Medical Image Analysis 44 (2018) 156–176

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Group-wise similarity registration of point sets using Student's t-mixture model for statistical shape models

Nishant Ravikumar[a,b,*], Ali Gooya[a,c], Serkan Çimen[a,c], Alejandro F. Frangi[a,c], Zeike A. Taylor[a,b]

[a] CISTIB Centre for Computational Imaging and Simulation Technologies in Biomedicine, INSIGNEO Institute for in silico Medicine, United Kingdom
[b] Department of Mechanical Engineering, The University of Sheffield, United Kingdom
[c] Department of Electronic and Electrical Engineering, The University of Sheffield, United Kingdom

ARTICLE INFO

Article history:
Received 12 October 2016
Revised 11 July 2017
Accepted 25 November 2017
Available online 5 December 2017

Keywords:
Group-wise point set registration
Statistical shape models
Student's t-mixture model
Expectation-maximisation (EM)

ABSTRACT

A probabilistic group-wise similarity registration technique based on Student's t-mixture model (TMM) and a multi-resolution extension of the same (mr-TMM) are proposed in this study, to robustly align shapes and establish valid correspondences, for the purpose of training statistical shape models (SSMs). Shape analysis across large cohorts requires automatic generation of the requisite training sets. Automated segmentation and landmarking of medical images often result in shapes with varying proportions of outliers and consequently require a robust method of alignment and correspondence estimation. Both TMM and mrTMM are validated by comparison with state-of-the-art registration algorithms based on Gaussian mixture models (GMMs), using both synthetic and clinical data. Four clinical data sets are used for validation: (a) 2D femoral heads ($K = 1000$ samples generated from DXA images of healthy subjects); (b) control-hippocampi ($K = 50$ samples generated from T1-weighted magnetic resonance (MR) images of healthy subjects); (c) MCI-hippocampi ($K = 28$ samples generated from MR images of patients diagnosed with mild cognitive impairment); and (d) heart shapes comprising left and right ventricular endocardium and epicardium ($K = 30$ samples generated from short-axis MR images of: 10 healthy subjects, 10 patients diagnosed with pulmonary hypertension and 10 diagnosed with hypertrophic cardiomyopathy). The proposed methods significantly outperformed the state-of-the-art in terms of registration accuracy in the experiments involving synthetic data, with mrTMM offering significant improvement over TMM. With the clinical data, both methods performed comparably to the state-of-the-art for the hippocampi and heart data sets, which contained few outliers. They outperformed the state-of-the-art for the femur data set, containing large proportions of outliers, in terms of alignment accuracy, and the quality of SSMs trained, quantified in terms of generalization, compactness and specificity.

1. Introduction

Statistical shape models (SSMs) have found widespread use in a variety of medical image analysis applications in recent years such as segmentation (Patenaude et al., 2011; Castro-Mateos et al., 2015), shape-based prediction of tissue anisotropy (Lekadir et al., 2014), quantitative shape analysis and classification for computer-aided-diagnosis (Styner et al., 2004; Shen et al., 2012; Gooya et al., 2015b), to name a few. Their primary challenge has persistently been the availability of training sets of sufficient size, necessary to adequately describe anatomical shape variability observed across different demographic and diagnostic populations. A training set of segmentations delineating the structure of interest in medical images, a fundamental requirement for training SSMs, is typically generated manually or semi-automatically. This can be laborious and prohibitive when analysing 3D structures from large cohorts. In the past, various solutions have been proposed such as, merging pre-exiting SSMs trained from different cohorts (Pereañez et al., 2014), generating artificial variations in shape using synthetic transformations (Koikkalainen et al., 2008) to enrich the data set with a higher degree of variability and employing automatic techniques to generate the required training set of segmentations, which is of particular interest in this study. The major challenges with this approach are the potential inclusion of outliers

* Corresponding author.
E-mail addresses: mta08nr@sheffield.ac.uk, mta08nr@shef.ac.uk (N. Ravikumar), a.gooya@sheffield.ac.uk (A. Gooya), s.cimen@sheffield.ac.uk (S. Çimen), a.frangi@sheffield.ac.uk (A.F. Frangi), z.a.taylor@sheffield.ac.uk (Z.A. Taylor).

https://doi.org/10.1016/j.media.2017.11.012

and the presence of missing information in the segmentations, as a result of variable image resolution and quality, motion artefacts, pathology-induced intensity inhomogeneities, among others. Consequently, in order to facilitate large-scale statistical shape analysis of anatomical structures using automated pre-processing techniques to generate the required training set, a robust framework capable of aligning and establishing anatomically valid correspondences across the group of shapes, is imperative. Such a framework forms the main contribution of this study.

'Shape' as defined by Kendall (1989), is the geometric information that remains once an object has been normalized with respect to rotation, scaling and translation. Various methods to represent this information and analyse the statistics of their variation across an ensemble of similar shapes have been proposed, some of which include point- or mesh-based discretisation (Cootes et al., 1995), implicit functions (signed distance maps) (Leventon et al., 2000), spherical harmonics (SPHARM) based parameterisation (Brechbühler et al., 1995; Gerig et al., 2001) and medial shape representation (Pizer et al., 2003; Styner et al., 2003). Among these, point-based representations of shape are the most prevalent for training SSMs, due to their simplicity and independence to topology. The latter property in particular is a desirable trait for anatomical structures, not afforded by some techniques such as SPHARM for example, which only permit shapes of spherical topology. Medial models are 'skeleton-like' representations which yield more compact shape descriptions than landmark-based approaches but utilise surface boundaries parametrised by SPHARM and consequently have identical topological constraints. Based on these factors, in this study we focus on point-based representations of shapes as the main purpose is to formulate a topology independent, automatic and robust framework for training SSMs.

Past approaches to automatic SSM generation have included: (1) a pair-wise, template-to-training set (or one-to-many) registration strategy where an atlas is non-rigidly registered to each training shape, thereby propagating the landmarks used to represent the atlas shape across the training set and establishing correspondences (Lorenz and Krahnstöver, 2000; Frangi et al., 2002); (2) population-based techniques based on minimum description length (MDL) (Davies et al., 2002, 2010) or entropy (equivalent to MDL) (Cates et al., 2007), which automatically estimate correspondences across training shapes by optimizing an objective function dependent on model quality; and (3) group-wise point set registration methods for jointly aligning a group of shapes and establishing correspondences across them (Hufnagel et al., 2008; Gooya et al., 2015a). A thorough review of various correspondence estimation approaches for training SSMs is provided in Heimann and Meinzer (2009). The third class of techniques is well-suited to automatic SSM generation as it combines the process of rigid shape alignment and correspondence estimation in a unified framework, unlike population-based methods where these two steps are often de-coupled. Furthermore, they can be imbued with inherent robustness to outliers and missing information in the data, through suitable stochastic formulations of the problem. Group-wise registration methods for establishing point correspondence across training shapes are in general preferable to pair-wise approaches as the established correspondences are not biased towards the template as with the latter. Outliers and missing information that may be present in the template shape are propagated to each training sample using pair-wise registration approaches, resulting in sub-optimal correspondences Consequently, iterative multi-template registration approaches are often employed together with majority voting/shape blending strategies (Frangi et al., 2002), to minimize bias towards any single template (this approach can however, be computationally very expensive for large data sets). Conversely, with group-wise methods, an unbiased mean shape is iteratively refined and jointly registered to each training sample, obviating the need for multi-template registration, reducing computational cost and resulting in the estimation of correspondences in an unbiased manner.

In the subsequent section we discuss relevant literature, focusing on probabilistic point set registration methods and their application to the construction of SSMs and statistical atlases.

1.1. Previous work

Registration of surfaces, curves or point sets and correspondence estimation is an open problem in computer vision and medical image analysis and has received significant attention over the past few decades. Early work in the field includes the well known and widely used Generalized Procrustes (GP) (Gower, 1975) and iterative closest point (ICP) (Besl and McKay, 1992) algorithms and various extensions of the same, namely, soft-assign Procrustes (Rangarajan et al., 1997) and EM-ICP (Granger and Pennec, 2002; Hufnagel et al., 2008; Hermans et al., 2011), respectively. These techniques rely on point-based representations of shapes to align and establish correspondences across the same. The main limitations of the GP method are its requirement for correspondences to be determined prior to alignment and high sensitivity to outliers (as the Euclidean distance is minimised between shapes). The conventional ICP algorithm relies on establishing exact correspondences by identifying the closest point pairs in the shapes to be aligned. Although such an approach is computationally very efficient, it is also severely affected by the presence of outliers in the point sets being aligned, which may lead to the estimation of incorrect correspondences and consequently sub-optimal transformations. Additionally, ICP is also constrained by the need for the two shapes to be well-aligned initially, to satisfy the assumption that closest point pairs correspond to each other, a non-trivial problem in medical imaging applications.

Subsequent approaches have employed different types of features for registration and adopted a probabilistic-view of estimating correspondences, to address the limitations of nearest-neighbour based techniques (such as ICP). In such approaches, correspondence for each point on one shape is formulated as a weighted combination of all points on the other shape, where the weights/probabilities are derived from a probabilistic function of the pairwise distances (typically the squared Mahalanobis distance) between the shapes. These include: the robust point matching (RPM) method which utilises point/edge-based features, a soft-assign algorithm for establishing correspondence and deterministic annealing optimisation for rigid (Rangarajan et al., 1997) and non-rigid (Chui and Rangarajan, 2003) point matching; deformable surface registration algorithms based on currents (Vaillant and Glaunès, 2005; Durrleman et al., 2007); and others that employ diffeomorphic transformations in combination with local geometry descriptors such as integral volume (Gelfand et al., 2005) and surface curvatures (Wang et al., 2003).

Landmark-based approaches to shape registration and their probabilistic extensions are of particular interest as they are independent of topology. Additionally, the latter are tailored to address the challenges of missing information and varying degrees of outliers, common to medical-image derived point sets. In recent years, various probabilistic approaches to shape/point set registration have been formulated, such as: coherent point drift (CPD) (Myronenko and Song, 2010), a pair-wise method for rigid, affine and non-rigid registration; joint registration of multiple point clouds (JRMPC) (Evangelidis et al., 2014), which is analogous to a group-wise version of strictly rigid-CPD (i.e. only rotation and translation estimated, does not estimate global scaling during alignment); robust pair-wise point set registration using Gaussian mixture models (Jian and Vemuri, 2011), where the

point sets are represented as independent GMMs and are aligned by minimizing the L2-norm between them; and a variety of GMM-based group-wise, rigid/similarity (Granger and Pennec, 2002; Hufnagel et al., 2008; Gooya et al., 2015a) and non-rigid (Rasoulian et al., 2012; Wang et al., 2006; Chen et al., 2010) registration methods. In the context of training SSMs, which is of particular interest in this study, the recent work of Gooya et al. (2015a) is most relevant, as their method (sparse statistical shape models or SpSSM) was shown to produce SSMs of higher quality than a conventional GMM-based method, namely, EM-ICP, proposed by Hufnagel et al. (2008). SpSSM employs a symmetric Dirichlet prior for the mixture coefficients to enforce sparsity (sparsity level is a user specified parameter, $s_p \in [0, 1]$) and identify and prune out mixture components with low probability in explaining the observed data. Such an approach starts from a maximal mean model, with a high density of points, which are subsequently removed as the registration progresses and the probability of model points drops below a threshold enforced by the specified sparsity level. The pruning process for the removal of such model points is achieved via quadratic programming, using a generalised sequential minimal optimiser. Consequently, the number of mixture components used for a given data set is selected over a continuous rather than discrete search space.

We restrict our attention to probabilistic similarity registration methods (i.e. a similarity transformation parametrised by 7 degrees of freedom in 3D: rotation, scaling and translation, is estimated) for point sets, as the main application of interest is the automatic construction of SSMs from medical images, requiring: independence to topology, robustness to outliers, ability to accommodate missing information and recover large similarity transformations in the presence of significant variations in shape. Additionally, group-wise registration approaches are of particular interest as they are able to estimate the desired similarity transformations and establish correspondences in an unbiased manner, a limitation of non-rigid pair-wise approaches employed in a one-to-many registration strategy.

1.2. Current work and contributions

In this study, we propose a Student's t-mixture model (TMM) based group-wise, rigid registration framework for unstructured point sets, to exploit the robustness of t-distributions to outliers and harness the generative nature of probabilistic model-based registration, to accommodate missing data. Most existing probabilistic point set registration approaches employ conventional GMMs as in Hufnagel et al. (2008) and Rasoulian et al. (2012) or mixture models that combine Gaussian components with a weighted uniform distribution component designed to model noise and outliers that may be present in the data, as in CPD (Myronenko and Song, 2010) and JRMPC (Evangelidis et al., 2014). Although the latter have been shown to outperform the former in the presence of outliers, a drawback of approaches such as CPD and JRMPC stems from the need for manually tuning the weight that controls the influence of the uniform distribution component relative to the Gaussian components when modelling data. Consequently, prior knowledge of the degree of noise and outliers present in the data being registered is often necessary. As this information is typically unavailable, a framework that is inherently robust to noise and outliers is desirable. TMM-based methods offer a suitable solution as demonstrated in two previous studies on pair-wise rigid (Gerogiannis et al., 2009) and non-rigid (Zhou et al., 2014) registration. We proposed two variants of TMM-based group-wise rigid registration recently, Ravikumar et al. (2016b) and Ravikumar et al. (2016a). In the former, a numerical approach was adopted to estimate the desired transformation parameters, via gradient-ascent optimisation, while in the latter these were estimated analytically by deriving closed-form expressions for the same. Furthermore, in Ravikumar et al. (2016a), we outlined a multi-resolution extension to the TMM algorithm (mrTMM), preliminary results of which showed significant improvement in registration accuracy and SSM quality relative to the single-resolution approach. In the present contribution we provide a complete set of derivations for estimating the mixture model and transformation parameters and conduct a comprehensive set of experiments on synthetic and clinical data to validate the proposed methods. We highlight their advantages over numerous state-of-the-art GMM-based approaches in terms of registration accuracy and the quality of SSMs trained. Further analyses of the proposed methods are also presented, investigating their convergence and ability to recover large rotational offsets.

2. Methods

As stated previously, Student's t-distributions (or t-distributions) are a robust alternative to Gaussian distributions when modelling data with outliers. A large body of literature exists on the use of GMMs for a variety of applications in medical image analysis, such as point set registration and generation of SSMs. There are however relatively few studies that investigate the use of a mixture of t-distributions for the same. TMMs have been used previously for clustering noisy data and shown to outperform GMMs due to their robust nature (Peel and McLachlan, 2000; Svensén and Bishop, 2005). Consequently, by formulating a t-mixture model (TMM) based group-wise registration framework to approximate the joint probability density of a group of point sets (representing shapes) and align them to a common reference frame, estimation of the desired transformations and soft-correspondences (across the group of shapes) is achieved with correspondingly greater degree of robustness to outliers. Such a group-wise framework allows for the unbiased estimation of a mean shape/mean model which is iteratively refined and aligned to each sample shape in a group. The estimated transformations aligning the mean model to each sample shape are subsequently used to robustly align all sample shapes in the group and the estimated mean model is used to compute the soft-correspodences (as in Hufnagel et al., 2008) necessary to train SSMs by PCA. In the context of training SSMs, optimal registration results in optimal models and consequently we argue that incorporating a higher degree of robustness in the registration step yields SSMs of higher quality.

2.1. Student's t-distribution

The Student's t-distribution $\mathcal{S}$, a generalization of the Gaussian distribution, can be expressed as an infinite mixture of (scaled) Gaussians, with identical means $\boldsymbol{\mu}$ but different covariances. t-distributions have heavier tails than Gaussians and thus are inherently more robust when fitting to data containing outliers (Bishop, 2006). For the multivariate case (Eq. (1b)), a t-distribution is derived by imposing a Gamma distribution $\mathcal{G}$ as a prior on the covariance $\boldsymbol{\Sigma}$[1] of a multivariate Gaussian distribution $\mathcal{N}$ and marginalising out the scaling weights u drawn from $\mathcal{G}$. This is achieved by evaluating the integral shown in Eq. (1a).

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

$$\mathcal{S}(\mathbf{x}|\boldsymbol{\mu}, \sigma^2, \nu) = \frac{\Gamma(\frac{\nu+D}{2})}{\Gamma(\nu/2)(\pi\nu\sigma^2)^{D/2}[1+\frac{\Delta^2}{\nu}]^{\frac{\nu+D}{2}}} \tag{1b}$$

[1] Here and throughout this paper we assume isotropic covariance i.e. $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$.

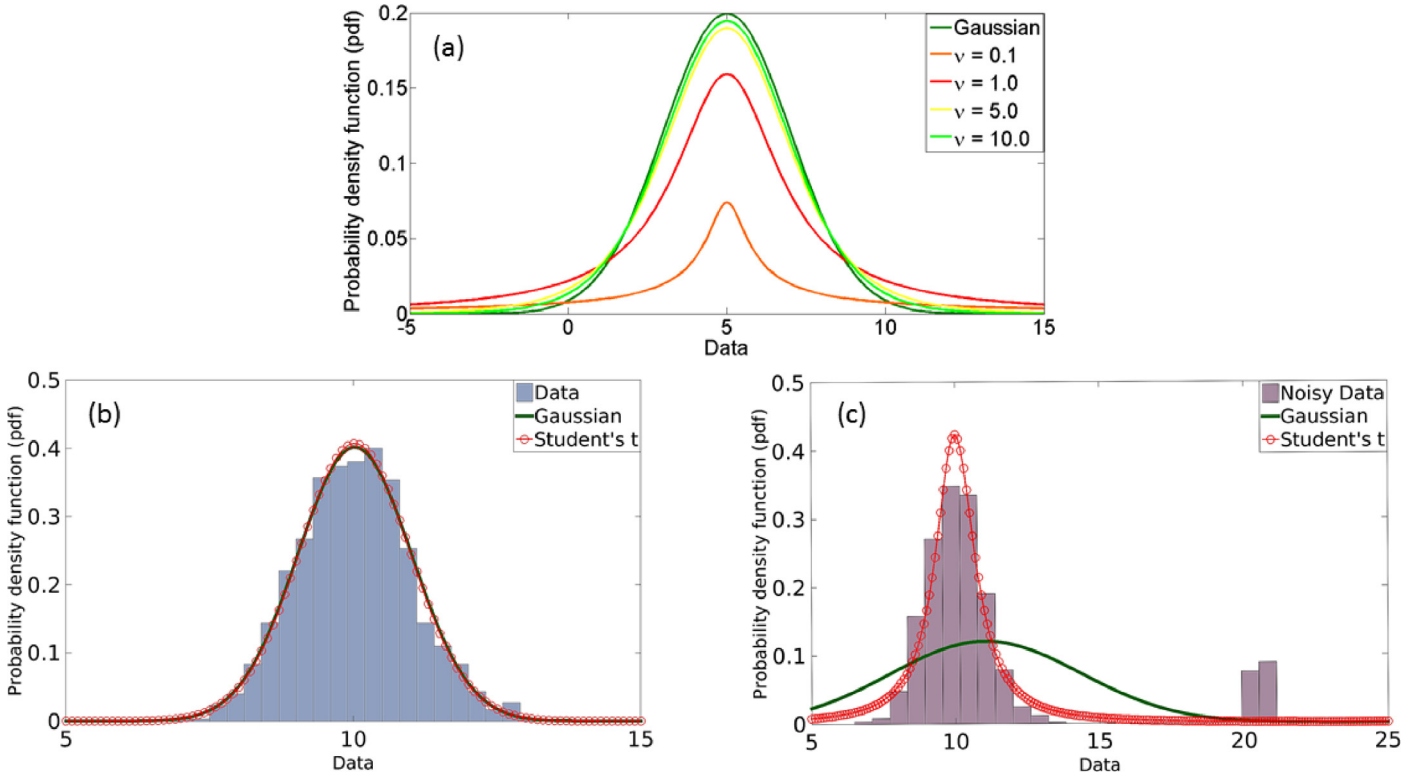

Fig. 1. (a) Plot depicting the influence of ν on the shape of t-distributions, showing increasing similarity to overlaid Gaussian distribution with increase in its magnitude. Maximum likelihood fits of univariate Gaussian and Student's t-distribution to uncorrupted data (b) and data corrupted by random noise (c) overlaid on their respective histogram distributions.

$$\Delta^2 = \frac{(\mathbf{x}-\boldsymbol{\mu})^T(\mathbf{x}-\boldsymbol{\mu})}{\sigma^2} \tag{1c}$$

where Δ^2 represents the squared Mahalanobis distance evaluated between the observed data $\mathbf{x}$ and a t-distribution centred at $\boldsymbol{\mu}$ with variance σ^2, Γ represents the gamma function and D the dimensionality of the data.

The t-distribution is parametrised by ν, which represent the number of degrees of freedom that control the shape of the distribution and its heavy-tails. In the limit $\nu \to \infty$ the t-distribution tends towards Gaussian behaviour and the effect of varying ν, for the univariate case, is further illustrated in Fig. 1a. To demonstrate the robust nature of t-distributions, univariate data sampled from a normal distribution and subsequently corrupted by the inclusion of random outliers, were fitted with a Gaussian and Student's t-distribution. The resulting probability density function (pdf) estimates for both distributions are very similar for data without outliers (Fig. 1b). However, as illustrated by Fig. 1c, the response of the Gaussian distribution is heavily distorted for data containing outliers while the t-distribution remains relatively unchanged and centred on the original, true mean value.

2.2. Group-wise point-set registration using TMMs

Mixture models are a weighted linear combination of probabilistic components, often used to approximate complex data distributions. Group-wise point set registration using mixture models is analogous to clustering data, except the data i.e. points representing each shape in the group, are considered to be transformed observations sampled from the model. Consequently, transformations that align the data are treated as model parameters (similar to the mean, variance and degrees of the mixture components).

Assuming a training set of K shapes ($k = 1 \dots K$) are transformed observations of a mixture model, with associated transformations $\mathcal{T}_k$, where $\mathbf{x}_{ki} \in \mathcal{X}_k$ represents the ith point ($i = 1 \dots N_k$) on the kth shape, there exists a t-distribution centred at $\mathcal{T}_k\boldsymbol{\mu}_j$, from which it is sampled. Additionally, all points $\mathbf{x}_{ki}$ on all K shapes in the group are assumed to be i.i.d. Henceforth subscript ($j = 1 \dots M$) is used to represent mixture components, $\boldsymbol{\mu}_j \in \mathcal{M}$ represents the centroids of the model $\mathcal{M}$ and $\mathcal{T}_k \in \mathbf{T}$ represents the similarity transformation, parametrised by rotation $\mathbf{R}_k$, scaling s_k and translation $\mathbf{t}_k$, that aligns the mean model $\mathcal{M}$ to the kth shape in the training set and $\mathcal{X}_k$ represents the set of all points on the kth shape in the training set. The conditional probability of a data point being sampled from a mixture component can thus be expressed as in Eq. (2a).

$$p(\mathbf{x}_{ki}|\mathcal{T}_k, \boldsymbol{\mu}_j, \sigma^2, \nu_j) = \mathcal{S}(\mathbf{x}_{ki}|\mathcal{T}_k\boldsymbol{\mu}_j, \sigma^2, \nu_j) \tag{2a}$$

$$p(\mathbf{x}_{ki}|\mathcal{T}_k, \mathcal{M}, \sigma^2, \Upsilon, \Pi) = \sum_{j=1}^{M} \pi_j \mathcal{S}(\mathbf{x}_{ki}|\mathcal{T}_k\boldsymbol{\mu}_j, \sigma^2, \nu_j) \tag{2b}$$

The conditional density for any data point $\mathbf{x}_{ki}$ on a training shape being sampled from the M-component mixture of t-distributions can subsequently be formulated using the sum rule of probability, as shown in Eq. (2b). Here $\nu_j \in \Upsilon$ represents the set of all degrees of freedom parameters in the mixture and $\pi_j \in \Pi$ represents the set of all mixture coefficients. Next, assuming that all data points on a training shape are independent and identically distributed (i.i.d) the joint probability density for all N_k points on the kth shape can be expressed as the product of the individual conditional densities, as described in Eq. (3a). Here $\{\mathcal{M}, \sigma^2, \Upsilon, \Pi\} \in \Theta$ represents the set of all model parameters.

$$p(\mathcal{X}_k|\mathcal{T}_k, \Theta) = \prod_{i=1}^{N_k} p(\mathbf{x}_{ki}|\mathcal{T}_k, \Theta) \tag{3a}$$

$$\log(p(\mathbf{X}|\mathbf{T},\Theta)) = \sum_{k=1}^{K} \log(p(\mathcal{X}_k|\mathcal{T}_k,\Theta)) \tag{3b}$$

$$\mathbf{T},\Theta = \arg\max_{\mathbf{T},\Theta} \log[p(\mathbf{T},\Theta|\mathbf{X})] \tag{3c}$$

Finally the log likelihood function of the complete training set $\mathcal{X}_k \in \mathbf{X}$ can be expressed in similar fashion assuming the K training shapes are i.i.d (Eq. (3b)). The optimal set of unknown parameters denoted $\Psi = \{\Theta, \mathbf{T}\}$ can be interpreted as those that maximise the posterior probability given by Eq. (3c) or conversely the log-likelihood in Eq. (3b). There is however, no closed-form solution to maximising Eq. (3b) and consequently, the conditional expectation of the complete data log-likelihood $\mathbf{Q}$ (refer to Eq. (4b)) is maximised iteratively with respect to each of the unknown parameters Ψ using the expectation-maximisation (EM) framework, based on Bayes' theorem. $\mathbf{Q}$ is derived (similar to Peel and McLachlan, 2000) by computing the conditional expectation of the complete data log-likelihood $\mathbf{L}$ (refer to Eq. (4a)) and treating the membership of data points $\mathbf{x}_{ki}$ to mixture components, and the covariance scaling weights of the latter, as latent variables $z_{kij} \in \mathbb{Z}$ and $u_{kij} \in \mathbb{U}$ respectively. The likelihood function is derived as a product of the marginal density of $\mathbb{Z}$, the conditional density of $\mathbb{U}$ given $\mathbb{Z}$, and the observed data $\mathbf{X}$ given $\mathbb{Z}$ and $\mathbb{U}$.

$$\begin{aligned}\mathbf{L}(\Psi) &= \log(p(\mathbf{X},\mathbb{U},\mathbb{Z}|\Psi)) = \log(p(\mathbb{Z}|\Psi)) + \log(p(\mathbb{U}|\mathbb{Z},\Psi)) \\ &\quad + \log(p(\mathbf{X}|\mathbb{U},\mathbb{Z},\Psi))\end{aligned} \tag{4a}$$

At the $(t+1)$th EM-iteration the current conditional expectation of the complete data log-likelihood, given the previous iteration's estimate for the model parameters Ψ^t, is expressed as:

$$\begin{aligned}&\mathbf{Q}(\Psi^{t+1}|\Psi^t) \\ &= \sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{j=1}^{M}\Bigg[P_{kij}^t\Bigg[\log(\pi_j) - \log\left(\Gamma\left(\frac{\nu_j}{2}\right)\right) + \frac{\nu_j}{2}\log\left(\frac{\nu_j}{2}\right) \\ &+ \frac{\nu_j}{2}\Bigg[\left(\log(U_{kij}^t) - U_{kij}^t\right) + \psi\left(\frac{\nu_j^t + D}{2}\right) - log\left(\frac{\nu_j^t + D}{2}\right)\Bigg] \\ &- \frac{D}{2}\log(2\pi) - \frac{1}{2}\log(\sigma^6) \\ &+ \frac{D}{2}\log(U_{kij}^t) - \frac{U_{kij}^t}{2}[1 + \frac{\left(\mathbf{x}_{ki} - \mathcal{T}_k\boldsymbol{\mu}_j\right)^T(\mathbf{x}_{ki} - \mathcal{T}_k\boldsymbol{\mu}_j)}{\sigma^2}]\Bigg]\Bigg],\end{aligned} \tag{4b}$$

where, Γ is the Gamma function, P_{kij} represents the posterior probability of an observed data point $\mathbf{x}_{ki}$ being drawn from a mixture component centred at $\boldsymbol{\mu}_j$ with ν_j degrees of freedom and U_{kij} represents the scaling weights of the equivalent Gaussian distribution (i.e. these are derived from the expression of multi-variate t-distributions as an infinite mixture of scaled Gaussians as discussed in Section 2.1). The EM algorithm iteratively alternates between two steps:

(1) In the expectation (E)-step, the product of the conditional expectations of the two latent variables $\mathbb{Z} = \{z_{kij}\}$ and $\mathbb{U} = \{u_{kij}\}$, are computed given an estimate of the unknown parameters Ψ. This results in a corrected set of posterior probabilities $P_{kij}^{\star}$ (as shown in Eq. (5c)), which represent robust correspondence probabilities between points on each shape and the mixture centroids. These are subsequently employed in the M-step to update estimates for the unknown parameters Ψ. On the $(t+1)$th EM-iteration, the expectations of the latent variables are computed as follows:

$$E_{\Psi^{(t)}}(z_{kij}|\mathbf{x}_{ki}) = P_{kij}^{(t)} = \frac{\pi_j \mathcal{S}(\mathbf{x}_{ki}|\mathcal{T}_k\boldsymbol{\mu}_j, \sigma^2, \nu_j)}{\sum_{j=1}^{M}\pi_j \mathcal{S}(\mathbf{x}_{ki}|\mathcal{T}_k\boldsymbol{\mu}_j, \sigma^2, \nu_j)} \tag{5a}$$

$$E_{\Psi^{(t)}}(u_{kij}|\mathbf{x}_{ki}, z_{kij}=1) = U_{kij}^{(t)} = \frac{\nu_j + D}{\nu_j + \Delta_{kij}^2} \tag{5b}$$

$$E_{\Psi^{(t)}}(z_{kij}|\mathbf{x}_{ki})E_{\Psi^{(t)}}(u_{kij}|\mathbf{x}_{ki}, z_{kij}=1) = P_{kij}^{\star(t)} = P_{kij}^{(t)}U_{kij}^{(t)} \tag{5c}$$

Δ_{kij}^2 is the squared Mahalanobis distance, defined previously in Eq. (1c).

(2) The maximisation (M)-step involves estimation of the model and transformation parameters by maximising $\mathbf{Q}$ (refer to Eq. (4b)) with respect to each unknown parameter, sequentially. Expressions for updating all parameters except ν_j are derived analytically (shown in Eqs. (6) – (12)). Differentiating $\mathbf{Q}$ with respect to ν_j results in a non-linear equation that is solved using an iterative root finding technique such as Newton's method, for each component in the mixture.

$$\boldsymbol{\mu}_j^{(t+1)} = \frac{\sum_{k=1}^{K}\sum_{i=1}^{N_k} P_{kij}^{\star(t)}(s_k^{-1}\mathbf{R}_k^T(\mathbf{x}_k - \mathbf{t}_k))}{\sum_{k=1}^{K}\sum_{i=1}^{N_k} P_{kij}^{\star}} \tag{6}$$

$$\sigma_{(t+1)}^2 = \frac{1}{ND}\sum_{k=1}^{K}\sum_{i=1}^{N_k}\sum_{j=1}^{M} P_{kij}^{\star(t)}\|\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j - \mathbf{t}_k\|^2 \tag{7}$$

In Eqs. (7) and (9), $N = \sum_{k=1}^{K}\sum_{i=1}^{N_k}\sum_{j=1}^{M} P_{kij}$ is the total number of data points in the training set. From an implementation point of view it is important to note, in Eqs. (6) and (7), $\boldsymbol{\mu}_j^{(t+1)}$ and $\sigma_{(t+1)}^2$ are updated using current estimates for the transformation parameters, i.e. $\mathcal{T}_k^{t+1}$, which are updated prior to the mixture model parameters Θ, at each EM-iteration. As there exists no closed-form expression to estimate $\nu_{j=1..M} \in \Upsilon$, they are computed by solving Eq. (8) using Newton's method:

$$\begin{aligned}&-\psi\left(\frac{\nu_j}{2}\right) + \log\left(\frac{\nu_j}{2}\right) + 1 + \frac{\sum_{k=1}^{K}\sum_{i=1}^{N_k} P_{kij}^{(t)}\left(\log(U_{kij}^{(t)}) - U_{kij}^{(t)}\right)}{\sum_{k=1}^{K}\sum_{i=1}^{N_k} P_{kij}^{(t)}} \\ &+ \psi\left(\frac{\nu_j^{(t)} + D}{2}\right) - \log\left(\frac{\nu_j^{(t)} + D}{2}\right) = 0\end{aligned} \tag{8}$$

The maximum likelihood estimate for mixture coefficients π_j is given by:

$$\pi_j^{(t+1)} = \frac{\sum_{k=1}^{K}\sum_{i=1}^{N_k} P_{kij}^{(t)}}{N} \tag{9}$$

The transformation parameters for each shape in the training set are estimated as follows:

$$\texttt{Rotation}: \mathbf{R}_k^{(t+1)} = \mathbf{USV}^T \tag{10}$$

$$\texttt{Scaling}: s_k^{(t+1)} = \frac{\text{tr}\left\{\sum_{i=1}^{N_k}\sum_{j=1}^{M} P_{kij}^{\star(t)}(\boldsymbol{\mu}_j - \mathbf{m}_k)(\mathbf{x}_{ki} - \mathbf{d}_k)^T\mathbf{R}_k\right\}}{\text{tr}\left\{\sum_{i=1}^{N_k}\sum_{j=1}^{M} P_{kij}^{\star(t)}(\boldsymbol{\mu}_j - \mathbf{m}_k)(\boldsymbol{\mu}_j - \mathbf{m}_k)^T\right\}} \tag{11}$$

$$\texttt{Translation}: \mathbf{t}_k^{(t+1)} = \mathbf{d}_k^{(t+1)} - s_k \mathbf{R}_k \mathbf{m}_k^{(t+1)} \tag{12}$$

In Eq. (10) **U,V** are unitary matrices estimated by singular value decomposition of a matrix $\mathbf{C}_k$, computed as:

$$\mathbf{C}_k^{(t+1)} = \sum_{i=1}^{N_k} \sum_{j=1}^{M} P_{kij}^{\star(t)} (\mathbf{x}_{ki} - \mathbf{d}_k)(\boldsymbol{\mu}_j - \mathbf{m}_k)^T, \tag{13}$$

and $\mathbf{S}$ is a diagonal matrix given by, $\mathbf{S} = diag(1, 1, det(\mathbf{UV}^T))$, used to enforce estimation of strictly orthogonal rotation matrices $\mathbf{R}_k$, whilst avoiding reflections (similar to Gooya et al., 2015a). In Eqs. (11) – (13), $\mathbf{d}_k$ and $\mathbf{m}_k$ represent weighted centroids/barycenters expressed as shown in Eqs. (14) - (15).

$$\mathbf{d}_k^{(t+1)} = \frac{\sum_{i=1}^{N_k} \sum_{j=1}^{M} P_{kij}^{\star(t)} \mathbf{x}_{ki}}{\sum_{i=1}^{N_k} \sum_{j=1}^{M} P_{kij}^{\star(t)}} \tag{14}$$

$$\mathbf{m}_k^{(t+1)} = \frac{\sum_{i=1}^{N_k} \sum_{j=1}^{M} P_{kij}^{\star(t)} \boldsymbol{\mu}_j}{\sum_{i=1}^{N_k} \sum_{j=1}^{M} P_{kij}^{\star(t)}} \tag{15}$$

Derivations for all parameters presented in Eqs. (6) – (12) are included in Appendix A. The EM algorithm is summarised by the pseudo-code presented in Algorithm 1.

Algorithm 1 TMM.

```
Inputs: Group of shapes X_{k=1..K}, number of mixture components M, max.iterations
Outputs: Set of similarity transformations T_k, aligned soft-correspondences, mean shape M
INITIALIZATION
Initialize M, σ² using K-means clustering, π_j = 1/M and ν_j = 3.0
All π_j = 1/M and ν_j = 3.0
procedure EM(X_k, M, σ², Υ, Π, T_k)                ▷ EM initialized
    while Iteration < max.iterations do
        Compute P*_kij = P_kij U_kij                ▷ E-step
        Update R_k, s_k, t_k                        ▷ M-step
        Update M, σ², Π and Υ                       ▷ M-step
    end while
    return M, σ², Υ, Π, T_k
end procedure
```

2.3. Multi-resolution registration

Registration algorithms in general, often suffer from convergence to local minima, resulting in sub-optimal solutions. In image registration, this has been addressed previously by adopting a hierarchical multi-resolution registration approach that operates in a coarse-to-fine fashion and thereby reduces the chances of local minima entrapment (Rueckert et al., 1999). Frangi et al. (2002) proposed a multi-resolution non-rigid B-spline registration framework for automatic landmarking (and correspondence estimation) of multi-object shape ensembles via an atlas-to-training-set registration strategy, for the purpose of training SSMs. Such an approach however, can be computationally expensive in the case of large data sets and requires construction of an unbiased atlas. A group-wise multi-resolution approach is novel in the context of point set registration and was proposed in our recent work (Ravikumar et al., 2016a), although, a similar approach (multi-scale EM-ICP) was proposed previously by Granger and Pennec (2002). The main differences between multi-scale EM-ICP and our method are — (a) the former is a pair-wise registration approach while mrTMM is group-wise. The latter consequently enables estimation of a mean shape, correspondences and transformations, in an unbiased manner; (b) multi-scale EM-ICP assumes uniform priors on the matches while mrTMM revises estimates for the mixture coefficients at each iteration; (c) in order to reject outliers, multi-scale EM-ICP chooses an ad-hoc threshold on the Mahalanobis distance and assigns a null weight for model points farther away, while no such threshold needs to be defined for mrTMM as it is inherently robust to outliers, due to its constituent heavy-tailed t-distributions; and (d) in multi-scale EM-ICP the 'scene' point set is decimated at each scale (or variance) and the latter is reduced with each iteration using an annealing scheme. With mrTMM the âsceneâ point sets are left untouched and the mean modelâs density is increased adaptively at each successive resolution. While the former may be suitable for pair-wise registration applications, it could lead to over-fitting of the model in the context of group-wise registration, as the main benefit of the approach arises from starting at a high scale, leading to substantial decimation of the 'scene' point set. Additionally, such an approach would reduce the degree of shape variability captured by the SSMs trained following registration (the main application of interest in this study). Consequently, mrTMM is more suitable for our application.

By embedding the TMM-based registration framework within a multi-resolution scheme (abbreviated as mrTMM), the influence of local minima during registration is reduced. mrTMM begins with a low density mean model (i.e. few mixture components) which is iteratively refined and upsampled at each successive resolution, through a process of 'adaptive sampling' from the mixture components. The transformations computed at each level are used to initialize the subsequent resolution and the overall model variance is decreased at each successive level by populating the mean model with new points. This reduction in model variance at each successive resolution refines the estimated transformations and improves registration accuracy. 'Adaptive sampling' to increase mean model density is achieved by imposing a multinomial distribution over the estimated mixture coefficients π_j and generating random samples S from those t-components in the mixture model that have a high probability in explaining the observed data, i.e. s^n new model points are sampled from the jth mixture component, subject to the constraints $\sum_{j=1}^{M} s_j^n = S$ and $\sum_{j=1}^{M} \pi_j = 1$. S is a user-specified parameter and in this study we fixed $S^r = M^{r-1}$, where r represents the current resolution level, for all experiments (i.e. M is doubled at each successive resolution). The number of new model points sampled from each mixture component is described by Eq. (16a). Random samples are drawn from a zero-centered Gaussian distribution and an inverse χ^2-distribution with ν_j degrees of freedom, to generate new model points. This is because t-distributed random variables can conveniently be expressed as shown in Eq. (16b), where $\boldsymbol{\mu}_j^n$ represents the nth model point sampled from $\boldsymbol{\mu}_j$, the centroid of the jth mixture component.

$$p(s_j^n | \pi_j, S) = \frac{S!}{\prod_{n=1}^{S} s_j^n!} \prod_{j=1}^{M} \pi_j \tag{16a}$$

$$\boldsymbol{\mu}_j^n = \boldsymbol{\mu}_j + \mathcal{N}(0, \sigma^2) \sqrt{\frac{\nu_j}{\chi^2(\nu_j)}} \tag{16b}$$

Such a multi-resolution approach reduces the influence of local minima on the registration process, which may be introduced during model initialisation by k-means clustering (or similar processes) or during estimation of the model parameters. The mrTMM framework is further described by Algorithm 2.

Algorithm 2 mrTMM.

Inputs: Group of shapes $\mathcal{X}_{k=1..K}$, number of mixture components M, max.resolutions, max.iterations
Outputs: Set of similarity transformations $\mathcal{T}_k$, aligned soft-correspondences, mean shape $\mathcal{M}$

INITIALIZATION
Initialize $\mathcal{M}, \sigma^2$ using K-means clustering, π_j = 1/M and ν_j = 3.0
procedure EM($\mathcal{X}_k, \mathcal{M}, \sigma^2, \Upsilon, \Pi, \mathcal{T}_k$) ▷ EM initialized
  while Resolution < max.resolutions **do**
    while Iteration < max.iterations **do**
      Compute $P^{\star}_{kij} = P_{kij} U_{kij}$ ▷ E-step
      Update $\mathbf{R}_k, s_k, \mathbf{t}_k$ ▷ M-step
      Update $\mathcal{M}$, σ^2, Π and Υ ▷ M-step
    end while
    Compute s^n_j to be sampled from each μ_j
    Adaptively sample M new model points
    Re-initialize all $\pi^{new}_j = 1/M_{new}$ and $\nu^{new}_j = 3.0$, ν^{old}_j retained
    return $\mathcal{M}^{new}, \sigma^2_{old}, \Upsilon^{new}, \Pi^{new}, \mathcal{T}^{old}_k$
  end while
end procedure

2.4. SSM generation and model-fitting

Following alignment of a group of shapes using the proposed methods, the estimated soft-correspondences are used to train SSMs by PCA. The resulting eigenvectors $\Phi \in \mathbb{R}^{D\times M\times m}$ (where m is the number of modes of variation retained), represent the principal axes of the shape space and eigenvalues λ_m describe the proportion of the total variation in shape described by each corresponding mode of variation. All SSMs trained in this study retained eigenmodes that describe 95% of the total variation in shape across each corresponding group. The process of fitting the trained models to new data involves two steps: (1) mixture-fitting and (2) SSM-fitting. The former is first used to align the new shape to the trained mean model and establish correspondences. This step is analogous to pair-wise registration, where the learnt mixture model parameters, apart from the variance, (i.e. mixture centroids, coefficients and degrees of freedom) remain fixed, as the trained mean model is iteratively aligned to the test shape. The estimated correspondences are subsequently projected to the trained SSM (refer to Appendix B for more details), to obtain model-predicted estimates for new shapes.

2.5. Evaluation and assessment

The proposed methods are validated using both synthetic and clinical data, and compared with four state-of-the-art point set registration methods: rigid-CPD, SpSSM, JRMPC and a group-wise GMM-based method (denoted GMM) similar to EM-ICP proposed by Hufnagel et al. (2008). Our implementation of GMM, however, explicitly estimates the variance of the mixture model at each iteration as the registration progresses, while EM-ICP heuristically decreases the same with each successive iteration. The difference between GMM and SpSSM lies in the estimation of the mixture coefficients. The former employs classical maximum likelihood estimation, while the latter, as discussed in Section 1.1, uses a conjugate prior and opts for Bayesian estimation. The original pair-wise, rigid-CPD algorithm (Myronenko and Song, 2010) is used for comparison with the synthetic data as instances in the group are transformed and modified versions of the raw bunny point set. For the clinical data however, we opt for a group-wise version of CPD to enable direct comparison with the other methods.

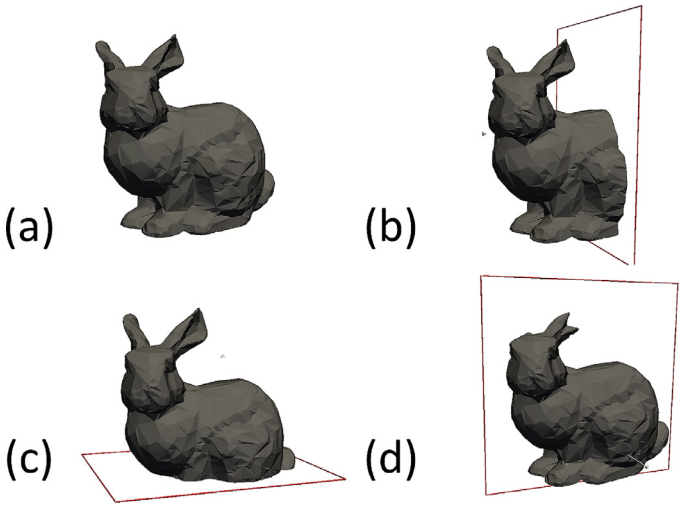

Fig. 2. 3D bunny data set: (a) decimated original surface mesh; sample (a) cropped along: (b) yz-plane, (c) xz-plane and (d) xy-plane.

The synthetic data set was generated using the 3D Stanford bunny point set.[2] It comprises of the original point set and three modified and transformed copies, generated as follows: (1) original bunny point set was cropped along the xy-, yz- and xz-planes to generate three distinct samples with missing information at different spatial locations (depicted in Fig. 2(b–d)), (2) rigid transformations (i.e. only translation and rotation) were subsequently applied to the cropped samples and (3) all four point sets were finally corrupted by the addition of uniformly distributed outliers and Gaussian noise to varying degrees. Table 1 describes the degree of noise and outliers applied to each sample in both synthetic data sets and their corresponding ground truth transformations. The resulting data sets are depicted in Fig. 3. The synthetic data were generated in this manner to evaluate the ability of the proposed methods to accurately align shapes in the presence of: (1) large rotational offsets with minimal overlap between samples, (2) missing information and (3) varying degrees of outliers.

Four clinical data sets were also used to validate the proposed methods: (a) `Femur`: 2D set of ($K = 1000$) femoral heads segmented automatically from dual energy X-ray absorptiometry (DXA) images (depicted in Fig. 4), using Hologic Apex 3.2, a software frequently employed in the clinic. DXA images and their segmentations were acquired in a previous study (McCloskey et al., 2007); (b) `Hippocampus_Ctrl`: 3D set of hippocampi segmented automatically from T1-weighted magnetic resonance (MR) images of healthy subjects ($K = 50$); (c) `Hippocampus_MCI`: 3D set as for (b), but acquired from patients diagnosed with MCI ($K = 28$) (examples shown in Fig. 5). For (b) and (c), MR images were acquired as part of the VPH-DARE@IT project[3] and the automatic segmentation tool, based on shape-constrained deformable models, was provided by Philips Research Laboratories, Hamburg, Germany (Zagorchev et al., 2011, 2016); and (d) `Heart`: 3D set of hearts comprising the epicardium and endocardium for both left and right ventricles ($K = 30$). Training segmentations for the heart were generated from short-axis MR images of healthy subjects ($K = 10$), patients diagnosed with pulmonary hypertension (PH, $K = 10$) and patients diagnosed with hypertrophic cardiomyopathy (HCM, $K = 10$), as part of a previous study (Albà et al., 2014). With (d), samples from all three diagnostic groups were pooled into a single

[2] The Stanford 3D scanning repository. Available at: http://graphics.stanford.edu/data/3Dscanrep/.
[3] http://www.vph-dare.eu/.

Table 1
Rigid transformations and degree of outliers used to generate bunny data set.

Sample	N_k	R_k^g	t_k^g (cm)	Gaussian noise(%)	Random outliers(%)
1	2420	-	-	13	2.50
2	1883	x: 64°, y: 22.50°	x: 0.20, y: −0.30, z: 0.50	9	4
3	1889	y: 50°, z: 20°	x: −0.25, y: 0.35, z: −0.15	14	6
4	1658	z: 60°, x: 18°	x: −0.10, y: −0.50, z: 0.40	11	5

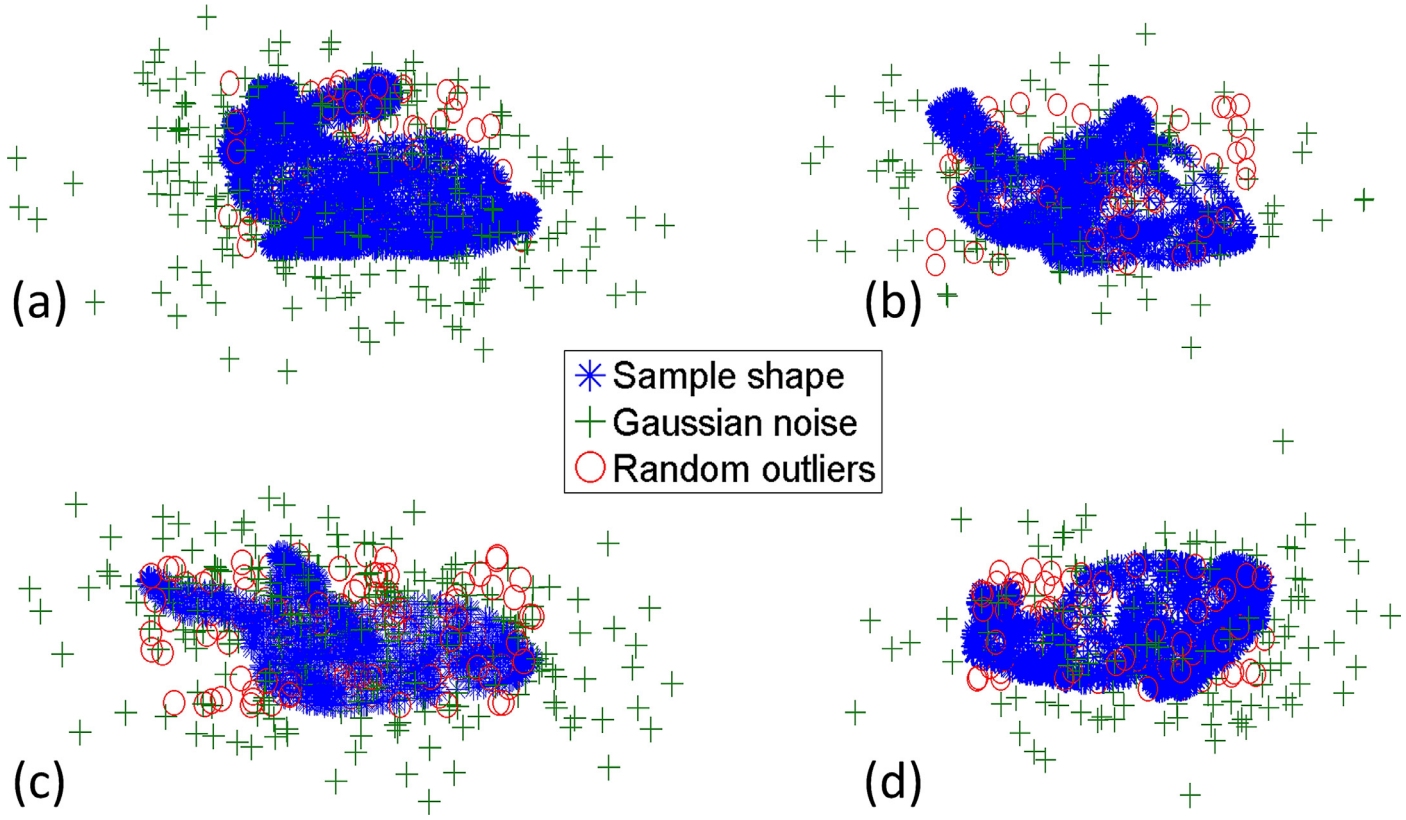

Fig. 3. Transformed bunny data set comprising four samples (blue). Samples (b–d) generated by rigidly transforming sample (a). All samples corrupted by varying proportions of Gaussian noise (green) and uniformly-distributed outliers (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

data set and used to validate the proposed methods. This strategy is adopted for the heart data set to assess the ability of the proposed methods to accurately align and generate high quality SSMs in the presence of significant variations in geometry across training shapes. Additionally, the heart comprising both ventricles and the endo-/epi-cardium is used for validation due to its topological complexity (i.e. not homeomorphic to a sphere), to illustrate the independence of the proposed framework to topology. Comparisons are made with the state-of-the-art in terms of registration accuracy for both synthetic and clinical data, and in terms of the quality of SSMs trained, using the latter.

2.5.1. Registration accuracy

As ground truth transformations were available for the synthetic data set, registration accuracy was assessed by evaluating the root-mean-squared-error (RMSE) (similar to Evangelidis et al., 2014; Huynh, 2009) of the estimated rotation matrices, relative to the ground truth rotations. RMSE was computed as described by Eq. (17a), where $||\cdot||_F$ denotes the Frobenius norm and $\mathbf{R}_k^g$ and $\mathbf{R}_k$ represent the ground truth rotation applied to the kth sample in the group and the corresponding rotation matrix estimated for the sample, respectively. As the synthetic data set is generated by rigidly transforming the bunny point set (denoted sample 1, refer to Table 1), the estimated rotations for the remaining samples (relative to the mean) are transformed to the coordinate frame of sample 1 to enable direct comparison with their corresponding ground truths. This is achieved by computing the product of $\mathbf{R}_{k=2,3,4}$ and the inverse of the rotation estimated for sample 1, denoted $\mathbf{R}_1^T$ in Eq. (17a). Additionally, the intrinsic distance between the estimated and ground truth rotations (Huynh, 2009) were also evaluated as shown in Eq. (17b), for easier interpretation of the rotation errors (θ_{err}), in terms radians/degrees.

$$RMSE = ||\mathbf{R}_k^g - \mathbf{R}_k\mathbf{R}_1^T||_F \tag{17a}$$

$$\theta_{err} = \arccos\left[\frac{\mathrm{tr}\ ((\mathbf{R}_k^g(\mathbf{R}_k\mathbf{R}_1^T)^T)-1}{2}\right] \tag{17b}$$

Alignment accuracy was also evaluated for all four clinical data sets, using the Hausdorff (HD) and mean surface distance (MSD) metrics. HD and MSD measures (formulated as shown in Appendix B) were computed between the aligned soft-correspondences estimated for each sample in the group and the corresponding mean shape estimated for the group. For the data set of hippocampi, alignments were performed independently for the healthy and MCI samples (as separate SSMs are desired) and consequently, registration accuracy was evaluated separately for the two hippocampi groups.

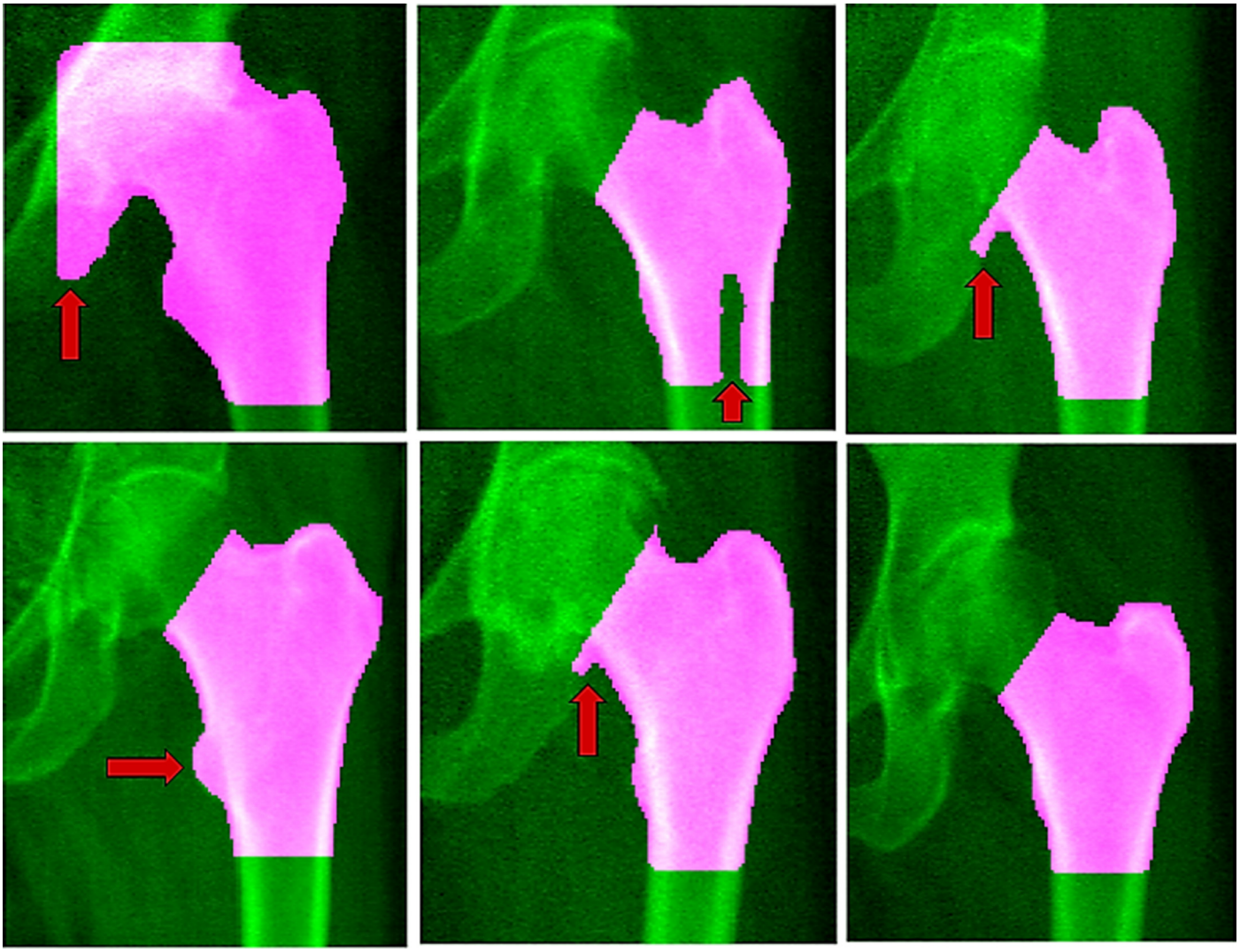

Fig. 4. Raw DXA images from the femur data set overlaid with their respective boundary masks. Red arrows indicate regions with over- or under-segmented boundaries, which result in point sets with varying degrees of outliers.

2.5.2. SSM quality

The quality of SSMs trained using each clinical data set is assessed based on their generality, specificity and compactness (Davies et al., 2010). Generalisation and specificity errors are evaluated using the MSD metric, computed between the ground truth test shape and the corresponding model-predicted shape. These measures were computed for models trained using: single- and multi-resolution TMM, SpSSM, a group-wise variant of CPD and GMM. We employ our own implementation of group-wise CPD (gCPD) rather than JRMPC as the latter estimates strictly rigid transformations, rather than the desired similarity transformations. Furthermore, gCPD is preferred to the original pair-wise approach to enable direct comparison with the other methods investigated. SpSSM and GMM are purely Gaussian-based and differ in the manner of estimation of the mixture coefficients, while gCPD incorporates a uniform distribution component in the mixture model. Together they represent a range of GMM-based, rigid, group-wise point set registration techniques proposed in recent years, suitable for validating the advantage of the proposed t-mixture model based methods.

Generalisation quantifies the ability of SSMs to reconstruct unseen shapes i.e. samples excluded from the training set. Compactness measures their ability to describe variation in shape across a group, with a minimal set of parameters i.e. the fewest modes of variation. It can also formally be defined as the cumulative sum of eigenvalues associated with the modes of variation. It is also crucial for SSMs to generate anatomically plausible instances and consequently, this is used as measure of their quality known as specificity (Davies et al., 2010).

In order to avoid over- or under-fitting to data, it is necessary to balance model complexity with performance. To this end, we identified the optimal (or rather suitable) number of mixture components (M_{opt}) for each clinical data set by conducting ten-fold cross validation experiments, evaluating the quality of SSMs trained. It is important to note that we chose M_{opt} by considering a trade off between reconstruction accuracy and model complexity and consequently, M_{opt} is not the true optimal model complexity for each data set. For the generalization experiments, 10 unseen test shapes from the same cohort (as the training samples) were used for the femur data set. For the hippocampi, 10 test shapes for both healthy and MCI groups were generated automatically from a separate database of MR images (also acquired as part of the VPH-DARE@IT project[4]), using a different state-of-the-art segmentation tool based on geodesic information flows (GIF parcellation)

[4] http://www.vph-dare.eu/.

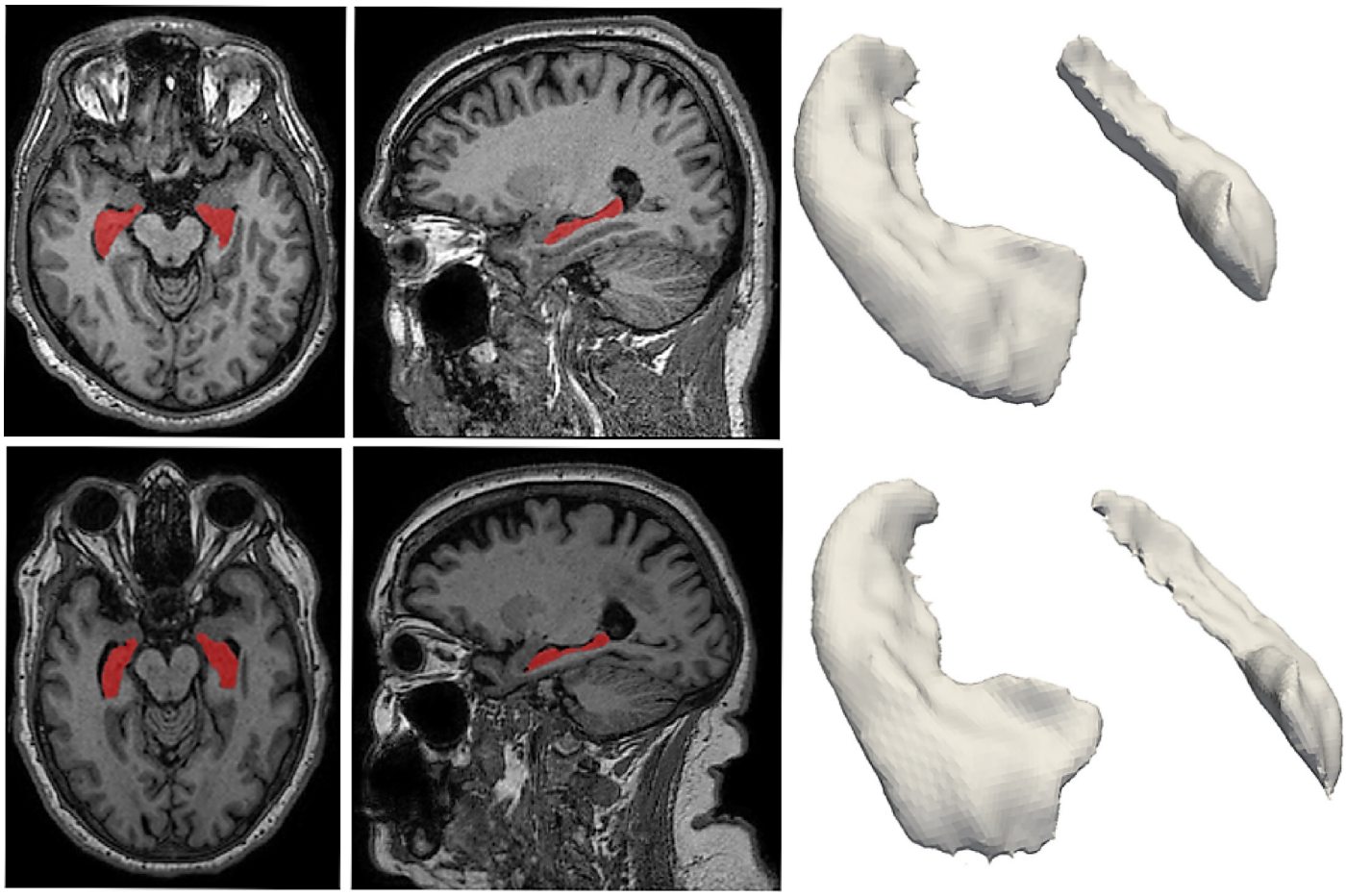

Fig. 5. Hippocampi automatically segmented from MR images of a healthy subject (top row) and MCI patient (bottom row). Axial and saggital view of segmentations overlaid on their respective raw images are shown in the left and centre columns respectively and the surfaces generated from these are depicted in the column on the right.

(Cardoso et al., 2015; Prados Carrasco et al., 2016). This method of validation for the hippocampi was selected to evaluate the ability of the proposed framework to characterise unseen shapes generated using a different protocol, to better emulate a real clinical scenario. Finally, for the heart data set 10 unseen shapes were selected, comprising three samples from the PH and HCM patient groups respectively, and four from the healthy cohort.

The quality of SSMs trained using the identified M_{opt} for each data set, are also assessed with respect to the number of modes of variation by leave-one-out full-fold cross validation. This second set of cross validation experiments assesses the quality of the trained models to characterise unseen shapes from the same cohort as the training sets. It is important to note that, correspondences were estimated jointly across both training and test shapes in all experiments evaluating generalization with respect to the number of modes of variation. Specificity and compactness are also assessed in this manner by leave-one-out full-fold cross-validation. In the former case, two shapes are randomly sampled from the trained SSMs by sampling PCA scores from a Gaussian distribution, within the range of valid parameters $[-3\sqrt{\lambda_m}, 3\sqrt{\lambda_m}]$ (where λ_m represents the eigenvalue of the mth mode of variation), using progressively increasing number of eigenmodes, similarly to Davies et al. (2010) and Gooya et al. (2015a). Subsequently, their mean surface distance to each left out case from the training set is computed. Compactness is assessed by plotting the cumulative sum of the percentage of variation (computed using the estimated eigenvalues) described by each eigenmode, against the modes of variation.

Table 2
RMSE values computed between estimated and ground truth rotations for 3D bunny data set.

Method	Sample 2	Sample 3	Sample 4	Mean RMSE
CPD	0.1781	0.1021	0.1841	0.1548 ± 0.05
SpSSM	1.5133	0.0944	0.0700	0.5592 ± 0.83
GMM	1.2156	0.8260	1.2786	1.1067 ± 0.24
JRMPC	1.8541	0.0022	0.0011	0.6191 ± 1.07
TMM	0.0232	0.0260	0.0287	**0.026 ±0.003**
mrTMM	0.0012	0.0031	0.0016	**0.002 ±0.001**

3. Results and discussion

3.1. Registration accuracy

3.1.1. Synthetic data

Table 2 summarizes the alignment accuracy of each method investigated, for the synthetic data. The RMSE values for each transformed sample (i.e. samples (b–d) in Fig. 3) indicate that the proposed methods, TMM and mrTMM, achieve significantly lower registration errors than CPD, SpSSM and GMM across all three samples. Although JRMPC shows good robustness and achieves marginally lower errors for samples 3 and 4 (relative to TMM and mrTMM), it is unable to recover the rotation for sample 2 and results in significantly higher errors. The number of model points (user-specified parameter) was set to 940, which is 50% of the median cardinality of the synthetic data set, for each group-wise registration method investigated. In the case of CPD, a pair-wise ap-

Table 3
Intrinsic rotation errors evaluated in terms of radians (Rad.) and degrees (Deg.) between estimated and ground truth rotations.

Method	Sample 2		Sample 3		Sample 4		Mean Error	
	Rad.	Deg.	Rad.	Deg.	Rad.	Deg.	Rad.	Deg.
CPD	0.128	7.327	0.072	4.139	0.155	8.886	0.118 ± 0.04	6.784 ± 2.42
SpSSM	1.129	64.69	0.067	3.824	0.049	2.838	0.415 ± 0.62	23.78 ± 35.43
GMM	0.903	51.74	0.623	35.70	0.943	54.05	0.823 ± 0.17	47.16 ± 9.99
JRMPC	1.5103	86.53	5×10^{-4}	0.029	0.002	0.134	0.50 ± 0.87	28.90 ± 49.91
TMM	0.015	0.838	0.015	0.887	0.019	1.107	**0.016** ± **0.002**	**0.944** ± **0.14**
mrTMM	0.001	0.06	0.0024	0.139	0.001	0.066	**0.001** ± **5.8** $\times$**10**$^{-4}$	**0.09** ± **0.04**

proach, this is determined by the number of points used to represent the 'moving' point set. As outlined previously in Section 1.2, JRMPC and CPD require a user-specified parameter that controls the weight of the uniform distribution component in the mixture model and consequently the degree of robustness of the model to noise and outliers. Values in the range of (0.1–0.9) were tested and the value returning the lowest registration errors (reported in Table 2) was considered optimal for the data set. Mean RMSE errors are computed across samples and used to compare each registration method (using a paired sampled t-test, considering a significance level of 1%). TMM and mrTMM achieve significantly lower errors, highlighted in bold in Table 2.

The RMSE values presented in Table 2 and the intrinsic rotation errors in Table 3 indicate that SpSSM and JRMPC are able to recover the applied rotations to a high degree of accuracy for samples 3 and 4, but fail to do so for sample 2. This can be attributed to the smaller capture range of group-wise GMM-based methods in general, in comparison to CPD and the proposed TMM-based methods. CPD is able to recover synthetic rotations to a moderate degree of accuracy and shows good robustness to noise and outliers. It is interesting to note that JRMPC was shown to outperform CPD in Evangelidis et al. (2014) when the applied synthetic rotations were of lower magnitude (maximum of 30°). This supports our observation that although JRMPC shows good robustness to noise and outliers, it lacks the ability to recover large rotational offsets. This hypothesis was further validated by conducting an additional experiment using the bunny data set where the cropped samples (refer to Fig. 2(b–d)) were rotated to larger degrees, without the inclusion of synthetic noise and outliers. The proposed methods were able to recover rotations in the range of $[-60°, 60°]$, applied along multiple axes, while CPD and JRMPC failed to do so. This is visually described and verified by the images shown in Fig. 6. The ability of the proposed methods and the state-of-the-art to recover the applied translations was also evaluated. A description of this analysis together with the translation errors estimated, is provided in Appendix B.

The proposed TMM-based methods offered substantial improvements over GMM-based approaches in the synthetic data experiments, as they are more robust to noise and outliers, have a wider capture range for recovering rotational offsets, do not require any prior knowledge of the degree of outliers present in the data and correspondingly, require fewer user-specified parameters than CPD and JRMPC. Rotation errors evaluated in terms of radians and degrees (Table 3) complement and are consistent with the RMSE values reported in Table 2. The proposed methods achieved significantly lower angular errors, evaluated between the estimated and ground truth rotations, relative to the state of the art. Additionally, in the presence of significant outliers and large rotational offsets, mrTMM offers substantial improvement over TMM in this regard.

Table 4
Alignment accuracy for MCI-hippocampi data set.

Method	M	`Hippocampus_MCI` ($K = 28$)	
		HD (mm)	MSD (mm)
gCPD	1280	**3.32 ± 1.17**	**0.61 ± 0.16**
SpSSM	1906	5.80 ± 1.11	0.80 ± 0.16
GMM	1280	5.60 ± 1.12	0.86 ± 0.20
TMM	1280	**3.39 ± 1.25**	**0.62 ± 0.17**
mrTMM	1280	**3.30 ± 1.32**	**0.58 ± 0.16**

Table 5
Alignment accuracy for control-hippocampi data set.

Method	M	`Hippocampus_Ctrl` ($K = 50$)	
		HD (mm)	MSD (mm)
gCPD	1280	**3.38 ± 1.06**	**0.62 ± 0.13**
SpSSM	787	6.25 ± 1.32	0.94 ± 0.24
GMM	1280	7.56 ± 1.22	0.88 ± 0.22
TMM	1280	**3.26 ± 1.02**	**0.62 ± 0.14**
mrTMM	1280	**3.25 ± 1.10**	**0.61 ± 0.14**

3.1.2. Clinical data

Alignment accuracy of the proposed methods was also compared with the state of the art using clinical data. The raw point sets from each data set are shown in Fig. 7 (a,e,i,m). Fig. 7 (b,f,j,n) depict the mean shapes estimated for each group using mrTMM and Fig. 7 (c,g,k,o) and Fig. 7 (d,h,l,p) represent the corresponding aligned shapes and soft-correspondences, respectively, for each clinical data set. The significant level of outliers present in the femur data set is evident in Fig. 7 (m), while Fig. 7 (n) demonstrates the ability of mrTMM (similar result obtained for TMM as well) to estimate a valid mean shape in the presence of such outliers. Alignment accuracy was quantified using the HD and MSD measures, presented in Tables 4,5 for the hippocampi, Table 6 for the heart and Table 7 for the 2D-femur data sets (where M represents the number of mixture components used, for each method).

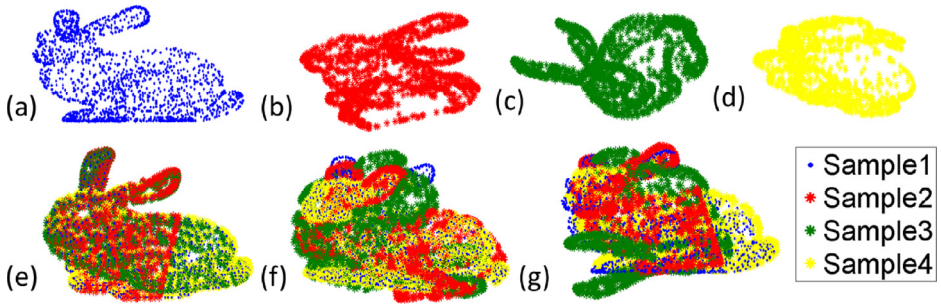

Fig. 6. Experiment investigating capture range of registration methods. Sample (a) original bunny point set, (b) point set in (a) rotated by 60° about x-axis and −60° about y-axis, (c)point set in (a) rotated by 60° about y- and z-axes and (d) point set in (a), rotated by −60° about z-axis and 60° about x-axis. (e) point sets aligned using mrTMM, (f) point sets after alignment using JRMPC and (g) point sets aligned using pair-wise CPD.

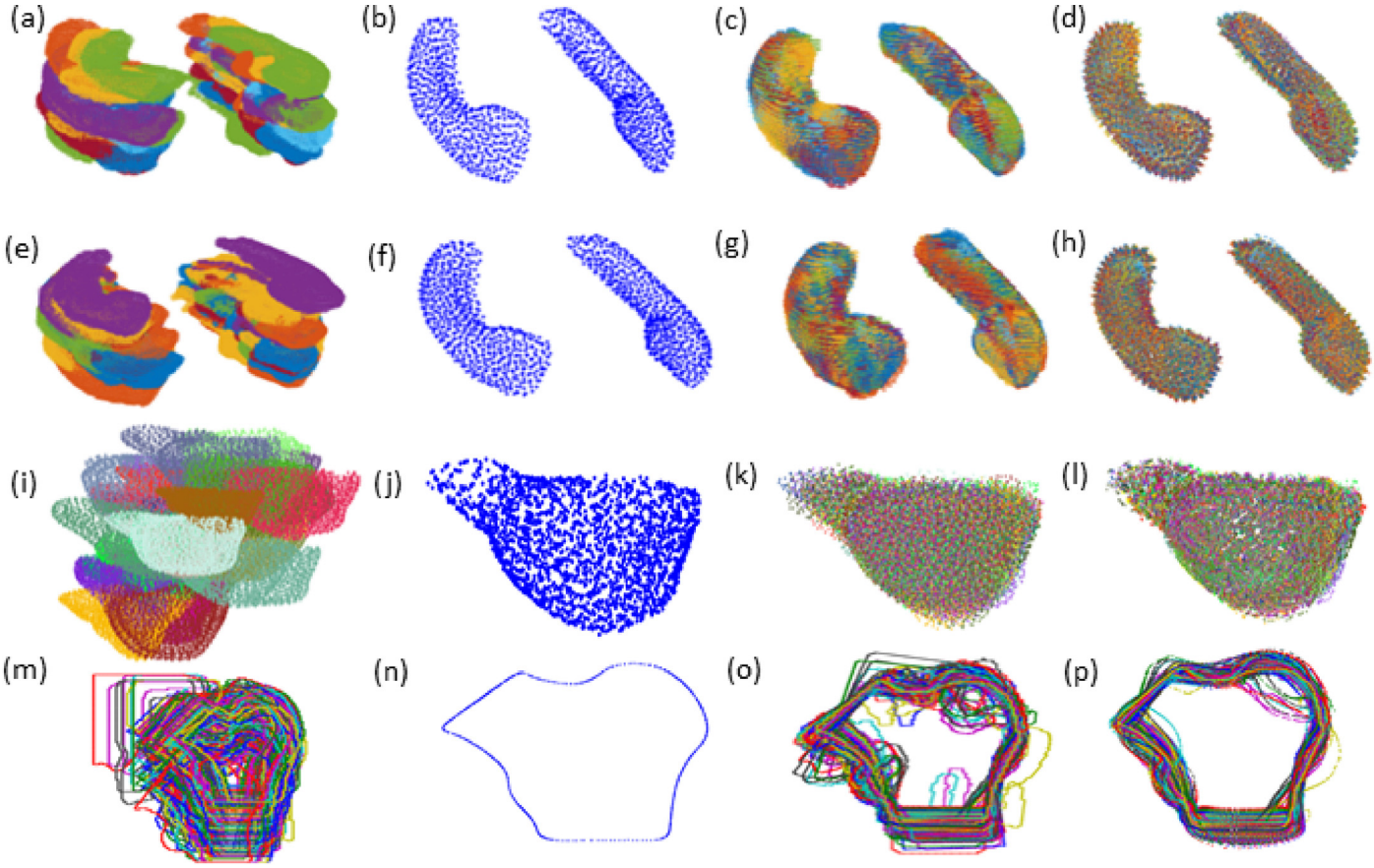

Fig. 7. MCI-hippocampi (first row), control-hippocampi (second row), heart data set (third row) and femur data set (fourth row). First column: Raw point sets prior to alignment; second column: estimated mean shapes; third column: aligned shapes; and fourth column: aligned soft-correspondences (using mrTMM).

Table 6
Alignment accuracy for heart data set.

Method	M	`Heart` ($K = 30$)	
		HD (mm)	MSD (mm)
gCPD	2560	**17.51 ±3.73**	**2.80 ±0.66**
SpSSM	2191	32.41 ± 9.96	4.10 ± 1.09
GMM	2560	32.41 ± 10.32	4.07 ± 1.17
TMM	2560	**15.45 ±3.96**	**2.80 ±0.67**
mrTMM	2560	**15.74 ±4.30**	**2.68 ±0.62**

Table 7
Alignment accuracy evaluated for femur data set.

Method	M	`Femur` ($K = 1000$)	
		HD (mm)	MSD (mm)
gCPD	1280	34.07 ± 2.98	2.56 ± 1.09
SpSSM	1474	77.43 ± 4.37	3.16 ± 1.07
GMM	1280	78.20 ± 4.29	3.31 ± 0.99
TMM	1280	**9.60 ±4.82**	**2.23 ±0.90**
mrTMM	1280	**10.04 ±5.33**	**2.19 ±0.92**

Statistical significance of the computed HD and MSD values was assessed using a paired-sample t-test (considering a significance level of 1%). The proposed methods (TMM and mrTMM) and gCPD[5] were comparable in registration accuracy and outper-

[5] As mentioned previously in Section 2.5.2, gCPD is employed in place of JRMPC as the latter does not recover global scaling across point sets.

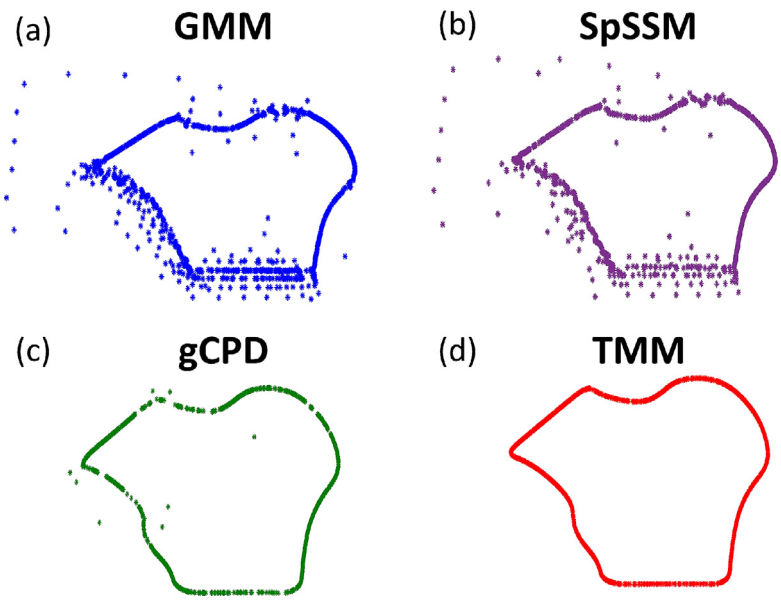

Fig. 8. Mean femur shapes estimated using: (a) GMM, (b) SpSSM, (c) gCPD and (d) TMM.

formed SpSSM and GMM in experiments conducted using both hippocampi and heart data sets. Significant improvements in accuracy are highlighted in bold in Tables 4 and 5 for the MCI and healthy groups, respectively. It should be noted however, gCPD requires an additional user-specified parameter (as with CPD), which controls the weight of the uniform distribution component in the mixture model and by extension the degree of robustness of the model to outliers. This parameter had to be tuned to identify the optimal value, based on alignment accuracy. Weights of 0.7 and 0.3 were found to produce the lowest HD and MSD values for MCI and healthy groups respectively, presented in Tables 4 and 5. While for the heart data set, 0.5 was found to be optimal (results presented in Table 6). The proposed methods were equally robust, showing marginal improvements in some cases, required fewer user-specified parameters and consequently, are better suited for automation.

Although the HD values reported in Tables 4 and 5 seem large for the hippocampi, it is important to note that they were computed between the aligned correspondences estimated for each sample and the mean shape estimated for the corresponding patient group (i.e. MCI or healthy). Consequently, they reflect the natural variation in hippocampal size present across samples in both patient groups. Furthermore, these are primarily concentrated at anterior and posterior ends of the hippocampi (and not in the middle section). The average lengths of the hippocampi were approximately 44 mm. and 45 mm. and width was 14 mm., for the MCI and healthy patient groups, respectively. The average length and width of the heart was 11 cm. and 10 cm., respectively. Finally, the average length and width of the 2D femoral heads was, 20cm. and 10cm., respectively.

However, with the femur data set, the proposed methods significantly outperformed all three GMM-based methods. This is attributed to the presence of significant levels of outliers in the femur shapes, which result in, the estimation of incorrect mean shapes (Fig. 8(a–c)), invalid correspondences, and consequently, the large HD and MSD values observed. TMM and mrTMM, however, are robust to the presence of such outliers and are able to estimate valid mean shapes and correspondences (as shown in Figs. 8(d) and 7(n,p)), achieving significantly lower HD and MSD values (summarized in Table 7). These experiments highlight the ability of the proposed approaches to remain robust to outliers and align groups of shapes to their respective mean shapes, more accurately than their GMM-based counterparts. They also complement subsequent experiments evaluating SSM quality, which reflect the accuracy of the correspondences estimated for each sample within each group, using the proposed methods and the state-of-the-art, i.e. more accurate correspondences result in improved preservation of shape variation across a group and consequently more descriptive SSMs. Thus both sets of experiments (i.e. evaluation of alignment accuracy and SSM quality) together reflect the registration quality of the proposed methods, relative to the state of the art.

3.2. Algorithm performance

3.2.1. Degrees of freedom

To quantitatively describe the influence of the degrees of freedom (ν) associated with TMM components on the robustness of the model to outliers, histograms depicting the range of values estimated for the synthetic and clinical data sets (following registration) are presented in Fig. 9(a–d). For the synthetic data set that contains a large proportion of noise and outliers (refer to Fig. 3 and Table 1) and missing data, the values estimated are concentrated in the range [2.1, 10] (as shown in Fig. 9(a)), conferring a greater degree of robustness to the registration process. A similar result is obtained for the femur data set (Fig. 9(b)) as it contains numerous samples with over-/under-segmented boundaries.

The heart and hippocampi data sets in contrast, contain few outliers and consequently, the values estimated for ν are distributed over a wider range, with high values indicating that the corresponding t-components behave similar to Gaussians (Heart: Fig. 9(c), MCI: Fig. 9(d)) as discussed previously in Section 2.1 and illustrated by Fig. 1(a). The flexible and robust nature of TMM-based registration is consequently attributed to the independent estimation of ν for each mixture component. Fig. 9(c) indicates that although the heart data set contains few visibly apparent outliers, the significant variation in geometry across the group (as a result of pathology) results in the estimation of low ν values for a greater number of mixture components, relative to the MCI-hippocampi data set (containing few outliers and moderate variation in shape across the group). This illustrates the role of ν in accommodating large variations in shape, while ensuring robust and accurate registration.

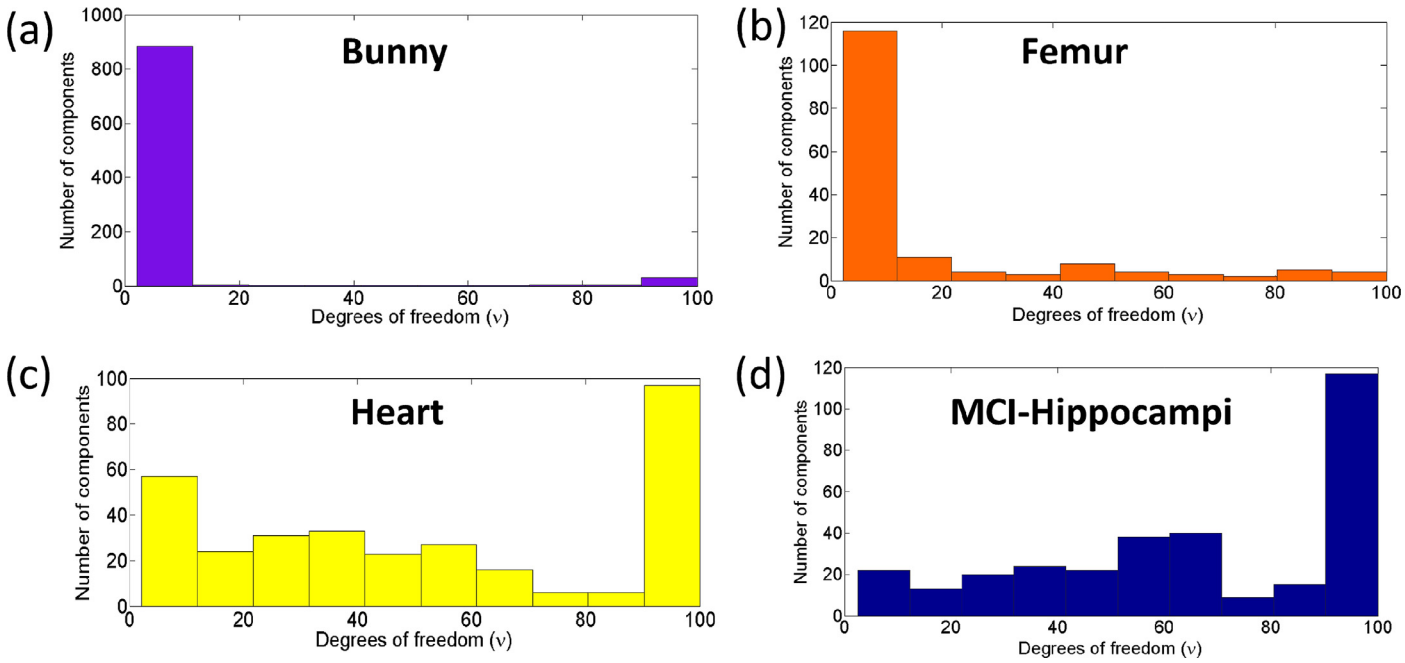

Fig. 9. Histograms of the degrees of freedom estimated for; (a) the synthetic data set with noise and outliers (M=940), (b) femur data set (M=160), (c) heart data set (M=320) and (d) set of hippocampi from MCI patients (M=320), following alignment using mrTMM.

Table 8
Run-time (minutes) for each data set aligned using TMM and mrTMM with M mixture components.

Method	Bunny ($M = 940$) ($K = 4$)	Hippocampus_MCI ($M = 2560$) ($K = 28$)	Hippocampus_Ctrl ($M = 2560$) ($K = 50$)	Heart ($M = 2560$) ($K = 30$)	Femur ($M = 1280$) ($K = 1000$)
TMM	8	158	289	45	1471
mrTMM	5	95	167	30	320

3.2.2. Convergence

The convergence of the proposed algorithms (TMM and mrTMM) are assessed based on the change in the mean shape, computed as $dM = ||\mathcal{M}^{new} - \mathcal{M}^{old}||_F / ||\mathcal{M}^{old}||_F$, across EM-iterations. This is illustrated by Fig. 10, where dM is plotted against the number of iterations. dM was formulated in this manner in order to define a common critical threshold/tolerance (10^{-3} was used for all experiments) to assess convergence for all data sets, thereby improving automation (and removing the need for identifying a unique threshold specific to each data set). Alternative convergence criteria may also be adopted by monitoring the change in, the log-likelihood, model variance or the estimated transformations, across successive EM-iterations.

In the case of mrTMM, convergence is assessed in this manner for each resolution level employed during registration. The plots depicted in Fig. 10 indicate that the convergence of TMM and the first resolution of mrTMM are similar to each other, for all data sets. Additionally, both TMM and mrTMM converge in fewer iterations for both hippocampi data sets (Fig. 10 (c,d)), relative to the remaining 3D data sets (Fig. 10 (a,b)). This is attributed to the presence of minimal outliers and moderate variation in shape across samples. Conversely, for the synthetic bunny data set (Fig. 10 (a)), containing significant proportions of outliers and large rotational offsets, both TMM and the first resolution of mrTMM require more iterations to converge, relative to the clinical data sets. However, beyond the first resolution (with mrTMM), subsequent levels converge quicker as evidenced by Fig. 10, improving computational efficiency as fewer iterations are required at higher model complexities (or number of mixture components M). These results are further verified by the run-times reported in Table 8 for all data sets, following alignment using TMM and mrTMM. From this we infer that mrTMM consistently improves computational efficiency compared to TMM, as fewer EM-iterations are required using the same number of mixture components (M). The code was implemented in MATLAB (R2014a) and tested on an Intel Xeon CPU (1.80GHz x 8) with 32GB RAM.

3.3. SSM quality

3.3.1. Generalisation

Generalisation errors were computed with respect to, the number of mixture components employed and subsequently, the modes of variation of the trained SSMs (using the identified optimal number of mixture components from the preceding experiments). Fig. 11 summarises these results for each clinical data set. Fig. 11(a) and (c) depict generalization errors evaluated with respect to the number of mixture components, for the MCI and healthy hippocampi data sets, respectively. The proposed methods perform comparably to GMM and gCPD while SpSSM achieves marginally lower errors for the MCI data. However, for the control-hippocampi, the proposed methods, GMM and gCPD outperform SpSSM and perform comparably to each other. While mrTMM offers some improvement over single-resolution TMM, as the training sets of hippocampi shapes contained no visibly apparent outliers, both proposed methods showed no significant difference in performance compared to GMM and gCPD. SpSSM achieves significantly lower errors than TMM for the MCI-hippocampi data set but shows no significant difference to mrTMM. As noted previously, in cases where the training data contains few outliers, the constituent t-distributions of a TMM behave similarly to Gaussians (due to es-

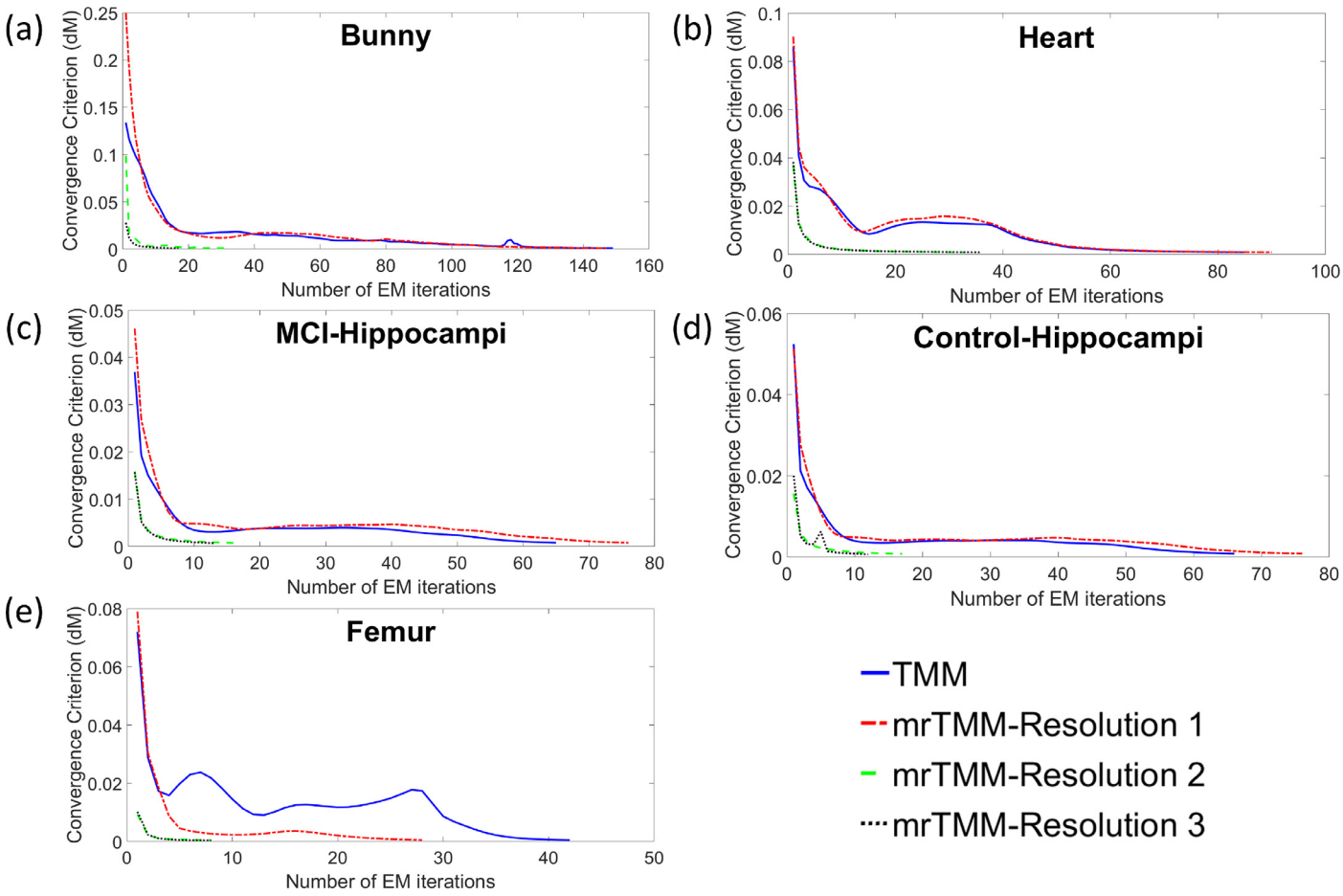

Fig. 10. Convergence of TMM and mrTMM algorithms: (a) Synthetic bunny data set containing significant outliers (refer to Table 1), using $M = 940$ mixture components; (b) Heart data set, using $M = 2560$ mixture components; (c) MCI-hippocampi data set, using $M = 2560$ mixture components; (d) Control-hippocampi data set, using $M = 2560$ mixture components; and (e) Femur data set, using $M = 1280$ mixture components.

timation of large values for the associated degrees of freedom ν, as shown in Fig. 9(d)) and consequently SSMs trained using both class of techniques are found to be of comparable quality.

A similar result is obtained for the heart data set, which contained few visibly apparent outliers, with gCPD, TMM and mrTMM significantly outperforming GMM and SpSSM, as illustrated by the generalization errors presented in Fig. 11(e). The heart data set comprised of samples with significant variations in geometry as a result of pathology (due to PH and HCM) which may be interpreted as missing information or outliers by the mixture model during registration (supported by the higher proportion of low ν values reported in Fig. 9(c), compared to the MCI-hippocampi data set, Fig. 9(d)). Consequently, lower errors achieved by gCPD, TMM and mrTMM are inferred to result from their robust nature, while GMM and SpSSM, lacking this quality, result in sub-optimal registration of the samples and by extension lower quality SSMs. Based on these experiments, the optimal number of mixture components was identified to be $M_{opt} = 2560$, for both hippocampi and heart data sets, using GMM, gCPD, TMM and mrTMM. With SpSSM a sparsity level of 0.1 was found to be optimal for the hippocampi and heart data sets, resulting in $M_{opt} = 1906$ for the MCI group, $M_{opt} = 2702$ for the healthy group and $M_{opt} = 2191$ for the heart. Generalisation errors evaluated with respect to the modes of variation for the MCI-hippocampi (Fig. 11(b)) show that while TMM performs comparably to GMM and SpSSM (with some marginal improvement), mrTMM and gCPD significantly outperform the same. A similar result is obtained for the control-hippocampi (Fig. 11(d)), with gCPD, TMM and mrTMM all providing substantial improvements to GMM and SpSSM. For the heart data set, (Fig. 11(f)) gCPD, TMM and mrTMM, once again offer marginal improvements over GMM and SpSSM.

Generalisation errors for the femur data set are presented in Fig. 11 (g) and (h), evaluated with respect to the number of mixture components and the modes of variation, respectively. As highlighted previously, the femur data set contained multiple training samples with significant outliers which adversely affected the quality of SSMs trained using the GMM-based approaches. Models trained using TMM and mrTMM on the other hand were robust to the presence of these outliers, resulting in significantly lower generalization errors compared to the state-of-the-art, when evaluated with respect to the number of mixture components (Fig. 11(g)). Fig. 11(h) suggests that the quality of SSMs generated are comparable across all methods, when evaluated with respect to the number of modes of variation. These results indicate that the proposed methods are able to reconstruct new shapes to a higher degree of accuracy. However, when correspondences are jointly estimated across all training and test shapes (as done with the generalization experiments evaluated with respect to the modes of variation) and SSM quality is subsequently evaluated by leave-one-out cross validation, the generalization-ability of the models is comparable across all methods. As the former set of generalization experiments (evaluated with respect to the number of mixture components), better emulate a real scenario, the improvement in reconstruction accuracy offered by both TMM and mrTMM compared to their GMM-counterparts, is compelling. Both TMM and mrTMM performed comparably across all generalization experiments con-

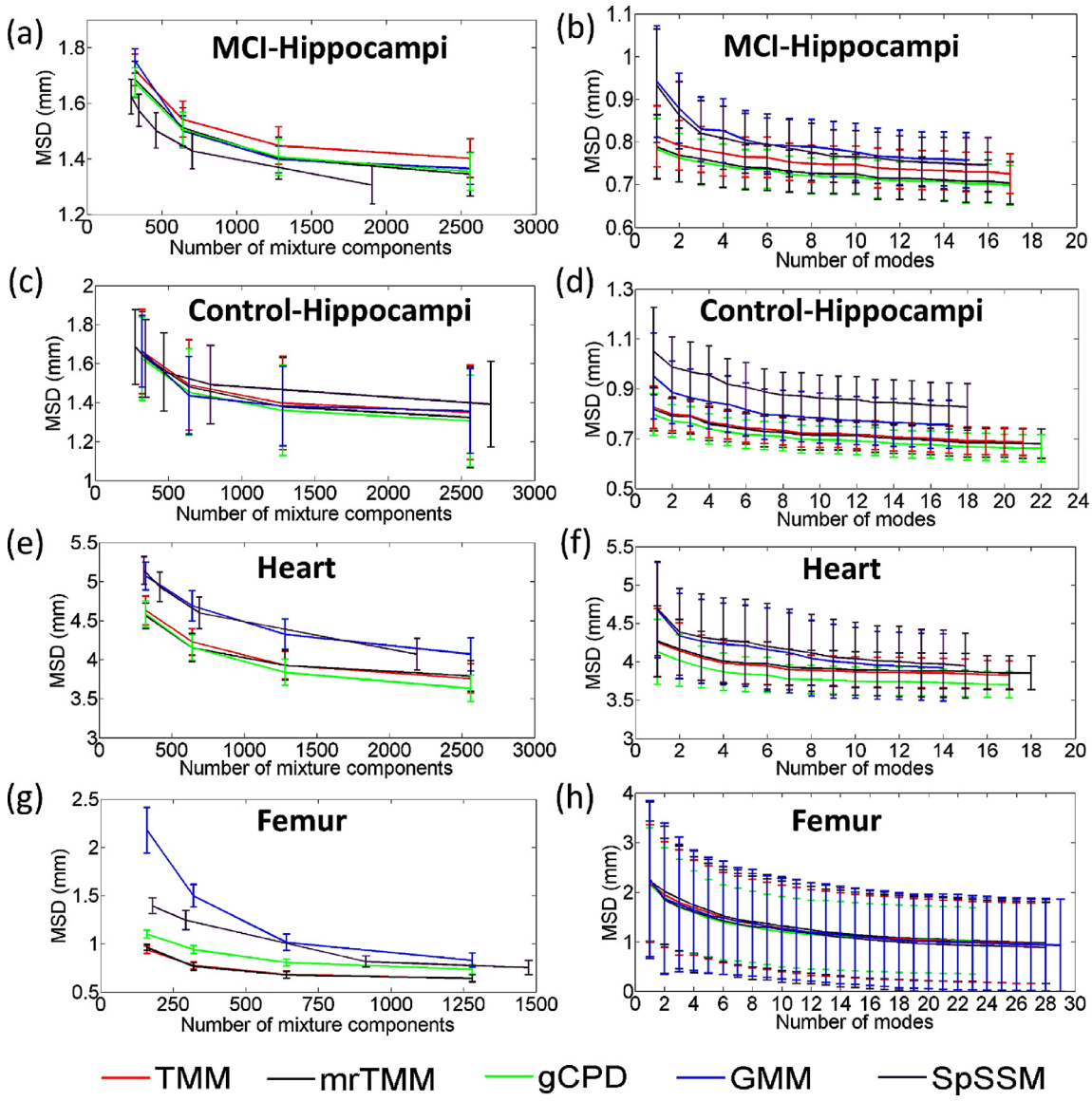

Fig. 11. SSM generalization errors evaluated with respect to number of mixture components (left column) and number of modes of variation (right column). (a,b) MCI-hippocampi, (c,d) control-hippocampi, (e,f) heart and (g,h) 2D-femur data set.

ducted, with mrTMM offering marginal improvements in some cases.

3.3.2. Specificity

The specificity errors presented in Fig. 12 (a), (b) and (c) show that models trained using the proposed methods and gCPD achieve significantly lower specificity errors than GMM and SpSSM and are inferred to generate shapes with a higher degree of anatomical plausibility, for both hippocampi, and the heart data sets. For the femur data set, as noted previously, wrong correspondences estimated using the GMM-based methods result in incorrect modes of variation. Consequently, shapes sampled randomly from the trained SSMs are implausible, resulting in the high specificity errors seen in Fig. 12(d). TMM, mrTMM and gCPD however, are robust to the presence of outliers in the training set and consequently achieve significantly lower specificity errors compared to the GMM and SpSSM. These results are consistent with those observed in the generalization experiments, indicating the superiority of the proposed methods when dealing with data containing outliers. As with the generalization experiments, the specificity of the models trained using TMM and mrTMM are similar. Although specificity experiments conducted using the femur data set indicate that TMM, mrTMM and gCPD are comparably robust and produce SSMs of similar quality, visual inspection of the modes of variation (and correspondingly of the model-predicted shapes) highlight the advantage of the proposed methods over gCPD, as illustrated by Fig. 13. The presence of large proportions of outliers in the training samples adversely affects the correspondences established using gCPD resulting in a mean shape and modes of variation that contain points offset from the true boundary of the femoral head (Fig. 13(a,b)). In comparison, both TMM and mrTMM are able to suppress the influence of such outliers and establish valid correspondences, resulting in plausible mean shapes and modes of variation (as shown in Fig. 13(c–f)).

3.3.3. Compactness

Cross-validation experiments revealed that SSMs generated using SpSSM, gCPD and both proposed methods were equally compact, for the MCI-hippocampi (Fig. 14(a)), while GMM generated more compact models. However, based on the generalization er-

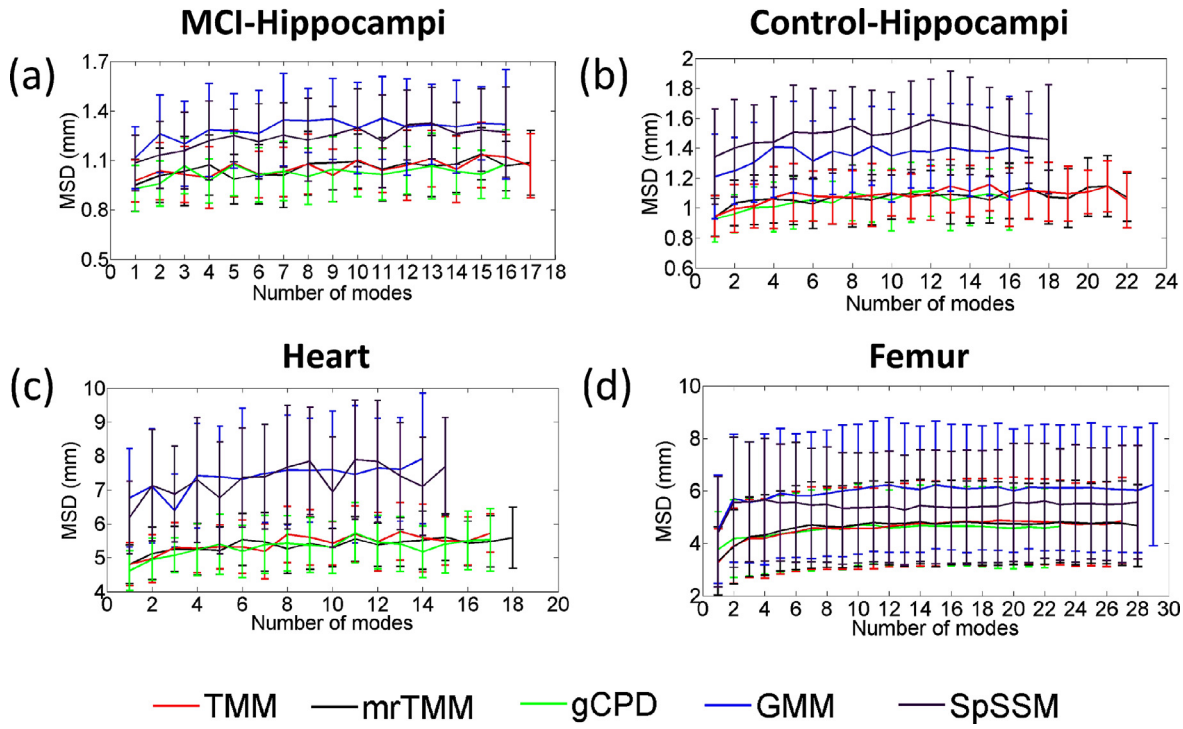

Fig. 12. SSM specificity errors evaluated with respect to the modes of variation for (a) MCI-hippocampi, (b) control-hippocampi, (c) heart, and (d) 2D-femur data sets.

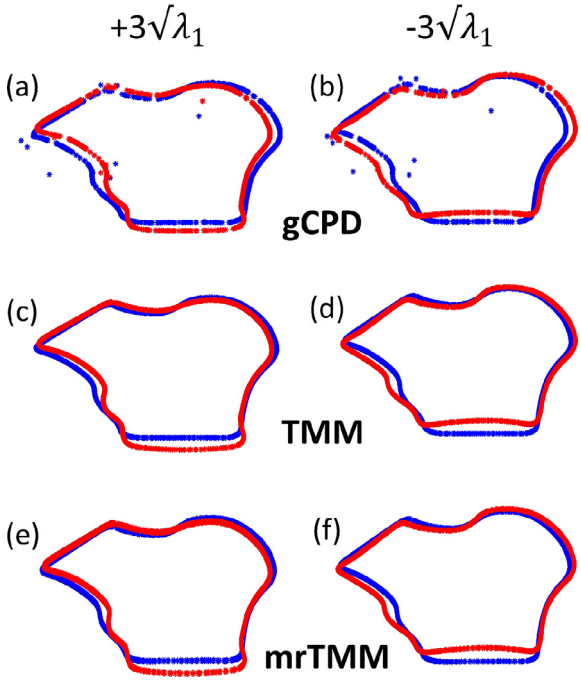

Fig. 13. First mode of variation for the 2D-femur data set (red) overlaid on the estimated mean shape (blue). SSMs were trained using: gCPD (a,b), TMM (c,d), and mrTMM (e,f). Here λ_1 denotes the first eigenvalue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

rors presented in Fig. 11(b) we note that improved compactness of GMM is at the expense of reduced generalization ability and model specificity (Fig. 12(a)). With the healthy hippocampi data set, both GMM and SpSSM produced models that were significantly more compact than gCPD and the proposed methods (Fig. 14(b)), however, once again at cost of lower generalization ability and model specificity (as illustrated by Figs. 11(d) and 12(b)). For the MCI-hippocampi data set (containing $K = 28$ samples), up to 17 eigenmodes were found to capture 95% of the total variation in shape found in the training set, while the control-hippocampi set (containing $K = 50$ samples) required up to 22 modes of variation. The heart data set ($K = 30$ samples) comprising instances of healthy subjects and patients diagnosed with PH and HCM, required up to 18 eigenmodes, using the proposed methods, with GMM and SpSSM generating more compact models (Fig. 14(c)). As with the hippocampi, although GMM and SpSSM generated more compact models, they suffered from reduced generalisation-ability and specificity. The variation in shape across the 2D femur data set containing $K = 1000$ samples was adequately captured by all methods, within 30 modes of variation and a similar trend is observed as with the remaining data sets, i.e. all three GMM-based methods produce more compact models than TMM and mrTMM, at the cost of higher generalisation and specificity errors.

The first mode of variation for the MCI- and control-hippocampi data sets are depicted in Figs. 15 and 16 for the heart. It is interesting to note that the first mode of variation for the presented heart-SSM describes a change in the shape and volume of the right ventricle, characteristic of pulmonary hypertension. The presented SSMs in Figs. 15 and 16 were trained using mrTMM and the optimal number of mixture components identified for each data set (refer to Section 3.3.1).

The improvement in SSM quality achieved using TMM and mrTMM, when dealing with noisy data, is demonstrated by the generalization and specificity experiments conducted in this study. Their ability to automatically align shapes in a robust fashion and reconstruct unseen shapes to a high degree of accuracy, can find application in large-scale studies investigating shape and morphological changes associated with pathological processes, as seen with dementia-related hippocampal changes, pulmonary hypertension and hypertrophic cardiomyopathy induced changes to ventricular morphology in the heart, among others. Additionally, the proposed methods can find use in intra-operative guidance applications requiring robust and automatic pose correction. The proposed methods can also be employed to initialize a subsequent

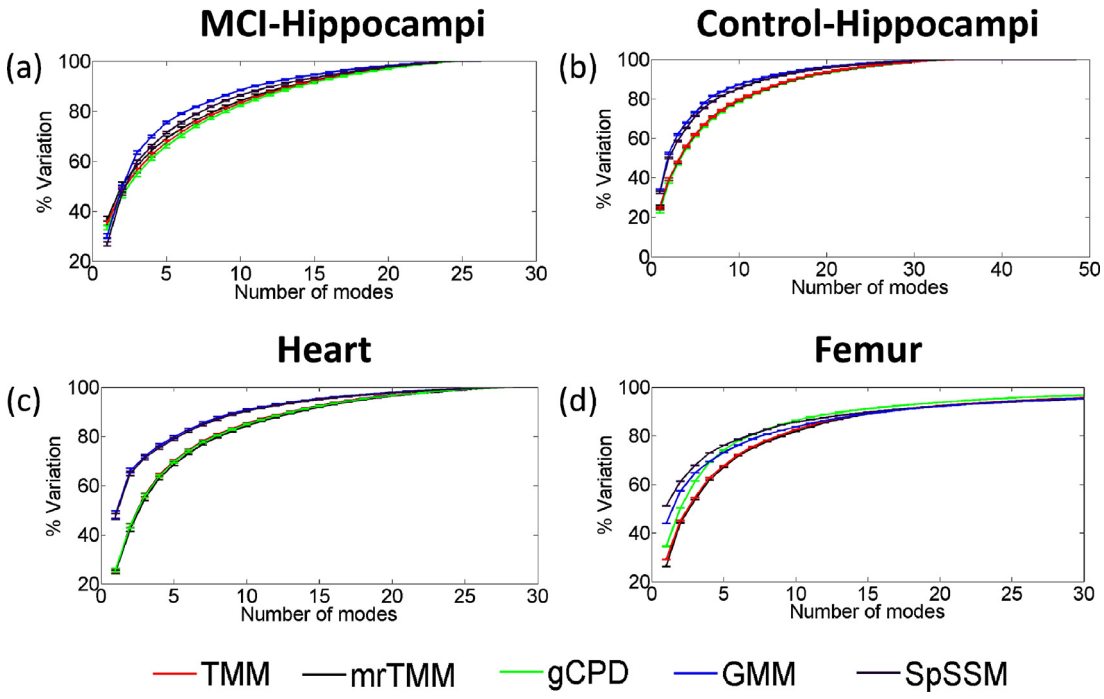

Fig. 14. SSM compactness assessed by plotting the cumulative sum of the variation % (expressed by each eigenmode), against the number of modes. (a) MCI-hippocampi, (b) control-hippocampi, (c) heart, and (d) 2D-femur data set.

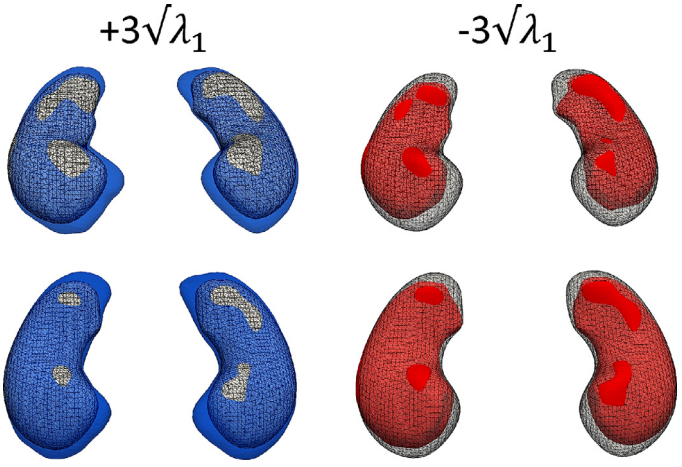

Fig. 15. First mode of variation for SSMs trained using mrTMM. Top row: MCI hippocampi, bottom row: healthy hippocampi. In all cases the overlaid surface mesh with visible edges represents the mean shape.

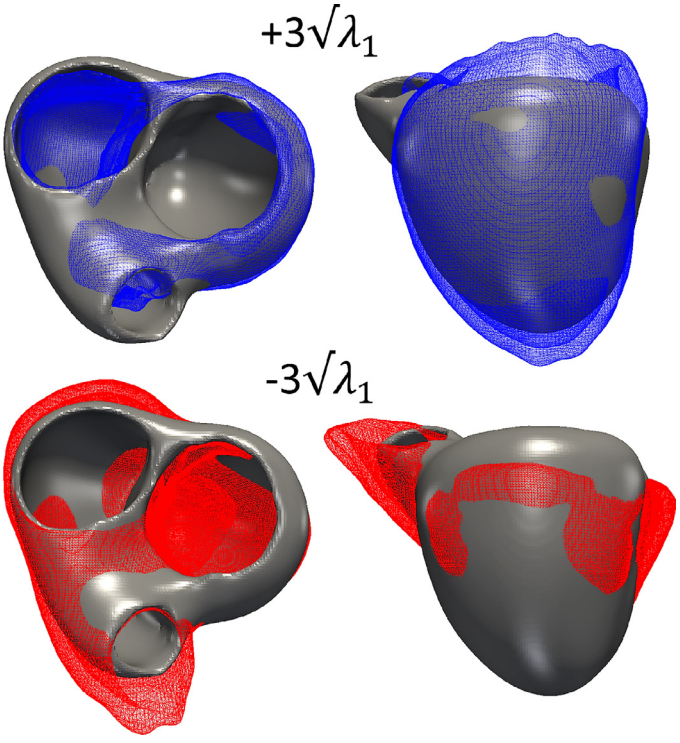

Fig. 16. First mode of variation of the heart-SSM trained using mrTMM, overlaid on the estimated mean shape (dark grey surface).

deformable registration step, often necessary to capture soft tissue deformations common to surgical procedures. Future work will look to extend the proposed techniques to a group-wise non-rigid registration framework.

4. Conclusions

The group-wise point set registration methods proposed in this study, namely, TMM and mrTMM, were shown to outperform state-of-the-art GMM-based techniques in terms of alignment accuracy (using both synthetic and clinical data). Although mrTMM offered significant improvement over TMM with synthetic data, their performance was comparable in most experiments involving clinical data. Cross-validation experiments evaluating SSM quality indicate that the proposed methods are comparable to the state-of-the-art for data containing few outliers, but significantly outperform the same in the presence of outliers (as in the femur data set). The robust nature of TMM and mrTMM makes them well-suited to automatic SSM generation, using automated techniques to generate the requisite training sets, as they are able to tolerate

missing data and the presence of significant proportions of outliers. Both methods are consequently, tailored for automatic shape analysis of large cohorts.

Acknowledgements

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 by the Engineering and Physical Sciences Research Council through the OCEAN project (EP/M006328/1). 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 hippocampi. The authors would also like to thank Mohsen Farzi, CISTIB, The University of Sheffield, UK, for providing access to the data set of segmented femoral heads.

Appendix A

M-step update equations for t-mixture model parameters, presented in Section 2.2 Eqs. (6) – (12), are derived by maximizing the complete data log-likelihood $\mathbf{Q}$ with respect to each model parameter as follows:

- Estimation of TMM centroids $\boldsymbol{\mu}_j$ at the $(t+1)$th EM-iteration:

$$\mathbf{Q}(\Psi^{t+1}|\Psi^t) = -\frac{1}{2}\sum_{k,i,j} P^{\star t}_{kij}\Delta_{kij} + O.T. \tag{18a}$$

$$\Delta_{kij} = \frac{(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j - \mathbf{t}_k)^T(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j - \mathbf{t}_k)}{\sigma^2} \tag{18b}$$

O.T. summarizes terms in $\mathbf{Q}$ independent of $\boldsymbol{\mu}_j$.

$$< \partial\mathbf{Q}, \partial\boldsymbol{\mu}_j >= \left[-\frac{1}{2}\sum_{k,i} P^{\star}_{kij}\Delta^{\boldsymbol{\mu}_j+\partial\boldsymbol{\mu}_j}_{kij}\right] - \left[-\frac{1}{2}\sum_{k,i} P^{\star}_{kij}\Delta^{\boldsymbol{\mu}_j}_{kij}\right] \tag{19a}$$

$$< \partial\mathbf{Q}, \partial\boldsymbol{\mu}_j >= \sum_{k,i} P^{\star}_{kij}\left[(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j - \mathbf{t}_k)^T s_k\mathbf{R}_k\right]\partial\boldsymbol{\mu}_j \tag{19b}$$

$$< \partial\mathbf{Q}, \partial\boldsymbol{\mu}_j >= 0 \Rightarrow \sum_{k,i} P^{\star}_{kij}\left[(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j - \mathbf{t}_k)^T s_k\mathbf{R}_k\right] = 0 \tag{19c}$$

$$\sum_{k,i} P^{\star}_{kij} s_k\mathbf{R}_k^T(\mathbf{x}_{ki} - \mathbf{t}_k) = \sum_{k,i} P^{\star}_{kij} s_k\mathbf{R}_k^T\mathbf{R}_k s_k\boldsymbol{\mu}_j \tag{19d}$$

$$\boldsymbol{\mu}_j = \frac{\sum_{k,i} P^{\star}_{kij} s_k^{-1}\mathbf{R}^T(\mathbf{x}_{ki} - \mathbf{t}_k)}{\sum_{k,i} P^{\star}_{kij}} \tag{19e}$$

- Estimation of model variance σ^2:

$$\frac{\partial\mathbf{Q}}{\partial\sigma^2} = \frac{\partial\sum_{k,i,j}\left[-\frac{P_{kij}}{2}[\log(\sigma^6)] - \frac{P^{\star}_{kij}}{2}[\Delta_{kij}]\right]}{\partial\sigma^2} = 0 \tag{20a}$$

$$\Rightarrow \sum_{k,i,j} -P_{kij}\frac{3}{\sigma} + P^{\star}_{kij}\frac{(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j - \mathbf{t}_k)^T(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j - \mathbf{t}_k)}{\sigma^3} = 0 \tag{20b}$$

$$\sigma^2 = \frac{\sum_{k,i,j} P^{\star}_{kij}(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j - \mathbf{t}_k)^T(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j - \mathbf{t}_k)}{3\sum_{kij} P_{kij}} \tag{20c}$$

- Estimation of translation $\mathbf{t}_k$:

$$< \partial\mathbf{Q}, \partial\mathbf{t}_k >= \left[-\frac{1}{2}\sum_{i,j} P^{\star}_{kij}\Delta^{\mathbf{t}_k+\partial\mathbf{t}_k}_{kij}\right] - \left[-\frac{1}{2}\sum_{i,j} P^{\star}_{kij}\Delta^{\mathbf{t}_k}_{kij}\right] \tag{21a}$$

$$< \partial\mathbf{Q}, \partial\mathbf{t}_k >= \sum_{i,j} P^{\star}_{kij}[(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j - \mathbf{t}_k)^T]\partial\mathbf{t}_k \tag{21b}$$

$$< \partial\mathbf{Q}, \partial\mathbf{t}_k >= 0 \Rightarrow \sum_{i,j} P^{\star}_{kij}(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j)^T = \sum_{i,j} P^{\star}_{kij}\mathbf{t}_k^T \tag{21c}$$

$$\mathbf{t}_k = \frac{\sum_{i,j} P^{\star}_{kij}\mathbf{x}_{ki}}{\sum_{i,j} P^{\star}_{kij}} - s_k\mathbf{R}_k\frac{\sum_{i,j} P^{\star}_{kij}\boldsymbol{\mu}_j}{\sum_{i,j} P^{\star}_{kij}} \tag{21d}$$

Setting the first term as $\mathbf{d}_k$ and the second term as $\mathbf{m}_k$ (refer to Eqs. (14) - (15)) we get:

$$\mathbf{t}_k = \mathbf{d}_k - s_k\mathbf{R}_k\mathbf{m}_k \tag{21e}$$

- Estimation of strictly orthogonal rotation $\mathbf{R}_k$: Using the lemma outlined in Myronenko and Song (2010), the optimal rotation matrix maximises tr $(\mathbf{C}_k^T\mathbf{R}_k)$ where $\mathbf{C}_k$ represents a matrix given by Eq. (13).

$$\tilde{\mathbf{x}}_{ki} = \mathbf{x}_{ki} - \mathbf{d}_k, \tilde{\mathbf{m}}_{kj} = \boldsymbol{\mu}_j - \mathbf{m}_k \tag{22a}$$

Using eq. (4e) and (5a) we get:

$$\mathbf{Q}(\Psi^{t+1}|\Psi^t) \propto \sum_{i,j} P^{\star t}_{kij}(\tilde{\mathbf{x}}_{ki}^T\mathbf{R}_k\tilde{\mathbf{m}}_{kj}) \tag{22b}$$

$$\mathbf{Q}(\Psi^{t+1}|\Psi^t) \propto \sum_{i,j} P^{\star t}_{kij}\text{tr}\,[\tilde{\mathbf{m}}_{kj}\tilde{\mathbf{x}}_{ki}^T\mathbf{R}_k] \tag{22c}$$

As equation (.5c) must be maximised with respect to $\mathbf{R}_k$,

$$\mathbf{C}_k = \sum_{i,j} P^{\star}_{kij}\tilde{\mathbf{x}}_{ki}\tilde{\mathbf{m}}_{kj}^T \tag{22d}$$

$\mathbf{R}_k = \mathbf{USV}^T$, where $\mathbf{U}, \mathbf{V}$ are unitary matrices computed by singular value decomposition of $\mathbf{C}_k$ and $\mathbf{S} = diag(1, 1, det(\mathbf{UV}^T))$ is a diagonal matrix that prevents reflections.

- Estimation of scaling s_k:

$$\frac{\partial\mathbf{Q}}{\partial s_k} = -\frac{1}{2}\frac{\partial\sum_{i,j} P^{\star}_{kij}\Delta_{kij}}{\partial s_k} = 0 \tag{23a}$$

$$\sum_{i,j} P^{\star}_{kij}\frac{(\tilde{\mathbf{x}}_{ki} - s_k\mathbf{R}_k\tilde{\mathbf{m}}_{kj})^T(\mathbf{R}_k\tilde{\mathbf{m}}_{kj})}{\sigma^2} = 0 \tag{23b}$$

$$\sum_{i,j} P^{\star}_{kij}[(\tilde{\mathbf{x}}_{ki})^T(\mathbf{R}_k\tilde{\mathbf{m}}_{kj})] = s_k\sum_{i,j} P^{\star}_{kij}[\tilde{\mathbf{m}}_{kj}^T\mathbf{R}_k^T\mathbf{R}_k\tilde{\mathbf{m}}_{kj}] \tag{23c}$$

$$s_k = \frac{\text{tr}\,[\tilde{\mathbf{m}}_{kj}\tilde{\mathbf{x}}_{ki}^T]\mathbf{R}_k}{\text{tr}\,[\tilde{\mathbf{m}}_{kj}\tilde{\mathbf{m}}_{kj}^T]} = \frac{\text{tr}\,[\mathbf{C}_k^T\mathbf{R}_k]}{\text{tr}\,[\tilde{\mathbf{m}}_{kj}\tilde{\mathbf{m}}_{kj}^T]} \tag{23d}$$

- Estimation of degrees of freedom ν_j:

$$\mathbf{Q}(\Psi^{t+1}|\Psi^t) = \sum_{k,i,j} P^t_{kij}\Big[-\log\Gamma\left(\tfrac{\nu_j}{2}\right) + \tfrac{1}{2}\nu_j\log\left(\tfrac{\nu_j}{2}\right) + \tfrac{\nu_j}{2}\Big[\log(U^t_{kij}) - U^t_{kij} + \Psi\left(\tfrac{\nu_j+D}{2}\right) - \log\left(\tfrac{\nu_j^t+D}{2}\right)\Big]\Big] + O.T. \tag{24a}$$

O.T. summarizes terms in $\mathbf{Q}$ independent of ν_j.

$$\frac{\partial\mathbf{Q}}{\partial\nu_j} = -\Psi\left(\frac{\nu_j}{2}\right) + \log\left(\frac{\nu_j}{2}\right) + 1 + \frac{1}{\sum_{k,i} P^t_{kij}}\sum_{k,i} P^t_{kij}\left(\log\left(U^t_{kij}\right) - U^t_{kij}\right) + \Psi\left(\frac{\nu_j^t+D}{2}\right) - \log\left(\frac{\nu_j^t+D}{2}\right) = 0 \tag{24b}$$

Eq. (8)is solved using Newton's method to estimate the degrees of freedom ν_j.

Table B1
RMSE values computed between estimated and ground truth translations for the synthetic bunny data set.

Method	Sample 2	Sample 3	Sample 4
CPD	0.648	0.356	0.670
SpSSM	0.118	0.046	0.103
GMM	0.107	0.046	0.099
JRMPC	0.121	0.057	0.111
TMM	0.115	0.051	0.107
mrTMM	0.124	0.054	0.114

Appendix B

- Synthetic Data: Translation Errors The ability of the proposed methods and the state-of-the-art to recover the ground truth translations applied to generate the synthetic bunny data set (refer to Table 1) was evaluated. Translation errors were evaluated as the RMSE between the estimated and ground truth translations as described by Eq. (25). The resulting errors summarised in Table B.9 indicate that the proposed methods (TMM and mrTMM) and their GMM-based counterparts perform comparably (with the exception of CPD which resulted in substantially higher translation errors) in this regard.

$$RMSE = ||\mathbf{T}_k^g - (\mathbf{T}_k - \mathbf{T}_1)||_F \tag{25}$$

- SSM generation and model-fitting: The process of fitting the trained models to new data involves two steps: (1) mixture-fitting and (2) SSM-fitting. The former is first used to align the new shape to the trained mean model and establish correspondences. The estimated correspondences are subsequently projected to the trained SSM according to Eq. (26a), to obtain estimates of weights $\mathbf{b}$. Here, $\overline{\mathbf{x}}$ represents the mean shape vector. Vector $\mathbf{b}$ represents the set of parameters used to generate variations in shape and are used to reconstruct the new shape $\mathbf{x}^{new}$, using Eq. (26b). To reduce the influence of noise on shape reconstruction, the estimated PCA scores are constrained as: $|b_m| \leq 3\sqrt{\lambda_m}$, where λ_m represents the eigenvalue of the mth mode of variation.

$$\mathbf{b} = \Phi^T(\mathbf{x} - \overline{\mathbf{x}}) \tag{26a}$$

$$\mathbf{x}^{new} = \overline{\mathbf{x}} + \Phi\mathbf{b} \tag{26b}$$

- Distance Measures: Alignment accuracy was evaluated for all four clinical data sets, using the Hausdorff distannce (HD) and mean surface distance (MSD) metrics. These measures were evaluated as shown in Eq. (27a) - (27b), where $\mathbf{d_{min}}(A, B)$ denotes the minimum distance for each point in shape A to shape B. HD and MSD measures were computed between the aligned soft-correspondences estimated for each sample in the group and the corresponding mean shape estimated for the group.

$$HD = max(max(\mathbf{d_{min}}(A, B)), max(\mathbf{d_{min}}(B, A))) \tag{27a}$$

$$MSD = mean(mean(\mathbf{d_{min}}(A, B)), mean(\mathbf{d_{min}}(B, A))) \tag{27b}$$

References

Albà, X., Lekadir, K., Hoogendoorn, C., Pereanez, M., Swift, A.J., Wild, J.M., Frangi, A.F., 2014. Reusability of statistical shape models for the segmentation of severely abnormal hearts. In: International Workshop on Statistical Atlases and Computational Models of the Heart. Springer, pp. 257–264.

Besl, P.J., McKay, N.D., 1992. A method for registration of 3-d shapes. IEEE Trans. Pattern Anal. Mach. Intell. 14 (2), 239–256.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Brechbühler, C., Gerig, G., Kübler, O., 1995. Parametrization of closed surfaces for 3-d shape description. Comput. Vis. Image Underst. 61 (2), 154–170.

Cardoso, M.J., Modat, M., Wolz, R., Melbourne, A., Cash, D., Rueckert, D., Ourselin, S., 2015. Geodesic information flows: spatially-variant graphs and their application to segmentation and fusion. IEEE Trans. Med. Imaging 34 (9), 1976–1988.

Castro-Mateos, I., Pozo, J.M., Pereañez, M., Lekadir, K., Lazary, A., Frangi, A.F., 2015. Statistical interspace models (sims): application to robust 3d spine segmentation. IEEE Trans. Med. Imaging 34 (8), 1663–1675.

Cates, J., Fletcher, P.T., Styner, M., Shenton, M., Whitaker, R., 2007. Shape modeling and analysis with entropy-based particle systems. In: Biennial International Conference on Information Processing in Medical Imaging. Springer, pp. 333–345.

Chen, T., Vemuri, B.C., Rangarajan, A., Eisenschenk, S.J., 2010. Group-wise point-set registration using a novel cdf-based Havrda-Charvát divergence. Int. J. Comput. Vis. 86 (1), 111–124.

Chui, H., Rangarajan, A., 2003. A new point matching algorithm for non-rigid registration. Comput. Vis. Image Underst. 89 (2), 114–141.

Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J., 1995. Active shape models-their training and application. Comput. Vis. Image Underst. 61 (1), 38–59.

Davies, R.H., Twining, C.J., Cootes, T.F., Taylor, C.J., 2010. Building 3-d statistical shape models by direct optimization. IEEE Trans. Med. Imaging 29 (4), 961–981.

Davies, R.H., Twining, C.J., Cootes, T.F., Waterton, J.C., Taylor, C.J., 2002. A minimum description length approach to statistical shape modeling. IEEE Trans. Med. Imaging 21 (5), 525–537.

Durrleman, S., Pennec, X., Trouvé, A., Ayache, N., 2007. Measuring brain variability via sulcal lines registration: a diffeomorphic approach. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 675–682.

Evangelidis, G.D., Kounades-Bastian, D., Horaud, R., Psarakis, E.Z., 2014. A generative model for the joint registration of multiple point sets. In: European Conference on Computer Vision. Springer, pp. 109–122.

Frangi, A.F., Rueckert, D., Schnabel, J.A., Niessen, W.J., 2002. Automatic construction of multiple-object three-dimensional statistical shape models: application to cardiac modeling. IEEE Trans. Med. Imaging 21 (9), 1151–1166.

Gelfand, N., Mitra, N.J., Guibas, L.J., Pottmann, H., 2005. Robust global registration. In: Symposium on geometry processing, 2, p. 5.

Gerig, G., Styner, M., Shenton, M.E., Lieberman, J.A., 2001. Shape versus size: improved understanding of the morphology of brain structures. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 24–32.

Gerogiannis, D., Nikou, C., Likas, A., 2009. The mixtures of student??s t-distributions as a robust framework for rigid registration. Image Vis. Comput. 27 (9), 1285–1294.

Gooya, A., Davatzikos, C., Frangi, A.F., 2015. A Bayesian approach to sparse model selection in statistical shape models. SIAM J. Imaging Sci. 8 (2), 858–887.

Gooya, A., Lekadir, K., Alba, X., Swift, A.J., Wild, J.M., Frangi, A.F., 2015. Joint clustering and component analysis of correspondenceless point sets: application to cardiac statistical modeling. In: International Conference on Information Processing in Medical Imaging. Springer, pp. 98–109.

Gower, J.C., 1975. Generalized procrustes analysis. Psychometrika 40 (1), 33–51.

Granger, S., Pennec, X., 2002. Multi-scale em-icp: a fast and robust approach for surface registration. In: European Conference on Computer Vision. Springer, pp. 418–432.

Heimann, T., Meinzer, H.-P., 2009. Statistical shape models for 3d medical image segmentation: a review. Med. Image Anal. 13 (4), 543–563.

Hermans, J., Smeets, D., Vandermeulen, D., Suetens, P., 2011. Robust point set registration using em-icp with information-theoretically optimal outlier handling. In: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, pp. 2465–2472.

Hufnagel, H., Pennec, X., Ehrhardt, J., Ayache, N., Handels, H., 2008. 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.

Huynh, D.Q., 2009. Metrics for 3d rotations: comparison and analysis. J. Math. Imaging Vis. 35 (2), 155–164.

Jian, B., Vemuri, B.C., 2011. Robust point set registration using gaussian mixture models. IEEE Trans. Pattern Anal. Mach. Intell. 33 (8), 1633–1645.

Kendall, D.G., 1989. A survey of the statistical theory of shape. Stat. Sci. 87–99.

Koikkalainen, J., Tolli, T., Lauerma, K., Antila, K., Mattila, E., Lilja, M., Lotjonen, J., 2008. Methods of artificial enlargement of the training set for statistical shape models. IEEE Trans. Med. Imaging 27 (11), 1643–1654.

Lekadir, K., Hoogendoorn, C., Pereanez, M., Albà, X., Pashaei, A., Frangi, A.F., 2014. Statistical personalization of ventricular fiber orientation using shape predictors. IEEE Trans. Med. Imaging 33 (4), 882–890.

Leventon, M.E., Grimson, W.E.L., Faugeras, O., 2000. Statistical shape influence in geodesic active contours. In: Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on, 1. IEEE, pp. 316–323.

Lorenz, C., Krahnstöver, N., 2000. Generation of point-based 3d statistical shape models for anatomical objects. Comput. Vis. Image Underst. 77 (2), 175–191.

McCloskey, E.V., Beneton, M., Charlesworth, D., Kayan, K., de Takats, D., Dey, A., Orgee, J., Ashford, R., Forster, M., Cliffe, J., et al., 2007. Clodronate reduces the incidence of fractures in community-dwelling elderly women unselected for osteoporosis: results of a double-blind, placebo-controlled randomized study. J. Bone Miner. Res. 22 (1), 135–141.

Myronenko, A., Song, X., 2010. Point set registration: coherent point drift. IEEE Trans. Pattern Anal. Mach. Intell. 32 (12), 2262–2275.

Patenaude, B., Smith, S.M., Kennedy, D.N., Jenkinson, M., 2011. A Bayesian model of shape and appearance for subcortical brain segmentation. Neuroimage 56 (3), 907–922.
Peel, D., McLachlan, G.J., 2000. Robust mixture modelling using the t distribution. Stat. Comput. 10 (4), 339–348.
Pereañez, M., Lekadir, K., Butakoff, C., Hoogendoorn, C., Frangi, A.F., 2014. A framework for the merging of pre-existing and correspondenceless 3d statistical shape models. Med. Image Anal. 18 (7), 1044–1058.
Pizer, S.M., Fletcher, P.T., Joshi, S., Thall, A., Chen, J.Z., Fridman, Y., Fritsch, D.S., Gash, A.G., Glotzer, J.M., Jiroutek, M.R., et al., 2003. Deformable m-reps for 3d medical image segmentation. Int. J. Comput. Vis. 55 (2–3), 85–106.
Prados Carrasco, F., Cardoso, M.J., Burgos, N., Wheeler-Kingshott, C., Ourselin, S., 2016. Niftyweb: web based platform for image processing on the cloud. International Society for Magnetic Resonance in Medicine (ISMRM).
Rangarajan, A., Chui, H., Mjolsness, E., Pappu, S., Davachi, L., Goldman-Rakic, P., Duncan, J., 1997. A robust point-matching algorithm for autoradiograph alignment. Med. Image Anal. 1 (4), 379–398.
Rasoulian, A., Rohling, R., Abolmaesumi, P., 2012. Group-wise registration of point sets for statistical shape models. IEEE Trans. Med. Imaging 31 (11), 2025–2034.
Ravikumar, N., Gooya, A., Ćimen, S., Frangi, A.F., Taylor, Z.A., 2016. A multi-resolution t-mixture model approach to robust group-wise alignment of shapes. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, p. InPress.
Ravikumar, N., Gooya, A., Frangi, A.F., Taylor, Z.A., 2016. Robust group-wise rigid registration of point sets using t-mixture model. SPIE Medical Imaging. International Society for Optics and Photonics. 97840S–97840S.
Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J., 1999. Non-rigid registration using free-form deformations: application to breast mr images. IEEE Trans. Med. Imaging 18 (8), 712–721.
Shen, K.-k., Fripp, J., Mériaudeau, F., Chételat, G., Salvado, O., Bourgeat, P., Initiative, A.D.N., et al., 2012. Detecting global and local hippocampal shape changes in alzheimer's disease using statistical shape models. Neuroimage 59 (3), 2155–2166.
Styner, M., Gerig, G., Joshi, S., Pizer, S., 2003. Automatic and robust computation of 3d medial models incorporating object variability. Int. J. Comput. Vis. 55 (2–3), 107–122.
Styner, M., Lieberman, J.A., Pantazis, D., Gerig, G., 2004. Boundary and medial shape analysis of the hippocampus in schizophrenia. Med. Image Anal. 8 (3), 197–203.
Svensén, M., Bishop, C.M., 2005. Robust Bayesian mixture modelling. Neurocomputing 64, 235–252.
Vaillant, M., Glaunès, J., 2005. Surface matching via currents. In: Biennial International Conference on Information Processing in Medical Imaging. Springer, pp. 381–392.
Wang, F., Vemuri, B.C., Rangarajan, A., 2006. Groupwise point pattern registration using a novel cdf-based Jensen-Shannon divergence. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), 1. IEEE, pp. 1283–1288.
Wang, Y., Peterson, B.S., Staib, L.H., 2003. 3D brain surface matching based on geodesics and local geometry. Comput. Vis. Image Underst. 89 (2), 252–271.
Zagorchev, L., Meyer, C., Stehle, T., Kneser, R., Young, S., Weese, J., 2011. Evaluation of traumatic brain injury patients using a shape-constrained deformable model. In: International Workshop on Multimodal Brain Image Analysis. Springer, pp. 118–125.
Zagorchev, L., Meyer, C., Stehle, T., Wenzel, F., Young, S., Peters, J., Weese, J., Paulsen, K., Garlinghouse, M., Ford, J., et al., 2016. Differences in regional brain volumes two months and one year after mild traumatic brain injury. J. Neurotrauma 33 (1), 29–34.
Zhou, Z., Zheng, J., Dai, Y., Zhou, Z., Chen, S., 2014. Robust non-rigid point set registration using student's-t mixture model. PLoS ONE 9 (3), e91381.