Detecting Abnormal Mammographic Cases in Temporal Studies Using Image Registration Features

Robert Martí[1], Yago Díez[1], Arnau Oliver[1], Meritxell Tortajada[1], Reyer Zwiggelaar[2], and Xavier Lladó[1]

[1] Computer Vision and Robotics Group, University of Girona, Av. Lluís Santaló 17071, Spain
{marly,yago,aoliver,txell,llado}@atc.udg.edu
http://atc.udg.edu/~marly

[2] Department of Computer Science, Aberystwyth University, Ceredigion, SY23 3DB, Wales, UK
rrz@aber.ac.uk

Abstract. Image registration is increasingly being used to help radiologists when comparing temporal mammograms for lesion detection and classification. This paper evaluates the use of image and deformation features extracted from image registration results in order to detect abnormal cases with masses. Using a dataset of 264 mammographic images from 66 patients (33 normals and 33 with masses) results show that the use of a non-rigid registration method clearly improves detection results compared to no registration (AUC: 0.76 compared to 0.69). Moreover, feature combination using left and right breasts further improves the performance (AUC to 0.88) compared to single image features.

1 Introduction

The detection of abnormalities in mammographic images is an important research topic in breast image analysis. Initial approaches found in the literature [1] were based on the analysis of individual images alone in order to detect (CADe) and classifiy (CADx) microcalcificacions and masses. While microcalcification detection has achieved a sufficient maturity for clinical (and commercial) CAD systems, mass detection still has to improve in terms of specificity and sensitivity. This is mainly due to a larger shape variability and the intensity inhomogeneity of the lesion itself but also of the surrounding tissue which often hinders the detection and segmentation steps.

A way of improving abnormality detection performance is the use of various images from the same patient, similar to radiologists when reading mammographic cases. This has already been approached using contralateral (comparing left and right breasts), ipsilateral (CC and MLO) [2] or temporal [3,4] (same view at different time intervals) studies. Common approaches to compare various images are based on image registration to spatially correlate the images

H. Fujita, T. Hara, and C. Muramatsu (Eds.): IWDM 2014, LNCS 8539, pp. 612–619, 2014.

or to extract and match image features (i.e. nipple position, principal axes or salient regions). This paper belongs to the former set of approaches. The aim of this work is to investigate whether image registration results can be used for the detection of malignant cases in temporal images of the same patient. Temporal comparison has been chosen rather than contralateral or ipsilateral assuming that radiological findings are better detected by analysing breast evolution over time. Note that the aim is not to obtain a particular lesion detection or segmentation but to classify cases as normal or abnormal using solely image registration results. This is of particular interest for CAD systems as a pre-sorting step (classification of normal and abnormal cases) or as prior information for subsequent processing.

The image registration algorithm used in this paper is similar to the work in [5]. The algorithm is based on combining an affine transformation maximising a mutual information similarity measure with a non-rigid point correspondence approach based on a robust point matching algorithm. Intensity and deformation based features obtained from the registration are subsequently used in a machine learning framework to detect abnormal cases with lesions. The contribution of the paper is two-fold: the application of a non-rigid point based image registration algorithm to temporal full-field digital mammographic (FFDM) images, and the use of a machine learning framework with features extracted from the registration results for evaluating detection of malignant cases in temporal images.

2 Image Database

A total of 264 full-field digital (FFD) mammograms were used from 66 different women randomly chosen from a screening population. From those, 33 were normal, while 33 suffered from breast cancer with a visible malignant mass in one of the breasts. Mass area size ranged from 8 to 356 mm^2 with a median area of 64.12 mm^2. Each woman had two mammographic studies (acquired 1-2 years apart) and each study contained two medio-lateral oblique images. Cranio-caudal views were also available but were not used in this study (will be used in future work). Mammograms were acquired using a Selenia FFD mammography system, with resolution 70 micron per pixel, size 4096x3328 or 2560x3328, and 12-bit depth. As the aim is the temporal comparison of mammograms, each mammogram image was registered to its homonymous mammogram from the posterior studies, performing 132 registrations. The presence of masses was annotated by expert radiologists. This allowed us to distinguish between those registration instances containing masses from those not containing them. Hence, for a woman diagnosed with breast cancer we had a registration of the breast with the mass, referred to as *Abnormal with Lesion* (AL), and the registration of the breast without the lesion (healthy breast), referred to as *Abnormal* (A). For a normal case, both registrations were considered as *Normal* (N).

3 Methodology

The main goal here is to investigate whether image registration results can provide significant information in order to help detecting abnormal cases. The overall methodology includes an initial image pre-processing step, registration of temporal images, and mammogram (or patient) classification based on features extracted from the registration results. Figure 1 shows the general framework of the methodology used, while the following subsections provide more details on each step.

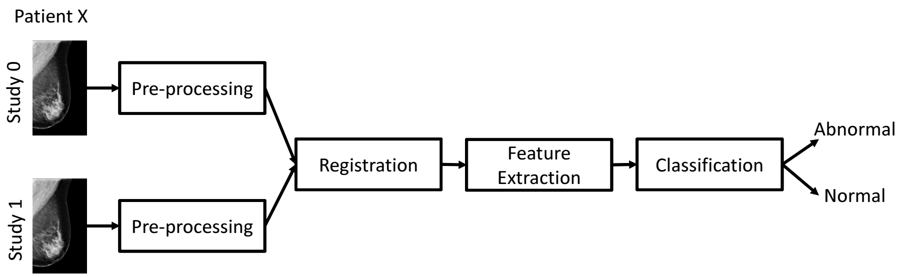


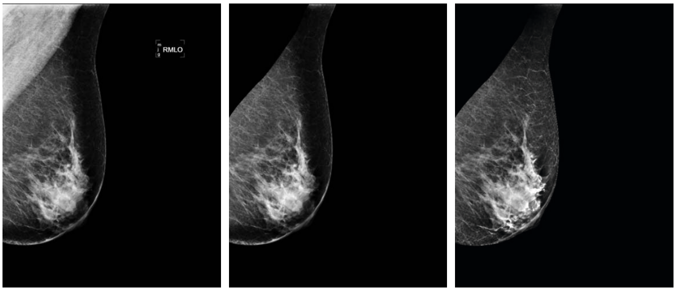
Fig. 1. Overview of the proposed methodology

3.1 Image Pre-processing

Image pre-processing is performed to minimise mammogram and breast variability and to facilitate the subsequent registration step. In that sense, the breast area is automatically segmented using simple thresholding and the pectoral muscle is removed [6]. In addition, a peripheral enhancement method is applied to compensate thickness variations in the breast periphery based on Tortajada et al. [7]. The method automatically restores the overexposed area by equalising the image using information from the intensity of non-overexposed neighbour pixels. The correction is based on a multiplicative model and on the computation of the distance map from the breast boundary. Finally, images are downsampled to half the size using bilinear interpolation in order to reduce computational cost. Figure 2 shows an example of the pre-processing steps described.

3.2 Image Registration

The registration methodology is based on robustly matching interest points in two mammographic images of the same view type. After an initial affine registration maximised by a mutual information metric, the registration algorithm



(a) (b) (c)

Fig. 2. Pre-processing: (a) Original mammogram (b) pectoral muscle removal, and (c) peripheral enhancement

extracts interest points found in the boundary and applies a robust point matching approach obtaining a non-linear transformation [5]. Salient points are defined by computing a maximal local curvature measure in the breast boundary. The point matching approach used here is based on the work of Zheng et al. [8], which uses shape contexts as the measure of point similarity and a graph matching formulation followed by relaxation labelling for obtaining the final point matches.

Point Matching. The robust point correspondence method is based on an iterative graph matching process in order to minimise correspondence errors [8]. Those errors are related to a cost matrix (C_{ij}) which describes the cost of matching one point i in one image (row i) with a point j in the second image (column j). The elements of this cost matrix are obtained using shape contexts [9]. Relaxation labelling is applied to the cost matrix in order to minimise ambiguous matchings. The optimal assignment of the points in the cost matrix is obtained using the Hungarian method, as in [9]. At the end of each iteration, the matched points are used for transforming one point set (p) in order to match the other (q). This transformation is based on a Thin-Plate Splines (TPS) transform, obtaining a smooth transformation between matched points. The transformed points p and q are used for building the cost matrix for the next iteration. The stopping criteria of the iterative process is usually stated in terms of a maximum number of iterations or when the number of matches does not change with respect to the last iteration.

Figure 3 shows an example of image registration of a normal and abnormal case, with the transformed moving image, the difference image and the deformation field magnitude. While differences in the deformation field are difficult to appreciate, structural dissimilarities in the difference image are highlighted, including the lesion in the abnormal case.

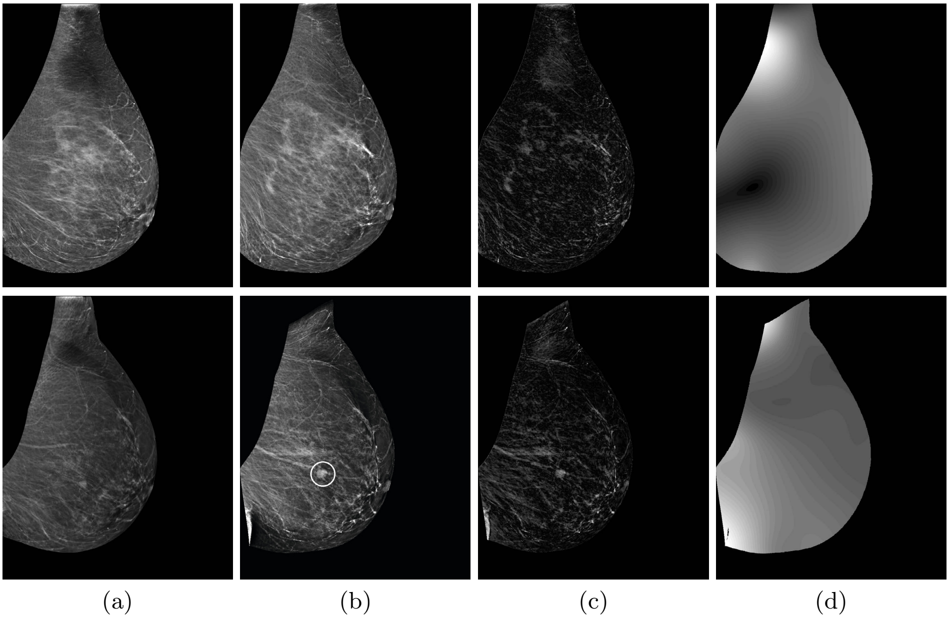


Fig. 3. Image Registration: (a) Fixed and (b) transformed moving mammograms, (c) image difference and (d) deformation field magnitude (brighter areas denote larger deformation). Top row shows a normal mammogram and bottom a mammogram with a lesion (white circle).

3.3 Image Features

From the registration results we extract three sets of features which are then used to classify a patient into normal or abnormal. The first feature set is computed from the difference image while the second set is extracted from the deformation field (the displacement experienced by each pixel normalised by the image size). In these two sets (difference image and deformation field) the features computed are the first five statistical moments of the intensity or deformation distribution. Finally, the third set of features is composed of various similarity measures commonly used in image registration computed between the fixed and moving images: root mean squared error, cross-correlation, entropy of the difference image and mutual information [10], having a total of 14 features.

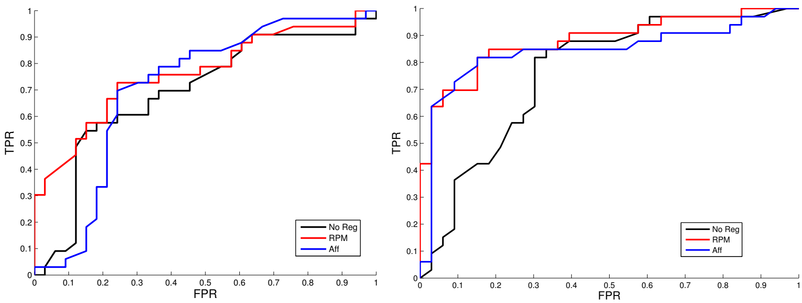
Feature Combination. The above described features are computed for each single temporal registration. As we are registering left and right temporal mammograms of the same patient independently, we also study the effect of combining the features hence obtaining a unique feature vector for each woman. The hypothesis is that this combination can help towards the classification as in normal cases those features are likely to be more stable compared to abnormal cases due to the development of breast cancer. Various simple combinations have been tested: mean, signed and absolute differences, and minimum and maximum. Experimental results evaluating the different combination approaches have shown that combining using the maximum value obtained the best results.

3.4 Classification

Features have been used in a Random Forest (RF) classifier in order to differentiate between normal and abnormal cases containing a mass. The parameters were experimentally set to 500 decision trees and a feature subset size of 3 features for each tree. Although other classifiers (such as SVM, Adaboost and KNN) and feature selection methods have been tested, RF obtained the best results overall. PRTools software has been used for the implementation [11]. All features have been normalised to a zero mean and unit standard deviation. A leave-one-woman-out validation approach has been used for testing.

4 Results

Figure 4 and Table 1 show classification results in terms of ROC curve (true positive rate (TPR) against false positive rate (FPR)) and area under the curve (AUC) when using the proposed algorithm (robust point matching (RPM)) compared to no registration (No Reg), and affine transformation using mutual information (Aff). Features are computed for two cases: for a single registration (*Single*) or combining left and right temporal features using the maximum of both features (*Combined*). For the single case, only one mammogram is used for feature extraction: the one with the mass for abnormal cases and left or right randomly selected for normal ones.



(a) (b)

Fig. 4. Abnormal classification ROC curves using features from robust point matching of the boundary points (RPM) and Affine algorithms also compared to no registration. (a) Single features; (b) combination using the maximum operation

Table 1. AUC for classification of abnormal cases. Features used in the classifier are obtained after no registration (No Reg), affine registration (Aff) or robust point matching of the boundary points (RPM). Single features are compared to their combination using the maximum operation.

	No Reg	RPM	Aff
Single	0.69	0.76	0.71
Combined	0.76	0.88	0.84

Regarding the ROC curves with single features, the use of RPM shows a clear improvement compared to no registration or even affine registration. This is also reflected in the AUC values (0.69 and 0.71 for No Reg and Aff compared to 0.76 for the RPM).

Regarding feature combination, it is also clear that results improve in all cases, including the no registration case. Differences are relevant with respect to the use of registration algorithms compared to no registration, although between Aff and RPM (0.84 vs 0.88) this difference is not that evident. This indicates that non-rigid registration improves classification results, however, further investigation should be carried out including other non-rigid algorithms. Regarding feature analysis it has been observed that features based on the intensity similarity (moments of the difference image and mutual information) show better discriminant properties than the rest of the features. However, with the inclusion of other registration algorithms this could change in favour of other features such as the deformation field.

5 Conclusions

A framework for classifying mammograms into normal and abnormal cases has been presented based on using image based features from temporal non-rigid image registration results. Feature combination between left and right breast has been shown to obtain better results in terms of ROC analysis compared to using single features alone. This indicates that combining features obtained in this fashion with other views such as CC has the potential of further improving the results. This combination will be part of the future work, as well as the evaluation of additional registration algorithms specially those based on intensity metric maximisation (i.e. B-splines and diffeomorphic demons) or the use of a larger and multi-center dataset of images.

Acknowledgements. The research leading to these results has been supported by the Spanish Science and Innovation grant no. TIN2012-37171-C02-01.

References

1. Oliver, A., Freixenet, J., Marti, J., Perez, E., Pont, J., Denton, E.R., Zwiggelaar, R.: A review of automatic mass detection and segmentation in mammographic images. Medical Image Analysis 14(2), 87–110 (2010)
2. Samulski, M., Karssemeijer, N.: Optimizing case-based detection performance in a multiview CAD system for mammography. IEEE Transactions on Medical Imaging 30(4), 1001–1009 (2011)
3. Marias, K., Behrenbruch, C.P., Parbhoo, S., Seifalian, A., Brady, M.: A registration framework for the comparison of mammogram sequences. IEEE Transactions on Medical Imaging 24(6), 782–790 (2005)
4. Díez, Y., Oliver, A., Lladó, X., Freixenet, J., Martí, J., Vilanova, J., Martí, R.: Revisiting intensity-based image registration applied to mammography. IEEE Transactions on Information Technology in Biomedicine 15(5), 716–725 (2011)
5. Martí, R., Raba, D., Oliver, A., Zwiggelaar, R.: Mammographic registration: Proposal and evaluation of a new approach. In: Astley, S.M., Brady, M., Rose, C., Zwiggelaar, R. (eds.) IWDM 2006. LNCS, vol. 4046, pp. 213–220. Springer, Heidelberg (2006)
6. Kwok, S.M., Chandrasekhar, R., Attikiouzel, Y., Rickard, M.: Automatic pectoral muscle segmentation on mediolateral oblique view mammograms. IEEE Transactions on Medical Imaging 23(9), 1129–1140 (2004)
7. Tortajada, M., Oliver, A., Martí, R., Vilagran, M., Ganau, S., Tortajada, L., Sentís, M., Freixenet, J.: Adapting breast density classification from digitized to full-field digital mammograms. In: Maidment, A.D.A., Bakic, P.R., Gavenonis, S. (eds.) IWDM 2012. LNCS, vol. 7361, pp. 561–568. Springer, Heidelberg (2012)
8. Zheng, Y., Doermann, D.: Robust point matching for nonrigid shapes by preserving local neighborhood structures. PAMI 28(4), 643–649 (2006)
9. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(4), 509–522 (2002)
10. Zitova, B.: Image registration methods: A survey. Image and Vision Computing 21(11), 977–1000 (2003)
11. Duin, R., Juszczack, P., Paclik, P., Pekalska, E., de Ridderand, D.M.J., Tax, D.: PRTools4, A Matlab Toolbox for Pattern Recognition. Delft University of Technology (2004)