Automatic Mammographic Registration: Towards the Detection of Abnormalities

Robert Marti[a], Reyer Zwiggelaar[a] and Caroline M E Rubin[b] *

[a]Division of Computer Science, University of Portsmouth, Portsmouth, UK
[b] Breast Screening Unit, Royal South Hants Hospital, Southampton, UK

Abstract. A novel method to obtain correspondence between landmarks when comparing pairs of mammographic images from the same patient is presented. Our approach is based on automatically establishing correspondence between linear structures (i.e. ducts and vessels) which appear in mammograms using robust features such as orientation, width and curvature extracted from those structures. Registration results are presented which show the potential use of the developed approach for the detection of abnormalities in mammograms.

1 Introduction

Detection of abnormal structures or architectural distortions in mammograms can be performed by comparing different images of the same patient, either the same breast taken at different times (temporal comparison) or using the left and right breast (contralateral comparison). This comparison is not straightforward due to additional dissimilarities between images which are related to patient movement, sensor noise, different radiation exposure and variation of breast compression. Therefore, in order to efficiently compare two mammograms and avoid non target dissimilarities, an initial alignment (also referred to as registration) must be carried out. Methods that are able to recover local deformation (e.g. [1]) rely on corresponding landmarks between images, which turns out to be the most difficult task and plays an important role in registration accuracy. Manual landmark generation is a tedious and time consuming task when the number of control points is large and, moreover, introduces variability. Automatic landmarking methods are, therefore, more suitable but also difficult to develop. Automatically extracted landmarks include breast boundary [2], pectoral muscle [2], salient regions [3] and crossings of horizontal and vertical structures [4]. This paper presents a novel method to establish image correspondence in mammograms based on matching their major linear structures. Establishing correspondence involves various steps: 1. identify linear structures (Sec. 2), 2. extract reliable information from those structures (Sec. 3), 3. obtain correspondence between the structures (Sec. 4) and 4. registration using a point based method [1]. Initial results are presented (Sec. 5) which show that the registration method can provide a framework for the detection and development of mammographic abnormalities.

2 Detection of Linear Structures

Our approach uses both linear structures from the breast boundary and from the breast region (i.e. ducts, vessels, fibrous tissue and pectoral muscle). A distinction between boundary and internal structures is made as boundary information is used to restrict the detection area of internal structures. This results in faster (applying detection to a smaller area) and more robust (avoiding mammographic artifacts) detection. In addition, the breast boundary provides valuable information about breast deformation and has been successfully used in mammographic registration. However, we believe that providing matching of internal structures will achieve a more accurate registration. We should note that the process of feature extraction and matching remains exactly the same for both types of structures, the only difference is that feature extraction and matching is performed separately ending up with two sets of matched points which are simultaneously used in the registration process. The breast boundary is extracted using a similar approach as in [5]. Thresholding and morphological opening (using an octagonal structuring element) is applied to the images. The breast boundary is obtained from segmenting the largest region in the image. Note that artificial boundaries (image edges) are not used, as they do not provide meaningful information.

For the detection of internal linear structures we use a non-linear line operator [6]. An example is shown in Fig. 1b. Once strength, direction and scale information have been obtained we perform different operations to facilitate the feature extraction process. First, we set a conservative threshold on the line strength image in order to remove background noise. Then non-maximum suppression [7] is applied which removes pixels with low intensity values compared to their neighbours along the normal of the linear structure. The line strength image after non-maximum suppression will be used to extract the width of the linear structures in the feature extraction section. Finally, a thinning operation will obtain the backbone of the most representative linear structures in the mammogram.

---

*E-mail: {robert.marti,reyer.zwiggelaar}@port.ac.uk

3   Feature Extraction

After obtaining the salient linear structures in both mammograms, feature extraction is needed in order to obtain descriptors of the structures to be used in the matching process. Corresponding linear structures in two mammograms can present large differences related to line strength and line continuity (due to different radiation exposure and breast compression) but width and orientation of the line and local curvature and branching points are more likely to be preserved and often are features used by radiologists when comparing mammograms. Therefore, features which take line length, end points and line strength into account turn out to be unreliable features to tackle the correspondence problem. In this paper we use local features such as curvature, width and orientation. The basic idea of our method is to extract characteristic points of linear structures determined by their maximal curvature within a local neighbourhood along the linear structure. An example of such points is shown in Fig. 1. It should be noted that most of the branching points and corner points are detected.

| (a) | (b) | (c) |
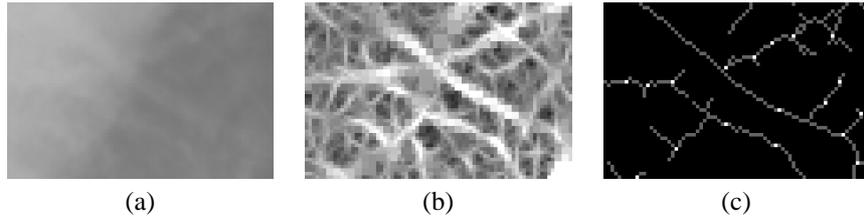
Figure 1. Feature extraction: (a) original mammogram (b) detection of linear structures and (b) extracted linear structures (grey) and curvature points (white).

Curvature values at each pixel are obtained with a similar approach as used in [8]. Curvature (or directional change) between two pixels p and q is defined by the scalar product of their normal vectors. Hence, the curvature measure of a given pixel p is obtained by computing the scalar product between p and its neighbouring pixels.

$$C_p = \frac{1}{N} \sum_{i=1}^{N} \exp(-d_{ip}^2)(1 - \cos(\phi_p - \phi_i)) \tag{1}$$

where ϕ_i is the angle of the normal at a pixel i. As we will be extracting curvature from binary thinned images, we assume unit vectors. N is the number of points in a local neighbourhood and d_{ip} is the Euclidean distance between points i and p. The distance factor is used here to weight the curvature of each point i, in order to incorporate a bias to points closer to p. Position, orientation and width are then extracted from those points and used in the matching process. Orientation is obtained directly from the thinned linear structures. Width is extracted after non-maximum suppression of the line strength images as described in Sec. 2. The width of a linear structure at a point is given by the number of pixels along the normal of the structure until a non-structure pixel is found.

4   Matching

The matching process of the two sets of feature points needs to consider the following assumptions: (a)*Non-rigid motion*: linear structures in mammograms suffer local distortions, therefore they may move independently and no geometrical relationship is established between neighbouring structures. (b)*Multiple matches*: a linear structure in one mammogram can match more than one structure in the other mammogram, and vice versa. (c)*Non-bijectivity*: a linear structure in one mammogram may not have a corresponding linear structure in the other, and vice versa. (d)*Localisation:* After global breast misalignment is removed, matched linear structures lie in approximately the same area in both mammograms; the localisation area M.

We adopt here a similar but more general approach than the one used in [4]. We denote the set of feature points from both mammograms, as $\{a_i | 1 \leq i \leq N_i\}$ and $\{b_j | 1 \leq j \leq N_j\}$, where N_i and N_j are the number of feature points used, which may not be the same. Subsequently, we build a distance matrix (DM) in which each position $DM(i, j)$ describes the normalised distance between features of points a_i and b_j. Hence, a low value means good matching between points. The use of the distance matrix structure fullfills the first three assumptions: independent motion (matched points a_i, b_j do not imply matching a_{i+1}, b_{j+1}); a point a_i may have multiple matched points b_j; and a point in either mammogram may remain unmatched. Satisfying the last assumption, localisation, position $DM(i, j)$ will only have a finite value if points a_i and b_j are in the same localisation area in both mammograms. This assumption can only be stated if both mammograms are globally aligned, that is, global deformation (i.e.

rotation, translation, scale and shear) is removed. Therefore, we initially register mammograms maximising a mutual information measure using an affine transformation (which removes the dependency on patient positioning). The combination of these aspects ensures that the developed method only matches linear structures in the same area and ignores those that have moved too much or are not matched. The normalised distance is determined by three components. The first distance is the Euclidean distance (D_E) between point coordinates. Coordinates of one of the points are transformed (T_α) using the parameters obtained from the registration mentioned earlier. The second distance is the orientation difference between two points (D_ϕ). The third distance is the width difference between two points (D_w) normalised using the maximum width of all the linear structures (W). The normalised distance is given by

$$DM(i,j) = D_E + D_\phi + D_w = \frac{|T_\alpha(a_i) - b_j|}{M} + \frac{1 - \cos(\phi_i - \phi_j)}{2} + \frac{|w_i - w_j|}{W}. \qquad (2)$$

A multi-level approach is used here to ensure a spread of control points over the whole image and to improve the accuracy of the global registration process. At the first level, we align the full images obtaining the transformation parameters α_1. Subsequently, we move to the second level dividing each mammogram in six rectangular sub-images and again align each sub-image to its corresponding sub-image. Note that transformation parameters are carried through each level, assuming that each sub-image at lower levels would suffer a different transformation but it would be related to the deformation on the higher level. Assuming this, we speed up the optimisation process as well as avoiding local minima situated away from the optimum solution. Once the last level i is reached, transformation parameters α_i in each sub-image on that level establish a correspondence for structures within each sub-image. Extracting the local best matches in each sub-image assures that a minimum number of matches will be present in each sub-image, providing an homogeneous point distribution over the mammograms.

5 Results

As mentioned earlier, matched points can be used as control points in mammographic registration using a point based method such as thin plate splines [1]. Previous work [9] showed that registration results using our method (in both temporal and contralateral cases) are comparable to manual registration. For a normal mammogram, the difference between registered images should be minimal, whereas for an abnormal case is likely to highlight the abnormal region. The latter case is shown in Fig. 2, where a spiculated lesion can be seen in the difference image after contralateral registration (the numbers indicate matched point pairs).
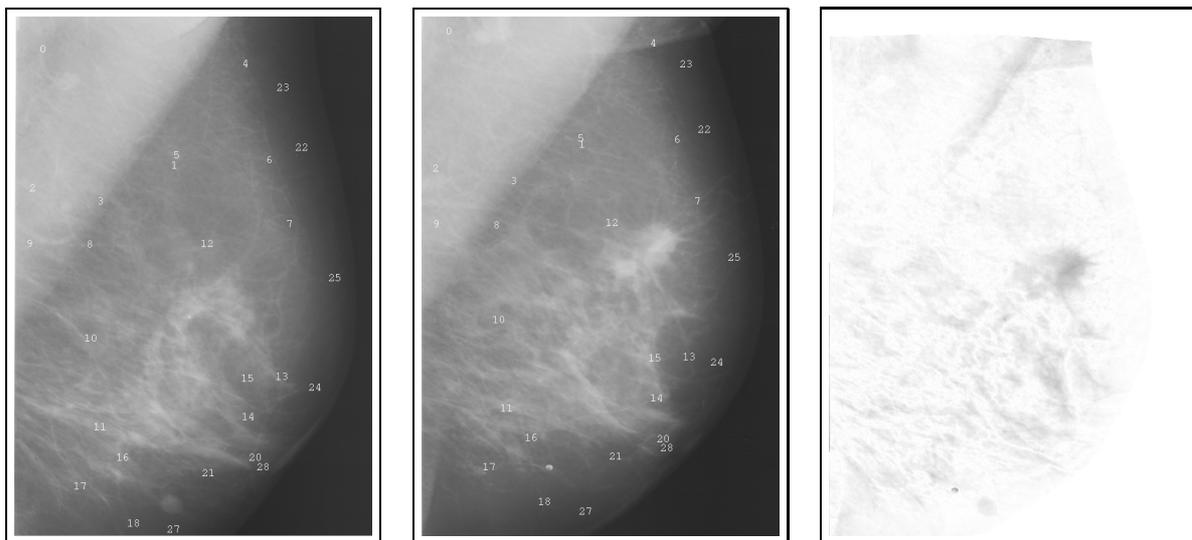
Figure 2. Correspondence in contralateral mammograms (from left to right): reference image (left breast), warped image (right breast), difference between registered image and reference.

Registration has been performed in a dataset containing pairs of temporal mammograms taken at different screening rounds (with approximately 3 years between each pair - 1993, 1996 and 1999 -). The dataset contains 22 patients with 12 of them presenting abnormalities (structural distortions or spiculated lesions) in the last screening round. We believe that the use of features extracted from the difference image could be used to discern between normal

and abnormal mammograms. For mammograms presenting abnormalities in the last screening major differences are likely to appear, whereas similar results to normal mammograms are expected in mammograms from previous years. All the above expectations are confirmed in the results shown in Fig. 3, which plots features extracted from the thresholded difference image in both normal and abnormal cases after registration. For the normal case, similar results are obtained from the three different rounds, although decreasing mean values are obtained through the years which could be due to changes in breast density (increase of the fatty tissue). On the other hand, abnormal mammograms in the last screening year (when the abnormality was diagnosed) show high dissimilarities compared to previous years and to normal mammograms. It should be noted that in early screenings (1993 and 1996) a similar trend to the normal case is obtained for the abnormal mammograms, although slightly higher values are obtained. Features used in Fig. 3 are the normalised area and perimeter of the maximum region in the difference image. Other features have been tested, obtaining similar but less pronounced results.
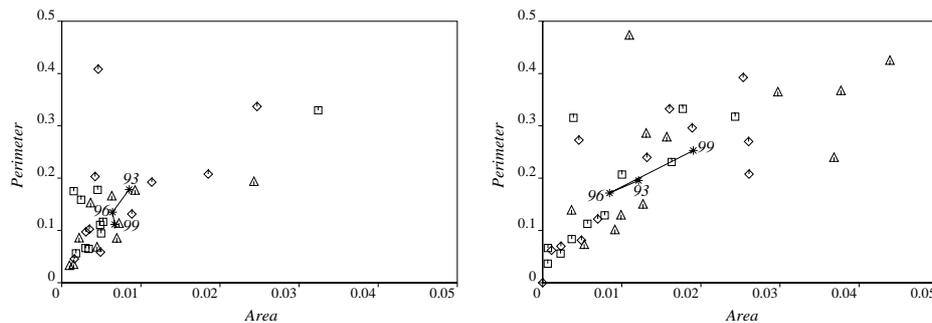
Figure 3. Features from registration of normal (left) and abnormal (right) mammograms in several screening rounds: 1999 ($\triangle$), 1996 ($\square$) and 1993 ($\diamond$), and the mean value for each year ($*$).

6   Conclusions

The work presented here describes a novel approach to solve the problem of extraction of reliable features in mammographic images and establishes correspondence between them in pairs of mammograms. Local features based on scale, orientation and position have been used. Results have shown that analysis of the difference image after registration provides meaningful information which can be used for classification of abnormal mammograms. In addition, results indicate that assessment of mammogram changes due to structural abnormalities developed over time is feasible using registered images. This would lead to early detection of abnormalities and could be used as a training tool for radiologists. Future work will be focused on further exploitation of the registration information, adding new features and incorporating them into classification schemes. In addition, similar experiments using temporal registration will be performed and compared to the contralateral case.

Acknowledgements

References

1. F. Bookstein. "Principal warps: thin-plate splines and the decomposition of deformations." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11(6)**, pp. 567–585, 1989.
2. N. Karssemeijer & G. T. Brake. "Combining single view features and asymmetry for detection of mass lesions." *Digital Mammography, Nijmegen* pp. 95–102, 1998.
3. S. Kok-Wiles, M. Brady & R. Highnam. "Comparing mammogram pairs for the detection of lesions." *Digital Mammography, Nijmegen* pp. 103–110, 1998.
4. N. Vujovic & D. Brzakovic. "Establishing the correspondence between control point in pairs of mammographic images." *IEEE Transactions on Image Processing* **6(10)**, pp. 1388–1399, 1997.
5. M. Y. Sallam. *Image unwarping and difference analysis: a technique for detecting abnormalities in mammograms.* Ph.D. thesis, University of South Florida, 1997.
6. R. Dixon & C. Taylor. "Automated asbestos fibre counting." *Inst. Phys. Conf. Ser* **44**, pp. 178–185, 1979.
7. M. Sonka, V. Hlavac & R. Boyle. *Image Processing, Analysis and Machine Vision.* Chapman and Hall Publishing, 1993.
8. J. Deschênes & D. Ziou. "Detection of line junctions and line terminations using curvilinear features." *Pattern Recognition Letters* **21**, pp. 637–649, 2000.
9. R. Marti, R. Zwiggelaar & C. Rubin. "Automatic registration of mammograms based on linear structures." 17th *International Conference on Information Processing in Medical Imaging* pp. 162–168, 2001.