Image based registration between full X-ray and spot mammograms: Analysis of registration accuracy in subgroups

S. Said[a], P. Clauser[b], N.V. Ruiter[a], P.A.T. Baltzer[b], and T. Hopp[a]

[a]Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[b]Medical University of Vienna, Vienna, Austria

ABSTRACT

The most typical type of cancer among women is breast cancer. Despite the crucial role that digital mammography plays in the early identification of breast cancer, many tumors could not be discriminated on mammography, especially in women with dense breast tissue. Contrast-enhanced magnetic resonance imaging (CE-MRI) of the breast is routinely used to find lesions that are invisible on mammography. MRI-guided biopsies must be used to further analyze these lesions. But MRI-guided biopsy is highly priced, time-consuming, and not frequently accessible. In our earlier work, we introduced a novel method using two methods of registration: biomechanical and image-based registration to transfer lesions from MRI to spot mammograms to allow X-ray guided biopsy. In this paper, we focus on enhancing and developing the image-based registration between full and spot mammograms and analyzing a correlation between the accuracy of our method and features such as views, location of lesion, breast area, size of lesion in each modality, and age. Results for 48 patients from the Medical University of Vienna are provided. The median target registration error is 20.9 mm and the standard deviation is 23.9 mm.

Keywords: Additional Image Similarity Metrics, Combined Image Similarity Metrics, Sub-group Analysis

1. INTRODUCTION

Medical imaging using various modalities has been utilized regularly to aid clinicians in qualitative diagnosis for early detection of breast cancer. MRI has the potential to detect lesions that are not evident using conventional imaging like X-ray mammography or ultrasound. However, detecting lesions by MRI can be challenging when treating breast illness clinically. Further investigations are required because malignancy is discovered in around every second to third of these lesions. Depending on the level of suspicion, these lesions may need to be monitored or a biopsy performed. An MRI-guided biopsy is crucial in these particular circumstances, according to Spick et al.[1] However, MRI guided biopsies are expensive, and a recent assessment found that there is a serious global lack of MRI-guided breast treatment.[2]

Based on our earlier work,[3] a novel method for a clinical process has been introduced that would enable the transfer of MRI-visible lesions to spot mammograms. It will provide the ability to use X-ray guided biopsy instead. It consists of two registration methods and has been tested using clinical datasets: biomechanical model-based registration between MRI and X-ray mammograms[4,5] followed by an image-based registration between full and spot mammograms.[6] In this paper, we focus on developing the image-based registration using more image similarity metrics and considering a combination of more than one image similarity metric to enhance robustness. By the combination of image similarity metrics, we are using K-means to cluster the estimated position of the spot mammogram within the full mammogram that has been calculated by the different image similarity metrics into two classes, thereby taking the majority vote. We are averaging those detected positions. We tested our method using considerably more clinical datasets than in our earlier work in order to evaluate the robustness of a variety of views, patient characteristics, and manufacturers. We also evaluate the correlation between the accuracy of our method and features such as views, location and size of lesion, breast area, and age.

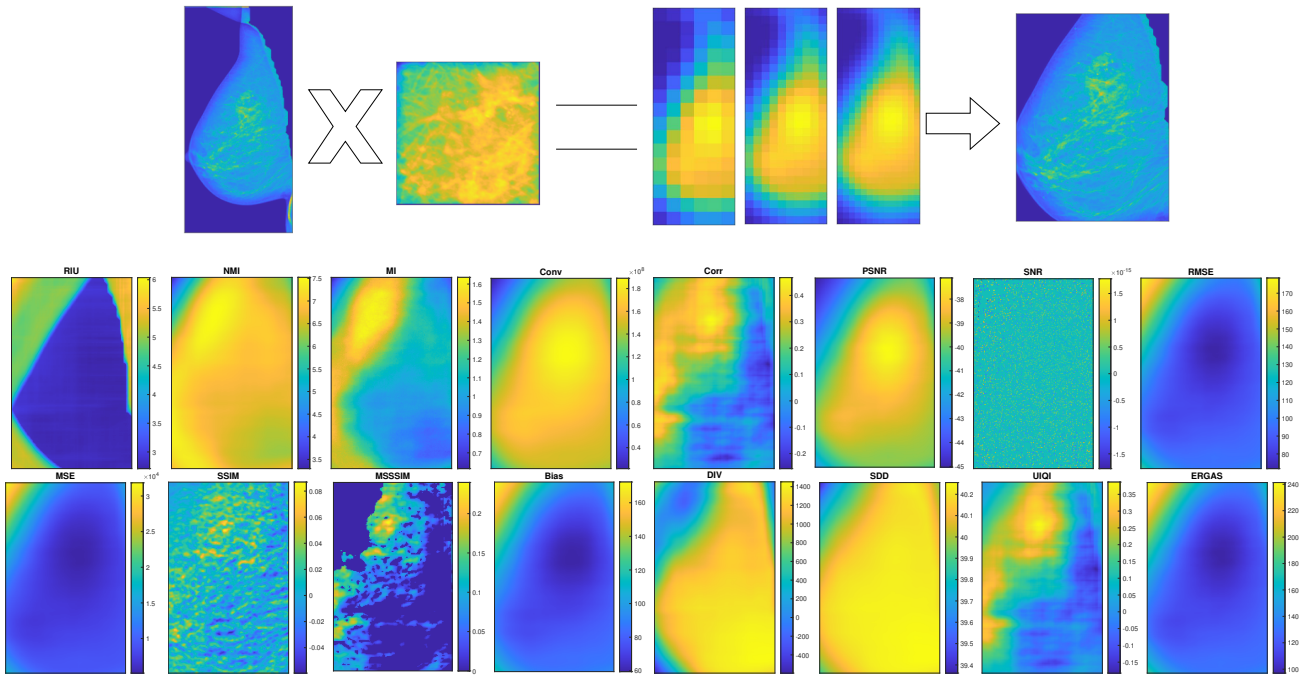


Figure 1. Two stages of image-based registration: getting ROI (first row). Applying different image similarity metrics while sliding spot mammogram over ROI (second and third row).

2. METHODS

We used the same method of image-based registration that was presented before[6] which uses two stages of image similarity metrics. In the first stage, the region of interest (ROI) is found by convolving the spot mammogram (moving image) with the full mammogram (reference image) as shown in Figure 1 (first row). The second stage is applying image similarity metrics while sliding the spot mammogram over the ROI to define the correct alignment with the spot mammograms in Figure 1 (second and third row).

In comparison to our earlier work, more image similarity metrics have been investigated. The first method is structural similarity (SSIM) and multi-scale structural similarity (MS-SSIM). SSIM is responsible for extracting three features; luminance, image contrast, and structure while MS-SSIM supports more variations such as image resolution.[7] The second method is bias, difference in variance (DIV), and standard deviation difference (SDD). Bias is the difference between the mean of each image while DIV is the difference between the variance of each image.[8] The third method is universal objective image quality index (UIQI), which is a combination of calculating three factors; loss of correlation, contrast distortion, and luminance distortion.[9] The fourth method is Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS) and root mean square error (RMSE) between the two images. ERGAS is summarizing the error that shows global results in the reference image (full X-ray mammograms).[8]

We extract the optimal values of the calculated image similarity metrics and register the moving image to the reference image using the optima of the image similarity metrics as reference point with the best match of the moving image. We are clustering the positions of the optima values of the most robust image similarity metrics into two classes using K-means to get the majority vote. We are averaging those detected positions of the best agreement according to the respective metric. The combined metric is denoted as combMat as shown in Figure 2 (a). Based on our available clinical datasets, the most robust image similarity metrics are MSE, correlation coefficient, convolution, SSIM, ratio image uniformity (RIU), UIQI, and bias.

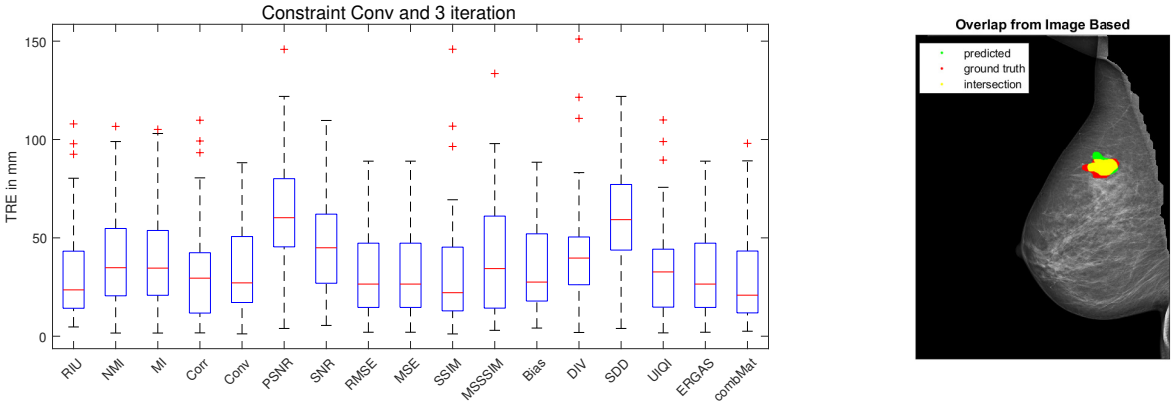
3. RESULTS

We tested our methods using 48 patient datasets from the Medical University of Vienna. The datasets consist of two categories since the spot mammograms could be taken in prone or upright position. We used 46 datasets taken in the prone position while only two datasets in the upright position were available. For the prone position, there are six options; the full X-ray mammograms could be cranio-caudal (CC) or mediolateral oblique (MLO) views, while the spot mammograms could be cranio-caudal (CC) or mediolateral (ML) or lateromedial (LM) views. The full X-ray mammograms have been taken with devices from several manufacturers such as Siemens Mammomat Inspiration, Siemens Mammomat Revelation, Sectra MicroDose L30 Mammography, FUJIFILM with two different models, Lorad Selenia, Sectra Imtec AB, Philips Digital Mammography, and Senographe Essential with two different editions. As a landmark, we are calculating the centers of gravity of freehand lesion annotations which has been performed by an experienced radiologist using Slicer 3D.[10] The target registration error (TRE) has been calculated as the Euclidean distance between the center of gravity of the lesion predicted in the registered moving image and the center of gravity of the annotated lesion in the reference image to evaluate our image-based registration method.

The analysis in Figure 2 (a) shows that the combined metrics could be considered the most robust. The median TRE of 48 patients is 20.9 mm. The lower and the upper quarters are 11.9 mm and 43.3 mm, respectively while the minimum and the maximum values are 2.5 mm and 98 mm, respectively. In all the other metrics with similar median such as RIU and SSIM, the number of outliers is larger. Also, in the cases without outliers, the median TRE is larger. An example of the annotation is shown in Figure 2 (b). The red lesion represents the lesion that is marked by the radiologist and the green lesion represents the predicted lesion based on the annotation in the spot mammogram using image-based registration. The yellow color represents the overlap of the two lesions.

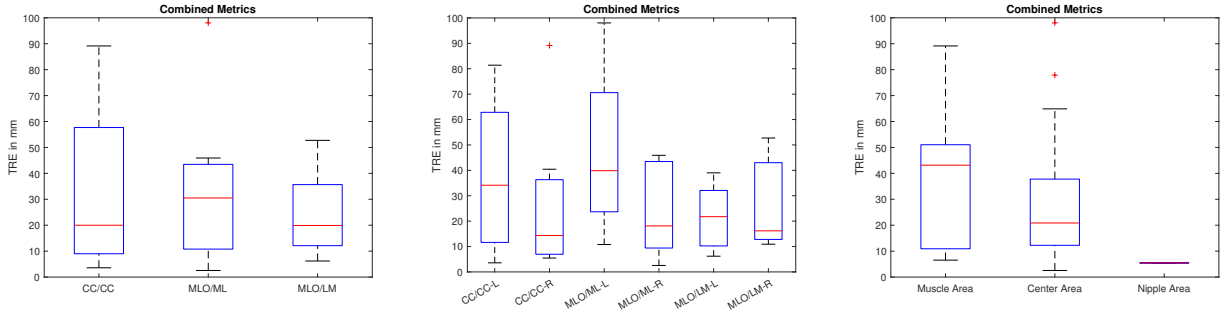
We analyzed our method using five features to investigate if subgroups are correlating with the TRE of our method. The first sub-class we analyzed, are the views of the full X-ray and spot mammograms. The views include the variance in the projection angle which may differ in full X-ray and spot mammograms and left and right breasts. Based on our available datasets, we focus on three views out of the twelve options which are CC/CC, MLO/ML, and MLO/LM for full X-ray and spot mammograms, respectively for the prone position, as 40 out of 48 clinical datasets fall in these views. It is observed that CC/CC (15 datasets) and MLO/LM (15 datasets) have nearly the same median while MLO/ML (10 datasets) is not working so well with our method as shown in Figure 3 (a). We divided these three classes of views into left and right breasts in which the number of datasets in each class is eight, seven, four, six, eight, and seven, respectively as shown in Figure 3 (b). For the left breasts in the three views, it shows that the median TRE is always higher than the median TRE of right breasts. One reason may be that the information that is recorded in the metadata regarding the field of view rotation relative to the physical detector which is 180° for left breasts and 0 ° for right breasts.

The second sub-class is the location of lesions in the full X-ray mammograms. We divided the locations into three main classes as shown in Figure 3 (c) at the back of the breast near the muscle in the posterior direction (11 datasets), in the center of the breast (36 datasets), and at the front of the breast near to the nipple in the anterior direction (1 dataset). The center area is the region that works better for our method, while it is expected that the lesions near to the muscle are hard to register. For the lesions near to the nipple, the dataset is likely too small for a conclusion. There are additionally three sub-classes that have an independent or weak correlation with the accuracy of our method as shown in Figure 4. We analyzed all of the next three-sub classes based on a polynomial fitting curve (first degree) to evaluate how the datasets are distributed compared to the TRE of our registration method. It is observed that the breast area has a weak correlation with the TRE of our method with a Pearson correlation coefficient (r) of 0.21 as shown in Figure 4 (a). The minimum area in our dataset is $5.6\text{x}10^3$ mm^2 while the maximum area is $3.1\text{x}10^4$ mm^2. For the independent parameters of our method, the first aspect is the lesion size in the X-ray and spot mammograms. Since both annotations of X-ray and spot mammograms are correlating, we only present here the results of the size of the lesion in X-ray mammograms with an r of 0.007 as shown in Figure 4 (b). The second aspect is the age as shown in Figure 4 (c), which ranges from 29 to 78. It is also considered an independent parameter of our method with an r of 0.07.



(a) (b)

Figure 2. Analysis of different image similarity metrics with 48 patients in CC/MLO and CC/ML/LM views for full and spot mammograms, respectively. Distribution of the TRE for all tested image similarity measures (a). Example of the predicted lesion (green), ground truth (red), and overlap with the ground truth (yellow) in one of the cases using image-based registration.



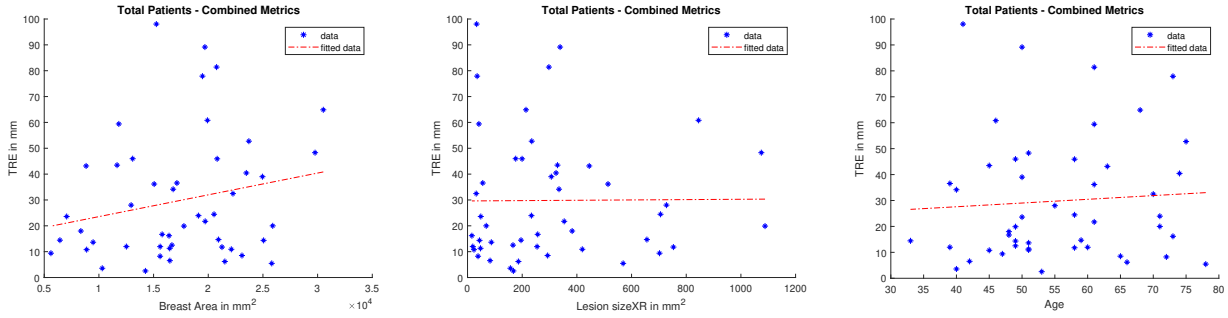
(a) (b) (c)

Figure 3. The first two sub-classes of having a correlation with the accuracy of our method. The first sub-class is the views category which consists of three main views CC/CC, MLO/ML, and MLO/LM for full X-ray/spot mammograms in the prone position (a). Results splitted into left and right breasts (b). The second sub-class is the location of lesions in X-ray mammograms which consists of three main categories near to the muscle in the posterior direction (Muscle Area), the center of the breast (Center Area), and near to the nipple in the anterior direction (Nipple Area) (c).

4. DISCUSSION AND CONCLUSION

We presented a follow-up study on the earlier proposed method for image-based registration between full and spot mammograms. For robustifying our method, we combined multiple image similarity metrics and clustered them into two classes using K-means. We analyzed our method using 48 patients. Our proposed method provides promising results with a median TRE of 20.9 mm. Even with a lot more diverse datasets, a similar TRE could be achieved compared to our earlier paper. Though, there is still a high error in some cases which we would like to investigate in future. By the subgroup analysis, we identified that these cases are mostly when a lesion is close to the muscle or when the views in full and spot mammograms are different. We believe that more investigation into the combined metrics and taking into consideration the difference in views between MLO/ML might improve these cases.

We found a relation to the accuracy of our method with imaging and patient features. Our analysis shows that the accuracy is correlating with views and location of the lesion in X-ray mammograms while it is only weakly correlating with breast area. Also, there is no correlation with size lesion and age. Our clustering method and the analysis in combination with the earlier presented image registration method will provide a robust alignment



(a) (b) (c)

Figure 4. The second three sub-classes of having no / weak correlation with the accuracy of our method. For weak correlation, the first sub-class is the breast area (a). For no correlation, the first sub-class is the size of lesions in X-ray mammograms (b). The second sub-class is the age (c).

between the full and spot mammograms. By having a robust method, in combination with the earlier presented registration workflow, it will provide the ability to use X-ray guided biopsy instead of MRI-guided interventions.

REFERENCES

[1] Spick, C. and Baltzer, P. A. T., "Diagnostic utility of second-look US for breast lesions identified at MR imaging: Systematic review and meta-analysis," *Radiology* **273**(2), 401–409 (2014). PMID: 25119022.

[2] Clauser, P., Mann, R., Athanasiou, A., Prosch, H., Pinker, K., Dietzel, M., Helbich, T. H., Fuchsjäger, M., Camps-Herrero, J., Sardanelli, F., Forrai, G., and Baltzer, P. A. T., "A survey by the European Society of Breast Imaging on the utilisation of breast MRI in clinical practice," *European Radiology* **28**, 1909—1918 (May 2018).

[3] Said, S., Clauser, P., Ruiter, N. V., Baltzer, P. A. T., and Hopp, T., "Image registration between MRI and spot mammograms for X-ray guided stereotactic breast biopsy: preliminary results," in [*Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling*], Linte, C. A. and Siewerdsen, J. H., eds., **11598**, 354 – 361, International Society for Optics and Photonics, SPIE (2021).

[4] Hopp, T., Baltzer, P., Dietzel, M., Kaiser, W., and Ruiter, N., "2D/3D image fusion of X-ray mammograms with breast MRI: Visualizing dynamic contrast enhancement in mammograms," *International Journal of Computer Assisted Radiology and Surgery* **7**, 339–48 (06 2011).

[5] Hopp, T., Dietzel, M., Baltzer, P., Kreisel, P., Kaiser, W., Gemmeke, H., and Ruiter, N., "Automatic multimodal 2D/3D breast image registration using biomechanical FEM models and intensity-based optimization," *Medical Image Analysis* **17**(2), 209–218 (2013).

[6] Said, S., Clauser, P., Ruiter, N., Baltzer, P. A. T., and Hopp, T., "Image based registration between full X-ray and spot mammograms for X-ray guided stereotactic breast biopsy," in [*Medical Imaging 2022: Image-Guided Procedures, Robotic Interventions, and Modeling*], Linte, C. A. and Siewerdsen, J. H., eds., **12034**, 614 – 621, International Society for Optics and Photonics, SPIE (2022).

[7] Wang, Z., Simoncelli, E., and Bovik, A., "Multiscale structural similarity for image quality assessment," in [*The Thrity-Seventh Asilomar Conference on Signals, Systems Computers, 2003*], **2**, 1398–1402 Vol.2 (2003).

[8] Cetin, M. and Musaoglu, N., "Merging hyperspectral and panchromatic image data: qualitative and quantitative analysis," *International Journal of Remote Sensing* **30**(7), 1779 – 1804 (2009).

[9] Wang, Z. and Bovik, A., "A universal image quality index," *IEEE Signal Processing Letters* **9**(3), 81–84 (2002).

[10] Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., Buatti, J., Aylward, S., Miller, J. V., Pieper, S., and Kikinis, R., "3D Slicer as an image computing platform for the Quantitative Imaging Network," *Magnetic Resonance Imaging* **30**(9), 1323–1341 (2012). Quantitative Imaging in Cancer.