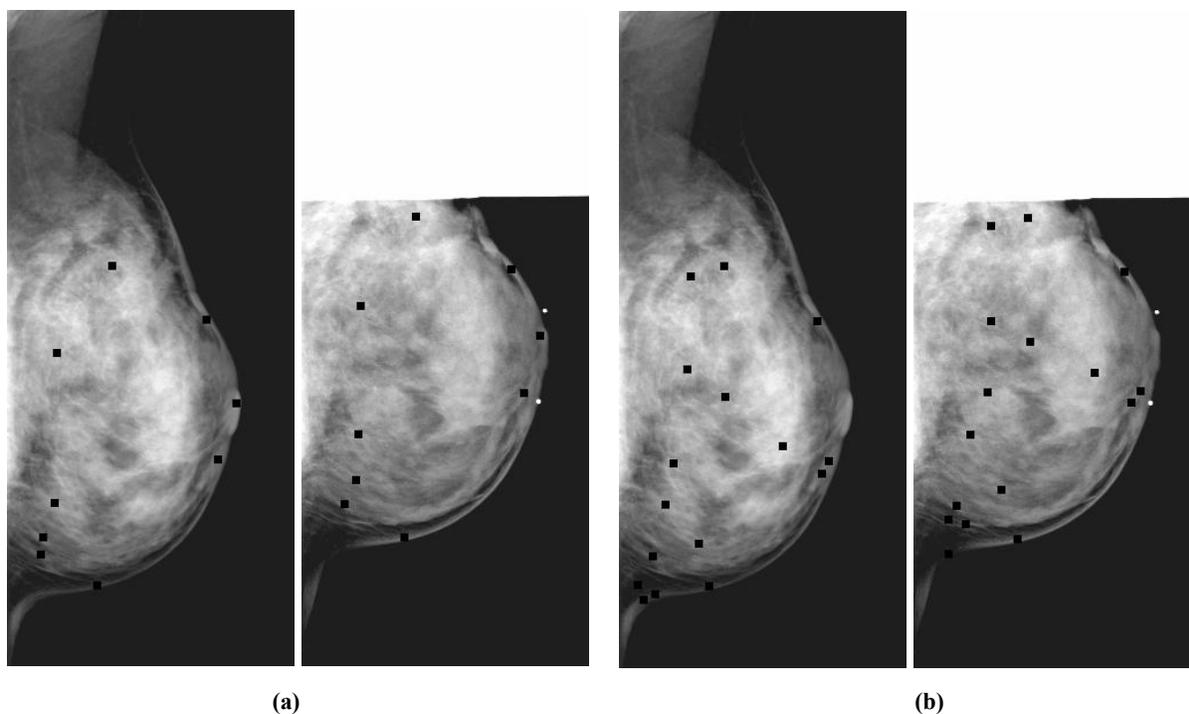


Figure 2: Correspondences established using signature (a) and texture (b) based methods between a mammogram (left image in the image pair) and a DBT projection (right image in the pair). The upper portion of the DBT projection images is occluded by the collimator.

The signature based method automatically generated an average of 12.2 correspondences per image pair, out of which 10.5 (range 2-22) correspondences were also manually identified by the observers. The texture based method generated 18.6 correspondences on average, out of which 14.7 (range 6-29) correspondences were also manually identified. Out of 72 image pairs considered, the algorithms yielded zero correspondences in only one instance.

We have tracked correspondences on the breast border and found that for both methods the number of correspondences was similar (6.8 vs. 7.3 for the signature and texture based methods, respectively). In the breast interior, the texture based method yielded on average twice as many correspondences as the signature based method (7.4 vs. 3.7, respectively). The number of correspondences between a mammogram and a DBT projection as a function of the DBT projection angle and results are shown in Figure 3.

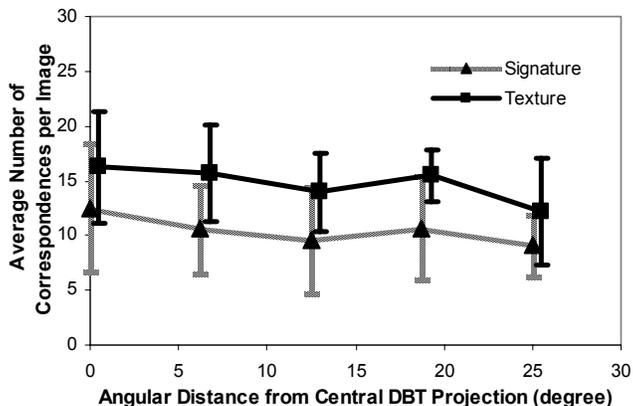


Figure 3: The mean and standard deviation (SD) of the number of correspondences as a function of the angular distance between a DBT projection and the mammogram. Error bars show \pm one SD.

The average displacement error for automatically established correspondences was 1.33 mm, for the signature based method, and 1.80 mm, for the texture based method. These displacement errors were averaged over manually identified correspondences. The displacement error for the correspondences in the breast interior were on average slightly higher than the error on the breast edge (2.2 mm vs. 1.1 mm, for the signature based method, and 2.1 mm vs. 1.5 mm, for the texture based method). Figure 4 shows the average displacement error as a function of the DBT projection angle. Figure 5 summarizes the surrogate reliability as a function of the DBT projection angle. We defined the surrogate

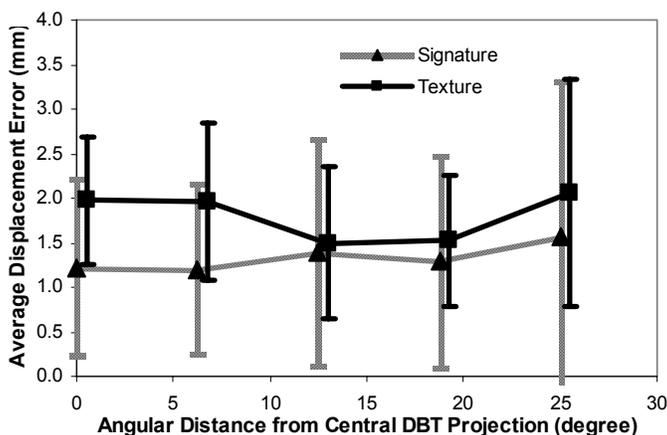


Figure 4: The mean and SD of the displacement error as a function of the angular distance between a DBT projection and the mammogram. The displacement error was computed as the difference between the position of an automatically established correspondence and its manually confirmed position. Error bars show \pm one SD.

reliability in Section 2.6 as the ratio of the number of manually and automatically established correspondences. The overall surrogate reliability for all images in this study was 0.82 (0.85 and 0.79 for the signature and texture based methods, respectively).

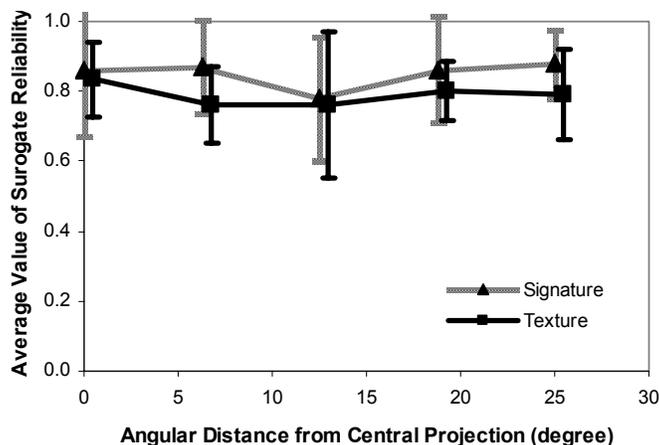


Figure 5: The mean and SD of the displacement error as a function of the angular distance between a DBT projection and the mammogram. The displacement error was computed as the difference between the position of an automatically established correspondence and its manually confirmed positions, Error bars show \pm one SD.

4. DISCUSSION

The proposed methods are successful in establishing correspondences between mammograms and DBT projections. The overall average displacement error was 1.6 mm. In our previous study of temporal series of mammograms, the average displacement error for the signature based method was 5.5 mm,⁷ The observed difference in the algorithm performance could be attributed to differences in breast positioning, spatial resolution, and technological improvements over the last decade.[‡] The texture based method has been presented for the first time in this study.

The signature-based method yields slightly higher accuracy but fewer correspondences compared with the texture based method (1.33 mm displacement error and 10.5 correspondences on average for the signature based method vs. 1.80 mm displacement error and 14.8 correspondences on average for the texture based method). We observed that both methods performed similarly for the correspondences located on the breast border. In the breast interior, however, the texture based method found more correspondences with similar accuracy when compared with the signature based method (2.2 mm displacement error and 3.7 correspondences on average for the signature method vs. 2.1 mm error and 7.4 correspondences for the texture method). We believe that the algorithms perform slightly better on the breast border, than in the breast interior, because the breast border is a prominent topological feature with strong similarity for both signature and texture characteristics.

Both methods yield satisfactory accuracy. We computed the average displacement error below 2 mm; we consider this result encouraging assuming that it is the goal to detect tumors smaller than 10 mm in diameter. Such tumors represent 25% of all tumors detected in the U.S. between 1995 and 1999,¹⁹ and are associated with 100% 5-year survival rate (as reported by the Surveillance, Epidemiology, and End Results (SEER) Program).²⁰ Although we expected a progressive increase in the error with the DBT acquisition angle, with the smallest error at the central projection, the accuracy of the algorithms was independent of the angle. The most likely reason for this result is that the effects of variation in breast positioning are more dominant than the effects of variation in the acquisition angle.

[‡] Our previous study was performed in 1997 using digitized films, while this study used digital mammograms and DBT images.

Note that the two methods do not exhibit significant correlation with respect to the number of established correspondences, as demonstrated in Figure 6; the corresponding Pearson coefficient of correlation is 0.42. Figure 7 shows the histogram of the number of established correspondences for the two methods. The shape of histograms is unimodal, suggesting that the methods do not favor any particular type of images from the 72 analyzed pairs. In addition, from our experience with the visual evaluation, the established correspondence points are relatively uniformly distributed throughout the breast region, indicating no spatial bias. This is a desirable property for achieving high accuracy in defining positions anywhere within the breast.

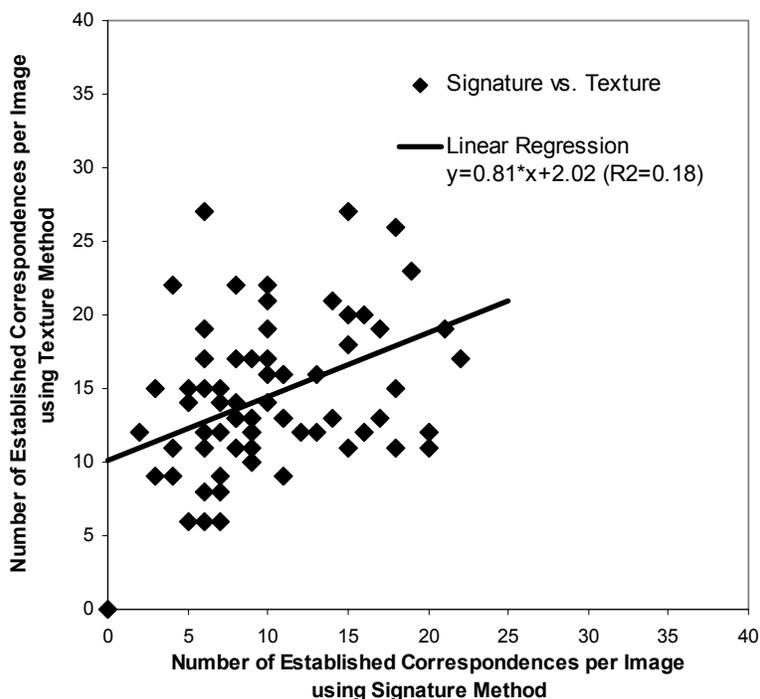


Figure 6: Scatter plot of the number of established correspondences using the signature and texture based methods.

The mean and SD values shown in Figure 4 have been computed using on average 180 correspondences from clinical images of four women. Although the set of correspondences represents a relatively large sample, we realize that the number of women in the study is small. Images from 52 women were available, however, our evaluation approach prohibited use of large number of cases; it took on average 45 minutes to visually inspect each image pair, requiring about 13.5 hours per woman.

To further optimize our method for establishing correspondence, it is necessary to develop an automated evaluation approach. We are currently investigating the use of synthetic images. We have developed a 3D computer model of breast anatomy,²¹ which could be used to generate synthetic images with known correspondences. A clinically plausible alternative would be to utilize breast images with identified lesions. The correspondence method would then be used to estimate position of a lesion in one image based on the position of the lesion in the other image relative to the established point correspondences. The evaluation would be performed by comparing the true with the estimated position of the lesion, rather than comparing individual point correspondences. We could not, however, justify the use of such a simplified method before we have gained confidence in accuracy of individual correspondences.

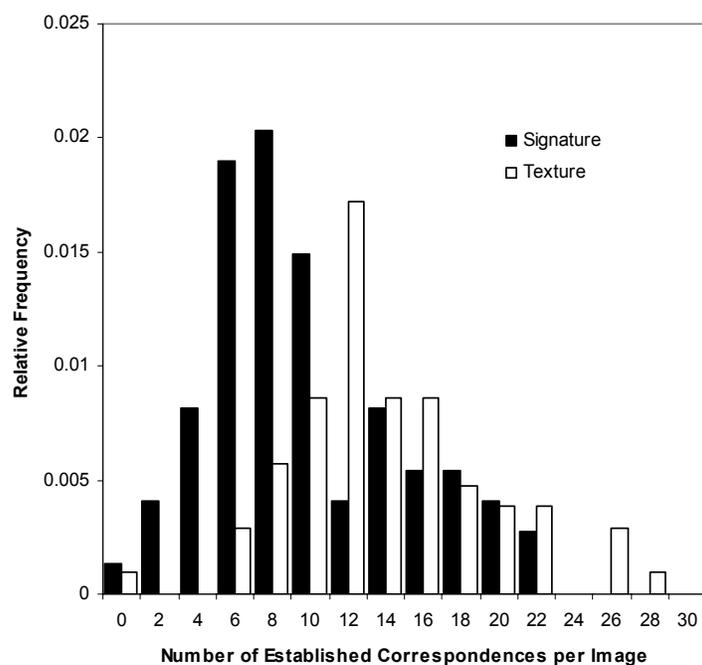


Figure 7: Histograms of the number established correspondences using the signature and texture based methods.

5. CONCLUSIONS

Using clinical data we have evaluated a method for establishing correspondence between mammograms and DBT projection images. The proposed methods are successful in establishing correspondences between mammograms and DBT projections; the average displacement error was 1.6 mm. On average 12.2 correspondences were generated per image pair. The algorithm performance was independent of DBT acquisition angle. No spatial bias has been observed, thus allowing for high accuracy in defining positions anywhere within the breast. To the best of our knowledge, this study represents the first analysis of an automated method for establishing correspondence between clinical breast images with controlled variations.

ACKNOWLEDGMENT

This work was supported in part by National Institutes of Health/National Cancer Institute Program Project Grant P01-CA85484 and by Opex Corporation Continuing Education Program. The authors are grateful to Dr. Ann-Katherine Carton for her help with the acquisition of digital breast tomosynthesis images.

REFERENCES

1. Niklason LT, Christian BT, Niklason LE, et al. Digital tomosynthesis in breast imaging. *Radiology*. 1997;205(2):399-406.
2. Rafferty EA. Tomosynthesis: New Weapon in Breast Cancer Fight. *Decisions in Imaging Economics*. April 2004 2004;17(4).
3. Sallam MY, Bowyer KW. Registering time sequences of mammograms using a two-dimension image unwarping technique. Paper presented at: Proc of 2nd Int Workshop on Dig Mam, 1994; York, England.

4. Sallam MY, Bowyer KW. Registration and difference analysis of corresponding mammogram images. *Medical Image Analysis*. 1999;3(2):103-118.
5. Timp S, Kerssemeijer N. Use of temporal features to improve mass detection. Paper presented at: Proc 7th Annual Int Workshop on Dig Mam, 2004; Chapel Hill, NC.
6. Timp S, van Engeland S, Karssemeijer N. A regional registration method to find corresponding mass lesions in temporal mammogram pairs. *Med Phys*. 2005;32(8):2629-2638.
7. Vujovic N, Brzakovic D. Establishing the correspondence between control points in pairs of mammographic images. *IEEE Transaction on Image Processing*. 1997;6:1388-1399.
8. Richard FJP, Bakic PR, Maidment ADA. Mammogram registration: a phantom-based evaluation of compressed breast thickness variation effects. *IEEE Transaction of Medical Imaging*. 2006;25(2):188-197.
9. Chan HP, Wei J, Sahiner B, et al. Computer-aided Detection System for Breast Masses on Digital Tomosynthesis Mammograms: Preliminary Experience. *Radiology*. 2005;237:1075-1080.
10. Reiser I, Nishikawa RM, Giger ML, et al. Computerized Mass Detection for Digital Breast Tomosynthesis Directly from the Projection Images. *Med. Phys*. 2006;33 482-491.
11. Chan HP, Wei J, Zhang Y, Helvie MA, Moore RH, Kopans DB. Digital Breast Tomosynthesis mammography: Computer-Aided Mass detection by fusion of tomosynthesis and 3D mass likelihood information. Paper presented at: RSNA Scientific Assembly, 2006; Chicago.
12. Conant EF, Schnall MD, Weinstein S. Multimodality screening for breast cancer in a high risk population. Paper presented at: RSNA, 2004; Chicago, IL.
13. Vujovic N. *Registration of Time-Sequences of Random Textures with Applications to Mammogram Follow-up* [Ph.D. dissertation]. Bethlehem, PA: Electrical Engineering and Computer Science Dept., Lehigh University; 1997.
14. Burt PJ. "Fast filter transforms for image processing,". *Comput. Graph. Image Process*. 1981;16:20-51.
15. Kories R, Zimmerman G. A versatile method for the estimation of displacement vector fields from image sequences. Paper presented at: IEEE Workshop on Motion: Representation and Analysis, 1986.
16. Laws KI. *Textured Image Segmentation*, USC; 1980.
17. Miller PI, Astley SM. Classification of Breast Tissue by Texture Analysis. *Image and Vision Computing*. 1992;10:277-282.
18. Bakic PR, Albert M, Brzakovic D, Maidment ADA. Mammogram synthesis using a 3D simulation. II. Evaluation of synthetic mammogram texture. *Medical Physics*. 2002;29(9):2140-2151.
19. Elkin EB, Hudis C, Begg CB, Schrag D. The effect of changes in tumor size on breast carcinoma survival in the U.S.: 1975-1999. *Cancer*. 2005;104(6):1149 - 1157.
20. Hankey BF, Ries LA, Edwards BK. The Surveillance, Epidemiology, and End Results Program: a national resource. *Cancer Epidemiol Biomarkers Prev*. 1999;8:1117-1121.
21. Bakic PR, Albert M, Brzakovic D, Maidment ADA. Mammogram synthesis using a 3D simulation. I. Breast tissue model and image acquisition simulation. *Medical Physics*. 2002;29(9):2131-2139.