A Probabilistic Approach for the Simultaneous Mammogram Registration and Abnormality Detection

Mohamed Hachama, Agnès Desolneux, and Frédéric Richard*

MAP5, University Paris 5,
45, rue des Saints Pères,
75006 Paris, France
{hachama, desolneux, richard}@math-info.univ-paris5.fr

Abstract. In this paper, we present a new method for simultaneously registering mammograms and detecting abnormalities. We assume that pixels can be divided into two classes: normal tissue and abnormalities (lesions). We define the registration constraints as a mixture of two distributions which describe statistically image gray-level variations for both pixel classes. The two distributions are weighted at each pixel by the probability of abnormality presence. Using the Maximum A Posteriori, we estimate the registration transformation and the probability map of abnormality presence at the same time. We illustrate the properties of our technique with some experiments and compare it with some classical methods.

1 Introduction

Mammograms are often interpreted by comparing left and right breast images or different mammograms of a same patient. Mammogram comparisons help radiologists to identify abnormalities and determine their clinical significance [1]. In the CAD context, image comparisons are not straightforward. The registration of images must be carried out to compensate for some normal differences that can cause high false-negative rates in abnormality detection schemes [2].

Several researchers have used the subtraction of registered images as a comparative means by which to detect abnormalities [3]. The obtained asymmetry image is then thresholded to extract suspicious regions. Thus, the success of the detection task depends on the preliminary registration process.

On the other hand, the registration problem is usually expressed as a minimization of an energy composed of a regularization term and a similarity term. Usually, similarity criteria rely on some assumptions about gray-level dependencies between images [4], which are not valid in the presence of abnormalities. The registration can be improved by including in the model some knowledge about these abnormalities, as it was done in [5, 6, 7] and for the optical flow estimation in [8].

In this paper, we present a mixture-based technique where pixels are classified into a normal tissue class and an abnormality class. The registration constraints are then defined as a mixture of two distributions describing gray-level characteristics of the two classes and which are weighted at each pixel by the probability of abnormality presence. The main feature of our model is the possibility to combine image registration and the detection of abnormalities, so as to take proper advantage of the dependence between the two processes.

The mixture-based technique and its mathematical formulation are presented in Section 2. In Section 3, we illustrate the method behavior on some examples and compare it with some classical techniques.

2   Method

Let I and J be two images defined on a discrete grid Ω_d associated to $\Omega = [0,1]^2$ and called respectively source image and target image. Image coordinates are matched using transformations ϕ which map Ω_d into itself. We assume that lesions may be present in the images. Let L be the lesion map which associates to each pixel of Ω_d its probability to belong to a lesion in I or J. Assuming that all variables are realizations of some random fields, Bayes rule can be expressed as:

$$p(\phi, L|I, J) = \frac{p(I, J| L, \phi)\, p(\phi)\, p(L)}{p(I, J)}.$$

For the sake of simplicity, we have assumed in the above formula that the deformation ϕ and the lesion map L are independent (i.e. $p(\phi, L) = p(\phi)p(L)$). We can estimate the pair (ϕ, L) as the solution of the Maximum A Posteriori (MAP):

$$(\widetilde{\phi}, \widetilde{L}) = \arg\max_{(\phi, L)}\ p(I, J|\phi, L)\, p(\phi)\, p(L).$$

To ensure that the transformations remain smooth, we assume that they arise from the Gibbs distribution:

$$p(\phi) = \frac{1}{Cst}e^{-H_d(\phi)}\ , \tag{1}$$

where H_d is a discrete elasticity potential [9] (a continuous version is given by Equation (5)). We also assume that the lesion map arises from a Gibbs distribution:

$$p(L) = \frac{1}{Cst}e^{-R_d(L)}\ , \tag{2}$$

where R_d is a discrete energy of regularization. We use in this paper an energy restricting the amount of abnormal pixels in the images via a real parameter α_L:

$$R_d(L) = \alpha_L \sum_{x \in \Omega_d} L(x)\,.$$

More specific terms should be defined to describe the spatial configurations of each type of lesion. We will investigate the use of such energies in the future.

In order to define the likelihood $p(I, J|\phi, L)$, we assume that, given the transformation ϕ, the probability of the pair of images (I, J) depends only on the registered images $I_\phi = I \circ \phi$ and J, and that pixels are independent. Hence, we can write

$$p(I, J|\phi, L) = \prod_x p(I_\phi(x), J(x)|L(x)).$$

The probability of the pair $(I_\phi(x), J(x))$ depends on the class of the pixel x. Each class is characterized by a probability distribution, denoted by p_N for the normal tissue and p_L for the lesion. Thus, the probability distribution $p(I_\phi(x), J(x)|L(x))$ can be defined as a mixture of the two class distributions

$$p(I_\phi(x), J(x)|L(x)) = (1 - L(x))p_N(I_\phi(x), J(x)) + L(x)p_L(I_\phi(x), J(x)). \quad (3)$$

The normal tissue class. We assume that image differences generated by normal tissue have a discrete centered Gaussian distribution with variance σ^2:

$$p_N(I_\phi(x), J(x)) = \frac{1}{Cst} \exp(-\frac{\mid I_\phi(x) - J(x) \mid^2}{2\sigma^2}),$$

The lesion class. For the sake of simplicity, we assume that a lesion is present in the target image J. We simply characterize the lesion as a region which is brighter in the target image than it is in the source image. Hence, we get the following distribution

$$p_L(I_\phi(x), J(x)) = \begin{cases} 0 & \text{if } I_\phi(x) > J(x) \\ Cst & \text{otherwise,} \end{cases}$$

Numerical resolution

Up to now, we have formulated a Bayesian registration model in a discrete setting. We now transform the discrete model into a continuous model so as to be able to use variational resolution techniques. First, using the negative-log function, we rewrite the MAP estimate as an energy minimization problem. Then, we define a continuous version of the obtained energy by interpolating all functions by the finite element method and replacing sums on the pixel grid Ω_d by integrals on Ω. Thus, we have to minimize the energy:

$$E(\phi, L) = H(\phi) + R(L) - \int_\Omega \log(p(I_\phi(x), J(x))) \, dx, \quad (4)$$

where the probability distribution $p(I_\phi(x), J(x))$ is the obtained continuous version of the mixture distribution given by Equation (3). $H(\phi)$ is the elasticity potential defined as

$$\sum_{i,j=1,2} \int_\Omega [\lambda \frac{\partial u_i(x)}{\partial x_i} \frac{\partial u_j(x)}{\partial x_j} + \mu(\frac{\partial u_i(x)}{\partial x_j} + \frac{\partial u_j(x)}{\partial x_i})^2] dx, \quad (5)$$

where $u = \phi - id$, and λ and μ are the Lame elasticity constants.

The term $R(L)$ is the following energy:

$$R(L) = \alpha_L \int_\Omega L(x) \, dx \, .$$

As in [10, 6], we use a gradient descent algorithm on the energy E and finite elements to approximate solutions of the minimization problem.

3   Results

3.1   Experiment 1: Comparison Results

We illustrate the characteristics of this mixture-based technique by comparing its performance with two other registration techniques. The first one is the minimization the Sum of Square Differences (SSD). The second one is a registration technique proposed in [6], which is related to M-estimation in robust Statistics. We apply the algorithms to the pair of bilateral mammograms (21, 22) of the MIAS database [11], for which the target image contains an asymmetric density (bright circular region at the bottom of Image 1(b). Registrations obtained with the SSD and the M-estimation techniques tend to incorrectly match the lesion with the bright tissue in the source image and thus reduce image differences due to the lesion (Images 1(d),1(e)). This is corrected by the mixture-based technique where images are correctly registered while differences due to the lesion are preserved (Image 1(f)).

In order to test the detection performance of the mixture-based technique, we compare a lesion binary image obtained by thresholding the lesion map $\tilde{L}$, to the ones obtained with the SSD and the M-estimation methods by thresholding the images of differences. Figures 1(g)-(i) show the binary lesion images obtained with the three techniques for the same amount of abnormal pixels. We can notice that the mixture-based method reduces the number of false-positives.

For evaluating and comparing the three algorithms without the influence of a threshold value, we have presented on Figure 2 the FROC curves obtained with the three methods. The FROC curve plots the sensitivity (fraction of detected true positives, calculated by using the expert segmented image) as a function of the number of false positives. For the mixture-based technique, we have presented the FROC curve obtained with $\alpha_L = 0.1$. When using different values of the weight α_L, we have obtained similar FROC curves.

As observed on Figure 2, the FROC curve associated to the mixture-based method is the highest. So, the detection by the mixture-based technique is more sensitive. For instance, for 10000 false positive pixels (2% of image pixels), the detection rate grows from 0.632 for the SSD and 0.627 for the M-estimation based method, to 0.947 for the mixture based method.

3.2   Experiment 2: The Prior Lesion Term

In this experiment, we study the influence of the weight associated to the regularization term $R_L(L) = \alpha_L \int_\Omega L(x)$ of the lesion map L. We use the pair
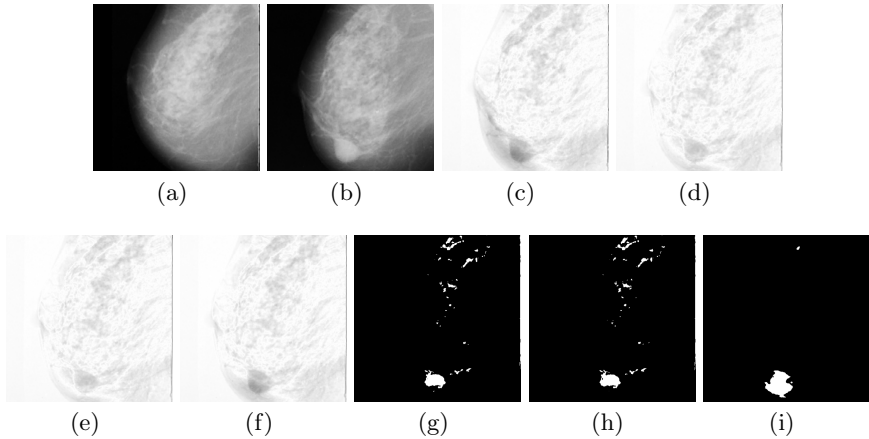
Fig. 1. Registration of bilateral mammograms. (a) Source image I, (b) Target image J, (c) The difference between the images before registration. The difference between the images after the registration using (d) the SSD method, (e) the M-estimation method, (f) the mixture-based method. Detection results containing 4180 pixels obtained with (g) the SSD method, (h) the M-estimation based method, (i) the mixture-based method.
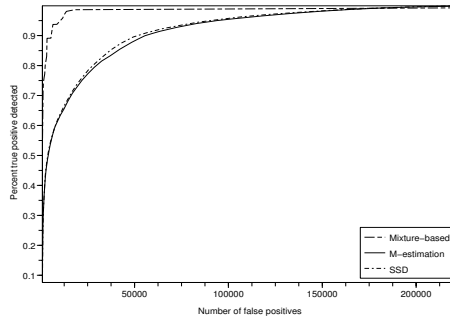

Fig. 2. FROC curves for the three detection methods

$(117 - 118)$ of the MIAS database (Images 3(a) and 3(b)) ; the images are segmented and the registration is initialized using a geometric approach based on the matching of the contours [2].

We have applied the mixture-based technique to the pair of images (3(c), 3(d)), using the lesion class distribution described in Section 2, for different values of the weight α_L. Results are presented in Figure 4.

As shown on Figure 4, we can use the weight α_L to limit the quantity of lesion pixels present in the image. High values of this weight restricts the amount of lesion pixels. However, the lesion map contains isolated pixels which, clearly, do not belong to any lesion. The use of the sum of the lesion map as the prior potential does not take this into account. More sophisticated terms should take
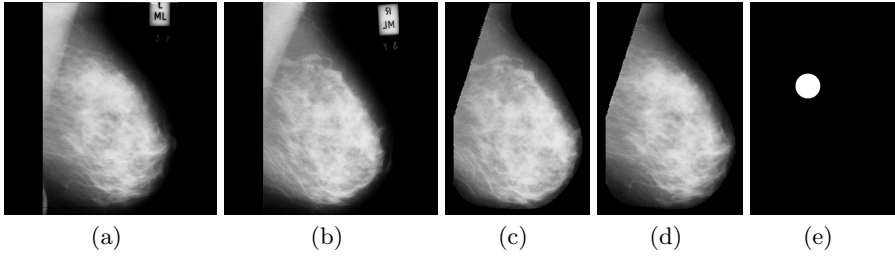
Fig. 3. Registration of bilateral mammograms. (a) Source image I, (b) Target image J, (c) the segmented and pre-registered source image, (d) the segmented target image, (e) the expert-segmented image.
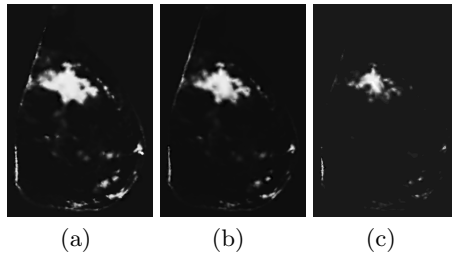

Fig. 4. The influence of the weight α_L associated to the prior on the lesion map. The lesion map obtained for (a) $\alpha_L = 0$, (b) $\alpha_L = 0.001$, (c) $\alpha_L = 0.01$.

into account the morphology of the lesion depending on its type: masses, calcifications, architectural distortions, ... In the future, we will investigate the design of prior terms adapted to each type of lesion.

3.3   Experiment 3: The Lesion Class Distribution

In this experiment, we use the same image pair of Figure 3. We apply the mixture-based method with different lesion class distributions.

First example. If we have no information about the photometric characteristics of the lesion, we should use an uniform distribution:

$$p_{\mathrm{L}}(I_\phi(x), J(x)) = \frac{1}{Cst}.$$

Second example. As explained in Section 2, one can also suppose that a lesion is just a region in one image that is more bright that its correspondent in the second image. If we assume that the lesion is present in the target image, we get:

$$p_{\mathrm{L}}(I_\phi(x), J(x)) = \begin{cases} 0 & \text{if } I_\phi(x) > J(x) \\ Cst & \text{otherwise,} \end{cases}$$

Third example. If we have more precise information about the gray-level values of lesion pixels, we can use a probability distribution of the form:

$$p_{\mathrm{L}}(I_\phi(x), J(x)) = \begin{cases} 0 & \text{if } I_\phi(x) > J(x) \\ \frac{1}{Cst} \exp(-\frac{(J(x)-m)^2}{2\sigma^2}), & \text{otherwise,} \end{cases}$$

where m is the mean value of the lesion brightness and σ its standard deviation. In this experiment, m and σ are determined using the images and the expert-segmented lesion image ($m = 215$ and $\sigma = 5$). More generally, one can use a full database to estimate these parameters. Detection results (lesion maps) obtained with these three terms are shown on Figure 5.
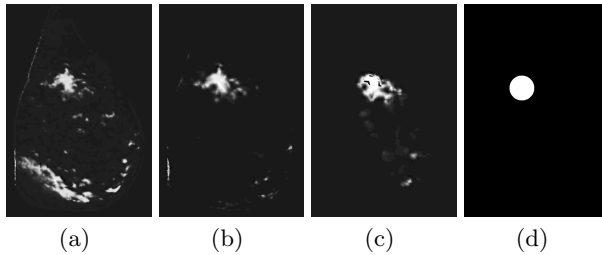

(a)   (b)   (c)   (d)

Fig. 5. Detection results with different lesion class distributions. Lesion map obtained with: (a) the first model, (b) the second model, (c) the third model. (d) Expert-segmented lesion.

As shown on Figure 5(a), when using an uniform model (which corresponds to the case when the lesion can be present either in the target image or in the source one), the algorithm tends to consider all asymmetric regions as lesions. The detection results are improved by using more information. If we suppose that the lesion is present in the target image, we can use the second distribution which produces better results. In practice, this is the case for the detection of the apparition, or change, of a lesion in a temporal sequence. In the third case, we have more precise information about the gray-level values of lesion pixels in the form of a Gaussian distribution with a known mean value and standard-deviation. With the third distribution, we get the best detection results: false positive are reduced and the lesion map is concentrated on the real lesion. In practice, one can estimate the parameters of the Gaussian distribution from a database.

4   Conclusion

We have presented a method for simultaneously registering mammograms and detecting abnormalities. Thanks to a combined approach, the mixture-based method improves the mammogram registration and reduces the false-positives rate for the lesion detection. In the future, we will focus on the design of lesion

models for different types of lesions, and the estimation of the distribution parameters for both lesion and normal tissue classes. Furthermore, we will test the mixture-based method on a mammogram database.

References

1. L. Tabar and P. Dean, *Teaching atlas of mammography*, Thieme Inc., Stuttgart, 1985.
2. C. Graffigne F. Richard, "An image matching problem for the registration of temporel or bilateral mammogram pairs," *Proc. of the 5th International Workshop on Digital Mammography (toronto, Canada)*, june 2000.
3. M.Y. Sallam and K. Bowyer, "Registration and difference analysis of corresponding mammogram images," *Medical Image Analysis*, vol. 3, no. 2, pp. 103–118, 1999.
4. A. Roche, G. Malandain, and N. Ayache, "Unifying Maximum Likelihood Approaches in Medical Image Registration," *International Journal of Computer Vision of Imaging Systems and Technology*, vol. 11, pp. 71–80, 2000.
5. D. Hasler, L. Sbaiz, S. Susstrunk, and M. Vetterli, "Outlier modeling in image matching," *IEEE Trans. on Patt. Anal. and Match. Intell.*, pp. 301–315, 2003.
6. F. Richard, "A new approach for the registration of images with inconsistent differences," in *Proc. of the Int. Conf. on Pattern Recognition, ICPR*, Cambridge, UK, 2004, vol. 4, pp. 649–652.
7. M. Hachama, F. Richard, and A. Desolneux, "A mammogram registration technique dealing with outliers," in *IEEE International Symposium on Biomedical Imaging ISBI'06, Arlington, Virginia, USA*, April 2006.
8. A. Jepson and M.J. Black, "Mixture models for optical flow computation," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 1993, pp. 760–761.
9. F. Richard, "A comparative study of markovian and variational image-matching techniques in application to mammograms," *Pattern Recognition Letters*, vol. 26, no. 12, pp. 1819–1829, 2005.
10. F. Richard and L. Cohen, "Non-rigid image registration with free boundary constraintes: application to mammography," *Journal of Computer Vision and Image Understanding*, vol. 89(2), pp. 166–196, 2003.
11. J. Suckling, J. Parker, and D. Dance, "The MIAS digital mammogram database," *In Proc. of the 2nd Int. Workshop on Digital Mammography, England*, july 1994.