Deformable Density Matching for 3D Non-rigid Registration of Shapes

Arunabha S. Roy[1], Ajay Gopinath[1], and Anand Rangarajan[2,★]

[1] Imaging Technologies Laboratory, GE Global Research Center, Bangalore, India
[2] Department of CISE, University of Florida, Gainesville, FL, USA

Abstract. There exists a large body of literature on shape matching and registration in medical image analysis. However, most of the previous work is focused on matching particular sets of features—point-sets, lines, curves and surfaces. In this work, we forsake specific geometric shape representations and instead seek probabilistic representations— specifically Gaussian mixture models—of shapes. We evaluate a closed-form distance between two probabilistic shape representations for the general case where the mixture models differ in variance and the number of components. We then cast non-rigid registration as a deformable density matching problem. In our approach, we take one mixture density onto another by deforming the component centroids via a thin-plate spline (TPS) and also minimizing the distance with respect to the variance parameters. We validate our approach on synthetic and 3D arterial tree data and evaluate it on 3D hippocampal shapes.

1 Introduction

The need for shape matching occurs in diverse sub-domains of medical image analysis. Whenever a biomedical image is segmented or parsed into a set of shapes, the need for shape analysis and comparison usually arises [1]. In brain mapping for example [2], we frequently require the comparison of cortical and subcortical structures such as the thalamus, putamen etc. extracted from subject neuroanatomical MRI images. Image databases often use shape features and here the need is to index and query the shape database. In MR angiography, the complex network of blood vessels in the brain can be represented as trees or graphs and need to be compared across subjects. And in cardiac applications, if heart chamber information is available and extracted as a set of shapes, the wall tracking problem requires us to solve for shape correspondences in the cardiac cycle [3].

The need for shape matching is followed by a need for good shape representations. When shape features are extracted from medical images, they can be represented using an entire gamut of representations—points, line segments, curves, surfaces, trees, graphs and hybrid representations. What should be noted here is that an inferential stage is present in any shape representation. That is,

---

the raw features extracted from the underlying images are then converted into one of the above mentioned shape representations. Consequently, once a shape representation is adopted, we are immediately faced with the problem of *robustness* when we seek to compare two shapes. Some of the raw features present in (or extracted from) one image may not be present in the other. A second problem we face is that there may be no or poor *correspondences* between the features. It is for this reason that most methods seek to fit curves or surfaces to the data. Once such representations are fitted, there is no need to seek for correspondences at the point feature level. However, methods that rely on fitting curves and surfaces to the data face a more difficult robustness problem. How do you match one shape consisting of 10 curves to another shape consisting of 15 curves?

For these reasons, we elect to go the probabilistic route. Beginning with raw point features, we fit probability density functions to the feature vectors [4]. Each feature set is converted into a probability density function. The advantage here is that we can now compare probability density functions rather than the original images or other feature representations extracted from the images. The robustness problem is alleviated since the density functions are compared at all locations in $\mathbb{R}^3$ and we are not faced with the problem of matching incommensurate entities such as one curve in one shape to two curves in the other. There is no correspondence problem since we are not conducting comparisons at the point feature level. Shape comparison by matching the density functions between shapes also has the advantage that the point feature counts (in the two feature sets) can differ considerably.

We summarize our new method as follows: i) We fit Gaussian mixture models to point features extracted from the two images. A standard maximum likelihood expectation-maximization (EM) approach is used for this step. ii) We derive a *closed-form distance* between the two Gaussian mixture models by comparing the probability density functions at each point in $\mathbb{R}^3$. iii) Since the problem of minimizing this distance w.r.t. non-rigid deformations is ill-posed, we add a thin-plate spline regularization term to the cost. iv) We use a standard conjugate-gradient (CG) optimization strategy to minimize the above objective function. It should be stressed that we minimize the objective function w.r.t. both the deformation parameters and the variance parameters of the Gaussian mixture model. The result is a deformable density matching (DDM) method which seeks to register two shapes by moving the density function parameters of one shape until the first density closely approximates the other.

2   Previous Work

There exists an enormous literature on feature matching. We restrict our focus to methods that seek to fit and/or match probabilistic representations of shape features. The joint clustering and matching (JCM) approach [5] begins as we do with fitting Gaussian mixture models to feature point-sets. However, instead of minimizing a distance between the two density functions, they seek to augment

the mixture model objective by linking a diffeomorphism between the centroid parameters of the mixture model. Consequently, they are forced to keep two sets of cluster centers in correspondence which we do not require. The methods in [6] and in [7] seek to convert a point matching problem into an image matching problem which is somewhat similar to density matching. However, the methods make no attempt to fit a probabilistic shape representation via maximum likelihood (or equivalent) to the features. Essentially, both methods convert the sparse feature set into dense images and then employ an image matching strategy. Furthermore, the method in [6] uses a deformation field parametrization which has to be applied at each point in $\mathbb{R}^3$ and this is computationally expensive compared to our approach. The method in [7] does not use a deformation field parametrization but the main differences are that they restrict their approach to point matching and make no attempt to fit a density function to the data via maximum likelihood or minimize their distance w.r.t. the centroid and variance parameters. Instead, they apply a thin-plate spline (TPS) on the original, noisy data. Perhaps the method that is closest in spirit to our approach is the recent work in [4]. They minimize the Jensen-Shannon divergence between the feature point-sets w.r.t. a non-rigid deformation. The Jensen-Shannon divergence cannot be computed in closed form for a mixture model and is estimated from the data using a law of large numbers approach. In sharp contrast, our distance between the two densities is in closed form and we apply the deformation parametrization to the centroidal parameters of the Gaussian mixture model.

3   Theory

3.1   Maximum-Likelihood Model for Shape Representation

As mentioned in the Introduction, the first step in our overall method is probabilistic shape representation based on the raw shape features.

The notation used in this paper is as follows. The two sets of input shape features are denoted by $\{X_i^{(1)} \in \mathbb{R}^d, i \in \{1,\ldots,N_1\}\}$ and $\{X_j^{(2)} \in \mathbb{R}^d, j \in \{1,\ldots,N_2\}\}$. The maximum likelihood approach assumes that the shape features $\{X_i^{(1)}\}$ and $\{X_j^{(2)}\}$ are independent and identically distributed (i.i.d.). The features of shape $X^{(1)}$ are represented by a Gaussian mixture model [8]

$$p(\mathbf{x}|\theta^{(1)}) = \sum_{a=1}^{K_1} \Omega_a^{(1)} \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma_a|^{\frac{1}{2}}} \exp\{-\frac{1}{2}(\mathbf{x} - \mu_a)^T \Sigma_a^{-1}(\mathbf{x} - \mu_a)\} \qquad (1)$$

(where $\mathbf{x} \in \mathbb{R}^3$) and that of shape $X^{(2)}$ is represented by a second Gaussian mixture model with parameter set $\theta^{(2)} = \{\Omega_\alpha^{(2)}, \nu_\alpha, \Xi_\alpha\}$. Constraints on $\Omega^{(\cdot)}$ are $\{\Omega_a^{(\cdot)} > 0, \sum_{a=1}^{K_1} \Omega_a^{(\cdot)} = 1\}$ where the superscript index can be either 1 or 2 corresponding to the two shapes respectively. Both probability density functions define measures on location with $\mathbf{x} \in \mathbb{R}^3$. We see that the set of model parameters for shape $X^{(1)}$ is $\theta^{(1)} = \{\Omega_a^{(1)}, \mu_a, \Sigma_a, a \in \{1,\ldots,K_1\}\}$ and

$\theta^{(2)} = \{\Omega_\alpha^{(2)}, \nu_\alpha, \Xi_\alpha, \alpha \in \{1, \ldots, K_2\}\}$ for shape $X^{(2)}$. Since $\{X_i^{(1)}\}$ and $\{X_j^{(2)}\}$ are assumed to be i.i.d., the likelihood of the set of features of $X^{(1)}$ is

$$p(\{X_i^{(1)}\}|\theta^{(1)}) = \prod_{i=1}^{N_1} \sum_{a=1}^{K_1} \Omega_a^{(1)} \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma_a|^{\frac{1}{2}}} \exp\{-\frac{1}{2}(X_i^{(1)} - \mu_a)^T \Sigma_a^{-1}(X_i^{(1)} - \mu_a)\}$$
(2)

where $\{X_i^{(1)}\}$ is the set of instances of $X^{(1)}$. For both shapes, we fit the model parameters by minimizing the negative log-likelihood objective function of the above mixture model w.r.t. the model parameters. In the experiments, we specialize to the case where the occupancy probabilities are uniform ($\Omega_a^{(1)} = \frac{1}{K_1}$) and where we have one isotropic covariance matrix ($\Sigma^2 = \sigma^2 I_3$), for the entire shape. This is done for reasons of computational efficiency. The objective function for this reduced version is minimized over its parameter set ($\{\mu_a\}, \sigma$) to obtain the model representation of the feature set.

We use the well known expectation-maximization (EM) algorithm [8] for the above minimization. While fitting mixture models is computationally difficult in the general case, in our special case of uniform occupancy probabilities and isotropic covariances it is not as difficult. Also, this computation is done offline for all the shapes once we fix the number of centroids (K_1 and K_2). Model selection for mixture models needs to be performed to fix the number of centroids.

3.2   A Closed-Form Distance Measure Between Two Gaussian Mixtures

We now derive the distance measure between the two probabilistic shape representations. Since this distance measure can be derived in closed form for the most general Gaussian mixture model case, we present this below. (Other distance measures like Kullback-Leibler cannot be derived in closed form for the Gaussian mixture model.) We hasten to add that the general mixture model can be quite unwieldy in practice and that specializations of the sort considered in the paper—like isotropic covariances and uniform occupancy probabilities—are very useful. Below we derive the distance measure as the squared pointwise difference between the two Gaussian mixture models with parameter sets ($\theta^{(1)}$ and $\theta^{(2)}$) integrated over $\mathbb{R}^d$ (where $d = 3$). [We have dropped terms that do not depend on $\theta^{(2)}$ since it is only $\theta^{(2)}$ that is deformed during the optimization.]

$$D[p(\mathbf{x}|\theta^{(1)}), p(\mathbf{x}|\theta^{(2)})] = \int_{\mathbb{R}^d} [p(\mathbf{x}|\theta^{(1)}) - p(\mathbf{x}|\theta^{(2)})]^2 dx \propto$$
$$-\sum_{a=1}^{K_1}\sum_{\alpha=1}^{K_2} \frac{2\exp\{-\frac{1}{2(\sigma^2+\xi^2)}||\mu_a - \nu_\alpha||^2\}}{K_1 K_2(\sigma^2+\xi^2)^{\frac{3}{2}}} + \sum_{\alpha=1}^{K_2}\sum_{\beta=1}^{K_2} \frac{\exp\{-\frac{1}{4\xi^2}||\nu_\alpha - \nu_\beta||^2\}}{2^{\frac{3}{2}}K_2^2\xi^3}. \quad (3)$$

This is the final expression for the distance function used in this paper. The number of centroids and the variances of the two models can be different.

3.3   Deformable Density Matching

We now turn to the description of the deformation model. We assume that the parameters $\theta^{(1)}$ of shape 1 are held fixed and that the parameters $\theta^{(2)} = (\{\nu_\alpha\}, \xi)$ (comprising the centroids and variance) of shape 2 are deformed so that $p(\mathbf{x}|\theta^{(2)})$ approaches $p(\mathbf{x}|\theta^{(1)})$.

   We use the familiar thin-plate spline (TPS) deformation model [9] for the centroids. That is, the action of the deformation takes the centroids $\{\nu_\alpha\}$ to new locations $\{\tilde{\nu}_\alpha \overset{\text{def}}{=} A\nu_a + \sum_{\beta=1}^{K_2} K(v_\alpha, \nu_\beta) Q_2 \gamma_\beta\}$ where A is the unknown (4×3) affine matrix, the TPS kernel $K(\nu_\alpha, \nu_\beta) = -||\nu_\alpha - \nu_\beta||$ in 3D, Q_2 is the $K_2 \times (K_2 - 4)$ part of the QR decomposition of ν (in homogeneous coordinates) and γ is the unknown $(K_2 - 4) \times 3$ matrix of deformation parameters. To avoid degenerate solutions which include all permutations of ν when $K_1 = K_2$, we add a deformation regularization term $\lambda \operatorname{trace}(\gamma^T Q_2^T K Q_2 \gamma)$ to the objective function in (3) with λ being a regularization parameter. In addition a regularization term $\lambda_A \operatorname{trace}[(A - I)^T (A - I)]$ is added to prevent reflections and unphysical affine transformations. We use a standard nonlinear conjugate-gradient (CG) algorithm (with line search) on the affine and deformation parameters (A, γ) and a 1-D search on the variance ξ^2. The variance parameter is updated separately from the remaining parameters and is not allowed to abruptly change.

4   Results

### 4.1   Synthetic Example: Sphere to Ellipsoid Density Matching

To showcase our approach of density matching, we begin with a synthetic example. We generate a sphere and an ellipsoid with 600 points each. We run a mixture model EM algorithm on the ellipsoid and sphere representing them with 120 and 60 centroids respectively. We then warp the ellipsoid using a Gaussian radial basis function (GRBF) using the 120 centroids as the centers for the GRBF. Subsequently, we run the deformable density matching (DDM) on the two synthetic datasets. The goal is to deform the sphere so that it matches the warped ellipsoid. Here the fixed variance σ^2 was set to the mean of the cluster variances resulting from the EM algorithm ($\sigma = 0.04$) and a 1-D search was performed over the parameter ξ, obtaining $\xi = 0.06$ as the optimum. The initial overlay of the sphere and the warped ellipsoid, the overlay of the centroids after matching and the final overlay of the warped sphere (using the recovered deformation parameters) are shown in Figure 1. Our results clearly demonstrate that the lack of correspondences at the point level as seen in the final shape overlay in Figure 1 is no deterrent to recovering the shape of the deformed ellipsoid.

4.2   Validation of Recovered Deformation: Bending and Stretching

Our approach to validation is based on comparing the recovered deformation vectors against the true (synthetically generated) deformation vectors generated at
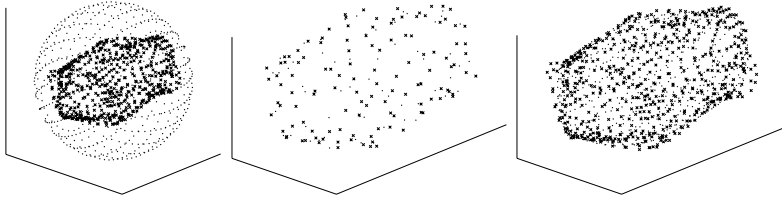
Fig. 1. Left: Initial overlay of sphere and warped ellipsoid, Middle: Cluster centers of the two shapes after density matching and Right: Overlay after registration
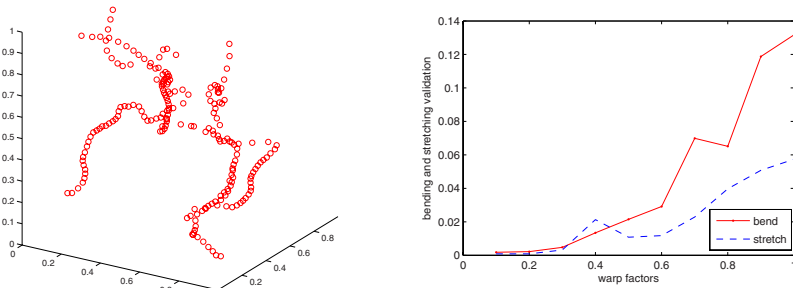
Fig. 2. Left: Vessel tree structure; Right: Validation over 30 noise trials, 10 warps

a set of vessel tree structure feature points [Figure 2]. Since we are not performing point matching, we cannot validate our results by using prior correspondence information. We begin with a point-set X consisting of about 200 points. A deformation field is applied to the point-set X to obtain a deformed point-set Y. We remove 50% of the points in Y to get a reduced set and this is done so that the two mixture densities have a large discrepancy in the number of centroids. Then we match X to the reduced Y using deformable density matching using 60 and 30 centroids for the two sets respectively. The recovered TPS parameters are used to compute the estimated displacement vector at each point in X. We compare the estimated displacement to the true displacement by separately plotting the stretching and bending components of the mean-squared displacement error: $\frac{1}{N_2}\sum_i||\hat{u}_i - u_i||^2 = \frac{1}{N_2}\sum(||\hat{u}_i|| - ||u_i||)^2 + \frac{1}{N_2}\sum_i 2||\hat{u}_i||||u_i||(1 - \frac{\hat{u}_i^T u_i}{||\hat{u}_i||||u_i||})$ where $\hat{u}_i$ is the estimated deformation, u_i is the true deformation and we have separated the total error into stretching (first term) and bending (second term) components. We executed 30 noise trials at different TPS warp factors w ranging from 0.1 to 1 in steps of 0.1. The TPS warp factor specifies that the TPS coefficients be uniformly generated from the interval $[-w, w]$. Higher the warp factor, greater the degree of deformation. The medians of the two error measures are shown in Figure 2, illustrating the difference between bending and stretching. The bending errors are more sensitive and increase more rapidly with w.
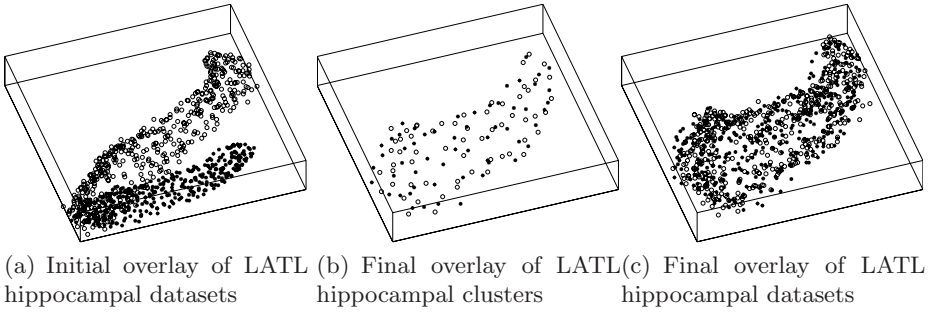
(a) Initial overlay of LATL (b) Final overlay of LATL(c) Final overlay of LATL
hippocampal datasets   hippocampal clusters   hippocampal datasets

Fig. 3.

(a) Initial overlay of RATL (b) Final overlay of RATL
hippocampal datasets   hippocampal clusters

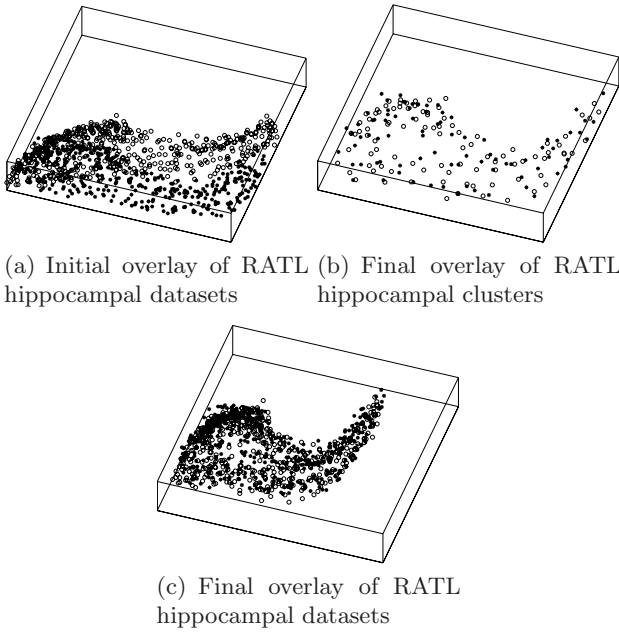(c) Final overlay of RATL
hippocampal datasets

Fig. 4.

4.3   Evaluation on 3D Hippocampal Datasets

The general problem of building a hippocampal atlas is a clinical application
requiring multiple registrations of shapes from different patients. Here we use
deformable density matching to register 3D hippocampal point-set pairs from a
database of 60 cases. We showcase results on two pairs of patients scheduled for
left and right anterior temporal lobectomy (LATL and RATL) respectively. The
point-sets consisting of a few hundred points each are clustered into about 80
centroids, and we set $\lambda = 0.01$. After matching, we overlay the two point-sets by
warping one set onto the other using the recovered TPS parameters. The initial
overlay of the data, the final overlay of the centroids and the final overlay of the

warped datasets are shown in Figures 3 and 4 for LATL and RATL respectively. In both cases, the final overlay shows that we have achieved good registration. We are currently looking at entropic registration measures for more objective evaluation of all pairwise registrations across our database.

5 Discussion

In summary: i) We used a maximum likelihood EM approach to fit centroids and variances to raw feature data. ii) We determined a closed-form distance between two Gaussian mixture models. iii) We recovered a TPS deformation of the centroids of one shape while also allowing the shape variance parameter to change in order to best match the two shape representations. iv) Finally, we applied the recovered TPS deformation to the original data and showed that shape matching was possible even though there were few correspondences at the point feature level.

Our goal was to obtain a robust shape distance with very few free parameters and with the deformation parameters only affecting the model parameters and not specified over $\mathbb{R}^3$. The TPS model only warps the centroids and the variance parameters are allowed to move until the best fit is achieved. There are four free parameters in our model: K_1, K_2 and the regularization parameters λ, λ_A. Since the process of fitting probabilistic representations is offline, we can take recourse to model selection in order to fit K_1 and K_2. The regularization parameters can be learned via a Bayesian MAP procedure. There also do not appear to be any technical barriers to generalizing the TPS deformation parametrization to a diffeomorphism and optimizing over the occupancy probabilities as well.

References

1. Davies, R.H., Twining, C., Cootes, T.F., Taylor, C.J.: A minimum description length approach to statistical shape modelling. IEEE Trans. Med. Imag. 21, 525–537 (2002)
2. Toga, A., Mazziotta, J.: Brain Mapping: The Methods. Academic Press, London (1996)
3. Tagare, H.: Shape-based nonrigid correspondence with application to heart motion analysis. IEEE Transactions on Medical Imaging 18(7), 570–579 (1999)
4. Wang, F., Vemuri, B.C., Rangarajan, A., Schmalfuss, I.M., Eisenschenck, S.J.: Simultaneous registration of multiple point-sets and atlas construction. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3953, pp. 551–563. Springer, Heidelberg (2006)
5. Guo, H., Rangarajan, A., Joshi, S.C.: 3-D diffeomorphic shape registration on hippocampal data sets. In: Duncan, J.S., Gerig, G. (eds.) MICCAI 2005. LNCS, vol. 3750, pp. 984–991. Springer, Heidelberg (2005)
6. Glaunes, J., Trouve, A., Younes, L.: Diffeomorphic matching of distributions: A new approach for unlabelled point-sets and sub-manifolds matching. In: IEEE Conf. on Computer Vision and Pattern Recognition, vol. 2, pp. 712–718 (2004)
7. Jian, B., Vemuri, B.C.: A robust algorithm for point set registration using mixture of Gaussians. In: Intl. Conf. on Computer Vision (ICCV), pp. 1246–1251 (2006)
8. McLachlan, G.J., Basford, K.E.: Mixture models: inference and applications to clustering. Marcel Dekker, New York (1988)
9. Wahba, G.: Spline models for observational data. SIAM, Philadelphia, PA (1990)