# A MAMMOGRAM REGISTRATION TECHNIQUE DEALING WITH OUTLIERS

*Mohamed Hachama, Frédéric Richard, and Agnès Desolneux*

MAP5, UMR CNRS 8145, University Paris Descartes.
{Hachama, Richard, Desolneux}@math-info.univ-paris5.fr

## ABSTRACT

In this paper, we present a new method for registering images in presence of abnormalities. By abnormalities, we mean variations of image intensity which are due to pathologies and cannot be corrected by registration. Our approach consists of characterizing them as outliers. This characterization is obtained in a Bayesian framework, by defining registration constraints as mixtures of distributions which describe statistically image gray-level variations on both inlier and outlier pixels. Thanks to an outlier map weighting these mixture distributions, we can also take proper advantage of some prior knowledge about the lesion location. We use synthetic images and mammograms to illustrate the properties of the method and to compare it with some classical ones.

## 1. INTRODUCTION

Image Registration is an active topic of research, which is being motivated by a wide variety of biomedical applications [1]. The registration of bilateral mammogram pairs or temporal sequences has been used to characterize breast asymmetry or evolution, and to detect lesions [2, 3].

Given a source image $I$ and a target image $J$, the registration problem is usually formulated in terms of an inverse problem [4, 5, 3] which consists of finding an image coordinate change $\phi$ minimizing an energy $E$ of the form:

$$E(\phi) = R(\phi) + S(I, J, \phi). \quad (1)$$

The energy $E$ is composed of a regularization term $R$ and a similarity term $S$. The later plays the role of a registration constraint, being as low as the geometric deformation $I \circ \phi$ of $I$ and the target image $J$ are "similar".

The definition of the similarity criterion $S$ relies on the nature of gray-level dependencies between images. For instance, the Sum of Square Differences (SSD) is a common criterion which is used whenever gray-level values are approximately the same in images to be registered. Thanks to recent works on criteria inspired from Information Theory [6], models can now be adapted to register sets of images having a wide variety of dependencies. However, the adequacy of the criterion depends on the validity of some assumptions about the gray-level dependencies which may not always be satisfied. In medical applications, such situations currently arise when pathologies are present in images. For instance, breast cancers often appear in mammograms as asymmetries between left and right breasts (e.g. unilateral bright regions) [7]. Such outliers are not consistent with gray-level dependency assumptions. They may distort registration constraints and cause registration errors.

In [8], F. Richard proposed a registration technique which takes into account outliers, by down-weighting their influence in the computation of registration constraints. The similarity criterion used in [8] is related to M-estimation criteria of robust statistics [9] which were also applied for the optical flow computation [10]. A more general approach is the use of mixture models, as it was already done in the framework of optical flow estimation by Jepson and Black [11] and in the framework of image registration by Hasler et al. [12].

In this paper, we present a mixture-based technique for the registration of mammograms. The main feature of our model is the use of an outlier map which weights the mixture distributions components at each pixel, and thus enables us to take proper advantage of some knowledge about the lesion location (in [11], the weights are pixel independent). Furthermore, we define the mixture model for pairs of image intensities and not only for their differences as done in [12, 8].

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

## 2. THE MIXTURE-BASED TECHNIQUE

### 2.1. Bayesian formulation

Let $I$ and $J$ be two images of the same size $(M, N)$, having gray-level values in the set $\{0, ..., 255\}$ and defined on a discrete grid $\Omega_d = \{(\frac{i}{M-1}, \frac{j}{N-1}), (i, j) \in \{0, ..., M-1\} \times \{0, ..., N-1\}\}$ associated to $\Omega = [0, 1]^2$. Image coordinates are matched using deformations $\phi$ which map $\Omega_d$ into itself. Registering $I$ and $J$ means: find a function $\phi$ such that the deformed image $I_\phi = I \circ \phi$ is similar to the target image $J$.

Our formulation follows the Bayesian framework for im-

This work was supported by the grant "ACI Young Researchers 2003 (French Ministry of Research)" No. 9060.

age analysis laid out in [13]. Assuming that images and transformations are realizations of some random fields, Bayes rule can be expressed as

$$p(\phi|I,J)\ \alpha\ p(I,J|\phi)\,p(\phi). \tag{2}$$

The Bayes rule allows us to write the posterior distribution $p(\phi|I,J)$ in terms of the prior $p(\phi)$ and of the likelihood $p(I,J|\phi)$. To ensure that the transformations remain smooth, we assume that they arise from the Gibbs distribution:

$$p(\phi) = \frac{1}{Z} e^{-H_d(\phi)}\ , \tag{3}$$

where $Z$ is a normalization constant, and $H_d$ is a discrete elasticity potential [14]. We can estimate the transformation $\phi$ as the solution of the Maximum A Posteriori (MAP), which is the value of $\phi$ maximizing $p(\phi|I,J)$:

$$\widetilde{\phi} = \arg\max_{\phi}\ (\,p(I,J|\phi)\ p(\phi)\,). \tag{4}$$

### 2.2. Outlier modeling

We assume that outliers are present in the images (lesions in mammograms). We define an outlier map $L(x)$ which associates to each pixel $x$ of $\Omega_d$ its probability to be an outlier. In practice, the map $L$ can be a preliminary estimation of outlier positions which may be computed from results of some filtering operations on the images $I$ and $J$. In the mammography literature, there is a wide choice of lesion detection and enhancement filters that we plan to use for the computation of $L$. However, in this work, we will either assume that $L(x)$ is known (lesion segmented by an expert) or random so as to simulate the case when no prior information about outlier locations is available.

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

$$p(I,J|\phi) = \Pi_x p(I_\phi(x), J(x)). \tag{5}$$

The probability of the pair $(I_\phi(x), J(x))$ depends on the class of the pixel $x$ (inlier vs outlier). Both classes are characterized by a probability distribution, denoted by $p_{\rm in}$ for the inlier class and $p_{\rm out}$ for the outlier class. The probability $p(I_\phi(x), J(x))$ is defined as a mixture of the two class distributions

$$\begin{aligned} &p(I_\phi(x), J(x)) = \\ &\quad (1-L(x))p_{\rm in}(I_\phi(x), J(x)) + L(x)p_{\rm out}(I_\phi(x), J(x)). \end{aligned} \tag{6}$$

In this mixture model, the value of the outlier map $L$ at location $x$ is used to weight both class distributions. Let us now give further details about distributions $p_{\rm in}$ and $p_{\rm out}$.

**The inlier model.** In an ideal situation, gray-level values of registered images should be exactly the same at inlier positions. But, in practice, these gray-level values usually differ because of acquisition noise, and possible slight misalignments. Assuming that these variations at inlier positions have a discrete gaussian distribution with mean 0 and variance $\sigma^2$, we can define $p_{\rm in}$ as

$$p_{\rm in}(I_\phi(x), J(x)) = \frac{1}{\rm Cst}\ \exp(-\frac{|\ I_\phi(x) - J(x)\ |^2}{2\sigma^2}\ ), \tag{7}$$

**The outlier model.** The definition of the outlier distribution is a difficult task. Outliers are located at pixels where images differ significantly, even if images are perfectly registered. In biomedical context, such a situation typically occurs when a lesion is present in an image and not in the other one. If we have no specific information about the outlier distribution, we can assume that an outlier point of an image can be associated to any pixel of the other image. In this case, we may assume that $p_{\rm out}$ is an uniform distribution:

$$p_{\rm out}(I_\phi(x), J(x)) = \frac{1}{(256)^2}.$$

In some cases, lesions are seen as almost uniform areas (e.g. mass lesions). If such a lesion is present in the target image $J$, we can use a distribution depending only on the target, which is of the form:

$$p_{\rm out}(I_\phi(x), J(x)) = \frac{1}{\rm Cst}\exp(-\frac{(J(x)-m)^2}{2\sigma'^2}), \tag{8}$$

where $m$ is the lesion mean gray-level, and $\sigma'$ its standard deviation.

### 2.3. Numerical resolution

Up to now, we have formulated a Bayesian registration model in a discrete setting in order to avoid the difficulty of defining probability distributions on infinite dimensional or non linear spaces [13]. Now, we transform formally the discrete model into a continuous model so as to be able to use variational resolution techniques. First, we rewrite the MAP estimate (Equation (4)) as the minimization of the negative-log function

$$E_d(\phi) = -\log(p(\phi)) - \log(p(I,J|\phi)).$$

Then, using the decomposition in Equation (5) and the Gibbs distribution in Equation (3), we get

$$E_d(\phi) = H_d(\phi) - \sum_{x\in\Omega_d} \log(p(I_\phi(x), J(x))) + K,$$

where $K$ is a constant value. Next, following approaches in [5, 14], we define a continuous expression of this energy, by interpolating images and transformations with finite elements and replacing sums on the pixel grid $\Omega_d$ by integrals on $\Omega$:

$$E(\phi) = H(\phi) - \int_\Omega \log(p(I_\phi(x), J(x)))\ dx, \tag{9}$$

where the probability distribution $p(I_\phi(x), J(x))$ is given by the mixture distribution in Equation (6). $H(\phi)$ is the elasticity potential defined as

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

where $u = \phi - id$, and $\lambda$ and $\mu$ are the Lame elasticity constants. As in [3, 8], we use a gradient descent algorithm on the energy $E$ and finite elements to approximate solutions of the minimization problem.

## 3. RESULTS

In this section, we illustrate the characteristics of the mixture model, by comparing its registration results with those of an usual SSD-based technique and the M-estimator-based technique proposed in [8]. Note that the SSD technique is a particular case of the mixture-based technique, obtained when $L(x) = 0, \forall x$ (no outliers).

### 3.1. Illustration on synthetic images

In the first experiment, the study image is composed of four different constant regions. To generate the target image, the study was deformed and a white square was added at the center, so as to represent outliers. For the application of the mixture-based technique, we used for $L$ a random uniform map (i.e. the $L(x)$'s are independent and uniformly distributed on $[0, 1]$), a gaussian model centered at $0$ with a standard deviation equal to $15$ for the inlier distribution, and a gaussian model centered at $255$ with a standard deviation equal to $2$ (Equation (8)) for the outlier distribution. Figure 1 shows the results obtained with the three techniques.

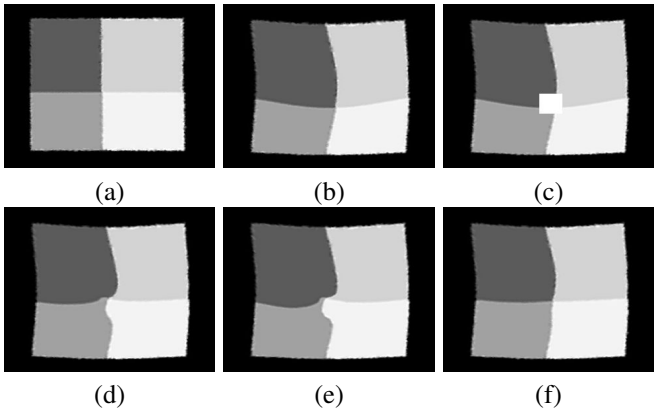


**Fig. 1**. A synthetic example. (a) Source image $I$, (b) Target image without outlier, (c) Target image $J$. Deformation of the source onto the target using (d) the SSD technique, (e) the M-estimation technique, (f) the mixture-based technique.

Using the SSD technique and the M-estimation technique, the source image is not correctly deformed at the center because of the presence of outliers. The algorithm tends to match brighter pixels of the source image and the white square in the target, resulting in a deterioration of the deformation in the center. This is corrected using the mixture-based technique which does not take into account the error generated by the white square. We think that, despite the use of a random lesion map, images are here correctly registered for mainly two reasons: (a) the outlier model $p_{\text{out}}$ is perfectly adapted to the outliers of this image ; (b) since deformations are smooth, a random lesion map has the same effect as a constant lesion map.

### 3.2. Mammogram registration

In the second experiment (Figure 2), we applied the algorithms to a pair of bilateral mammograms (case 21 of the MIAS database [15]), for which the target image contains a lesion (bright circular region at the bottom of Image 2 (b)). For the mixture-based technique, we used a centered gaussian model with a standard deviation equal to $15$ for the inlier distribution, an uniform distribution for the outliers and a binary lesion map (Figure 3 (a)) containing the exact location of the lesion. Results of the application of the three techniques are shown on Figure 2.

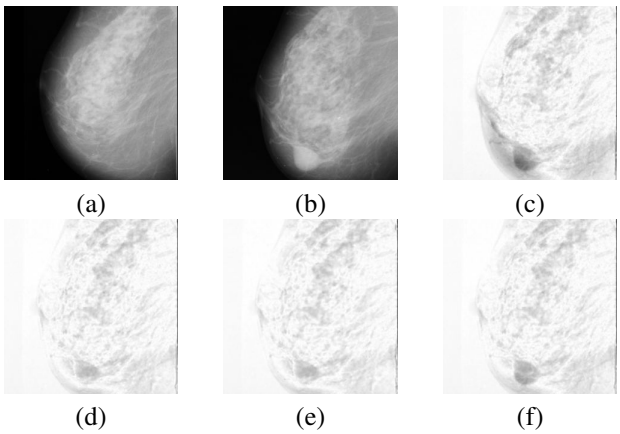


**Fig. 2**. Registration of bilateral mammograms. (a) Source image $I$, (b) Target image $J$, (c) Differences between images $I$ and $J$ before registration. Difference between registered images with (d) the SSD technique, (e) the M-estimation technique, (f) the mixture-based technique.

Registrations obtained with the SSD technique and with the M-estimation one tend to incorrectly match the lesion with the bright tissues in the source image and thus reduce image differences due to the lesion (Images 2 (d)-(e)). When using the mixture-based technique, images are correctly registered while differences due to the lesion are preserved (Image 2(f)).

To test the robustness of the mixture-based technique, we

did the same experiments with lesion maps corrupted by noise. Results are shown on Figure 3.

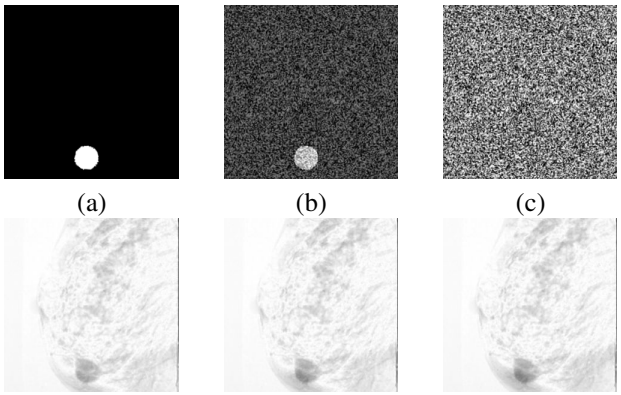

**Fig. 3**. Study of the robustness. (a) Binary lesion map, (b) Lesion map corrupted with a uniform noise of parameter $0.5$, (c) Random lesion map, (d) Result obtained with (a), (e) Result obtained with (b), (f) Result obtained with (c).

As observed on Figures 3(d) to (f), the registration by the mixture-based technique is robust to the noise that may corrupt the outlier map. With a noisy lesion map (b), the registration is somewhat altered (near the lesion) but differences due to the lesion are not reduced. With a fully random map, the registration remains correct on the lesion but becomes less accurate in the other parts of the image.

## 4. CONCLUSION

In this paper, we have presented a method for mammogram registration dealing with outliers. With some experiments, we have demonstrated that it improves the mammogram registration in the presence of lesions. Thanks to an outlier map, the new method allows us to take into account spatial and gray-level information about lesions which may be present in images. In the future, we will focus on how to design outlier maps for the different types of lesions, by adapting methods designed for the detection and enhancement of lesion in single image. Furthermore, we will adapt the model presented here to combine the estimation of the lesion map and of the deformation at the same time.

## 5. REFERENCES


[1] J. Maintz and M. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2(1), pp. 1–36, 1998.

[2] 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.

[3] 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.

[4] R. Bajcsy and M. Kovacic, "Multiresolution elastic matching," *Computer Vision Graphics and Image Processing*, vol. 46, pp. 1–21, 1989.

[5] M.I. Miller, G. Christensen, Y. Amit, and U. Grenander, "Mathematical textbook of deformable neuroanatomies," in *Proc. Natl. Acad. Sc.*, USA, 1993, pp. 11944–11948.

[6] 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.

[7] L. Tabar and P. Dean, *Teaching atlas of mammography*, Thieme Inc., Stuttgart, 1985.

[8] 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.

[9] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions*, Wiley, New York, 1986.

[10] N. Sebe and M.S. Lew, *Robust Computer Vision: Theory and Applications*, vol. 26, Series: Computational Imaging and Vision, 2003.

[11] 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.

[12] 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.

[13] D. Mumford, "The bayesian rationale for energy functionals," *Geometry-driven Diffusion in Comp. Vis., Kluwer Academic Publ*, vol. 46, pp. 141–153, 1994.

[14] F. Richard, "A comparative study of markovian and variational image-matching techniques in application to mammogram," *Pattern Recognition Letters*, vol. 26, no. 12, pp. 1819–1829, 2005.

[15] 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.