# A MULTI-VIEW APPROACH TO CONSENSUS CLUSTERING IN MULTI-MODAL MRI

*C. Andrés Méndez, Gloria Menegaz*

University of Verona
Department of Computer Science
Strada le Grazie, 15 - CA' Vignal
37134 Verona, Italy

*Paul Summers*

European Institute of Oncology
Department of Radiology
Via Ripamonti 435
20141 Milan, Italy

## ABSTRACT

It has been shown that the combination of multi-modal MRI images can improve the discrimination of diseased tissue. The fusion of dissimilar imaging data for classification and segmentation purposes however, is not a trivial task, as there is an inherent difference in information domains, dimensionality and scales. This work proposes a multi-view consensus clustering methodology for the integration of multi-modal MR images into a unified segmentation of tumoral lesions for heterogeneity assessment. Using a variety of metrics and distance functions this multi-view imaging approach calculates multiple vectorial dissimilarity-spaces for each MRI modality and makes use of cluster ensembles to combine a set of unsupervised base segmentations into an unified partition of the voxel-based data. The methodology is demonstrated in application to DCE-MRI and DTI-MR, for which a manifold learning step is implemented in order to account for the geometric constrains of the high dimensional diffusion information.

***Index Terms***— Cluster Ensembles, Clustering, Classification, Segmentation, DTI-MR, DCE-MRI

## 1. INTRODUCTION

Many tumors, such as human gliomas, are characterized by topographically heterogeneous histopathology or have locally evolved to different stages of tumor progression. The heterogeneity that exists within a single tumor requires not only a simple binary distinction between normal and pathologic tissue, but also the development of methods for the assessment and segmentation of subregions, which may help improve treatment planning and management.

Magnetic Resonance imaging is a capable of providing more complete tissue coverage. A discriminating strategy based on a single imaging modality however, is unable to universally differentiate normal from cancerous tissue, thus suggesting the use of a multi-modal view of the tissue for clinical assessment. The fusion of dissimilar imaging data for classification and segmentation purposes is not a trivial task as there is an inherent difference in information domains, dimensionality and scales [2].

Particularly interesting MR modalities for tissue characterization are Dynamic Contrast Enhanced MRI (DCE-MRI), that uses serial image acquisition after the intravenous injection of a contrast agent, and Diffusion Tensor Imaging (DTI-MR), sensitive to the directionality of the microscopic diffusion of water molecules in tissue. In DTI a diffusion tensor (DT), a $3 \times 3$ positive-definite symmetric matrix, is calculated for each voxel from measurements in several directions of diffusion-sensitized magnetic gradients. Each DT characterizes the directionality and magnitude of the anisotropic diffusion occurring in that particular voxel. Their mathematical definition restricts the DTs to lie on a manifold in $\mathbb{R}^6$. Tissue anisotropy and underlying geometry confine the local neighboring tensors to a more restricted submanifold in $\mathbb{R}^6$ [3]. The approach of writing 6 DT components as a feature vector is hindered by the non-linear nature of DT. To address this problem, there have been attempts to use manifold learning techniques such as ISOMAP [25]. Alternatively, kernel methods for manifold learning in DTI were initially used in Ref. [9].

Here we present a methodology for the integration of multi-modal MR images to an unsupervised segmentation of tumoral lesions for heterogeneity assessment. It is often the case that the objects to be clustered have multiple facets or views, each conveying information belonging to a different domain. In this work we extend this 'multi-view' notion to the calculation of multiple vectorial dissimilarity spaces for each MRI modality, which are then clustered to produce a set of base segmentations that are combined with a Cluster Ensemble strategy into an unified partition of the voxel data.

Cluster Ensembles (CE) address the problem of combining multiple *base clusterings* of the same set of objects into a single consolidated clustering [4]. Many clustering algorithms can be used to generate base clusterings, however it is desirable for them to have different biases, i.e. making different errors on new instances.The final partition is obtained with a consensus function that maps the set of base clusterings to an integrated final clustering [4].

To represent each MR modality in diverse spaces we relied on the concepts behind dissimilarity based representations (DBR) [6] in which objects are characterized through

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

pairwise dissimilarities instead of using an absolute characterization by a set of features. Using a variety of established metrics and distance functions we calculate several dissimilarity spaces for each MR modality. Kernel Principal Component Analysis (KPCA) was used as a non-linear manifold learning technique to address the geometric constraints of the DT data [8, 7]. A key difference to the way KPCA was employed in Ref. [9], where it was used for statistical analysis of groups using as input the 6 DT elements of each voxel or a given neighborhood in the form of a vector, here KPCA procedure is performed using as inputs the dissimilarity spaces calculated with the DT metrics that make use of the whole tensor information.

## 2. MATERIALS AND METHODS

### 2.1. Overview

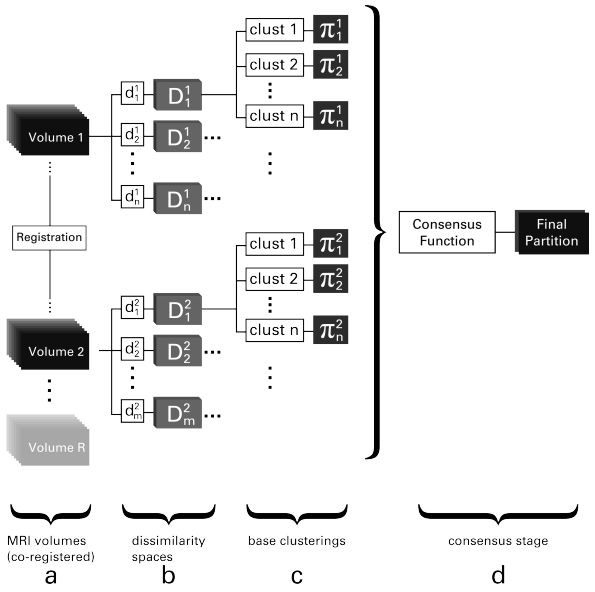



**Fig. 1**. The proposed multi-view methodology for cluster ensembles in multi-modal MRI. The initial multi-modal volumes are spatially registered (a), from each imaging volume an assortment of distance functions ($d_i^j$) is chosen and used to calculate a set of dissimilarity spaces ($D_i^j$) using voxel-wise relationships (b), a set of different clustering algorithms ($clust_{1,\ldots,n}$) are applied to each space and an ensemble of base clusterings ($\pi_i^j$) formed (c), a consensus function combines the base partitions into a final unified clustering (d).

The initial spatial registration problem between MRI modalities is not covered here as it is the subject of a wide variety of methods and frequently the strategy is problem-dependent. First a *Region of Interest* (ROI) is delineated around the detected lesion.Afterwards, the pairwise relationships between voxels in the ROI are calculated, creating dissimilarity vectorial spaces by DBR. Following this, a set of distinct clustering algorithms is used to partition each of the derived spaces. From a single space and clustering algorithm several base clusterings may be calculated.The labels from all the calculated base clusterings are arranged in an ensemble matrix which serves as input to a *consensus function*. This function evaluates the relationships between all the datapoints belonging to the diverse base clusterings and produces a unified similarity matrix which is later partitioned hierarchically to obtain the final unified result (Fig. 1).

### 2.2. Data Representation in Derived Vectorial Spaces

A critical step in our methodology is the data representation of each MRI modality in a diverse set of derived vectorial spaces. These spaces emphasize certain *views* or aspects from each MRI volume. A dissimilarity space is constructed as a square matrix where every voxel is represented by a row vector calculated by the dissimilarities to each other voxel. Let $X = \{x_1, \ldots, x_n\}$ be a voxel-based dataset. Given a dissimilarity function, a data-dependent mapping $D$ is defined as $D(\cdot, R) : X \rightarrow \mathbb{D}^n$ linking $X$ to a *dissimilarity space* [6]. Every object is described by an $n$-dimensional vector of distances between the object $x$ and all the elements of $X$, such that $D(x, X) = [d(x, x_1) \ldots d(x, x_n)]^T$. The core concept of this multi-view approach is to exploit the uniqueness and commonality of these derived dissimilarity spaces into a unified consensus clustering methodology.

#### 2.2.1. DTI-MR processing

The complexity of the DT data, belonging to a high dimensional geometrical manifold structure, requires a careful selection of methods that guarantee the correct use of the DT information. We chose as a starting point two common scalar measures, that is, the Fractional Anisotropy difference ($ds_{FA}$) and the Mean Diffusivity difference ($ds_{MD}$). Even though they reductionist, their widespread use makes them relevant and well studied in a clinical context. Of more theoretical utility are the measures that use the full tensor information. Measures based on Riemannian geometry take into account the constrain of the diffusion tensors to be positive definite. We considered the *geometric distance* ($dg$) proposed in [17], which belongs to this category and the *Log-Euclidean* metric $d_{LE}$, equivalent to the $d_{L2}$ metric of the logarithm of the tensors [18]. From the statistical category we included the distance function proposed in [19], ($d_{KL}$), based on the square root of the J-divergence (symmetrized Kullback-Leibler) between two Gaussian distributions corresponding to the diffusion tensors being compared. Finally, the Frobenius distance ($d_F$) and angular difference ($d_{ANG}$), although they ignore the actual structure and dependencies among tensors, they are included for evaluation considering their widespread use.

#### 2.2.2. DCE-MRI processing

We developed a distance function $D_{DCE}$ [11, 12] based on the adaptive dissimilarity index first proposed in [10]. For

two voxel-derived perfusion curves $S_1 = (u_1, \ldots, u_p)$ and $S_2 = (v_1, \ldots, v_p)$, the complete distance function $D_{DCE}$ for DCE-MRI derived perfusion curves is defined as follows:

$$D_{DCE}(S_1, S_2) = \frac{2}{1 + exp(k_{DCE}\textsc{Cort}(S_1, S_2))} dH(S_1, S_2) \quad (1)$$

where $S_1 = (u_1, \ldots, u_p)$ and $S_2 = (v_1, \ldots, v_p)$ are two voxel-derived perfusion curves sampled at time instants $(t_1, \ldots, t_p)$ [10]. CORT is the temporal correlation (Eq. 2) and $dH$ is the Hausdorff distance, which is used to measure the value-based similarity between perfusion curves.

$$\textsc{Cort}(S_1, S_2) = \frac{\sum_{i=1}^{p-1}(u_{(i+1)} - u_i)(v_{(i+1)} - v_i)}{\sqrt{\sum_{i=1}^{p-1}(u_{(i+1)} - u_i)^2}\sqrt{\sum_{i=1}^{p-1}(v_{(i+1)} - v_i)^2}} \quad (2)$$

In $D_{DCE}$ (Eq. 1) the parameter $k_{DCE}$ weights the relative contribution of both the value-based similarity and the similarity with respect to their behavior, computed respectively by $dH$ and CORT (Eq. 2).

## 3. RESULTS

We first examine the performance of the cluster ensemble method in simulated DTI data against segmentations based on individual metrics. As a second investigation, we examine the multi-modal combination of DTI and DCE data in a simulated tumour by way of different test cases.

### 3.1. Cluster Ensemble test with a syntehtic DTI dataset

The basis for the test consisted of a synthetic diffusion tensor field (Figure 2) where two zones were created with different orientations of the main eigenvector and a third one laid out along a semi-circular arc. This poses the challenge of discerning not only among diffusion characteristics but also requires the structure of each zone be taken into account. The zones composing the dataset were created to share certain diffusion characteristics among some of them while differing with others. The results are presented in Table 1. For most single metrics the use of KPCA as a manifold learning technique yielded a better score in the evaluation indices. More importantly, the use of our proposed ensemble methodology outperformed all the results obtained through individual segmentations. The test was repeated adding noise (SNR 14 and SNR 6) to the simulated diffusion signal, the results obtained in both cases outperformed the clusterings obtained by single metrics.

### 3.2. Tests with a synthetic multi-modal MRI model

We created a synthetic DTI-MR and DCE-MR dataset for validation, composed of four main zones designed to simulate the characteristics of white matter, vasogenic edema, infiltrated fibers and central tumoral region (Figure 3). We relied

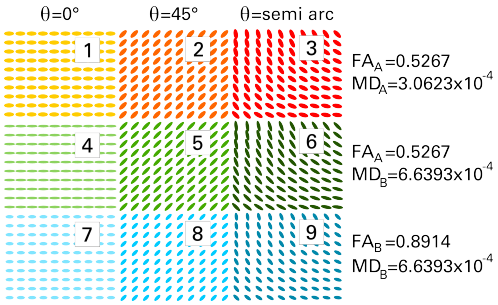



**Fig. 2**. Lay-out of the diverse zones that form the synthetic tensor field. The column on the right shows the mean scalar parameters, FA and MD (in $mm^2/s$).

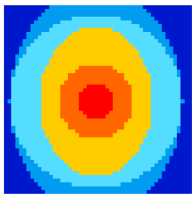
| Method | ARI | RI | CA |
|---|---|---|---|
| dsFA | 0.46 | 0.88 | 0.67 |
| dsFA+KPCA | 0.62 | 0.89 | 0.75 |
| dsMD | 0.34 | 0.81 | 0.54 |
| dsMD+KPCA | 0.56 | 0.89 | 0.67 |
| dANG | 0.33 | 0.84 | 0.44 |
| dANG+KPCA | 0.34 | 0.84 | 0.44 |
| dF | 0.93 | 0.96 | 0.91 |
| dF+KPCA | 0.87 | 0.97 | 0.91 |
| dg | 0.87 | 0.97 | 0.91 |
| dg+KPCA | 0.87 | 0.97 | 0.90 |
| dLE | 0.87 | 0.97 | 0.91 |
| dLE+KPCA | 0.80 | 0.95 | 0.96 |
| dKL | 0.81 | 0.95 | 0.83 |
| dKL+KPCA | 0.98 | 0.98 | 0.99 |
| DT components | 0.74 | 0.94 | 0.84 |
| DT components + KPCA | 0.54 | 0.86 | 0.74 |
| Cluster Ensemble | 0.87 | 0.97 | 0.91 |
| **Cluster Ensemble + KPCA** | **1** | **1** | **1** |

**Table 1**. Evaluation of the methodology with a syntetic DTI dataset, using single metrics and a Cluster Ensemble. The results are scored by means of the Rand Index (RI), Adjusted Rand Index (ARI) and Classification Accuracy (CA).

mainly on the measurements reported by [13, 22, 14]. Further, three subzones were simulated in the white matter zone, representing two differently oriented white matter tracts and a mutual crossing zone. In creating a corresponding synthetic DCE-MRI dataset functionally with the calculated diffusion tensor field we relied on the segmentation of a real glioma MRI volume from a pre-clinical mouse model acquired 35 days after inoculation with glioma cells. This model assumes a relation between the DTI and DCE information in co-registered voxels that might not exist in reality. However for the purposes of an initial methodological validation this presupposed relationship proves to be useful and valid.

### 3.3. Assessment

We hypothesized that for segmentation using our proposed method, the use of base clusterings derived from metrics that use the full DT information with manifold learning would outperform segmentation based on simpler scalar desriptors




| | FA | MD $(mm^2/sec)$ |
|---|---|---|
| White Matter 1 | 0.73 | $7.63\times10^{-4}$ |
| Crossing WM | 0.70 | $7.98\times10^{-4}$ |
| White Matter 2 | 0.72 | $7.81\times10^{-4}$ |
| Vasogenic Edema | 0.41 | $14\times10^{-4}$ |
| Infiltrated Region | 0.29 | $11\times10^{-4}$ |
| Central Tumor | 0.15 | $11\times10^{-4}$ |

**Fig. 3**. The different zones composing the synthetic DTI dataset (left). Table with the mean Fractional Anisotropy (FA) and Mean Diffusivity (MD) values for the corresponding zones (right).

of the DT data. Using the synthetic MRI dataset, distinct vectorial dissimilarity spaces were calculated using the DT metrics described in Sec. 2.2.1: $ds_{FA}$, $ds_{MD}$, $d_{ang}$, $d_F$, $d_g$, $d_{LE}$, $d_{KL}$. We included the 6 DT components as another input vector for the kernel manifold learning processing and subsequent clustering. The first fifty principal components of the DTI-derived spaces were obtained by virtue of the KPCA methodology using a Gaussian Kernel [8]. From the DCE-MRI side, the corresponding dissimilarity spaces were obtained using Eq. 1, varying the tuning parameter $k_{DCE}$ in equation 1 from 1 to 5. For the cluster ensemble we used three algorithms coming from different theoretical domains: the classic K-means, Support Vector Clustering (SVC) and clustering by Affinity Propagation (AP) [15, 16]. All the vectorial spaces were clustered with these algorithms to create a set of base clusterings. The resulting cluster ensemble was then used as input for the link-based consensus function [5]. The output similarity matrix was partitioned hierarchically, after which the final partition was assessed. For our analysis, we defined a set of test cases, as indicated in Table 2. The results are presented in Table 3. As hypothesized, the case using the DT metrics that make use of the full tensor information and manifold learning scored the best in the evaluation. In this case with manifold learning, our technique was able to differentiate both white matter zones and their mutual crossing area, a task where the other cases failed in varying degrees. It is worth noting the diverse results obtained on the three different WM zones in both cases. When using the Cluster Ensemble methodology without manifold learning the method is unable to differentiate the very subtle path variations between the two simulated WM tracts. The use of KPCA allows a finer discrimination based on the manifold to which the tensors belong, in this case a circular and an elliptical path, as well as the mutual crossing zone.

## 4. DISCUSSION

The creation of multiple vectorial spaces from each MRI modality allows each metric to focus on a specific characteristic or view of the multi-dimensional information conveyed by the MRI data using a dissimilarity representation. Our results show a consistent improvement in the assessment

| Method | $ds_{FA}$ | $ds_{MD}$ | $d_{ang}$ | $d_F$ | $d_g$ | $d_{LE}$ | $d_{KL}$ | DT elements |
|---|---|---|---|---|---|---|---|---|
| Case 1 | x | x | x | x | x | x | x | |
| Case 2 | x | x | x | x | x | x | x | x |
| Case 3 | x | x | x | | x | x | x | |
| Case 4 | x | x | x | | x | x | x | x |
| Case 5 | x | x | | | x | x | x | |
| Case 6 | x | x | | | x | x | x | x |

**Table 2**. Definition of the diverse test cases used for evaluation of the methodology with synthetic datasets.

| Method | CA | RI | ARI |
|---|---|---|---|
| Case 1 | 0.719 | 0.826 | 0.515 |
| Case 1 KPCA | 0.864 | 0.915 | 0.750 |
| Case 2 | 0.6848 | 0.801 | 0.489 |
| Case 2 KPCA | 0.842 | 0.884 | 0.662 |
| Case 3 | 0.7312 | 0.831 | 0.533 |
| Case 3 KPCA | 0.866 | 0.911 | 0.7398 |
| Case 4 | 0.726 | 0.833 | 0.529 |
| Case 4 KPCA | 0.852 | 0.900 | 0.7080 |
| Case 5 | 0.7352 | 0.837 | 0.543 |
| Case 5 KPCA | 0.910 | 0.925 | 0.8275 |
| Case 6 | 0.7312 | 0.834 | 0.532 |
| Case 6 KPCA | **0.986** | **0.991** | **0.973** |

**Table 3**. Results of the evaluation performed with synthetic datasets. The results are evaluated by means of the Classification Accuracy (AC), Rand Index (RI) and Adjusted Rand Index (ARI).

scores when using KPCA as a manifold learning technique, a consideration taken specifically to address the geometric structure of the DT data. This outcome shows the importance of including problem-specific knowledge in choosing the appropriate set of metrics or dissimilarity functions. An obvious drawback of this method is the computational cost of obtaining the kernel matrix and the standard eigendecomposition in KPCA, involving a time complexity of $\mathcal{O}(n^3)$. However, various optimization methods have been introduced for kernel methods in general [33, 26, 27, 29] and specifically for KPCA [30, 31, 32, 28]. As presented in Table 3, different combinations of base segmentations derived from contrasting sets of metrics and dissimilarity measures may lend themselves to different results.The definition of a general method to choose the appropriate metrics for different problems remains as an open issue and is problem-dependent.

## 5. CONCLUSION

With this work we have demonstrated that the use of Consensus Clustering techniques in multi-modal medical image segmentations is a promising strategy for assessing the heterogeneity of tumoral regions. Although the choice of the appropriate metrics is problem-dependent and requires further research, our method circumvents some the problems that arise when combining metrics into a single vectorial space or in multiple kernel learning strategies, such as the weighting decision for each individual component [34, 12].

## 6. REFERENCES


[1] D. Yang, Y. Korogi, T. Sugahara, "Cerebral gliomas: prospective comparison of multivoxel 2D chemical-shift imaging proton MR spectroscopy, echoplanar perfusion and diffusion-weighted MRI.," Neuroradiology, vol. 44, no. 8, pp. 656–66, Aug. 2002.

[2] K. M. McMillan, B. P. Rogers "An objective method for combining multi-parametric MRI datasets to characterize malignant tumors," Med.Phys., v.34, no.3,p.1053,2007.

[3] O. Pasternak, R. Verma, N. Sochen, and P. Basser, "On what manifold do diffusion tensors live," MICCAI Workshop, 2008.

[4] J. Ghosh and A. Acharya, "Cluster ensembles," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 1, no. 4, pp. 305–315, Jul. 2011.

[5] N. Iam-On, T. Boongoen, and S. Garrett, "Refining pairwise similarity matrix for cluster ensemble problem with cluster relations," Disc. Science, pp. 222–233, 2008.

[6] E. Pekalska, P. Paclik, and R. P. W. Duin, "A generalized kernel approach to dissimilarity-based classification," J. of Mach. Learning Research, v.2, p. 211, 2002.

[7] B. Schölkopf, A. Smola, and K.R. Muller, "Kernel principal component analysis,". Artificial Neural Networks -ICANN97, Berlin, pp. 583-588, 1997.

[8] K. W. Jorgensen and L. K. Hansen, "Model Selection for Gaussian Kernel PCA Denoising," Trans. on N. Networks and L. Sys.,, v. 23, no. 1, pp. 163-168, 2012.

[9] P. Khurd, R. Verma, and C. Davatzikos, "Kernel-based manifold learning for statistical analysis of diffusion tensor images.," Inf.Proc. in M.Im., v.20,pp.581-93,2007.

[10] A. D. Chouakria and P. N. Nagabhushan, "Adaptive dissimilarity index for measuring time series proximity," Adv. in D. An. and Class., v.1, no.1, pp.5-21, 2007.

[11] C.A. Méndez, F. Pizzorni Ferrarese, P. Summers, G. Menegaz,"Multimodal MRI-based tissue classification in breast ductal carcinoma," Int. Symp. on Biom. Imaging (ISBI 12), pp. 142-145, 2012.

[12] C.A. Méndez, F. Pizzorni Ferrarese, P. Summers, G. Petralia, G. Menegaz, "DCE-MRI and DWI Integration for Breast Lesions Assessment and Heterogeneity Quantification," Int. Journal of Biomedical Imaging vol. 2012, Art. ID 676808

[13] T. Inoue, K. Ogasawara "Diffusion tensor imaging for preoperative evaluation of tumor grade in gliomas.," Cl. neur. and neurosurgery, v.107, no.3, pp.174-80, 2005.

[14] U. Wieshmann and C. Clark, "Reduced anisotropy of water diffusion in structural cerebral abnormalities demonstrated with DTI," M. Res. Im, v.17, no.9, 1999.

[15] J. Yang, V. Estivill-Castro, S.K. Chalup, "Support Vector Clustering Through Proximity Graph Modelling", Proc. 9th Intl Conf. Neural Inf. Proc., 898–903, 2002.

[16] B. J. Frey, D. Dueck, "Clustering by Passing Messages Between Data Points", Science, Vol. 315 no. 5814 pp. 972–976, 2007

[17] P. G. Batchelor, M. Moakher, D. Atkinson, "A rigorous framework for diffusion tensor calculus.," Mag. Res. in Med., vol.53, no. 1, pp.221-5, 2005.

[18] V. Arsigny, P. Fillard, X. Pennec, "Log-Euclidean metrics for fast and simple calculus on diffusion tensors.," Mag. Res. in Med.,vol. 56, no. 2, pp. 411-21, 2006.

[19] Z. Wang and B.C. Vemuri. "DTI segmentation using an information theoretic tensor dissimilarity measure". IEEE Trans. on Med. Im., v.24, pp.1267-1277, 2005.

[20] T.H.J.M. Peeters, P. R. Rodrigues, A. Vilanova, and B. M. ter Haar Romeny, "Analysis of Distance/Similarity Measures for Diffusion Tensor Imaging," Visualization and Processing of Tensor Fields, 2009, pp. 113138.

[21] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Fast and simple calculus on tensors in the log-Euclidean framework." Medical Image Computing and Computer-Assisted Intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention, vol. 8, no. Pt 1, pp. 11522, Jan. 2005.

[22] K. Morita, H. Matsuzawa, Y. Fujii, R. Tanaka, I. L. Kwee, and T. Nakada, "Diffusion tensor analysis of peritumoral edema using lambda chart analysis indicative of the heterogeneity of the microstructure within edema.," Journal of neurosurgery, vol. 102, no. 2, pp. 33641, Feb. 2005.

[23] O. Soderman and B. Jonsson, "Restricted Diffusion in Cylindrical Geometry," Journal of Magnetic Resonance, Series A, vol. 117, no. 1, pp. 9497, Nov. 1995.

[24] A. Barmpoutis, B. Jian, and B. C. Vemuri, "Adaptive kernels for multi-fiber reconstruction." Information processing in medical imaging : proceedings of the ... conference, vol. 21, pp. 33849, Jan. 2009.

[25] P. Khurd, R. Verma, and C. Davatzikos, "On Characterizing and Analyzing Diffusion Tensor Images by Learning their Underlying Manifold Structure," 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW06), pp. 6161, 2006.

[26] B. Scholkopf, S. Mika, C. Burges, P. Knirsch, K. R. Muller, G. Ratsch and A. Smola, "Input space versus feature space in kernel-based methods", IEEE Transactions on Neural Networks 10(5), 10001017, 1999.

[27] C. J. C. Burges and B. Scholkopf, Improving the accuracy and speed of support vector learning machines, Advances in Neural Information Processing Systems,Vol. 9, eds. M. Mozer, M. Jordan and T. Petsche (MIT Press, Cambridge, MA, 1997), pp. 375381.

[28] M. E. Tipping, "Sparse kernel principal component analysis", in NIPS 2000: Neural Information Processing Systems, eds.T. K.Leen,T. G.Dietterich and V.Tresp (MIT Press, 2000), pp. 633639.

[29] X. Jiang, Y. Motai, R. R. Snapp and Xingquan Zhu, Accelerated kernel feature analysis, CVPR 1 (2006) 109116.

[30] T.J. Chin and D. Suter, "Incremental kernel principal component analysis", IEEE Transactions on Image Processing 16(6),16621674. 2007

[31] Y.Washizawa, "Subset kernel PCA for pattern recognition", in Proceedings of the IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops), 2010)

[32] D. Zhang andW. Shi, "An improved kernel principal component analysis for large-scale data set", Lecture Notes in Computer Science 6064,916, 2010

[33] Joachims, T. "Making large-scale Support Vector Machine learning practical" Neural Information Processing Systems, 1999.

[34] Rodrigues P.R., Vilanova A., and Twellmann T., "Adaptive Distance Learning Scheme for Diffusion Tensor Imaging using Kernel Target Alignment," MICCAI Workshop on Computational Diffusion MRI (CDMRI), New York, United States, 2008, pp. 148158.