

A Multi-resolution T-Mixture Model Approach to Robust Group-Wise Alignment of Shapes

Nishant Ravikumar^{1,2}(✉), Ali Gooya^{1,3}, Serkan Çimen^{1,3},
Alejandro F. Frangi^{1,3}, and Zeike A. Taylor^{1,2}

¹ CISTIB Centre for Computational Engineering
and Simulation Technologies in Biomedicine,
INSIGNEO Institute for in Silico Medicine, Sheffield, UK
mta08nr@shef.ac.uk

² Department of Mechanical Engineering, The University of Sheffield, Sheffield, UK

³ Department of Electronic and Electrical Engineering,
The University of Sheffield, Sheffield, UK

Abstract. A novel probabilistic, group-wise rigid registration framework is proposed in this study, to robustly align and establish correspondence across anatomical shapes represented as unstructured point sets. Student's t-mixture model (TMM) is employed to exploit their inherent robustness to outliers. The primary application for such a framework is the automatic construction of statistical shape models (SSMs) of anatomical structures, from medical images. Tools used for automatic segmentation and landmarking of medical images often result in segmentations with varying proportions of outliers. The proposed approach is able to robustly align shapes and establish valid correspondences in the presence of considerable outliers and large variations in shape. A multi-resolution registration (mrTMM) framework is also formulated, to further improve the performance of the proposed TMM-based registration method. Comparisons with a state-of-the-art approach using clinical data show that the mrTMM method in particular, achieves higher alignment accuracy and yields SSMs that generalise better to unseen shapes.

1 Introduction

Generative probabilistic model-based point set registration techniques are used in various medical image analysis applications, such as landmark-based image registration, statistical shape model (SSM) generation, correction of incomplete image segmentations, and 3D shape reconstruction from 2D projection images. Existing probabilistic approaches to rigid [1–3] and non-rigid [4] group-wise point set registration are based on Gaussian mixture models (GMMs) which, while affording efficient solutions for associated model parameters, lack robustness to outliers. An elegant solution to this limitation is to adopt a t-mixture model (TMM) formulation, which is inherently more robust due to TMMs' so-called heavy tails. Such an approach forms the main contribution of this study. Additionally, we propose a novel multinomial distribution-based, multi-resolution

extension that encapsulates the group-wise TMM registration and involves a process of adaptive sampling from the components of the TMM at each resolution level.

Aligning a group of medical image-derived unstructured point sets is particularly challenging due to the presence of missing information, varying degrees of outliers, and unknown correspondences. Probabilistic approaches offer a solution, by casting the registration problem as one of probability density estimation, where each sample shape in a group is assumed to be a transformed observation of a mixture model, leading to the joint inference of unknown correspondences and desired spatial transformations across the group [3].

Pair-wise point set registration using TMMs has been proposed in two previous studies [5, 6]. However, group-wise registration methods are, in general, preferable as they provide an unbiased solution to the registration problem [1]. Use of TMMs in the latter context was recently proposed in [7], wherein rigid transformation parameters were estimated numerically by gradient ascent optimisation. In the present work, closed-form expressions are derived for the transformation parameters, which afford significant improvement in computational efficiency, and a further multi-resolution extension is formulated.

Estimation of valid correspondences and invariance to rigid transformations across a group of shapes are necessary for training SSMs by principal component analysis (PCA). However, the overall process is challenged by the need for cumbersome pre-processing (typically manual) during training set generation. This is prohibitive when learning models from large-scale image databases. We aim to ameliorate this challenge by means of robust methods for jointly establishing correspondence and aligning point sets in the presence of outliers, which are therefore compatible with fully automated segmentation and landmarking tools.

The proposed single- and multi-resolution methods are validated by comparison with a state-of-the-art GMM-based approach [2], based on registration accuracy and quality of SSMs generated.

2 Methods

2.1 Group-Wise Rigid Registration Using TMMs

Student's t-distributions \mathcal{S} are derived as an infinite mixture of scaled Gaussians, where the precision scaling weights, denoted u , are treated as latent variables drawn from a Gamma distribution \mathcal{G} :

$$\mathcal{S}(\mathbf{x}|\boldsymbol{\mu}, \Sigma, \nu) = \int_0^\infty \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \Sigma/u)\mathcal{G}(u|\nu/2, \nu/2)du. \quad (1)$$

t-distributions are a generalisation of Gaussians with heavy tails that result from finite values for the associated degrees of freedom ν [8]. We argue that point set generation from medical images is affected by outliers and, consequently, TMMs are well suited to align and establish correspondence across such point sets.

Optimal TMM and registration parameters are estimated by maximising their posterior probability conditioned on the observed data. A tractable solution to this problem is formulated by maximising the complete data log likelihood, with respect to the unknown parameters Ψ via expectation-maximisation (EM). For a group of K point sets denoted $\mathcal{X}_k \in \mathbf{X}$, to be aligned using an M component mixture model, the complete data log likelihood is expressed as Eq. (2). In Eq. (2), $\mathbb{U} = \{u_{kij}\}$, $\mathbb{Z} = \{z_{kij}\}$ represent the sets of latent variables associated with each component in the TMM. The former scales the precision of the equivalent Gaussian distribution, while the latter is a binary vector specifying the unique membership of the observed data (\mathbf{X}) to components in the mixture model. Subscript $j = 1 \dots M$ is used to represent mixture model components while $i = 1 \dots N_k$ is used to represent N_k data points that belong to the k^{th} shape in the data set.

$$\log p(\mathbf{X}, \mathbb{U}, \mathbb{Z} | \Psi) = \log(p(\mathbb{Z} | \Psi)) + \log(p(\mathbb{U} | \mathbb{Z}, \Psi)) + \log(p(\mathbf{X} | \mathbb{U}, \mathbb{Z}, \Psi)) \quad (2)$$

This results in an iterative estimation alternating between evaluating the conditional expectation of the latent variables, given an estimate of the M component mixture model parameters $\Theta = \{\boldsymbol{\mu}_j, \sigma^2, \nu_j, \pi_j\}$ and rigid registration parameters $\mathbf{T} = \{\mathcal{T}_{k=1}^K\}$, and updating these parameters $\Psi = \{\Theta, \mathbf{T}\}$. This leads to estimation of the product of the conditional expectations of the two latent variables, $P_{kij}^* = E(z_{kij} | \mathbf{x}_{ki}) E(u_{kij} | \mathbf{x}_{ki}, z_{kij} = 1)$, in the E-step. The posterior probabilities given by $E(z_{kij} | \mathbf{x}_{ki})$ describe the responsibility of the j^{th} mixture component with mean $\boldsymbol{\mu}_j$, variance σ^2 , degrees of freedom ν_j and mixing coefficient π_j , in describing the i^{th} point on the k^{th} shape in the group, given by \mathbf{x}_{ki} . Subsequently, the M-step maximises the conditional expectation of the complete data log likelihood with respect to each of the unknown parameters, sequentially. M-step equations to update estimates of the model Θ and rigid transformation parameters \mathcal{T}_k are analytically derived. The latter are presented in (3a, b, c):

$$\mathbf{t}_k = \left[\sum_{i=1}^{N_k} \sum_{j=1}^M P_{kij}^* \mathbf{x}_{ki} \right] \left[\sum_{i=1}^{N_k} \sum_{j=1}^M P_{kij}^* \right]^{-1} - s_k \mathcal{R}_k \left[\sum_{i=1}^{N_k} \sum_{j=1}^M P_{kij}^* \boldsymbol{\mu}_j \right] \left[\sum_{i=1}^{N_k} \sum_{j=1}^M P_{kij}^* \right]^{-1}, \quad (3a)$$

$$\mathcal{C}_k = \left[\sum_{i=1}^{N_k} \sum_{j=1}^M P_{kij}^* [(\mathbf{x}_{ki} - \mathbf{t}_k) \boldsymbol{\mu}_j^T] \right] \left[\sum_{i=1}^{N_k} \sum_{j=1}^M P_{kij}^* \boldsymbol{\mu}_j \boldsymbol{\mu}_j^T \right]^{-1}, \quad (3b)$$

$$s_k = [Tr \{ \sum_{i=1}^{N_k} \sum_{j=1}^M P_{kij}^* \mathcal{R}_k^T (\mathbf{x}_{ki} - \mathbf{t}_k) \boldsymbol{\mu}_j^T \}] [Tr \{ \sum_{i=1}^{N_k} \sum_{j=1}^M P_{kij}^* \boldsymbol{\mu}_j \boldsymbol{\mu}_j^T \}]^{-1}. \quad (3c)$$

While the degrees of freedom ν_j are estimated numerically, expectations in the E-step and estimates for the remaining model parameters are derived analytically, similar to [3, 8]. \mathcal{C}_k represents a real covariance matrix from which the orthogonal rotation matrix \mathcal{R}_k is computed by singular value decomposition similar to [2], \mathbf{t}_k represents the translation, and s_k represents the scaling. $\{\mathcal{R}_k, s_k, \mathbf{t}_k\}$ together form the rigid transformation \mathcal{T}_k mapping the current estimate of the mean model to the k^{th} shape.

2.2 Multi-resolution Registration by Adaptive Sampling

A multi-resolution approach (mrTMM), wherein the density of the mean model is increased (and consequently, the model variance is decreased) at each successive resolution, was formulated to further improve the performance of the proposed method. We argue that such a framework reduces the influence of local minima which may be introduced during initialisation (by k-means clustering) and estimation of the mean model, and thereby improves registration accuracy and quality of SSMS trained. Multi-resolution schemes are often employed in image registration [9]. A probabilistic approach to multi-resolution registration of point sets, however, is novel to the best of our knowledge.

Increase in mean model density is achieved through adaptive sampling by imposing a multinomial distribution over the estimated mixture coefficients π_j and drawing N_s random samples from a subset of M TMM components, i.e. those with high responsibilities in explaining the observed data. The number of new model points s_j sampled from the j^{th} mixture component is described by the multinomial distribution as $p(s_j|\pi_j, N_s) = \text{Mult}(s_1 \dots s_M | \pi_j, N_s)$. New model points are generated by drawing random samples from a zero-centered multivariate Gaussian distribution and an inverse chi-squared distribution (with ν_j degrees of freedom), since t-distributed random variables can be expressed as $\mathbf{s}_j = \boldsymbol{\mu}_j + \mathcal{N}(0, \sigma^2) \sqrt{\nu_j / \chi^2(\nu_j)}$. The registration proceeds in a hierarchical coarse-to-fine fashion, with the mean model increasing in density at each successive level and the rigid transformations estimated at each level initialising the subsequent resolution.

3 Results

Two clinical data sets were used for validation: (1) a 2D set of 100 femur boundaries segmented automatically using a clinically employed software (Hologic Apex 3.2) from dual-energy X-ray absorptiometry images of healthy subjects; (2) a 3D set of 30 caudate nuclei, segmented automatically [10] from T1-weighted MR images¹ of healthy subjects. Point sets were generated from these segmentations using a marching cubes-based surface extraction algorithm. Alignments were then performed, and correspondences computed using the TMM and mrTMM approaches (results from the latter shown in Fig. 1(a–e)). Finally, SSMS were constructed (example presented in Fig. 1(f, g)) from the estimated correspondences using PCA. The process was repeated using the state-of-the-art sparse statistical shape model (SpSSM [2]) approach for comparison. SpSSM was chosen as in [2] the authors demonstrated its ability to generate SSMS of higher quality than a traditional GMM-based approach [3].

Comparisons were made on the basis of (1) rigid alignment accuracy of the estimated probabilistic correspondences across each group of shapes, and (2) quality of the corresponding SSMS. The former were quantified by Hausdorff

¹ Public database: <http://www.brain-development.org> (c) Copyright Imperial College of Science, Technology and Medicine 2007. All rights reserved.

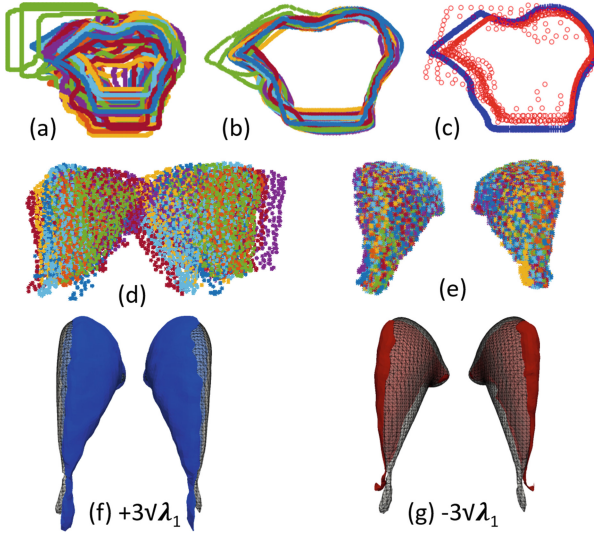


Fig. 1. Top row: (a) raw 2D femur point sets prior to alignment, (b) aligned probabilistic correspondences using mrTMM, (c) estimated mean shapes using mrTMM (blue) and SpSSM (red); Middle row: (d) raw 3D caudate point sets prior to alignment, (e) aligned probabilistic correspondences using mrTMM; Bottom row: first mode of variation of caudate SSM trained using mrTMM, overlaid on mean shape (transparent outline), (f) $+3\sqrt{\lambda_1}$ and (g) $-3\sqrt{\lambda_1}$. (Color figure online)

(HD) and quadratic surface distance (QSD), computed between the estimated mean shapes and all transformed samples in each corresponding group; the latter from the generality of the trained models in five-fold, leave-one-out, cross validation experiments. For each method tested, the optimal number of mixture components M_{opt} was identified by plotting their respective generalisation errors against the number of mixture components employed to align each clinical data set. M_{opt} values for both data sets and each registration method are summarised in Table 1. The corresponding average alignment accuracy of the estimated probabilistic correspondences are also presented in the table. SSMs trained using the identified M_{opt} values for both data sets were also assessed by evaluating their generality and specificity with respect to the number of modes of variation through full-fold cross-validation experiments.

The gain in computational efficiency, relative to [7], afforded by the proposed algorithm was assessed by aligning the caudate data set using 300 mixture components. A substantial improvement in speed was achieved using the proposed method: execution time: 317 s vs. 987 s. Furthermore, in [7], numerical estimation of the rigid transformation parameters required an additional user specified parameter (i.e. optimiser step size) which may need to be tuned to different data sets. This parameter is obviated in the proposed algorithm and consequently analysis is more automated and its application on large data sets is more robust.

Registration Errors. The registration errors quantified in Table 1 for the caudate data set indicate that the proposed single- and multi-resolution TMM-based methods outperform SpSSM in terms of registration accuracy, with mrTMM being the most accurate. The femur data set included many instances of over-segmented boundary masks, making it heavy with outliers (as depicted in Fig. 1(a)). In this case, the TMM-based methods conferred significantly higher accuracy and were able to estimate mean femur shapes in a manner robust to these outliers. This is depicted in Fig. 1(c), which shows the adverse effect outliers have on the estimated mean shape using SpSSM and the advantage mrTMM affords in this respect. In the experiments conducted, the TMM-based methods consistently outperform SpSSM and are more robust for group-wise rigid point set registration. Statistical significance of the computed registration errors for TMM and mrTMM with respect to SpSSM were assessed using a paired-sample t-test, considering a significance level of 1%. Significant improvements on SpSSM are highlighted in bold in Table 1.

Table 1. Registration accuracy for clinical data sets evaluated using HD and QSD metrics at optimal mean model density M_{opt}

Registration method	3D Caudate Nuclei (K = 30)			2D Femur (K = 100)		
	M_{opt}	HD (mm)	QSD (mm)	M_{opt}	HD (mm)	QSD (mm)
SpSSM	1182	5.95 ± 1.36	1.05 ± 0.12	838	41.32 ± 3.54	2.90 ± 0.70
TMM	2560	4.60 ± 2.38	0.95 ± 0.15	640	11.07 ± 5.17	2.15 ± 1.11
mrTMM	2560	3.76 ± 1.85	0.93 ± 0.14	640	11.49 ± 5.48	2.18 ± 0.96

Additionally, by estimating valid mean shapes for the femur data set, the TMM-based methods were able to establish valid probabilistic correspondences (as shown in Fig. 1(b)). SpSSM on the other hand, was affected by the presence of outliers, resulting in the estimation of wrong mean shapes and correspondences, and consequently, incorrect modes of variation in the trained SSM.

SSM Quality. SSM generalisation errors evaluated with respect to number of mixture components are shown in Fig. 2(a, b). The proposed TMM and mrTMM approaches generated SSMs of both the caudate and femur that generalise better to unseen test shapes than does SpSSM. Generalisation errors with respect to number of modes of variation (for the first ten modes, using M_{opt} identified from the former experiments), evaluated by full-fold cross-validation, are depicted in Fig. 2(c, d). The first mode of variation achieved the lowest errors for the caudate, while two modes were optimal for the femur, using TMM and mrTMM.

As the caudate nuclei segmentations were of high quality, with no apparent outliers (from visual inspection), the performance of TMM and SpSSM were similar. In such cases, the constituent t-distribution components are little different

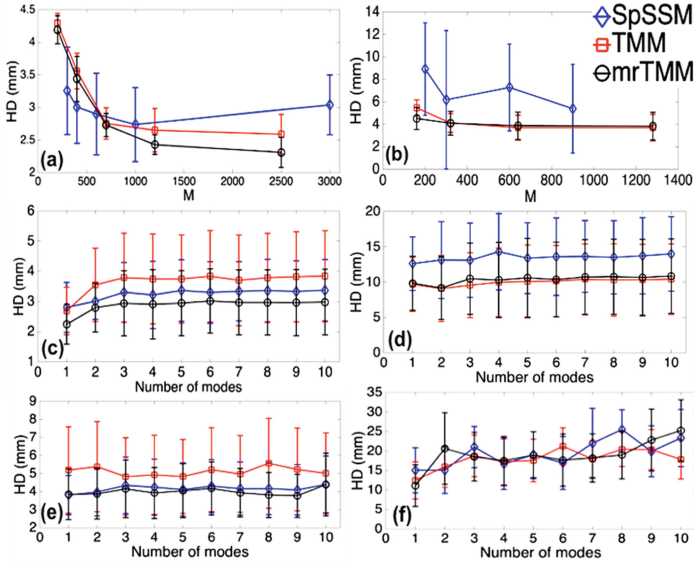


Fig. 2. SSM quality using SpSSM, TMM and mrTMM for caudate and femur data sets. Top row: generalisation errors (using HD) vs. number of mixture components M ; middle row: generalisation errors vs. number of modes of variation N ; bottom row: specificity errors vs. N . Left and right columns show results for caudate and femur data sets, respectively.

from Gaussian distributions. Nonetheless, for this set, mrTMM consistently outperformed both other methods in all generalisation experiments. The improvement afforded by mrTMM is also reflected in the specificity errors presented in Fig. 2(e).

For the femur data set, both TMM-based methods achieved significantly lower generalisation errors with respect to number of mixture components (Fig. 2(b)) and number of modes of variation (Fig. 2(d)); SpSSM estimated incorrect modes of variation in the femur shape, due to the wrong probabilistic correspondences established. This also resulted in higher specificity errors using SpSSM, as shown in Fig. 2(f).

4 Conclusion

Single- and multi-resolution TMM-based methods for group-wise rigid registration of unstructured point sets (such as derived from medical images) have been described and shown to be particularly advantageous over a state-of-the-art GMM-based method (SpSSM) for data containing outliers. For clinical data sets containing few or no outliers, the proposed methods showed some improvements over SpSSM, in terms of registration accuracy and quality of corresponding SSMs, though these were modest. Conversely, for clinical data with high levels of

outliers, the proposed methods significantly outperformed SpSSM and showed excellent robustness. In most cases, the multi-resolution extension (mrTMM) afforded further improvement over the single-resolution (TMM) formulation. The presented approaches should be especially advantageous for applications involving large data sets, which correspondingly require automated processing techniques that are robust to outliers.

Acknowledgments. This study was funded by the European Unions Seventh Framework Programme (FP7/2007 2013) as part of the project VPH-DARE@IT (grant agreement no. 601055) and partly supported by the Marie Curie Individual Fellowship (625745, A. Gooya). The authors would like to thank Dr. Fabian Wenzel, Philips Research Laboratories, Hamburg, Germany, for providing access to their fully automated tool, to segment the caudate nuclei.

References

1. Evangelidis, G.D., Kounades-Bastian, D., Horaud, R., Psarakis, E.Z.: A generative model for the joint registration of multiple point sets. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8695, pp. 109–122. Springer, Heidelberg (2014). doi:[10.1007/978-3-319-10584-0_8](https://doi.org/10.1007/978-3-319-10584-0_8)
2. Gooya, A., Davatzikos, C., Frangi, A.F.: A bayesian approach to sparse model selection in statistical shape models. *SIAM J. Imaging Sci.* **8**(2), 858–887 (2015)
3. Hufnagel, H., Pennec, X., Ehrhardt, J., Ayache, N., Handels, H.: Generation of a statistical shape model with probabilistic point correspondences and the expectation maximization-iterative closest point algorithm. *Int. J. Comput. Assist. Radiol. Surg.* **2**(5), 265–273 (2008)
4. Chen, T., Vemuri, B.C., Rangarajan, A., Eisenschenk, S.J.: Group-wise point-set registration using a novel CDF-based Havrda-Charvát divergence. *Int. J. Comput. Vis.* **86**(1), 111–124 (2010)
5. Gerogiannis, D., Nikou, C., Likas, A.: The mixtures of Students t-distributions as a robust framework for rigid registration. *Image Vis. Comput.* **27**(9), 1285–1294 (2009)
6. Zhou, Z., Zheng, J., Dai, Y., Zhou, Z., Chen, S.: Robust non-rigid point set registration using Student’s-t mixture model. *plos one* **9**(3), e91381 (2014)
7. Ravikumar, N., Gooya, A., Frangi, A.F., Taylor, Z.A.: Robust group-wise rigid registration of point sets using t-mixture model. In: SPIE Medical Imaging, pp. 97840S–97840S. International Society for Optics and Photonics (2016)
8. Peel, D., McLachlan, G.J.: Robust mixture modelling using the t distribution. *Stat. Comput.* **10**(4), 339–348 (2000)
9. Schnabel, J.A., et al.: A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations. In: Niessen, W.J., Viergever, M.A. (eds.) MICCAI 2001. LNCS, vol. 2208, pp. 573–581. Springer, Heidelberg (2001). doi:[10.1007/3-540-45468-3_69](https://doi.org/10.1007/3-540-45468-3_69)
10. Zagorchev, L., Meyer, C., Stehle, T., Wenzel, F., Young, S., Peters, J., Weese, J., et al.: Differences in regional brain volumes two months and one year after mild traumatic brain injury. *J. Neurotrauma* **33**(1), 29–34 (2016)