

The mixtures of Student's t -distributions as a robust framework for rigid registration

Demetrios Gerogiannis¹ Christophoros Nikou* Aristidis Likas

*University of Ioannina,
Department of Computer Science,
PO Box 1185, 45110 Ioannina, Greece*

Abstract

The problem of registering images or point sets is addressed. At first, a pixel similarity-based algorithm for the rigid registration between single and multimodal images is presented. The images may present dissimilarities due to noise, missing data or outlying measures. The method relies on the partitioning of a reference image by a Student's t -mixture model (SMM). This partition is then projected onto the image to be registered. The main idea is that a t -component in the reference image corresponds to a t -component in the image to be registered. If the images are correctly registered the distances between the corresponding components is minimized. Moreover, the extension of the method to the registration of point clouds is also proposed. The use of SMM components is justified by the property that they have heavier tails than standard Gaussians, thus providing robustness to outliers. Experimental results indicate that, even in the case of low SNR or important amount of dissimilarities due to temporal changes, the proposed algorithm compares favorably to the mutual information method for image registration and to the Iterative Closest Points (ICP) algorithm for the alignment of point sets.

Key words: image registration, point set registration, Gaussian mixture model, mixtures of Student's t -distribution, Expectation-Maximization (EM) algorithm.

1 Introduction

The goal of image registration is to geometrically align two or more images in order to superimpose pixels representing the same underlying structure. Image reg-

* asterisk indicates corresponding author

Email addresses: dgerogia@cs.uoi.gr (Demetrios Gerogiannis),
cnikou@cs.uoi.gr (Christophoros Nikou), arly@cs.uoi.gr (Aristidis Likas).

¹ This work was partially supported by Interreg IIIA (Greece-Italy) grant I2101005.

istration is an important preliminary step in many application fields involving, for instance, the detection of changes in temporal image sequences or the fusion of multimodal images. For the state of the art of registration methods we refer the reader to [43]. Medical imaging, with its wide variety of sensors (MRI, nuclear, ultrasonic, X-Ray) is probably one of the first application fields [24,1,15]. Other research areas related to image registration are remote sensing, multisensor robot vision and multisource imaging used in the preservation of artistic patrimony. Respective applications include the following of the evolution of pathologies in medical image sequences [28], the detection of changes in urban development from aerial photographs [20] and the recovery of underpaintings from visible/X-ray pairs of images in fine arts painting analysis [16].

The overwhelming majority of change detection or data fusion algorithms assume that the images to be compared are perfectly registered. Even slightly erroneous registrations may become an important source of interpretation errors when inter-image changes have to be detected. Accurate (i.e. subpixel or subvoxel) registration of single modal images remains an intricate problem when gross dissimilarities are observed. The problem is even more difficult for multimodal images, showing both localized changes that have to be detected and an overall difference due to the variety of responses by multiple sensors.

Since the seminal works of Viola and Wells [41] and Maes *et al.* [23], the maximization of the mutual information measure between a pair of images has gained an increasing popularity as a criterion for image registration [31]. The estimation of both marginal and joint probability density functions of the involved images is a key element in mutual information based image alignment. However, this method is limited by the histogram binning problem. Approaches to overcome this limitation include Parzen windowing [41,19], where we have the problem of kernel width specification, and spline approximation [39,25]. A recently proposed method relies on the continuous representation of the image function and develops a relation between image intensities and image gradients along the level sets of the respective intensity [33].

Gaussian mixture modeling (GMM) [5,26] constitutes a powerful and flexible method for probabilistic data clustering that is based on the assumption that the data of each cluster has been generated by the same Gaussian component. In [22], GMMs were trained off-line to provide prior information on the expected joint histogram when the images are correctly registered. GMMs have also been successfully used as models for the joint [14] as well as the marginal image densities [17], in order to perform intensity correction. They have also been applied in the registration of point sets [21] without establishing explicit correspondence between points in the two images. The parameters of GMMs can be estimated very efficiently through maximum likelihood (ML) estimation using the EM algorithm [8]. Furthermore, it is well-known that GMMs are capable of modeling a large variety of pdfs [26].

An important issue in image registration is the existence of outlying data due to temporal changes (e.g. urban development in satellite images, lesion evolution in medical images) or even the complimentary but non redundant information in pairs of multimodal images (e.g. visible and infrared data, functional and anatomical medical images). Although a large variety of image registration methods have been proposed in the literature only a few techniques address these cases [18,28,36].

The method proposed in this study is based on mixture model training. More specifically, we train a mixture model once for the reference image and obtain the corresponding partitioning of image pixels into clusters. Each cluster is represented by the parameters of the corresponding density component. The main idea is that a component in the reference image corresponds to a component in the image to be registered. If the images are correctly registered the sum of distances between the corresponding components is minimum.

A straightforward implementation of the above idea would consider models with Gaussian components. However, it is well known that GMMs are sensitive to outliers and may lead to excessive sensitivity when the number of data points is small. This is easily understood by recalling that maximization of the likelihood function under an assumed Gaussian distribution is equivalent to finding the least-squares solution which lacks robustness. Consequently, a GMM tends to over-estimate the number of clusters since it uses additional components to capture the tails of the distributions [4]. The problem of attaining robustness against outliers in multivariate data is difficult and increases with the dimensionality. In this paper, we consider mixture models (SMM) with Student's- t components for image registration. This pdf has heavier tails compared to a Gaussian [29]. More specifically, each component in the SMM mixture originates from a wider class of elliptically symmetric distributions with an additional parameter called the number of degrees of freedom. In this way, a more robust mixture model is employed than the typical GMM.

The main contributions of the proposed registration method are the following: (i) the histogram binning problem is overcome through image modeling with mixtures of distributions which provide a continuous representation of image density. (ii) Robustness to outlying pixel values is achieved by using mixtures of Student's t -distributions. The widely used method of maximization of the mutual information is outperformed. (iii) The method may be directly applied to vector valued images (e.g. diffusion tensor MRI) where standard histogram-based methods fail due to *the curse of dimensionality*. (iv) The proposed method is faster than histogram based methods where the joint histogram needs to be computed for every change in the transformation parameters.

Moreover, the registration problem is extended to the case of point sets where the nature of the problem is different since there is no spatial ordering contrary to image grids (e.g. pixelized images). Therefore, the difficulty consists in simultaneously estimating the transformation parameters and establishing correspondences

between points.

In the related literature of point set registration, the standard approach is the well known Iterative Closest Points (ICP) algorithm [3] and its variants [37,13,8,30]. In [9,10] a robust point matching algorithm is proposed relying on *soft-assign* [34] and an iterative optimization procedure. The *soft-assign* is based on a matrix whose entries describe the probability that a point of one set matches upon transformation to one of the other set. Mutual information was also used as a constraint [35] for point set matching under the above framework. Features extracted from the point sets are employed in [2,40], a kernel-based method is used in [38] and a method modeling the point sets by a GMM with constraints on the component centers is presented in [27]. Also, an approach to the construction of an atlas from multiple point sets is proposed in [42]. Finally, a work related to the herein proposed approach is presented in [21]. The authors propose to model the probability density function (pdf) of the points of the two sets by GMMs and estimate the transformation parameters through the minimization of an energy function describing the distance of the two GMMs. Our model completes this study by proposing a more robust framework for modeling the point sets.

The remainder of this paper is organized as follows. In section 2, the image registration method as a problem of minimizing distances between mixture models is presented. ML estimation of the parameters of a Student's t -mixture model and the generalization of the image registration method using SMMs are described in section 3, while the extension of the algorithm to the registration of point sets is described in section 4. Experimental results and comparison with the state of the art image registration method of maximization of the mutual information (MI) are provided in section 5. Results on the registration of point sets are also presented in this section. Finally, conclusions are drawn in section 6.

2 Image registration by minimization of the distance between mixture models

Let I_{ref} be an image of $N \times N$ pixels with intensities denoted as $I_{ref}(x^i)$, where x^i , $i = 1, \dots, N^2$, is the i^{th} pixel index. The purpose of rigid image registration is to estimate a set of parameters \mathcal{S} of the rigid transformation $T_{\mathcal{S}}$ minimizing a cost function $E(I_{ref}(\cdot), I_{reg}(T_{\mathcal{S}}(\cdot)))$ that, in a similarity metric-based context, expresses the similarity between the image pair. In the 2D case the rigid transformation parameters are the rotation angle and the translation parameters along the two axes. In the 3D case, there are three rotation and three translation parameters. Eventually, scale factors may also be included, depending on the definition of the transformation.

Consider, now, a partitioning of the reference image I_{ref} into K clusters (groups)

by training a mixture model with K components with arbitrary pdf $p(I(x); \Theta)$:

$$\phi(I_{ref}(x)) = \sum_{k=1}^K \pi_k p(I_{ref}(x); \Theta_k^{ref})$$

Therefore, the reference image is represented by the parameters Θ_k^{ref} , $k = 1, \dots, K$ of the mixture components. The partitioning of the image is described using the function $f(x) : [1, 2, \dots, N] \times [1, 2, \dots, N] \rightarrow \{1, 2, \dots, K\}$, where $f(x) = k$ means that pixel x of the reference image I_{ref} belongs to the cluster defined by the k^{th} component. Let us also define the sets of all pixels of image I_{ref} belonging to the k^{th} cluster:

$$P_k = \{x^i \in I_{ref}, i = 1, 2, \dots, N^2 | \delta(f(x^i) - k) = 1\}$$

for $k = 1, \dots, K$, where $\delta(x)$ is the Dirac function:

$$\delta(f(x^i) - k) = \begin{cases} 1, & \text{if } f(x^i) = k \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

The above mixture-based segmentation of the reference image is performed once, at the beginning of the registration procedure. The reference image I_{ref} is, thus, partitioned into K groups, generally, not corresponding to connected components in the image. This spatial partition is projected on the image to be registered I_{reg} , yielding a corresponding partition of this second image (i.e., the partitioning of the reference image acts as a mask on the image to be registered). Then, we assume that the pixel values of each cluster k in I_{reg} are modeled using a mixture component with parameters Θ_k^{reg} obtained from the statistics of the intensities of pixels in group k of I_{reg} .

In order to apply our method it should be possible to define a distance measure $D(\Theta_k^{ref}, \Theta_k^{reg})$ between the corresponding mixture components with pdf $p(I)$. Then the energy function we propose, is expressed by the weighted sum of distances between the corresponding components in I_{reg} and I_{ref} :

$$E(I_{ref}(\cdot), I_{reg}(T_S(\cdot))) = \sum_{k=1}^K \pi_k D(\Theta_k^{ref}, \Theta_k^{reg}) \quad (2)$$

where π_k is the mixing proportion of the k^{th} component:

$$\pi_k = \frac{|P_k|}{\sum_{l=1}^K |P_l|}$$

where $|P_k|$ denotes the cardinality of set P_k . If the two images are correctly registered the criterion in (2) assumes that the total distance between the whole set of

components would be minimum.

For a given set of transformation parameters \mathcal{S} , the total energy between the image pair is computed through the following steps:

- segment the reference image $I_{ref}(\cdot)$ into K clusters by a mixture model.
- for each cluster $k = 1, 2, \dots, K$ of the reference image:
 - project the pixels of the cluster onto the transformed image to be registered $I_{reg}(T_{\mathcal{S}}(\cdot))$.
 - determine the parameters Θ_k^{reg} of the projected partition of I_{reg} .
- evaluate the energy in eq. (2) by computing the distances between the corresponding densities.

In the case of GMMs, the above registration procedure can be applied as follows:

Consider the multivariate normal distributions $N_1(\mu_1, \Sigma_1)$ and $N_2(\mu_2, \Sigma_2)$ and denote $\Theta_i = \{\mu_i, \Sigma_i\}$, with $i = \{1, 2\}$, their respective parameters (mean vector and covariance matrix). The Chernoff distance between these distributions is defined as [12]:

$$C(\Theta_1, \Theta_2, s) = \frac{s(1-s)}{2} (\mu_2 - \mu_1)^T [s\Sigma_1 + (1-s)\Sigma_2]^{-1} (\mu_2 - \mu_1) + \frac{1}{2} \ln \left(\frac{|s\Sigma_1 + (1-s)\Sigma_2|}{|\Sigma_1|^s |\Sigma_2|^{1-s}} \right).$$

The Bhattacharyya distance is a special case of the Chernoff distance with $s = 0.5$:

$$B(\Theta_1, \Theta_2) = \frac{1}{8} (\mu_2 - \mu_1)^T \left[\frac{\Sigma_1 + \Sigma_2}{2} \right]^{-1} (\mu_2 - \mu_1) + \frac{1}{2} \ln \left(\frac{|\frac{\Sigma_1 + \Sigma_2}{2}|}{\sqrt{|\Sigma_1| |\Sigma_2|}} \right) \quad (3)$$

A representative GMM for the reference image can be obtained via the EM algorithm [5]. Therefore, the reference image is represented by the parameters $\Theta_k^{ref} = \{\mu_k^{ref}, \Sigma_k^{ref}\}$, $k = 1, \dots, K$ of the GMM components. After projecting the pixel groups of the reference image to obtain the corresponding groups in the registered image, the parameters Θ_k^{reg} can be estimated by taking the sample mean μ_k^{reg} and the sample covariance matrix Σ_k^{reg} :

$$\mu_k^{reg} = \frac{1}{|P_k|} \sum_{i=1}^{N^2} I_{reg}(T_{\mathcal{S}}(x^i)) \delta(f(x^i) - k) \quad (4)$$

and

$$\Sigma_k^{reg} = \frac{1}{|P_k|} \sum_{i=1}^{N^2} (\Delta I_k^i) (\Delta I_k^i)^T \delta(f(x^i) - k), \quad (5)$$

where $\Delta I_k^i = I_{reg}(T_S(x^i)) - \mu_k^{reg}$. The role of $\delta(f(x^i) - k)$ in eq. (4) and (5) is to determine the support (the pixel coordinates) for the calculation of the mean and covariance. These parameters are computed *on the image to be registered* for the pixel coordinates belonging to the k^{th} group *on the reference image*. This also implies a Gaussian mixture model for the components of I_{reg} . The total distance between the two images is computed using eq. (2), where the Bhattacharyya distance between the corresponding Gaussian components is considered as distance measure D .

However, GMMs are very sensitive to outlying data and their outcome is largely influence by the presence of pixels not belonging to the dominating model. In order to overcome this drawback of GMMs, we have employed in our registration method mixtures of Student's t -distributions. These mixtures are more robust to outliers as it is described in the next section.

3 Robust image registration with mixtures of Student's t -distributions

In what follows, we briefly present the properties of mixtures of Student's t -distributions (SMMs), as well as the ML estimation of their parameters using the EM algorithm. Then, we describe how SMMs can be employed as mixture models in the the general registration approach presented in the previous section.

3.1 ML estimation of mixtures of Student's t -distributions

A d -dimensional random variable X that follows a multivariate t -distribution with mean μ , positive definite, symmetric and real $d \times d$ covariance matrix Σ and has $\nu \in [0, \infty)$ degrees of freedom has a density expressed by:

$$p(x; \mu, \Sigma, \nu) = \frac{\Gamma\left(\frac{\nu+d}{2}\right) |\Sigma|^{-\frac{1}{2}}}{(\pi\nu)^{\frac{d}{2}} \Gamma\left(\frac{\nu}{2}\right) [1 + \nu^{-1}\delta(x, \mu; \Sigma)]^{\frac{\nu+d}{2}}} \quad (6)$$

where $\delta(x, \mu; \Sigma) = (x - \mu)^T \Sigma^{-1} (x - \mu)$ is the Mahalanobis squared distance and Γ is the Gamma function.

It can be shown that the Student's t distribution is equivalent to a Gaussian distribution with a stochastic covariance matrix. In other words, given a weight u following a Gamma distribution parameterized by ν :

$$u \sim \text{Gamma}(\nu/2, \nu/2). \quad (7)$$

the variable X has the multivariate normal distribution with mean μ and covariance

Σ/u :

$$X|\mu, \Sigma, \nu, u \sim N(\mu, \Sigma/u), \quad (8)$$

It can be shown that for $\nu \rightarrow \infty$ the Student's t -distribution tends to a Gaussian distribution with covariance Σ . Also, if $\nu > 1$, μ is the mean of X and if $\nu > 2$, $\nu(\nu - 2)^{-1}\Sigma$ is the covariance matrix of X . Therefore, the family of t -distributions provides a heavy-tailed alternative to the normal family with mean μ and covariance matrix that is equal to a scalar multiple of Σ , if $\nu > 2$ (fig. 1). A K -component

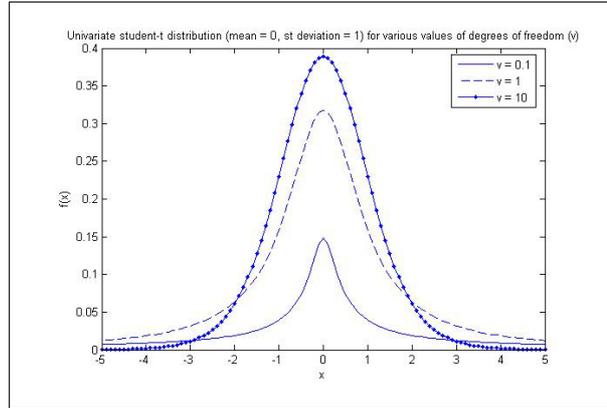


Fig. 1. The Student's t -distribution for various degrees of freedom. As $\nu \rightarrow \infty$ the distribution tends to a Gaussian. For small values of ν the distribution has heavier tails than a Gaussian.

mixture of t -distributions is given by

$$\phi(x, \Psi) = \sum_{i=1}^K \pi_i p(x; \mu_i, \Sigma_i, \nu_i) \quad (9)$$

where $x = (x_1, \dots, x_N)^T$ denotes the observed-data vector and

$$\Psi = (\pi_1, \dots, \pi_K, \mu_1, \dots, \mu_K, \Sigma_1, \dots, \Sigma_K, \nu_1, \dots, \nu_K)^T. \quad (10)$$

are the parameters of the components of the mixture.

A Student's t -distribution mixture model (SMM) may also be trained using the EM algorithm [29]. Consider now the complete data vector

$$x_c = (x_1, \dots, x_N, z_1, \dots, z_N, u_1, \dots, u_N)^T \quad (11)$$

where z_1, \dots, z_N are the component-label vectors and $z_{ij} = (z_j)_i$ is either one or zero, according to whether the observation x_j is generated or not by the i^{th} component. In the light of the definition of the t -distribution, it is convenient to view that the observed data augmented by the z_j , $j = 1, \dots, N$ are still incomplete because the component covariance matrices depend on the degrees of freedom. This is the reason that the complete-data vector also includes the additional missing data u_1, \dots, u_N . Thus, the E-step on the $(t + 1)^{\text{th}}$ iteration of the EM algorithm requires

the calculation of the posterior probability that the datum x_j belongs to the i^{th} component of the mixture:

$$z_{ij}^{t+1} = \frac{\pi_i^t p(x_j; \mu_i^t, \Sigma_i^t, \nu_i^t)}{\sum_{m=1}^K \pi_m^t p(x_j; \mu_m^t, \Sigma_m^t, \nu_m^t)} \quad (12)$$

as well as the expectation of the weights for each observation:

$$u_{ij}^{t+1} = \frac{\nu_i^t + d}{\nu_i^t + \delta(x_j, \mu_i^t; \Sigma_i^t)} \quad (13)$$

Maximizing the log-likelihood of the complete data provides the update equations of the respective mixture model parameters:

$$\pi_i^{t+1} = \frac{1}{N} \sum_{j=1}^N z_{ij}^t, \quad (14)$$

$$\mu_i^{t+1} = \frac{\sum_{j=1}^N z_{ij}^t u_{ij}^t x_j}{\sum_{j=1}^N z_{ij}^t u_{ij}^t}, \quad (15)$$

$$\Sigma_i^{t+1} = \frac{\sum_{j=1}^N z_{ij}^t u_{ij}^t (x_j - \mu_i^{t+1})(x_j - \mu_i^{t+1})^T}{\sum_{j=1}^N z_{ij}^{t+1}}. \quad (16)$$

The degrees of freedom ν_i^{t+1} for the i^{th} component, at time step $t+1$, are computed as the solution to the equation:

$$\log\left(\frac{\nu_i^{t+1}}{2}\right) - \psi\left(\frac{\nu_i^{t+1}}{2}\right) + 1 - \log\left(\frac{\nu_i^t + d}{2}\right) + \frac{\sum_{j=1}^N z_{ij}^t (\log u_{ij}^t - u_{ij}^t)}{\sum_{j=1}^N z_{ij}^t} + \psi\left(\frac{\nu_i^t + d}{2}\right) = 0 \quad (17)$$

where $\psi(x) = \frac{\partial(\ln\Gamma(x))}{\partial x}$ is the digamma function.

At the end of the algorithm, the data are assigned to the component with maximum responsibility using a maximum *a posteriori* (MAP) principle.

The Student's t -distribution is a heavy tailed approximation to the Gaussian. It is therefore, natural to consider the mean and covariance of the SMM components to approximate the parameters of a GMM on the same data as it was described in

the previous section. If the statistics of the images follow a Gaussian model, the degrees of freedom ν_i are relatively large and the SMM tends to be a GMM with the same parameters. If the images contain outliers, parameters ν_i are weak and the mean and covariance of the data are appropriately weighted in order not to take into account the outliers. Thus, the parameters of the SMM, computed on the reference image I_{ref} , are used as component parameters Θ_k^{ref} in a straightforward way as they generalize the Gaussian case by correctly addressing the outliers problem. After projection of the pixel groups of the reference image to their corresponding groups in the registered image, the parameters Θ_k^{reg} are computed using the sample mean (4) and the sample covariance matrix (5).

Once model inference is accomplished, the Bhattacharyya distance between the components of the Student's t -mixtures is minimized. The difference with respect to the GMM is that the covariance matrices are properly scaled by the Gamma distributed parameters u as it is defined in equations (7)-(8).

Finally, let us notice that the energy in (2) may be applied to both single and multi-modal image registration. In the latter case, the difference in the mean values of the distributions in (3) should be ignored, as we do not search to match the corresponding Student's t -distributions in position but only in shape. In that case, the distance in (3) becomes:

$$B(\Theta_1, \Theta_2) = \ln \left(\frac{|\frac{\Sigma_1 + \Sigma_2}{2}|}{\sqrt{|\Sigma_1| |\Sigma_2|}} \right) \quad (18)$$

which is equivalent to a correlation coefficient between the two distributions.

4 Robust registration of point sets with mixtures of Student's t -distributions

An extension of the registration algorithm to handle point sets is described in this section. Given two sets of points X and Y such that Y is derived from X after applying a rigid transformation T_S with parameters \mathcal{S} , that is $Y = T_S(X)$, the problem consists in estimating the transformation parameters from the two data sets without prior knowledge on any correspondence. In fact, in our formulation, there could be no exact correspondence at all due to noise or outlying points.

Let us denote $p(x)$ the density at a point $x \in X$ and assume that it is expressed by a GMM of M components:

$$p(x) = \sum_{j=1}^M \pi_j^x \mathcal{N}(x | \mu_j^x, \Sigma_j^x). \quad (19)$$

By the same assumption, the density at a point $y \in Y$ is given by another GMM:

$$q(y) = \sum_{j=1}^N \pi_j^y \mathcal{N}(y | \mu_j^y, \Sigma_j^y). \quad (20)$$

Considering the transformed point set distribution as $p_{R,t}(x)$, where R is the rotation matrix and t is the translation vector, that is

$$p_{R,t}(x) = \sum_{i=1}^M \pi_i^x \mathcal{N}(x | R\mu_i^x + t, R\Sigma_i^x R^T), \quad (21)$$

we seek to minimize the energy function:

$$D(p_{R,t}, q) = \int [p_{R,t}(z) - q(z)]^2 dz \quad (22)$$

with respect to R and t . More specifically, we seek to match the continuous shapes of the mixtures $p_{R,t}$ and q over their region of support. Equation (22) may be simplified:

$$D(p_{R,t}, q) = \int [p_{R,t}^2(z) + q^2(z) - 2p_{R,t}(z)q(z)] dz \quad (23)$$

The first two terms are invariant under rigid transformation and therefore, the above expression yields the maximum of the product of the two distributions over the whole sets of points. This is equivalent to maximizing the correlation between the pdfs. The cross term may be also expressed as[21]:

$$\int \int p_{R,t}(x)q(y)dx dy = \sum_{i=1}^M \sum_{j=1}^N \pi_i^x \pi_j^y \mathcal{N}(0 | R\mu_i^x + t - \mu_j^y, R\Sigma_i^x R^T + \Sigma_j^y) \quad (24)$$

meaning that given the i^{th} component from the first mixture and the j^{th} component from the second mixture, each term of the sum is evaluated as a Gaussian pdf with mean vector $R\mu_i^x + t - \mu_j^y$ and covariance matrix $R\Sigma_i^x R^T + \Sigma_j^y$ at $x = 0$.

Replacing the GMMs by the more robust SMMs in the above equations (19) and (20) leads to a better modeling of the point sets. Figures 2 and 3 illustrate the performance of a mixture of Student's t -mixture with respect to a standard GMM to model a 2D point set. In the original set, both methods correctly captured the shape of the data (fig. 2). On the other hand, when a small amount of outliers (5%) was present in the set the GMM failed to provide a satisfactory solution while the heavier tailed SMM correctly modeled the point sets (fig. 3). Thus, SMM seems to be a preferable model for density-based point set registration.

An alternative approach would be to provide a model for the outliers using a GMM with a background component or, generally, a probabilistic a model for false observations [29,6]. However, as it will be shown in the experimental results, if the background outliers are not uniformly or normally distributed this approach has its limitations.

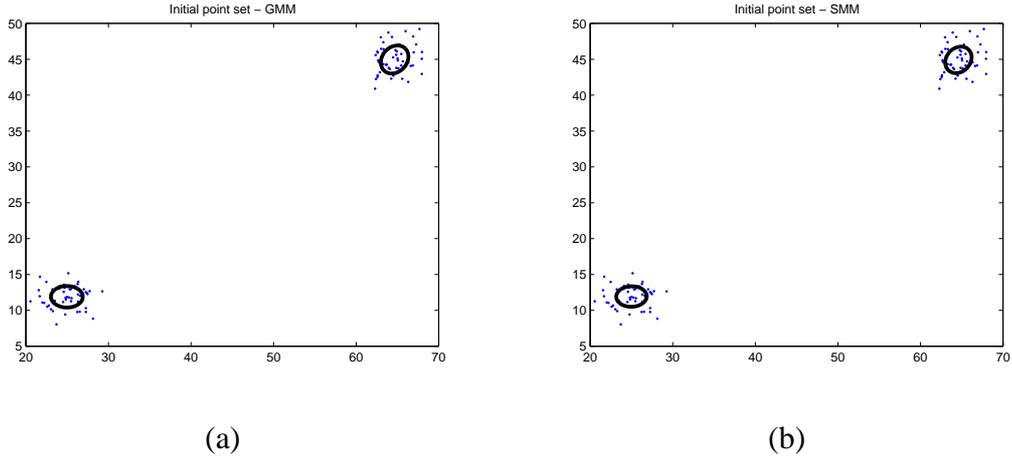


Fig. 2. A 2D point set and the obtained models (a) GMM and (b) SMM.

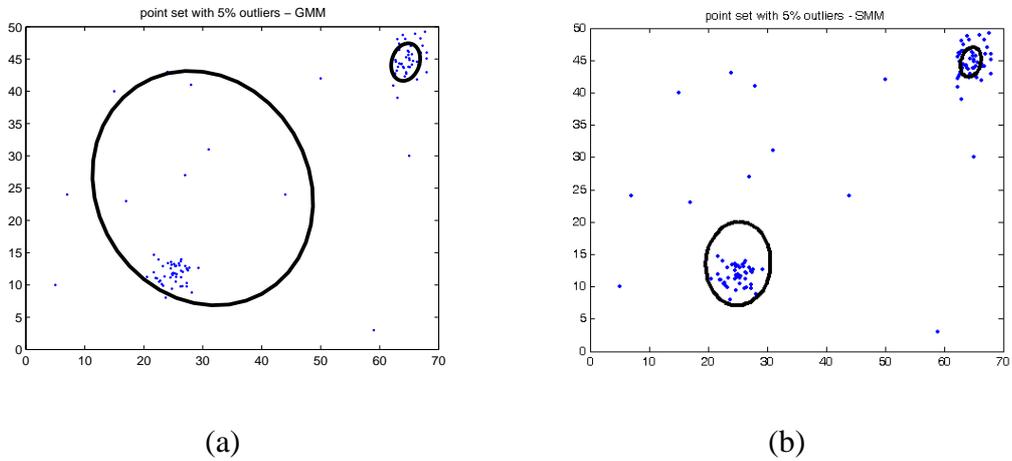


Fig. 3. The point set of figure 2 with 5% outliers and the obtained models by (a) GMM and (b) SMM. Notice that the GMM solution is affected by the outliers while the SMM is more robust.

Let us denote that the above formulas also apply for the registration of point sets using the mixtures of Student's t -distributions by properly computing the components mean vectors and covariance matrices following the definition of the distributions (7)-(8) and the respective EM algorithm described in section 3.

5 Experimental results

A large number of interpolations are involved in the registration process. The accuracy of the rotation and translation parameter estimates is directly related to the accuracy of the underlying interpolation model. Simple approaches such as the nearest neighbor interpolation are commonly used because they are fast and simple to implement, though they produce images with noticeable artifacts. More satis-

factory results can be obtained by small-kernel cubic convolution techniques. In our experiments, we have applied a cubic interpolation scheme, thus preserving the quality of the image to be registered.

The Matlab optimization toolbox was used to perform optimization. In particular we tested the algorithm with a derivative free optimization algorithm (simplex) and a Quasi-Newton algorithm (BFGS) with a numerical calculation of the derivatives. Notice that the methods mentioned perform only local optimization, thus depending the final result highly with the initial starting point. Global optimization methods may also be considered but they are highly time consuming.

5.1 Image registration

In order to evaluate the proposed method, we have performed a number of experiments in some relatively difficult registration problems. Registration errors were computed in terms of pixels and not in terms of transformation parameters. Registration accuracies in terms of rotation angles and translation vectors are not easily evaluated due to parameter coupling. Therefore, the registration errors are defined as deviations of the corners of the registered image with respect to the ground truth position. Let us notice that these registration errors are less forgiving at the corners of the image (where their values are larger) with regard to the center of the image frame.

At first, we have simulated a multimodal image registration example. The image in 4(a) is an artificial piecewise constant image. The image in 4(b) is its negative image. The image in 4(a) was degraded by uniformly distributed noise in order to achieve various SNR values (between 14 dB and -1 dB). The degraded images underwent several rigid transformations by rotation angles varying between $[0, 20]$ degrees and translation parameters between $[-15, 10]$ pixels. To investigate the robustness the proposed method to outliers we have applied the algorithm with $K = 3$ components considering both GMMs and SMMs, and 256 histogram bins in the case of the normalized MI. Figure 5 illustrates the average registration errors for the different SNR values. For each SNR, four different transformations were applied to the image and the average value of the registration error is presented. For comparison purposes, the performance of the MI method is also shown. As it can be observed, both the GMM and the SMM-based registration methods outperform the MI which fails when the SNR is low. Moreover, the heavier tailed SMM demonstrates better performance for considerable amounts of noise.

Furthermore, let us notice that the proposed energy function involving the Battacharyya distances is convex around the true minimum (fig. 6) as it is also the case for the MI [32].

An open issue in mixture modeling is the determination of the number of com-

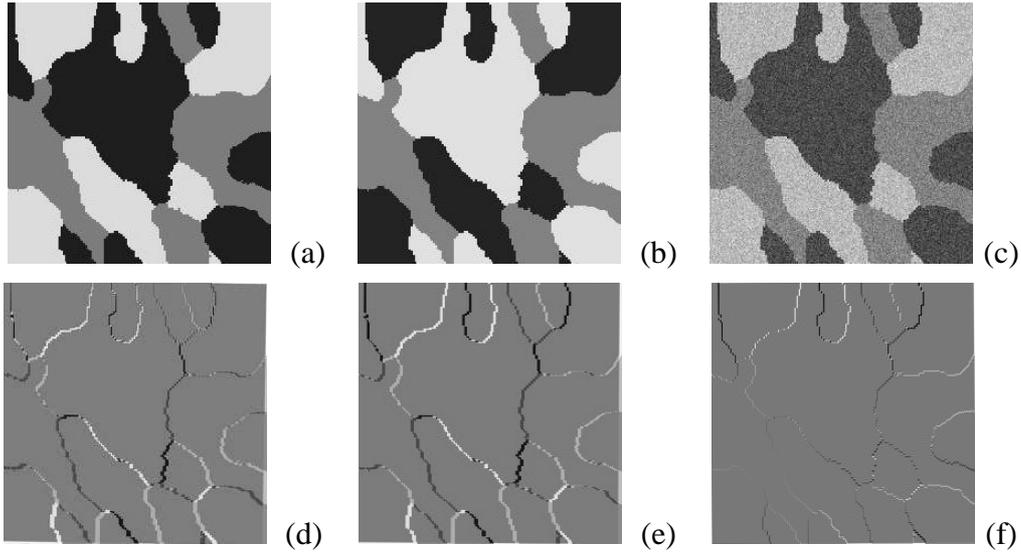


Fig. 4. (a) A three-class piecewise constant image with intensity values 30, 125 and 220, and (b) its negative image (corresponding values, 225, 130 and 35). (c) The image in (a) degraded by uniform noise at 14 dB . This image was then registered to the image in (b). The bottom line shows the registration errors for the compared methods. The ground truth solution is 0 deg for the rotation and zero translation (the original image). (d) MI, (e) GMM, (f) SMM. The errors present the difference between the noise free registered image and the reference image. the values are scaled for better visualization.

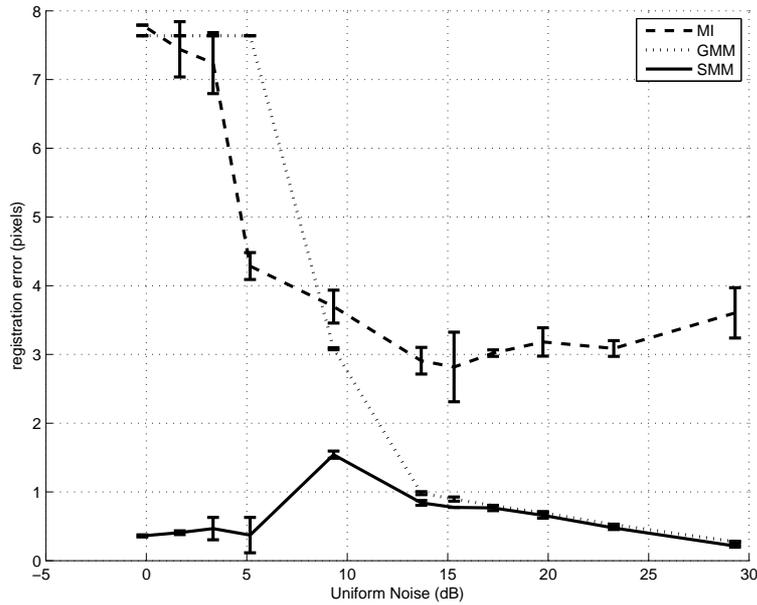


Fig. 5. Mean registration error versus signal to noise ratio (SNR) for the 3-class registration experiment of figure 4.

ponents. In our experiments, in the case of non artificial images, the number of components is unknown. If the number of components of the mixtures is neither to high (overfitting) nor to low (underfitting) with respect to the ground truth the

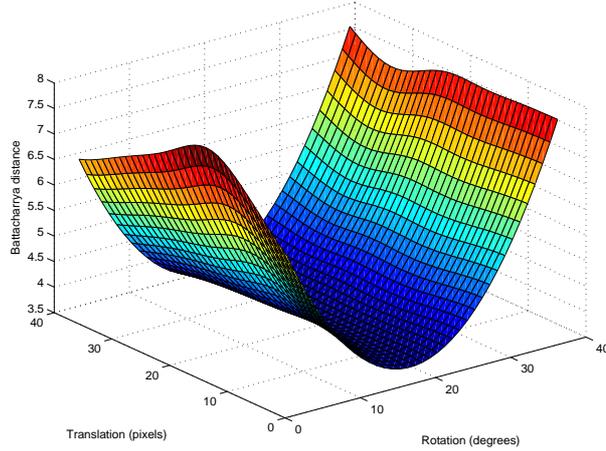


Fig. 6. The objective function in eq. (2) for the registration of the image of figure 4(a) with its counterpart rotated by 20 degrees and translated by 10 pixels.

Table 1

Statistics on the registration errors for the images in fig. 7 with varying number of mixture components. The errors are expressed in pixels.

Registration errors - Cell images						
	K	mean	st. dev.	median	max	min
MI	256 bins	3.663	0.957	4.019	4.25	1.461
SMM	2	3.157	0.009	3.153	3.178	3.150
SMM	3	2.955	0.636	3.148	3.178	1.146
SMM	4	2.956	0.604	3.159	3.101	1.146
SMM	5	2.953	0.640	3.152	3.177	1.132

registration accuracy is not affected by that parameter. In order to demonstrate it, we have performed the experiments involving non artificial images by varying the number of components in the experiments.

In that framework, the proposed registration method was tested on a multimodal image pair such as the cell images in fig. 7. The complimentary but not redundant information carried by the multimodal images increases the difficulty of the registration process. In both experiments we have applied 20 rigid transformations to one of the images, for each configuration of the transformation parameters, with rotation angles varying between $[0, 20]$ degrees and translation parameters between $[-15, 10]$ pixels.

The experiments in the case of the images in figure 7 were realized with the number of components varying from $K = 2$ to $K = 5$. For the MI we used 256 histogram bins. Table 1 summarizes the statistics on the registration errors. As it can be observed, the SMM method achieves highly better registration accuracy. Also, the number of components did not significantly affect the registration accuracy.

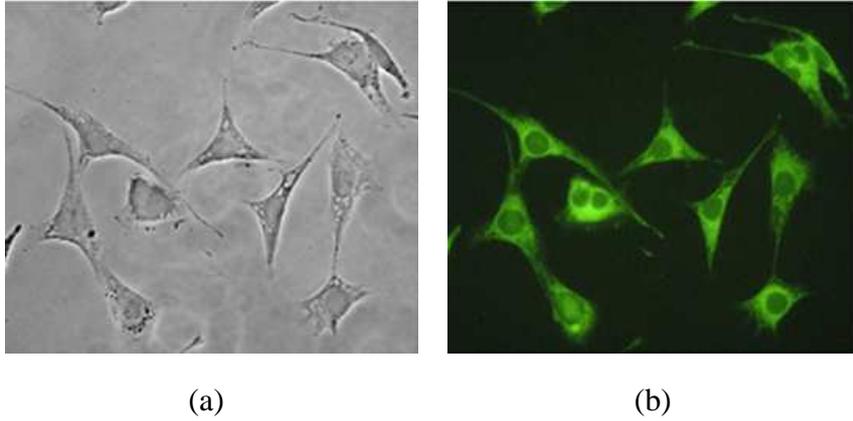


Fig. 7. A pair of NIH 3T3 electron microscope images (400x magnification) of rat cells under (a) normal and (b) fluorescent light.

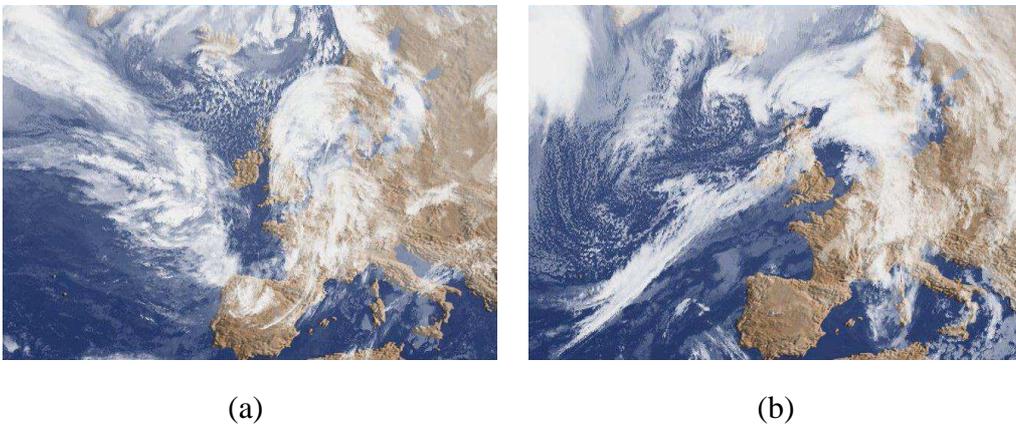


Fig. 8. (a) Image of Europe on 8 January 2007 at 01h00, provided by MeteoSat. (b) Image of Europe on 9 January 2007 at 01h00, provided by MeteoSat (by courtesy of Meteo-France). Notice the large amount of outliers (cloudy regions in different locations in the image pair) introducing important difficulties in the registration process.

A last experiment demonstrating the ability of the proposed SMM method to deal with outliers is the registration of a remotely sensed image pair. The meteorological images of Europe in fig. 8 were acquired at different dates. The image in fig. 8(b) underwent 20 rigid transformations for each parameter instance, with values of rotation angle uniformly sampled in the interval $[0, 20]$ degrees and translations between $[-15, 10]$ pixels. The experiments were realized with the number of components varying between $K = 2$ and $K = 6$ for GMM and SMM and 256 bins for the MI.

The large amount of clouds at different locations in the image pair introduce difficulties in the registration procedure. It is worth commenting that the MI method failed to register the images and systematically provided registration errors of the order of 6 pixels. The SMM method produced very small registration errors which are summarized in table 2.

Table 2

Statistics on the registration errors for the images in fig. 8 with varying number of mixture components. The errors are expressed in pixels.

Registration errors - Satellite images						
	K	mean	st. dev.	median	max	min
MI	256 bins	6.742	1.493	7.463	7.733	3.565
SMM	2	2.975	0.013	2.979	2.991	2.951
SMM	3	1.857	1.202	1.251	3.653	1.283
SMM	4	2.129	2.289	2.960	3.651	1.359
SMM	5	1.208	0.237	1.142	1.999	1.141
SMM	6	1.210	0.238	1.145	2.001	1.142

5.2 Registration of point sets

In order to evaluate the proposed point set registration method we have performed three types of experiments. At first, a 2D set of 600 points was generated from three different Gaussian distributions with means $(-16, 9)$, $(0, 5)$ and $(18, 9)$ and spherical covariance matrices with the standard deviation being 2 in each dimension. The point set underwent rotations varying between $[-90^\circ, 90^\circ]$ and translations varying between $[-100, 100]$ in both dimensions. In all of the cases the proposed algorithm provided solutions close to the true transformation parameters. The registration error was measured as the average distance between the points transformed by the true parameters and the points obtained by the estimated transformation. In all cases, the order of the registration error was approximately 10^{-6} . This experiment was repeated for increased number of non overlapping components and the previous results were confirmed.

A second experiment consisted in comparing the SMM not only to a typical GMM but also to a GMM having an extra background component (called GMMb) in order to model the outliers. This is a standard technique to capture the distribution of outliers and it is also proposed in [29,6]. We have observed that when the outliers are normally or uniformly distributed the performance of the two approaches (GMMb and SMM) is similar because the fourth component is a good model for outliers. However, if the outliers are *signal-dependent* the fourth component does not provide the optimal solution.

In our experiments, the previous point set was corrupted by outlying data from 1% up to 15%. Each of the three set of points was corrupted by a uniform noise having range the double of the initial range of the points generated by the respective component. By these means, the outliers are sparsely distributed around each component. Also, 1% extra outliers were globally added to make the problem more

challenging. For each configuration of the percentage of the outliers, 5 registration experiments were performed with random translation and rotation parameters. A representative example for 9% of points being contaminated is shown in figure 9. In figure 10, the results for the registration errors are summarized. As it can be observed, although the GMMb performs better than the standard GMM due to its background component, the SMM provides smaller registration errors consistently. This behavior is easily explained by the shapes of the ellipses in figures 9(b) and 9(c). Both the GMMb and the SMM estimated small covariances but in GMMb the orientations of the ellipses diverge more from the noise-free case. Finally, it is worth noticing that the standard ICP registration algorithm fails in all cases to provide an acceptable registration.

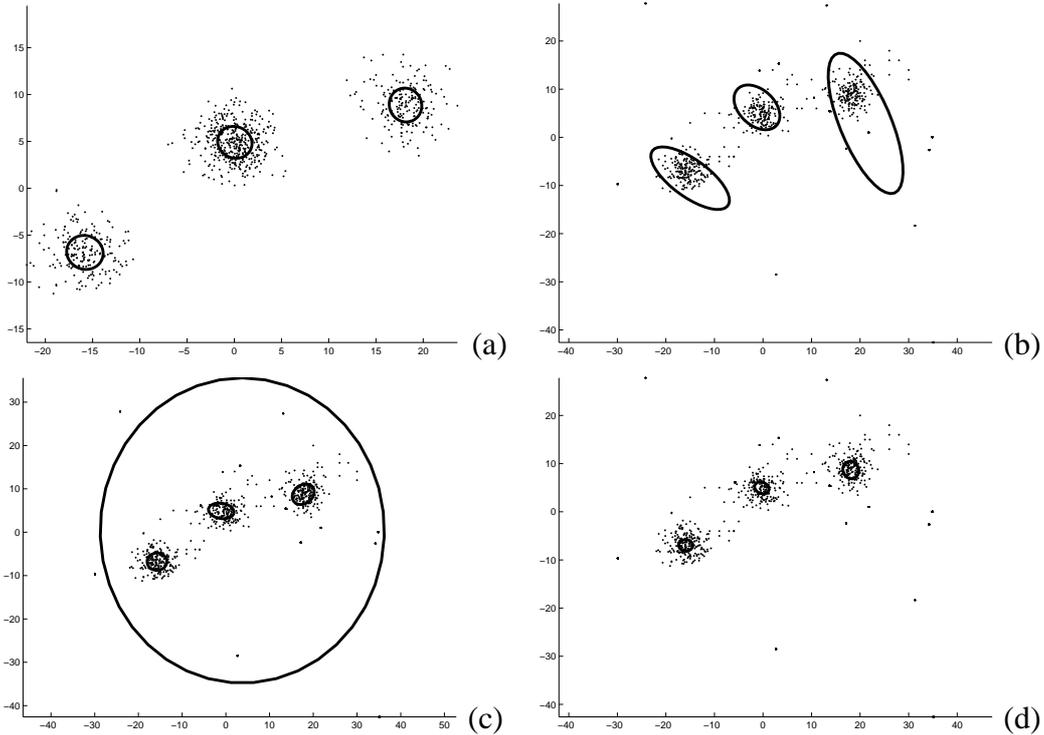


Fig. 9. Example of a set of points used in the experiments. (a) A point set (presented by dots) was generated by 3 Gaussians with means $(-16, 9)$, $(0, 5)$, $(18, 9)$ and spherical covariance matrices of standard deviation 2. The points were corrupted with 9% outliers. The resulting modeling of the noisy set by (b) a 3-component GMM, (c) a 4-component GMM with the fourth component modeling the distribution of outliers and (d) a 3-component SMM.

Finally, we have tested the efficiency of the proposed method to the registration of *shaped* or *structured* point sets, contrary to the scattered points of the previous example. This type of problems may come up from many computer vision applications such as comparison of trajectories in object tracking or shape discrimination and the presence of outliers makes registration difficult even if a good initialization is provided. To this end, we have applied the registration algorithm to data from the *Gaitor Bait 100* data base (as provided by the Department of Computer and Information Science and Engineering, University of Florida, USA,

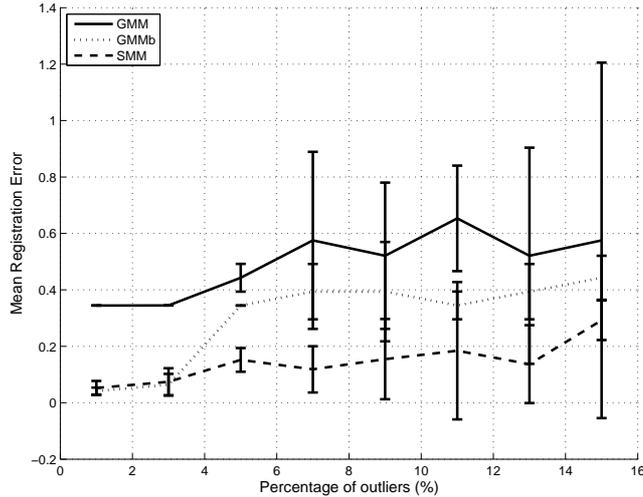


Fig. 10. Registration error as a function of outliers for the experiment presented in figure 9. (<http://www.cise.ufl.edu/>).

In this experimental setting, we begin by illustrating the differences of the compared methods (GMM and SMM) in capturing the data. At first, the same shape, was modeled by a GMM (fig. 11(a)) and an SMM (fig. 11(b)) both with $K = 30$ components. The methods employed the same initialization by the the K-means clustering algorithm. As it can be observed, both methods provided similar approximations. Consequently, the registration algorithm is not affected and the compared methods (GMM and SMM) provide equivalently good performances.

We then eliminated a certain amount of points by to simulate missing data and added outliers to the remaining points. In that case, we also used the the same K-means initialization which naturally provided a certain number of centers that captured the structure of the outliers. However, in any case, the SMM modeled the degraded data better than the GMM by eliminating the majority of erroneous centers, due to its heavier tails. A representative example is presented in figures 11(c) and 11(d) where the missing data percentage is 20% and the percentage of outliers is 10%. In these figures, one can observe that the GMM finally provided two noisy components of relatively large covariance. On the other hand, due to the heavier tails of the SMM components, not only more outlier points were absorbed by the components located on the fish shape, but also the erroneous component has smaller support. This is important in a registration procedure because the L_2 distance in eq. (24) will be less influenced in the case of the SMM, as indicated by the experiments that follow.

The original point set was artificially rotated, translated and corrupted by outliers at 15%. The transformed point set was then registered to its original, noise free counterpart. We have compared the proposed GMM and SMM algorithms with the ICP by initializing them from the ground truth. The results are summarized in table 3, where it is clear that both of the proposed methods (GMM and SMM) perform

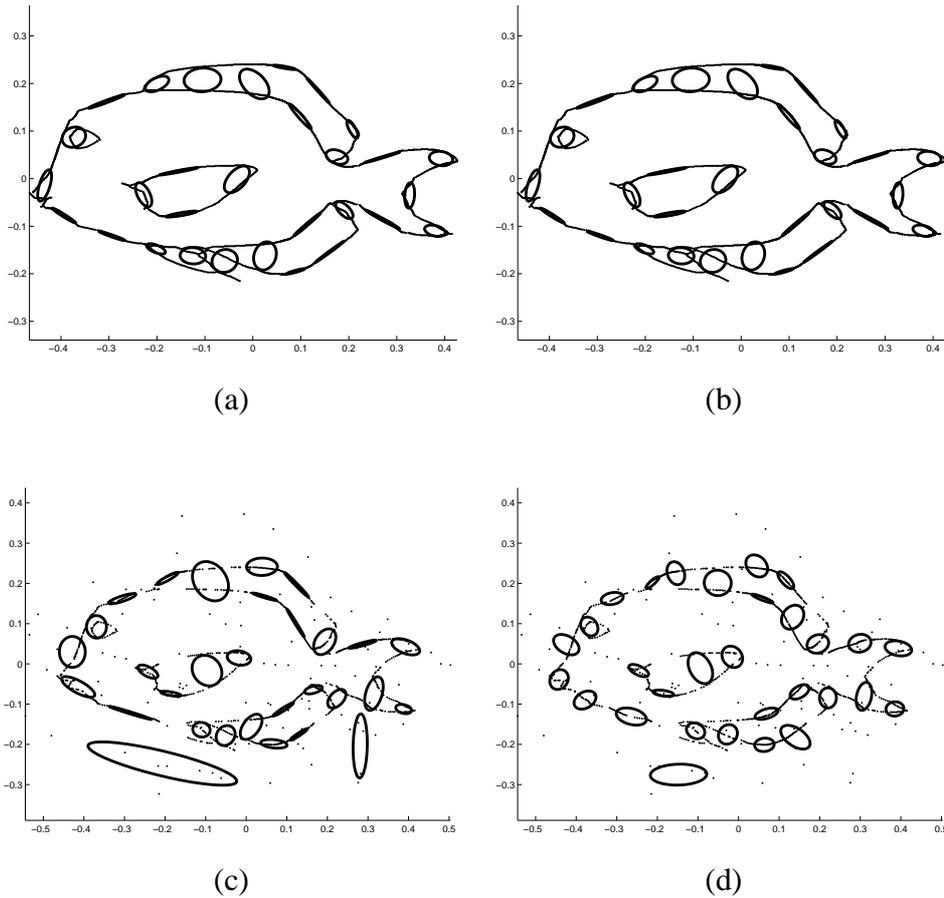


Fig. 11. Modeling of a *shaped* point set from the *Gator Bait 100* data base by (a) GMM with $K = 30$ components and (b) SMM with $K = 30$ components. Notice that the two models provided similar solutions. The bottom row shows the modeling of the point set with 20% missing points and 10% outliers by (c) GMM and (d) SMM. Notice that the solution of the SMM was less affected. In all cases the mixtures were similarly initialized using the K-means algorithm. The axes in (c) and (d) are normalized to the range of the outliers.

better than the ICP. Also, SMM is more accurate than the less robust GMM. It is worth noticing that the ICP algorithm, as it is sensible to initialization, is always trapped around the same minimum.

6 Conclusion

In this paper, we have shown how a mixture model consisting of Students's t components may be efficiently used for registering images and point sets. We have shown the effectiveness and accuracy of the proposed method especially with images presenting dissimilarities where the mutual information method fails to correctly register the two images. The same is confirmed for point sets where the standard ICP algorithm fails in such setups.

Table 3

Registration errors for the *shaped* point set of figure 11 when it is corrupted by 15% outliers.

Method	mean	st. dev.	median	max	min
ICP	40.3784	15.8546	43.6067	58.0508	10.3555
GMM ($K = 15$)	2.6950	1.5169	2.8450	5.1540	0.5894
SMM ($K = 15$)	2.1136	0.8052	1.8880	3.5104	1.2366
GMM ($K = 20$)	2.4334	1.1380	2.4886	4.5563	0.9656
SMM ($K = 20$)	1.9506	0.9084	2.0361	3.4830	0.5927

Let us also notice that Student's t -mixtures overcome the binning problem of histogram based methods and provide a continuous model of the image density. When successfully trained, they produce a sensible approximation of the pdf of the image intensity, by placing density components in a sensible *data-driven* way (i.e on intensity regions exhibiting high density). Although there is still the problem of specifying the number of components in finite mixture modeling, our experimental results indicated that our SMM-based method is robust from this point of view, provided that the number of components is neither very big (overfitting) nor very small (underfitting).

Vector valued images or point data are expected to benefit from this registration technique where the employment of high-dimensional joint histograms makes the use of standard methods prohibitive.

Important open questions for mixture-based registration are how the number of model components can be selected automatically [7] and which features, apart from image intensity, should be used. These are open issues of ongoing research [11].

References

- [1] I. Bankman, Handbook of medical image processing and analysis, Academic Press, 2000.
- [2] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (4) (2002) 509–522.
- [3] P. J. Besl, N. D. McKay, A method for registration of 3-D shapes, IEEE Transactions on Pattern Analysis and Machine Intelligence 14 (2) (1992) 239–256.
- [4] C. Bishop, M. Svensen, Robust Bayesian mixture modelling, Neurocomputing (2005) 69–74.
- [5] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.

- [6] S. S. Brandt, Maximum likelihood robust regression by mixture models, *Journal of Mathematical Imaging and Vision* 25 (1) (2006) 25–48.
- [7] S. Chen, H. Wang, B. Luo, Greedy EM algorithm for robust t-mixture modeling, in: *Proceedings of the 3rd IEEE International Conference on Imaging and Graphics (ICIG'04)*, Hong-Kong, China, 2004.
- [8] D. Chetverikov, D. Stepanov, P. Krsek, Robust euclidean alignment of 3d point sets: the trimmed iterative closest point algorithm, *Image and Vision Computing* 23 (2004) 299–309.
- [9] H. Chui, A. Rangarajan, A new algorithm for non-rigid point matching, in: *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR'00)*, vol. 2, 2000.
- [10] H. Chui, A. Rangarajan, A new point matching algorithm for non-rigid registration, *Computer Vision and Image Understanding* 89 (2003) 114–141.
- [11] C. Constantinopoulos, M. K. Titsias, A. Likas, Bayesian feature and model selection for Gaussian mixture models, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (6) (2006) 1013–1018.
- [12] K. Fukunaga, *Introduction to statistical pattern recognition*, Academic Press, 1990.
- [13] S. Granger, X. Pennec, Multi-scale em-icp: A fast and robust approach for surface registration, in: *Proceedings of the European Conference on Computer Vision (ECCV'02)*, 2002.
- [14] A. Guimond, A. Roche, N. Ayache, J. Meunier, Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections, *IEEE Transactions on Medical Imaging* 20 (1) (2001) 58–69.
- [15] J. V. Hajnal, D. L. G. Hill, D. J. Hawkes, *Medical image registration*, CRC Press, 2001.
- [16] F. Heitz, H. Maître, C. de Couessin, Event detection in multisource imaging: application to fine arts painting analysis, *IEEE Transactions on Acoustic, Speech and Signal Processing* 38 (4) (1990) 695–704.
- [17] P. Hellier, Consistent intensity correction of MR images, in: *Proceedings of the 2003 IEEE International Conference on Image Processing (ICIP'03)*, vol. 1, Barcelona, Spain, 2003.
- [18] M. Herbin, A. Venot, J. Y. Devaux, E. Walter, F. Lebruchec, L. Dubertet, J. C. Roucayrol, Automated registration of dissimilar images: application to medical imagery, *Computer Vision, Graphics and Image Processing* 47 (1989) 77–88.
- [19] G. Hermosillo, O. Faugeras, Dense image matching with global and local statistical criteria, in: *Proceedings of the 2001 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'01)*, vol. 1, Los Alamitos, USA, 2001.
- [20] J. R. Jensen, *Introductory digital image processing: a remote sensing perspective*, 3rd ed., Prentice Hall, 2004.

- [21] B. Jian, B. C. Vemuri, A robust algorithm for point set registration using mixture of Gaussians, in: Proceedings of the 10th International Conference on Computer Vision (ICCV'05), Beijing, China, 2005.
- [22] M. Leventon, E. Grimson, Multi-modal volume registration using joint intensity distributions, in: Proceedings of the Medical Image Computing and Computer Assisted Intervention Conference (MICCAI'98), Cambridge, Massachusetts, USA, 1998.
- [23] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, P. Suetens, Multimodality image registration by maximization of mutual information, IEEE Transactions on Medical Imaging 16 (2) (1997) 187–198.
- [24] J. B. A. Maintz, M. A. Viergever, A survey of medical image registration techniques, Medical Image Analysis 2 (1) (1998) 1–36.
- [25] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, W. Eubank, Nonrigid multimodality image registration, in: K. Hanson (ed.), Proceedings of the 2001 SPIE Medical Imaging Conference, vol. 4322, San Diego, USA, 2003.
- [26] G. McLachlan, Finite mixture models, Wiley-Interscience, 2000.
- [27] A. Myronenko, X. Song, M. Á. Carreira-Perpiñán, Non-rigid point set registration: Coherent Point Drift, in: B. Schölkopf, J. Platt, T. Hoffman (eds.), Proceedings of the 20th Conference on Advances in Neural Information Processing Systems (NIPS'06), MIT Press, Vancouver, Canada, 2006, pp. 1009–1016.
- [28] C. Nikou, F. Heitz, J. P. Armspach, Robust voxel similarity metrics for the registration of dissimilar single and multimodal images, Pattern Recognition 32 (1999) 1351–1368.
- [29] D. Peel, G. J. McLachlan, Robust mixture modeling using the *t*-distribution, Statistics and Computing 10 (2000) 339–348.
- [30] J. Phillips, R. Liu, C. Tomasi, Outlier robust ICP for minimizing RMSD, in: Proceedings of the 6th International Conference on 3D Digital Imaging and Modeling (3DIM'07), 2007.
- [31] J. Pluim, J. B. A. Maintz, M. Viergever, Mutual information based registration of medical images: a survey, IEEE Transactions on Medical Imaging 22 (8) (2004) 986–1004.
- [32] J. P. Pluim, M. J. B. A., M. A. Viergever, Interpolation artefacts in mutual information-based image registration, Computer Vision and Image Understanding 77 (2000) 211–232.
- [33] A. Rajwade, A. Banerjee, A. Rangarajan, A new method of probability density estimation with application to mutual information based image registration, in: Proceedings of the 2006 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, USA, 2006.

- [34] A. Rangarajan, H. Chui, F. L. Bookstein, The softassign procrustes matching algorithm, in: Proceedings of the 15th International Conference on Information Processing in Medical Imaging (IPMI'97). Lecture Notes in Computer Science., vol. 1230, 1997.
- [35] A. Rangarajan, H. Chui, S. Duncan, Rigid point feature registration using mutual information, *Medical Image Analysis* 4 (1999) 1–17.
- [36] A. Roche, X. Pennec, G. Malandain, N. Ayache, Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient information, *IEEE Transactions on Medical Imaging* 20 (10) (2001) 1038–1049.
- [37] S. Rusinkiewicz, M. Levoy, Efficient variants of the ICP algorithm, in: Proceedings of the 3rd International Conference on 3D Digital Imaging and Modeling (3DIM'01), 2001.
- [38] M. Taron, N. Paragios, M. Jolly, Registration with uncertainties and statistical modeling of shapes with variable metric kernels, *IEEE Transactions on Pattern Analysis and Machine Intelligence* Accepted for publication.
- [39] P. Thévenaz, M. Unser, Optimization of mutual information for multiresolution image registration, *IEEE Transactions on Image Processing* 9 (12) (2000) 2083–2099.
- [40] Z. Tu, S. Zheng, A. Yuille, Shape matching and registration by data-driven EM, *Computer Vision and Image Understanding* 109 (3) (2008) 290–304.
- [41] P. Viola, W. Wells III, Alignment by maximization of mutual information, *International Journal of Computer Vision* 24 (2) (1997) 137–154.
- [42] F. Wang, C. B. Vemuri, A. Rangarajan, M. I. Schmalfluss, J. S. Eisenschenk, Simultaneous nonrigid registration of multiple point sets and atlas construction, in: Proceedings of the European Conference on Computer Vision (ECCV'06), 2006.
- [43] B. Zitova, J. Flusser, Image registration methods: a survey, *Image and Vision Computing* 21 (11) (2003) 1–36.