

COMPARISON OF REGISTRATION METHODS USING MAMMOGRAPHIC IMAGES

Yago D ez, Arnau Oliver, Xavier Llad  and Robert Mart 

Computer Vision and Robotics Group.
University of Girona, Spain.

ABSTRACT

The detection of architectural distortions and abnormal structures in mammographic images can be based on the analysis of bilateral and temporal cases using image registration. This work presents a quantitative evaluation of eight state-of-the-art image registration methods applied to mammographic images. These methods range from a global and rigid transformation to local deformable paradigms using various metrics and multi-resolution approaches. The aim of this study is to assess the suitability of these methods for mammographic image analysis. Evaluation using temporal cases based on quantitative analysis gives an indication of the accuracy and robustness of the different algorithms. This work shows that local deformable paradigms (B-spline deformations) obtain the most accurate registration results.

Index Terms— Mammographic image analysis, image registration, algorithm evaluation.

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 using an image registration algorithm must be carried out. Although registration of mammographic images is regarded as an ill-posed problem where the perfectly registered image can never be obtained due to the projective nature of the images, it is still as an important research topic for the development of computer aided diagnosis (CAD) system and this has yet to be included into currently commercially available CAD systems. Image registration has been widely used in medical applications for quite a while now, see for instance the surveys of [1].

e-mail: {yago;aoliver;llado;marly}@eia.udg.edu. This work was supported by the Spanish Ministerio de Educaci n y Ciencia (no. TIN2007-60553) and by the CIRIT and CUR of DIUIE of Generalitat de Catalunya (no. 2008SALUT00029).

In that sense, the analysis of mammographic images is not an exception. Most of the published approaches on mammographic image registration use breast boundary information as it is relatively easy to extract and provides important information about the breast deformation [2]. Another group of approaches can be classified as being intensity based, where the deformation is recovered maximising a measure of similarity between images. These methods were reported to obtain robust results for global transformations [3], but could not account for severe local distortions. In addition to the breast boundary, information about the deformation of internal regions has also been used in several approaches [2, 4]. Consequently, the aim of this work is to review and evaluate the applicability of state-of-the-art intensity based image registration algorithms to mammographic image analysis. In next section, we briefly describe the image registration algorithms. Subsequently, registration results are presented, describing the data, quantitative evaluation experiments. Finally, discussions and conclusions are presented.

2. METHODOLOGY

The methodology used in this paper is based on applying an image registration method from those described below to a temporal case: two images of the same breast acquired at different time intervals. The result of the registration can be used for tasks such as visualization but also to detect changes or as a form of prior information for CAD systems. The example in Figure 1 illustrates the use of image registration for the detection of abnormalities. The moving image (an image acquired in the last screening round) shows a spiculated lesion in the central breast region which is not visible on the fixed image (previous screening). After performing the registration, the spiculated lesion is clearly visible in the difference image and the registration information can be used to analyze possible early signs of the lesion in the previous screening rounds. The following image registration methods are used in this work, divided into global and local methods.

2.1. Global Methods: Rigid and Affine

Rigid (translation and rotation) and affine transformations (allowing additional shearing for a total of 6 parameters in 2D)

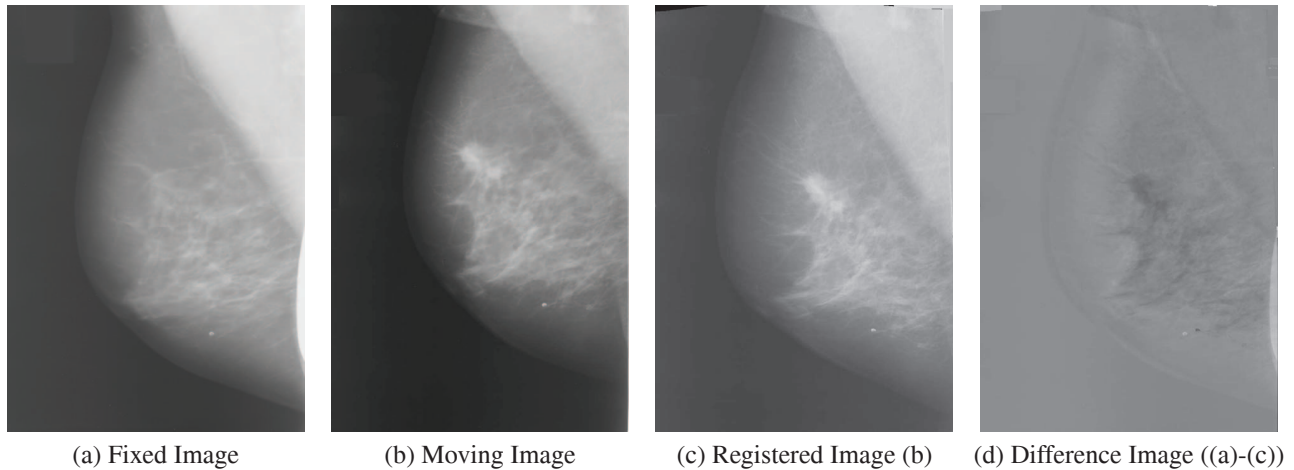


Fig. 1. Non-rigid registration example of mammographic images.

have been evaluated from the perspective of global methods, in which all pixels suffer the same transformation. Parameters are recovered by maximizing a similarity measure using an optimization approach. Mutual information (MI) and sum of squared differences (SSD) have been used as metrics and gradient descent method as the optimization algorithm [1].

2.2. Local Methods: B-Splines, Polyrigid and Demons

Local methods (also known as deformable registration) include methods where pixels are transformed locally, having a different transformation depending on their local similarity and position. Many methods and variations have been proposed under this assumption. In addition to the local computation of the metric, aspects such as regularization in order to ensure smoothness and continuity are usually incorporated. This regularizations can be implicit in the transformation or considered as a constraint added to the transformation function. Among these methods we have selected B-spline free form deformations (FFD) [5], Polyrigid transformations [6] and Thirion's Demons algorithm [7], due to its wide popularity in medical applications although not widely tested in x-ray mammographic images. B-spline FFD algorithm is based on deforming an image by modifying a mesh of control points following a maximisation of a similarity measure. These control points define a mesh of smooth and continuous B-Spline functions with the characteristic of having a limited support (modifying a control point only affects neighboring points). Polyrigid transformations were proposed as a novel type of transformations in order to provide a higher degree of flexibility compared to rigid transformations but a less deformable nature as for instance found in the B-splines formulation. They exhibit a locally rigid behavior and continuous and diffeomorphic properties by integrating the infinitesimal displacements of each rigid transformation into an ODE formulation. Finally, the Demons algorithm is based on view-

ing the registration as a diffusion process, inspired by optical flow formulation, where the diffusivity is related to the local characteristics of the image (ie. second order derivatives).

2.3. Multi-resolution and Algorithm Combination

Although these methods are often applied independently it is commonly accepted that results can be improved in terms of accuracy and robustness by using a multi-resolution (MR) approach or combining different approaches. The former is based on registering the images in a lower resolution, propagating parameter estimation into a higher resolution and registering again. This often avoids local minima in the parameter search space and reduces computational time. Combination exploits the benefits of the different methods, for instance using a global and a local method (i.e. affine registration with a B-spline deformation). In this case, the global method recovers for main pose and scale differences and the local method accounts for localized non-linear deformations.

3. RESULTS

This section shows evaluation results for the registration algorithms described in the previous section. The data used in this paper is a local database of 22 normal and abnormal patients with temporal information (images of the same breast taken at different time intervals usually two or three years). The mammograms were originally on film and scanned using a Lumisys scanner at a resolution of 50 microns and rescaled up to 200 microns for computational purposes. Evaluating the results of registration methods in mammographic images is not an easy task. In this paper we compute similarity metrics before and after registration to obtain an indication of how similar images are. A higher similarity is expected after image registration, and the method with the highest similarity

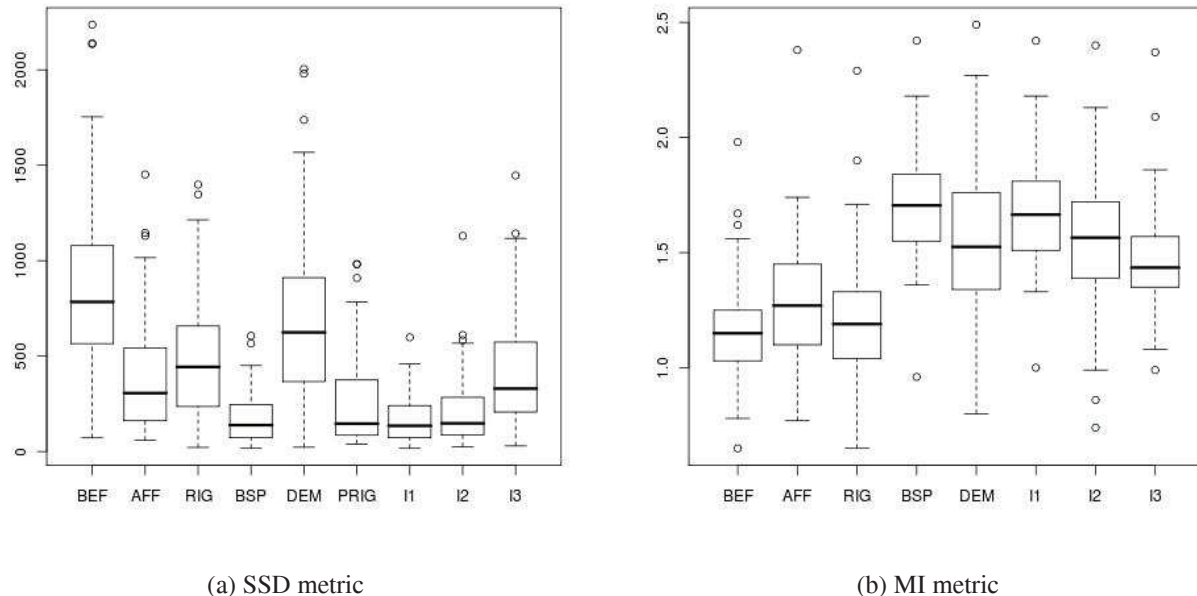


Fig. 2. Boxplots for Metric Evaluation. BEF = Before registration, AFF= Affine, RIG = Rigid, BSP = B-Splines FFD, DEM = Demons, PRIG= Polyrigid, I1 = MR Bsplines, I2 = Affine + Bsplines, I3 = MR Affine + MR Bsplines

would be expected to be the most accurate. However, metric does not always tell the full story as sometimes images that are "closer" in terms of metric functions are perceived to be more different. In order to analyze the correlation between similarity metric and visually correct registration, we also reviewed our methods using an observer study, where registration results were evaluated by 11 experts with a different degree of expertise in both computer vision and radiology. These observers had different degrees of experience in mammographic image analysis and medical practice: one expert radiologist, one trainee radiologist, and 9 computer vision experts with over 10 years experience (4), 5-10 years (3) and less than 5 years (2). In general, evaluation using similarity metrics agreed with observer results, except for the cases where the registration algorithm introduced unrealistic deformations (results not included due to space limitations).

3.1. Implementation Details

All registration methods have been implemented using the Insight Toolkit (itk) libraries, available at www.itk.org. Generally we implemented all algorithms in two versions, one using the SSD distance for the optimizer (*itk::RegularStepGradientDescentOptimizer*) objective function and another using the MI distance. This is not true for the Demons method (that does not use an optimizer as such) and for the Polyrigid method, where only an SSD version was provided due to computational efficiency. The code for the Polyrigid

method was obtained following the instructions provided in [6]. For practical reasons we fixed a maximum number of iterations for all methods. We considered a maximum of 1000 iterations for each registration, in improved methods these iterations were distributed between the two methods used or for all the multiresolution levels.

3.2. Metric Evaluation

Our assertion is that higher similarity metric means more similar images, hence, better registration. For all the images in the database we calculated SSD and MI metrics. We analyzed 125 registrations. For each case, images were registered using the methods described in section 2 (global and local). Figure 2 presents boxplot charts for the complete database for both metrics. The metric value (SSD or MI) used is computed after registration between the registered and the moving images. This value is related to the metric optimized by the registration algorithm (i.e. SSD values are related to registration methods which maximise using SSD and the same for MI).

All the methods used improved metric measurements in both distances. Concerning SSD distance, B-Splines seems to work best among individual methods although PolyRigid and Affine get good results too. The use of MR and combination of methods generally do better than individual ones. MR B-Splines is seen to be the best methods overall. We also observe that, concerning this metric, the difference between the two best observed methods (i.e. B-Splines with and with-

out multi-resolution) is small. However, the method that used multiresolution obtained much better rating in the observation study mentioned earlier. As for the MI metric, B-Splines methods obtain the best results and MR and method combination keep on doing generally better than individual ones. In this case, however, rigid and affine methods do not perform too well and the demons method obtains significantly better results compared with its SSD implementation.

3.3. Time Study

The time needed to execute an algorithm might be a limiting factor in certain contexts. In this section we study the time needed by the algorithms studied throughout the paper. Figure 3 presents the mean times for 25 executions of the algorithms (represented by bars) and the standard deviation for these execution time (depicted as error bars).

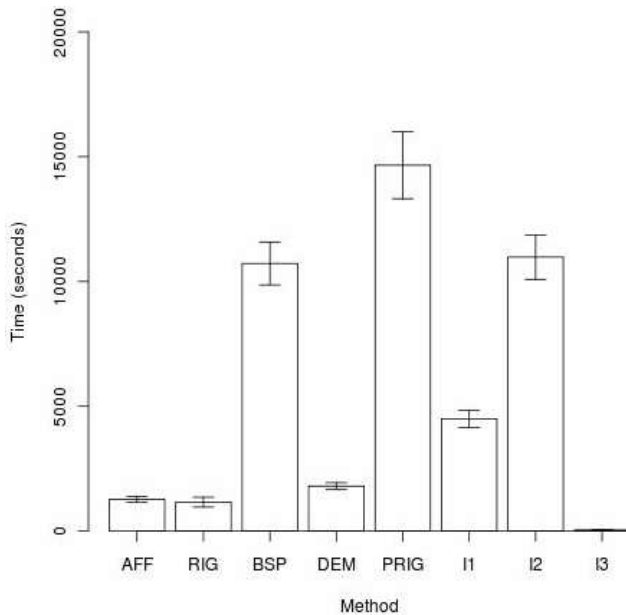


Fig. 3. Average time (bars) and standard deviation (error bars). AFF= Affine, RIG = Rigid, BSP = B-Splines FFD, DEM = Demons, PRIG= Polyrigid, I1 = MR Bsplines, I2 = Affine + Bsplines, I3 = MR Affine + MR Bsplines

We observe how Polyrigid registration is the slowest. B-Spline registration methods are also quite slow (as can be seen in the bars corresponding to the Bspline method alone as well as the combination of affine and B-Spline registration), although their running times are greatly reduced if multiresolution is used. Affine, Rigid and Demons methods are much faster, but not as fast as the combination of the combination of Affine and Flexible Multiresolution methods. We have observed how, if this method is allowed to run for a higher number of iterations not only does its running time grow but its results also improve. This improvement is however not too

significant, so, for the sake of concretion, we present only the data that we consider representative.

4. CONCLUSIONS

We have shown how state-of-the art registration methods are applicable to mamographic image registration. We obtained significant reductions in the metric measurements between images prior and after registration. B-spline based methods obtained the best results from this point of view. Concerning running times we have shown how using multiresolution helps reduce the time needed. The method that we consider best overall is the B-spline method that uses multiresolution, as it is quite fast and gets the best results overall in distance reduction. Future work will focus on adapting the B-spline registration methodology to contrast enhanced x-ray mammo-graphic images, taking into account the non-linear intensity relationship into the similarity metric.

5. REFERENCES

- [1] B. Zitova, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, October 2003.
- [2] Robert Marti, David Raba, Arnau Oliver, and Reyer Zwiggelaar, "Mammographic registration: Proposal and evaluation of a new approach," in *Digital Mammography / IWDM*. 2006, vol. 4046 of *Lecture Notes in Computer Science*, pp. 213–220, Springer.
- [3] S. van Engeland, P. Snoeren, J. Hendriks, and N. Karssemeijer, "A comparison of methods for mammogram registration," *IEEE Transactions on Medical Imaging*, vol. 22, no. 11, pp. 1436–1444, 2003.
- [4] K. Marias, C. P. Behrenbruch, S. Parbhoo, A. Seifalian, and M. Brady, "A registration framework for the comparison of mammogram sequences," *IEEE Transactions on Medical Imaging*, vol. 24, no. 6, pp. 782–790, 2005.
- [5] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast mr images.," *IEEE transactions on medical imaging*, vol. 18, no. 8, pp. 712–721, August 1999.
- [6] Vincent Arsigny, Xavier Pennec, and Nicholas Ayache, "Polyrigid and polyaffine transformations: A novel geometrical tool to deal with non-rigid deformations application to the registration of histological slices," *Medical Image Analysis*, vol. 9, no. 6, pp. 507–523, December 2005.
- [7] J. P. Thirion, "Image matching as a diffusion process: an analogy with maxwell's demons," *Medical Image Analysis*, vol. 2, no. 3, pp. 243–260, September 1998.