Generalised coherent point drift for group-wise registration of multi-dimensional point sets

Nishant Ravikumar[1,2] Ali Gooya[1,3], Alejandro F. Frangi[1,3] and Zeike A. Taylor[1,2]

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

Abstract. In this paper we propose a probabilistic approach to group-wise registration of unstructured high-dimensional point sets. We focus on registration of generalised point sets which encapsulate both the positions of points on surface boundaries and corresponding normal vectors describing local surface geometry. Richer descriptions of shape can be especially valuable in applications involving complex and intricate variations in geometry, where spatial position alone is an unreliable descriptor for shape registration. A hybrid mixture model combining Student's t and Von-Mises-Fisher distributions is proposed to model position and orientation components of the point sets, respectively. A group-wise rigid and non-rigid registration framework is then formulated on this basis. Two clinical data sets, comprising 27 brain ventricle and 15 heart shapes, were used to assess registration accuracy. Significant improvement in accuracy and anatomical validity of the estimated correspondences was achieved using the proposed approach, relative to state-of-the-art point set registration approaches, which consider spatial positions alone.

1 Introduction

Group-wise registration of hybrid shape representations, defined by spatial positions of points, with associated surface normal vectors (or directional data in general), is useful for a variety of medical imaging and intra-operative guidance applications. High-dimensional features are more descriptive than spatial positions alone, and improve discriminative capacity when establishing correspondences, due to the low probability of matching all features. Such a framework is thus tailored to registering shapes with detailed structural features that are unreliably captured by conventional probabilistic point set registration (PSR) approaches. The latter are in general, well-suited to registering shapes derived from medical images, due to their ability to accommodate varying shape cardinalities and missing data, independence to topology, and robustness to outliers.

While numerous probabilistic approaches to PSR based on spatial positions alone have been proposed [1,2,3,4,5], only two previous studies [6], [7], have attempted to cohesively model both position and surface normal orientation,

within a probabilistic registration framework. In [6], [7] the authors proposed variants of the iterative closest point registration algorithm for pair-wise rigid shape registration. Both methods model the joint probability density function (PDF) of points and their associated surface normals using a combination of Gaussian and Von-Mises-Fisher distributions, in an expectation maximisation (EM)-based mixture modelling framework. In this work we formulate a new hybrid mixture model- (HMM-) based approach to group-wise shape registration that combines Von-Mises-Fisher distributions for orientation data with Student's t distributions for position data. Such a scheme is aimed at leveraging both the effectiveness for complex geometry that hybrid point sets can afford with the robustness to outliers in position data imparted by Student's t distributions. We demonstrate its effectiveness in preserving fine, anatomically significant structural details in brain and cardiac data.

2 Hybrid mixture model-based group-wise registration

Group-wise registration of hybrid point sets comprising positions and directional data (such as surface normals) distributed on a unit sphere, is achieved using a HMM, where, the former are modelled using Student's t-distributions ($\mathcal{S}$), and the latter, using Von-Mises-Fisher distributions ($\mathcal{F}$).

For a group of $k = 1...K$ shapes to be registered, hybrid representations of the shapes $\mathbf{D}_k$ were first constructed from image segmentations, where each data point is a 6-dimensional vector denoted $\mathbf{d}_{ki} = [\mathbf{x}_{ki}, \mathbf{n}_{ki}]$. Here $\mathbf{x}_{ki}$ represents the spatial position of points defining the object boundary, and $\mathbf{n}_{ki}$ represents the associated surface normal vector, for the ith point, of the kth shape in the group. The resulting hybrid point sets were, subsequently, jointly registered and clustered by fitting an M-component HMM to the data. By assuming conditional independence between positions and their associated surface normals, for all data points (given their latent mixture memberships), their joint PDF can be approximated as a product of the individual conditional densities for position and direction. Consequently, by considering all hybrid data points $\mathbf{d}_{ki} \in \mathbf{D}_k$, from all K subjects, to be independent and identically distributed (i.i.d), the log-likelihood function is derived as shown in equation 1. This defines the cost function to be optimised with respect to the mixture model parameters, to jointly register and cluster the hybrid point sets $\mathbf{D}_k \in \mathbb{D}$. Here and throughout, π_j represents the coefficient of the jth component in the HMM. A tractable solution to maximising the log-likelihood is achieved by EM.

$$\log p(\mathbb{D}|\Theta_p, \Theta_n) = \sum_{k=1}^{K} \sum_{i}^{N_k} \log \sum_{j=1}^{M} \pi_j \mathcal{S}(\mathbf{x}_{ki}|\Theta_p) \mathcal{F}(\mathbf{n}_{ki}|\Theta_n) \tag{1}$$

Joint registration and clustering of such data was achieved over two stages: (1) group-wise rigid registration of the hybrid point sets $\mathbf{D}_k$ and mean template $\mathbf{M}$ construction; and (2) group-wise non-rigid registration, wherein, the mean template estimated in stage 1 was non-rigidly registered to each sample

in the group. The set of similarity transformations $\mathbf{T}_k$ estimated in stage 1, aligning the hybrid point sets to the estimated mean template, initialise the subsequent non-rigid registration step (stage 2) by correcting global pose differences across the data set. Stage 2 of the algorithm estimates non-rigid transformations mapping the mean template to each sample shape within the group. The parameters estimated for each of the $j = 1...M$ components of the HMM include: $\{\mathbf{m}_j^p, \sigma^2, \nu_j\} \in \Theta_p$, which represent mean spatial positions, their variance and the degrees of freedom, respectively, for the Student's t-distributions; and $\{\mathbf{m}_j^d, \kappa_j\} \in \Theta_n$, which represent the mean surface normal directions and concentration around the means, respectively, for the Fisher distributions.

Surface normal vectors represent directional data distributed over a unit sphere. Gaussian (GMMs) and Student's t-mixture models (TMMs), are inappropriate for clustering such data and consequently, a mixture of Von-Mises-Fisher (FMMs) distributions, also defined over the spherical domain, is used. Fisher distributions are fully defined by two parameters, namely, the mean direction $\mathbf{m}^d$ and a scalar concentration parameter κ, which is analogous to the precision of a Gaussian distribution. The latter describes the degree of concentration about the mean direction, i.e. high values for κ indicate high concentration along the mean direction [8]. The PDF of a 3-D unit vector $\mathbf{n}$ sampled from a Fisher distribution with mean direction $\mathbf{m}^d$ and concentration κ is expressed as, $p(\mathbf{n}|\mathbf{m}^d, \kappa) = \frac{\kappa}{4\pi \sinh(\kappa)} \exp^{\kappa \mathbf{m}^{d^T} \mathbf{n}}$. With GMMs and TMMs, the PDF of the observed data is evaluated using the squared Mahalanobis distance. This results in less accurate estimates than the cosine similarity measure, used by FMMs, for spherical data. FMMs are well-suited to model such data and were chosen for their simplicity and computational efficiency. The observed surface normals $\mathbf{n}_{ki} \in \mathbf{N}_k$, across a group of K shapes, are considered rotated observations sampled from Fisher distributions with mean directions $\mathbf{m}_j^d$ and concentrations κ_j. The joint PDF for all observed normals across the group (denoted $\mathbf{N}_k \in \mathbb{N}$) is evaluated as, $p(\mathbb{N}|\Theta_n) = \sum\limits_{k=1}^{K} \sum\limits_{i=1}^{N_k} \log \sum\limits_{j=1}^{M} \pi_j p(\mathbf{n}_{ki}|\mathbf{m}_j^d, \kappa_j)$.

$$Q(\Theta_n^{t+1}|\Theta_n^t) = \sum_{k=1}^{K} \sum_{i=1}^{N_k} \sum_{j=1}^{M} P_{kij}^{(t)} \log p(\mathbf{n}_{ki}|\mathbf{m}_j^d, \kappa_j) + \lambda_j(1 - \mathbf{m}_j^{d^T}\mathbf{m}_j^d) \quad (2a)$$

$$\frac{1}{\kappa_j} = \coth(\kappa_j) - \mathbf{m}_j^{d^T} \frac{\sum\limits_{k=1}^{K}\sum\limits_{i}^{N_k} P_{kij}^{(t)}\mathbf{n}_{ki}}{\sum\limits_{k=1}^{K}\sum\limits_{i}^{N_k} P_{kij}^{(t)}}, \mathbf{m}_j^d = \frac{\sum\limits_{k=1}^{K}\sum\limits_{i=1}^{N_k} P_{kij}^{(t)}\mathbf{n}_{ki}}{||\sum\limits_{k=1}^{K}\sum\limits_{i=1}^{N_k} P_{kij}^{(t)}\mathbf{n}_{ki}||} \quad (2b)$$

Maximum likelihood estimates for the Fisher distributions' parameters in the HMM are evaluated at each M-step, by maximising the expectation of the complete data log-likelihood Q (equation 2a), with respect to $\mathbf{m}_j^d$ and κ_j. Here P_{kij}^t represents the posterior probabilities evaluated in the E-step, at the tth EM iteration, which are employed to revise estimates of the FMM parameters, at the $(t+1)^{\text{th}}$ EM iteration. While $\mathbf{m}_j^d$ are estimated analytically, similarly to [8], no

closed-form solution exists for estimating concentrations κ_j. Past approaches employing FMMs have proposed empirically determined approximations [8] and/or adopted numerical estimation techniques. Here, fixed point iteration is used to solve the non-linear equation (2b) resulting from differentiating $Q(\Theta_n^{t+1}|\Theta_n^t)$ with respect to κ_j. $\mathbf{m}_j^d$ are estimated by maximising $Q(\Theta_n^{t+1}|\Theta_n^t)$ subject to the constraint $||\mathbf{m}_j^d||= 1$. This is achieved using a Lagrange multiplier λ_j and maximising the form of Q presented in equation 2a, with respect to $\mathbf{m}_j^d$. Revised estimates for the mean directions $\mathbf{m}_j^d$ are evaluated at each M-step of the algorithm, as shown in equation 2b. Additionally, parameters associated with the Student's t-distributions in the HMM, namely, $\mathbf{m}_j^p$, σ^2, are estimated analytically, while ν_j, are computed numerically, at each M-step, identically to [10].

2.1 Rigid registration

Group-wise rigid registration of hybrid point sets is achieved by maximising the joint likelihood of point positions and surface normal vectors, using EM. As the directional data (surface normal vectors) modelled is consistently oriented, and rotationally dependent, estimation of the desired rotations is driven by both spatial positions and their associated surface normal vectors. Consequently, the joint PDF of spatial positions $\mathbf{x}_{ki} \in \mathbf{X}_k$ and their associated normals $\mathbf{n}_{ki} \in \mathbf{N}_k$, across K shapes in a group (denoted, $\mathbf{X}_k \in \mathbb{X}$ and $\mathbf{N}_k \in \mathbb{N}$) is given by,

$$\log p(\mathbb{X}, \mathbb{N}|\Theta_p, \Theta_n, \mathbb{T}) = \sum_{k=1}^{K} \sum_{i=1}^{N_k} \log \sum_{j=1}^{M} \pi_j \mathcal{S}(\mathbf{x}_{ki}|\mathbf{T}_k(\mathbf{m}_j^p), \sigma^2, \nu_j) \mathcal{F}(\mathbf{n}_{ki}|\mathbf{R}_k \mathbf{m}_j^d, \kappa_j).$$

Correspondingly, the form of Q to be maximised, to estimate the desired similarity transformations $\mathbf{T}_k \in \mathbb{T}$ and model parameters Θ_p, Θ_n, is given by equation 3. Two variants of TMM-based rigid group-wise PSR were proposed recently [1], [10], which utilise spatial positions alone to recover global pose differences across a group of shapes. An extension to these is presented in this study, by incorporating Fisher distributions, to model surface normal vectors, and enable rigid alignment of shapes represented by hybrid feature vectors. A similar approach to estimating rotations based on both positions and surface normals was presented in [6] and [7] although for pair-wise shape registration.

$$Q(\Theta_p^{t+1}|\Theta_p^t) = \sum_{k,i,j=1}^{K,N_k,M} P_{kij}^{\star(t)} \left[- \frac{||\mathbf{x}_{ki} - s_k \mathbf{R}_k \mathbf{m}_j^p - \mathbf{b}_k||^2}{2\sigma^2} + \kappa_j (\mathbf{R}_k \mathbf{m}_j^d)^T \mathbf{n}_{ki} \right] \quad (3)$$

As unit vectors are invariant to translation $\mathbf{b}_k$ and scaling s_k, these parameters are estimated identically to [1]. Rotations meanwhile, are computed by maximising Q (equation 3) with respect to $\mathbf{R}_k$, and expressed as $\mathbf{R}_k^{(t+1)} = \mathbf{U}\mathbf{S}\mathbf{V}^T$. $P_{kij}^{\star(t)}$ represents the set of corrected posterior probabilities, estimated identically to [1]; $\mathbf{U}, \mathbf{V}$ are unitary matrices evaluated by singular value decomposition of the covariance matrix $\mathbf{C}_k^{(t+1)} = \sum_{i,j=1}^{N_k,M} P_{kij}^{\star(t)}[(\mathbf{x}_{ki} - \mathbf{d}_k)(\mathbf{m}_j^p - \mathbf{m}_k)^T + (\mathbf{n}_{ki}\mathbf{m}_j^{dT})]$; and $\mathbf{S}$ is a diagonal matrix given by, $\mathbf{S} = \text{diag}\,(1, 1, \det(\mathbf{U}\mathbf{V}^T))$ enforcing strictly or-

thogonal rotations, similarly to [1,3]. Additionally, $\mathbf{d}_k$ and $\mathbf{m}_k$ represent barycenters evaluated as: $\mathbf{d}_k^{(t+1)} = \dfrac{\sum\limits_{i,j=1}^{N_k,M} P_{kij}^{\star(t)} \mathbf{x}_{ki}}{\sum\limits_{i,j=1}^{N_k,M} P_{kij}^{\star(t)}}$, $\mathbf{m}_k^{(t+1)} = \dfrac{\sum\limits_{i,j=1}^{N_k,M} P_{kij}^{\star(t)} \mathbf{m}_j^p}{\sum\limits_{i,j=1}^{N_k,M} P_{kij}^{\star(t)}}$.

2.2 Non-rigid registration

The mean template estimated during the initial group-wise rigid registration step (comprising both mean spatial positions $\mathbf{m}_j^p$ and mean surface normal directions $\mathbf{m}_j^d$) is non-rigidly registered to each shape in the group. The desired non-rigid transformations are defined with respect to the template $\mathbf{M}$ as: $\mathbf{M} + v^k(\mathbf{M})$ (considering spatial positions $\mathbf{m}_j^p$ alone), where v is a displacement function mapping the template to the kth sample in the group. In [3] the authors show that by regularizing the norm of v, the resulting deformation field is constrained to be smooth. This is akin to employing a prior on the displacement field of the form $p(v) = \exp^{-\frac{\lambda}{2}\phi(v)}$, where $\phi(v)$ represents the regularization term and λ controls the trade-off between registration accuracy and smoothness of the deformation field. Incorporating this prior into the model results in a log-likelihood function expressed as, $\log p(\mathbb{X}|\Theta) = \sum\limits_{k=1}^{K} \sum\limits_{i=1}^{N_k} \log \sum\limits_{j=1}^{M} \pi_j \mathcal{S}(\mathbf{x}_{ki}|v^k(\mathbf{m}_j^p), \sigma^2, \nu_j) + \frac{\lambda}{2}\phi(v^k)$. Using variational calculus, [3] show that the function v, which maximises the log-likelihood, can be expressed as a linear combination of radial basis functions given by, $v^k(\mathbf{q}) = \sum\limits_{j=1}^{M} w_{kj} G(\mathbf{q} - \mathbf{m}_j^p)$. A similar approach for pair-wise registration of shapes using TMMs was proposed in [5], which is reformulated for group-wise registration in this study.

$$Q(\Theta_p^{t+1}|\Theta_p^t) = \sum_{i,j=1}^{N_k,M} -P_{kij}^{\star(t)} \frac{\|\mathbf{x}_{ki} - (\mathbf{m}_j^p + v^k(\mathbf{m}_j^p))\|^2}{2\sigma^2} + \frac{\lambda}{2} \text{Tr}\{\mathbf{W}_k^T \mathbf{G} \mathbf{W}_k\} \quad (4)$$

To register the mean template to the kth sample in the group, the cost function to be maximised with respect to the basis function weights $w_{kj} \in \mathbf{W}_k$, is expressed as shown in equation 4, where $\mathbf{G}$ represents the Gaussian kernel/Gram matrix, $M \times M$ in size, with elements: $G(\mathbf{m}_l^p, \mathbf{m}_m^p) = \exp^{-\|\frac{\mathbf{m}_l^p - \mathbf{m}_m^p}{2\beta}\|^2}$. Here β represents the width of the Gaussian kernel, which regulates the frequencies filtered out when fitting to data (as the Gaussian kernel acts as a low-pass filter) and consequently, controls the smoothness of the resulting deformation field. Both parameters λ and β, responsible for regulating the smoothness of the non-rigid transformations estimated, were fixed at 0.5, for all registration experiments conducted in this study. The basis function weights required to register the mean template to each sample in the group, were estimated similarly to [3].

Table 1: Non-rigid registration accuracy evaluated using clinical data sets.

Method	Neuro ($K = 27$)			Cardiac ($K = 15$)		
	$M = 500$	$M = 800$	$M = 1200$	$M = 500$	$M = 1000$	$M = 1500$
	MSD (mm)	MSD (mm)	MSD (mm)	MSD (mm)	MSD (mm)	MSD (mm)
HMM	**1.92 ± 0.13**	**1.58 ± 0.16**	**0.60 ± 0.22**	**4.20 ± 0.28**	**3.76 ± 0.30**	**3.63 ± 0.34**
TMM-NR	2.09 ± 0.12	1.68 ± 0.11	1.20 ± 0.20	4.55 ± 0.26	4.38 ± 0.29	4.33 ± 0.31
gCPD	2.04 ± 0.11	1.72 ± 0.10	1.24 ± 0.14	4.41 ± 0.23	4.23 ± 0.26	4.14 ± 0.28

3 Validation using clinical data

Two clinical data sets comprising: (a) ($K = 27$) brain ventricles (Neuro); and (b) ($K = 15$) heart ventricles (Cardiac); were used to assess registration accuracy. These data sets were chosen as they pose a significant challenge to establishing valid correspondences. This is due to the narrow separation between left and right lateral ventricles in the former, varied myocardial thickness in the latter, among other factors. Surface meshes were generated from segmentations using a marching cubes-based algorithm, following which hybrid point sets comprising positions and associated surface normals, were extracted. These were subsequently registered using HMM. The surface normals estimated for each sample in the data set, were consistently oriented outwards, prior to registration. The Neuro data set was automatically segmented from T1-weighted MRIs, using the approach proposed in [9]. Registration accuracy was quantified using the mean surface distance (MSD) metric and compared with two state-of-the-art approaches: (1) TMM-based group-wise non-rigid registration (TMM-NR) and (2) a group-wise variant of the coherent point drift method [3] (gCPD). These results are summarised in Table 1, with statistically significant improvements (assessed using paired-sample t-tests and a significance level of 5%) in accuracy highlighted in bold. The proposed HMM approach produced significantly lower registration errors than the point-based approaches (TMM-NR and gCPD) for both datasets, and for all mixture model sizes (M).

Correspondences estimated using HMM (top row), TMM-NR (middle row) and gCPD (bottom row), are depicted in Fig. 1, for two samples from the Neuro data set. Black arrows in these images indicate regions in the posterior horns of the lateral ventricles (missing in the majority of samples in the data set), which lead to non-homologous correspondences using TMM-NR and gCPD. HMM on the other hand (Fig. 1 top row) is robust in this regard and consequently, ensures greater anatomical validity in the estimated correspondences. Green arrows shown in Fig. 1 highlight another region in the inferior horns, for a sample in the data set, where, HMM affords significant improvement in the quality of correspondences established, over TMM-NR and gCPD. The images in columns 2 & 4 of Fig. 1, depict correspondences estimated for each sample (discussed in
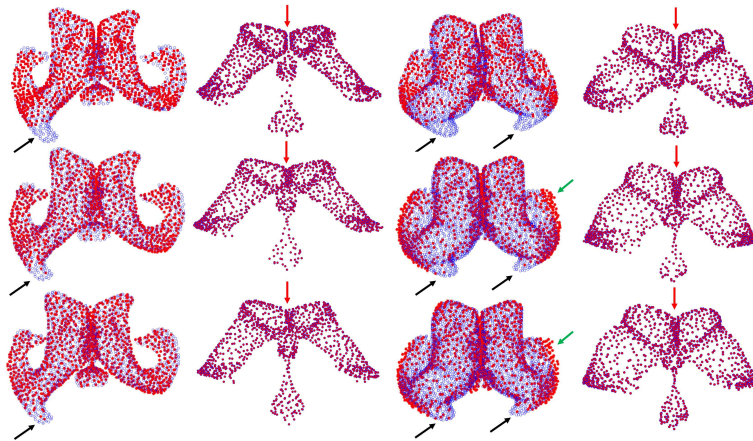
Fig. 1: Two registered samples from the `Neuro` data set. Results from HMM, TMM-NR, and gCPD are shown in the top, middle, and bottom rows, respectively. Columns 1 & 2 and 3 & 4 show results from samples 1 and 2, respectively. Registered shapes (red) overlaid on their corresponding raw samples (blue) are shown in columns 1 & 3. Estimated correspondences are depicted in columns 2 & 4. Black and green arrows indicate regions in the temporal horns where incorrect correspondences are predicted by TMM-NR and gCPD, whereas HMM-generated correspondences remain anatomically valid. Red arrows highlight HMM's effectiveness in preserving separation of the left and right ventricles, which TMM-NR and gCPD cannot achieve.

the preceding set of images), using HMM, TMM-NR and gCPD. Arrows shown in these images highlight the ability of HMM to preserve the separation between the left and right lateral ventricles, not afforded by TMM-NR and gCPD.

Similarly, for the `Cardiac` data set, images presented in Fig. 2 highlight the ability of HMM (top row) to preserve the thickness of the ventricular wall, by establishing correspondences with greater anatomical validity in the epi- and endocardium (see arrows). This is attributed to the higher discriminative power of HMM, which itself results from the richer description of local surface geometry afforded by use of surface normals plus positions, over the latter alone. The advantages identified for HMM from visual inspection of the estimated correspondences for both clinical data sets, is consistent with the registration errors summarised in Table 1. Consequently, the proposed approach is inferred to improve registration accuracy relative to state-of-the-art point-based methods.

4   Conclusions

The proposed group-wise registration approach preserved fine structural details and established correspondences with greater anatomical validity than its point-based counterparts. This was verified by quantitative evaluation of registration
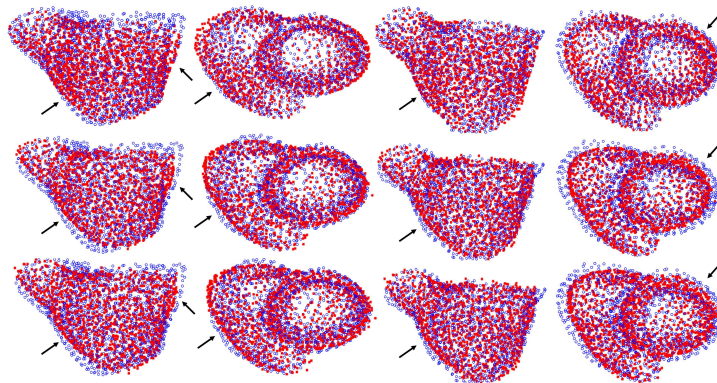
Fig. 2: Two registered samples from `Cardiac` data set. Results from HMM, TMM-NR, and gCPD are shown in the top, middle, and bottom rows, respectively. Columns 1 & 2 and 3 & 4 show results from samples 1 and 2, respectively. Black arrows indicate regions of the epicardium, where HMM establishes correspondences with greater anatomical validity than TMM-NR and gCPD.

accuracy and visual inspection of the established correspondences. Furthermore, the ability of the proposed framework to accommodate different shape cardinalities, its independence of topology, and robustness to outliers, is compelling for its use in various clinical domains.

References

1. Ravikumar, N., et al.: A multi-resolution t-mixture model approach to robust group-wise alignment of shapes. In MICCAI (pp. 142-149), Athens, (2016)
2. Hufnagel, H., et al.: Generation of a statistical shape model with probabilistic point correspondences and the expectation maximization-iterative closest point algorithm. INT J COMPUT ASS RAD, 2(5), 265-273, (2008)
3. Myronenko, A. and Song, X.: Point set registration: Coherent point drift. IEEE T PATTERN ANAL, 32(12), pp.2262-2275, (2010)
4. Chen, T., et al.: Group-wise point-set registration using a novel CDF-based Havrda-Charvt divergence. INT J COMPUT VISION, 86(1), pp.111-124, (2010)
5. Zhou, Z., et al.: Robust Non-Rigid Point Set Registration Using Student's-t Mixture Model. PloS one, 9(3), e91381, (2014)
6. Billings, S. and Taylor, R.: Iterative most likely oriented point registration. In MICCAI (pp. 178-185), Boston (2014)
7. Billings, S. and Taylor, R.: Generalized iterative most likely oriented-point (G-IMLOP) registration. INT J COMPUT ASS RAD, 10(8), pp.1213-1226, (2015)
8. Banerjee, A., et al.: Clustering on the unit hypersphere using von Mises-Fisher distributions. J MACH LEARN RES, pp.1345-1382, (2005)
9. Zagorchev, L., et al.: Differences in regional brain volumes two months and one year after mild traumatic brain injury. J NEUROTRAUM 33(1), 29-34, (2016)
10. Ravikumar, N., et al.: Robust group-wise rigid registration of point sets using t-mixture model. In SPIE Medical Imaging (pp. 97840S-97840S), San Diego, (2016)