2014 22nd International Conference on Pattern Recognition

A Kernel-based Representation to Support 3D MRI Unsupervised Clustering

D. Cárdenas-Peña, M. Orbes-Arteaga, A. Castro-Ospina, A. Alvarez-Meza, and G. Castellanos-Dominguez
Signal Processing and Recognition Group
Universidad Nacional de Colombia
Km. 9, vía al aeropuerto, Campus la Nubia, Caldas, Manizales
Email: {dcardenasp,morbesa,aecastroo,amalvarezme,cgcastellanosd}@unal.edu.co

***Abstract*—A new kernel-based image representation is proposed on this paper aiming to support clustering tasks on 3D magnetic resonances images. The approach establishes an effective way to encode inter-slice similarities, so that the main shape information is kept on a lower dimensional space. Additionally, a spectral clustering technique is employed to estimate a compact embedding space where natural groups are easily detectable. Proposed approach outperforms the conventional voxel-wise sum of squared differences on clustering the gender category. Additionally, a pair of eigenvectors describing accurately the subject age is found.**

I. Introduction

Brain Magnetic Resonance (MR) imaging plays an important role on many medical applications. For instance, the identification of changes and functions of a given brain structure along development, aging can help to model the evolution of pathologies such as dementia, Alzheimer and schizophrenia [1]. In electromagnetic source imaging, MR images are employed to build an electric conductivity model of the head, since it is known that a more realistic conductivity modelling enhances the accuracy of the activity reconstruction algorithms [2], [3]. The processing of such medical images for finding spatial characteristics as size, shape and location allows to construct representative anatomical models of a given population [4].

For such applications a segmentation of brain regions and structures is required. Nevertheless, such a task is not easy to perform, mainly due to image artifacts and low inter-structure contrast [5], [6]. Aiming to improve the segmentation accuracy, template-based techniques have rised, which take into account prior spatial distributions of brain structure shapes. To this end, such priors have to be provided as an atlas, which corresponds to a set of shape, intensity and/or functional models of structures [7].

However, such approaches suffer of three main drawbacks. Firstly, segmentation quality is highly dependant on the template-to-image registration algorithm performance. Secondly, given that a ground-truth segmented template is usually drawn by expert clinicians, provided template can produce errors on final segmentation results [8]. Finally, a unimodal distribution of the brain structure shape is imposed, which is not proved [7]. Therefore, multi-atlas segmentation schemes have been proposed recently to overcome such issues.

On such scheme, instead of one template, a set of atlases are registered to a query image. In this sense, the label for each voxel is estimated from the main trend over all atlases, so the influence of misregistered or mislabeled outlier templates is reduced. Nevertheless, for large atlas datasets, the computational burden of registering all templates to a query is highly increased. Moreover, anatomically unrepresentative atlases can bias the solution [2], [7], [8]. Therefore, there is a need for selecting a smaller, more representative subset of atlases from a large set.

In this regard, [2] computed specialized head models for a given subject based on demographic categories such as age, ethnicity, gender and head size. Results on electromagnetic source imaging showed an reduction on localization error against using synthetic and whole-dataset-computed atlases. On [7], an atlas stratification process is performed, which corresponds to find modes on the whole image set, specifically the mean shift algorithm is employed. Although clustering results are provided by means of multidimensional scaling, the original image space is not showed to be compact enough to guarantee the mean estimator to converge. Moreover, there is not an interpretation on the resulting modes. As another approach, [8] performed atlas selection based on a ranking computed from the similarity of the query image and all subjects on the dataset. Measures such as sums of squared differences (SSD), cross-correlation (CC) and mutual information (MI) are employed as similarity measures. Nevertheless, as [7], estimator convergence can not be guaranteed on such super-high-dimensional spaces.

For overcoming such a drawback, some approaches include a dimensionality reduction stage allowing algorithms to work on more compact spaces. For instance, authors in [1], [9], [10] reduce the number of features by grouping voxels into anatomical regions through the alignment of labeled atlases. Then a morphological pathology classification is performed using classical machine learning techniques. On the other hand, proposals as [11], [12] use high level features, such as volume/shape measurements of pre-labeled structures. Nevertheless, all those approaches require a previous alignment stage, so being constrained to the both aforementioned registration issues of computational burden and mis-registration.

In this paper, we proposed a new kernel-based representation for 3D MR images, which reduces the original feature space dimension while encoding the relationships between slices. Additionally, a spectral clustering technique is employed to build a low-dimensional, compact space where the natural groups become evident. Proposed approach is compared against the conventional MR image similarity voxel-

1051-4651/14 $31.00 © 2014 IEEE
DOI 10.1109/ICPR.2014.552

wise sum of squared differences on the task of clustering the demographic categories age and gender. Obtained results, shown in terms of cluster distributions and centroid distances, proved that our approach is able to unfold the subject age category and separate the gender category more accurately than baseline approach. Moreover, provided scatter matrices show that the use of spectral clustering allows to discover the inherent distribution of the data.

This paper is organized as follows: Section II contains the mathematical background of employed techniques and a description of the proposed image representation approach. Then, the experimental setup for testing the methodology and the analysis of resulting clusters is described on Section III. Finally, a discussion of obtained results and some concluding remarks can be found on Section IV.

II. Background

Provided a set of N MR images, $\mathcal{X} = \{\Psi_n \in \Omega : n = 1, \ldots, N\}$, to encode the affinity of a pair of images $\{\Psi_n, \Psi_m\}$, we employ the following kernel function

$$\begin{aligned}\zeta(\Psi_n, \Psi_m) &= f(\Psi_n, \Psi_m) \\ &= \langle \varphi(\Psi_n), \varphi(\Psi_m) \rangle ; m, n = 1, \ldots, N.\end{aligned}$$

where $\varphi(\cdot)$ is a function mapping from the original feature space (Ω) to a Reproduced Kernel Hilbert Space $(\mathbb{K}^q)$. In general, $|\Omega| \ll q$ and $q \to \infty$ are assumed. Nevertheless, through the so called "kernel trick", there is no need for computing $\varphi(\cdot)$ directly [13]. Therefore, a matrix $\mathbf{Z} \in \mathbb{R}^{N \times N}$, with elements $z_{nm} = \zeta(\Psi_n, \Psi_m) \in \mathbb{R}^+$, encoding pair-wise image similarity is estimated from the image set $\mathcal{X}$.

A. Kernel-based Image Representation

Specifically on this work, the original feature space corresponds to $\Omega = \mathbb{R}^{W \times H \times L}$. Therefore, each MR image can be read as an ordered set $\Psi = \{\mathbf{X}_i : i = 1, \ldots, L\}$, with $\mathbf{X}_i \in \mathbb{R}^{W \times H}$ holding $(W \times H)$-sized MR slices. By assuming smooth variations between adjacent slices on Ψ, the inter-slice relationship is encoded by means of the following kernel function:

$$\kappa(\mathbf{X}_i, \mathbf{X}_j) = \langle \varphi(\mathbf{X}_i), \varphi(\mathbf{X}_j) \rangle, i, j = 1, \ldots, L \qquad (1)$$

From Eq. (1), a symmetric matrix $\mathbf{K} \in \mathbf{R}^{L \times L}$ holds elements $k_{ij} = \kappa(\mathbf{X}_i, \mathbf{X}_j)$ with $k_{ij} \in \mathbb{R}^+$. Therefore, each Ψ_n can be represented by its pair-wise slice similarity matrix $\mathbf{K}_n$, yielding the computation of such matrix to a feature extraction stage.

B. Spectral Clustering

Given a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ composed of a set of N vertices $\mathcal{V} = \{v_n \in \mathbb{N} : n = 1, \ldots, N\}$ and a set of edges $\mathcal{E} = \{e_{nm} \in \mathbb{R}^+ : n, m = 1, \ldots, N\}$ linking them, the goal of spectral clustering is to find K disjoint subsets from $\mathcal{V}$. In this regard, $\boldsymbol{Z}$ can be seen as the weighting edge matrix of the undirected, fully connected graph $\mathcal{G}$, where each entry represents the similarity between each image pair, $e_{nm} = z_{nm} = \zeta(\Psi_n, \Psi_m)$.

Since the number of connected vertices in the graph $\mathcal{G}$ corresponds to the eigenvalue multiplicity of the normalized Laplacian matrix of $\boldsymbol{Z}$, such a matrix is defined as [14]:

$$\boldsymbol{L} = \boldsymbol{D}^{-1/2} \boldsymbol{Z} \boldsymbol{D}^{-1/2}$$

where $\boldsymbol{D} \in \mathbb{R}^{N \times N}$ is a diagonal matrix, known as the degree matrix with elements $d_{mm} = \sum_{n=1}^{N} z_{mn}$. Hence, a spectral decomposition of $\boldsymbol{L}$ is required. Specifically, if the K largest eigenvectors of $\boldsymbol{L}$ are stacked as columns on a matrix $\boldsymbol{Y} \in \mathbb{R}^{N \times k}$ and each row is scaled to have unit length, thus mapping the original data points into a unit hypersphere [15], data modes can be inferred by simple clustering techniques while enhancing latent distributions [16].

III. Experimental Setup

The proposed MRI similarity measure approach is tested on a well known image dataset and compared against the baseline mean squared error (MSE) similarity measure. The proposal is outlined on three main stages: *i*) the image preprocessing stage is aimed to reference all images in the dataset to the same intensity space and spatial framework, *ii*) the image representation stage is performed to code the high-dimensional information of each image on a lower feature space, and *iii*) the image embedding stage is employed to build up a new low dimensional space provided with a better interpretability and where inferences can be carried out easily.

A. Database

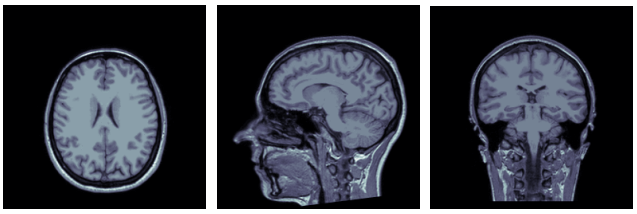
The `IXI` dataset is a brain imaging study, holding MR images from 575 normal subjects which age between 20 and 80 years. Subjects are provided with T1, T2, PD, DTI and angiogram volumes. The image sequences were acquired with three different scanners (Philips 1.5T, Philips 3T and a GE 3T), anonymised and converted to NIFTI format. Additionally, basic demographic information for each subject is included (age, gender, ethnicity, among others). The whole dataset is freely available online [1].

Since the target of the current paper is related to atlas construction, only the T1 sequences of $N = 322$ subjects (acquired with the GE 3T scanner) were taken into account. T1 sequences are composed of $256 \times 256 \times 150$-sized volumes with a voxel size of $0.9375 \times 0.9375 \times 1.2$mm. Thus, the considered subset is composed of 141 male, 175 female and 6 unknown subjects. Figure 1 shows an example of the MR image for a given subject along three different views.

B. Preprocessing

Two preprocessing steps are performed over all images of the considered dataset. Firstly, each image is registered to the MNI305 template by an affine transform so that the whole dataset is referenced to the Talairach space [17]. Due to the registering, each volume is resampled to $197 \times 233 \times 189$ size. Lastly, an intensity normalization procedure is performed by scaling each voxel intensity, so that the mean intensity

[1] `http://www.brain-development.org/`



(a) Axial (b) Sagittal (c) Coronal

Fig. 1. Volume sample from the IXI database. Subject 002

of the white matter is fixed to be 110. Both preprocessing steps, normalization and registrering, are performed with the Freesurfer image analysis suite, which is documented and freely available for download online [2].

C. Image Feature Extraction

For the sake of comparison, two image representation techniques are employed in the current work. The first one is a baseline where each voxel on the image is used as a feature. While the second one corresponds to the proposed approach where each image is represented by an Inter Slice Kernel (ISK), noted as $\mathbf{K}_i \in \mathbb{R}^{L \times L}$. Therefore, feature space dimensions of order 10^6 and 10^4 are achieved for the former and latter approaches, respectively.

Regarding the ISK-based feature extraction, the Gaussian kernel is employed to compute K_{ij} for each image:

$$\kappa\left(\mathbf{X}_i, \mathbf{X}_j\right) = \exp\left(\frac{-||\mathbf{X}_i - \mathbf{X}_j||_F^2}{2\sigma_\kappa^2}\right)$$

where $\sigma_\kappa \in \mathbb{R}^+$ is a scale parameter and notation $\|\cdot\|_F$ stand for the Frobenious norm.

Hence, two important issues have to be highlighted. Firstly, ISK can be computed along three different axes (namely: Axial, Sagittal and Coronal). Therefore, all axes are considered for subsequent analyses. Secondly, σ_κ parameter has to be tuned for every axis. In the concrete case, taking into account that $\lim_{\sigma_k \to 0} \text{Var}(\mathbf{K}(\sigma_k)) = 0$, $\lim_{\sigma_k \to \infty} \text{Var}(\mathbf{K}(\sigma_k)) = 0$ and an appropriate σ_κ value spans widely the values of $\mathbf{K}$, then, the following criterion finds the parameter maximizing the element-wise matrix variance:

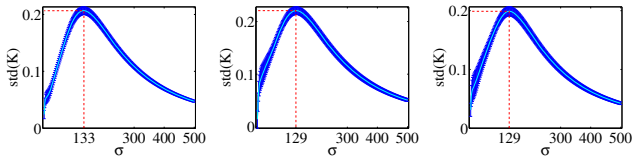
$$\sigma_\kappa^* = \arg\min_\sigma \text{Var}(\mathbf{K}(\sigma)). \tag{2}$$

Obtained tuning curves are shown in Figure 2 for the three considered axes. Mean and standard deviation are computed for a randomly selected subset of 30 MR images. Figure 3 shows an example of the proposed image representation for a given MR image along the axial, sagittal and coronal axes.

D. Unsupervised Learning

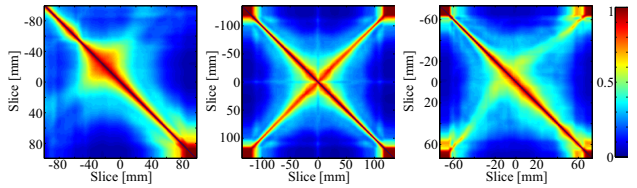
Starting from both aforementioned representation approaches, an embedding low dimensional space is built by using the aforementioned explained spectral decomposition.

[2] http://surfer.nmr.mgh.harvard.edu/



(a) Axial (b) Sagittal (c) Coronal

Fig. 2. Sigma tunning for Inter Slice Kernels along the three possible axes. Mean and standard deviation are plotted.



(a) Axial (b) Sagittal (c) Coronal

Fig. 3. Inter Slice Kernel similarity for a given image. All axes are measured in mm.

Visualization, clustering, and regression are performed over a new more compact space. For this specific task, the common Gaussian kernel is employed:

$$\zeta(\Psi_n, \Psi_m) = \exp\left(\frac{-d(\Psi_n, \Psi_m)^2}{2\sigma_\zeta^2}\right)$$

where $\sigma_\zeta \in \mathbb{R}^+$ is the scale parameter and $d(\Psi_n, \Psi_m) \in \mathbb{R}^+$ is the distance function between the n-th and m-th images. For the baseline image representation approach, that distance corresponds to the voxel-wise image euclidean norm, defined as:

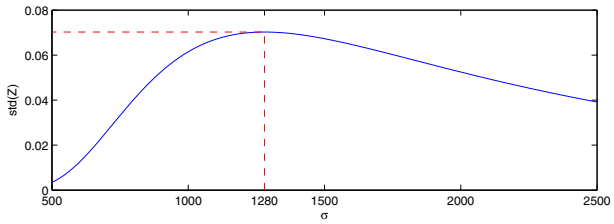
$$d_{VW}(\Psi_n, \Psi_m) = ||vec(\Psi_n) - vec(\Psi_m)||_2$$

while for the proposed representation, the distance is computed as the ISK matrix Frobenius norm as follows:

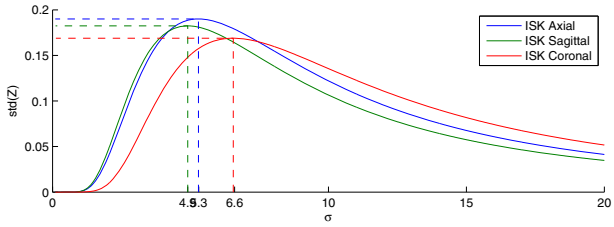
$$d_{ISK}(\Psi_n, \Psi_m) = ||K_n - K_m||_F$$

As said before in Section III-C, the Gaussian kernel parameter has to be appropriately tuned. To this end, the criterion proposed on Equation (2) is used to tune σ_ζ for all considered image representations.

Resulting kernels for considered representations are shown in Figure 5, where subjects are sorted by gender and age values. Although all dataset information is encoded on matrices and some small subsets can be identified, it is still hard to group subjects on categories as gender and age. Therefore, a PCA-based projection space is computed from the laplacian of above matrices. The four largest decomposition eigenvectors are shown in Figures 6 and 7. Obtained projection allows to come up the following statements: i) The first and second decomposition eigenvectors build a space able where the age is "unfolded/unwraped". ii) The fourth decomposition eigenvector decodes the gender category more accurately than remaining components.



(a) Voxel-Wise



(b) Inter Slice Kernel

Fig. 4. σ_ζ tuning curve for considered image representations: (a) Voxel-Wise approach and (b) Proposed ISK along the three possible axes.

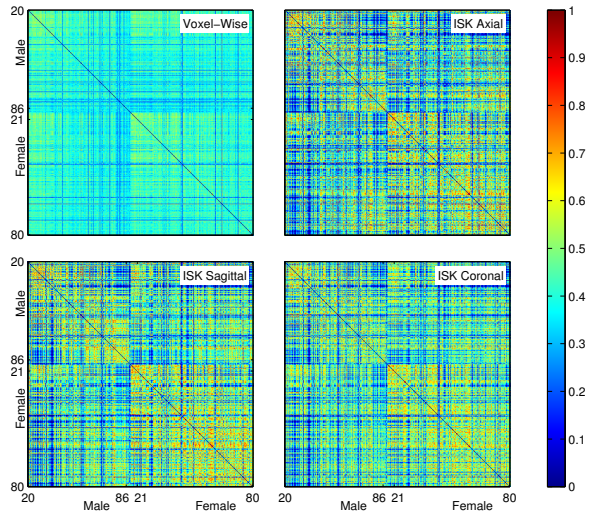


Fig. 5. IXI dataset kernel matrices for both considered image representations. Each row and column on the matrices corresponds to a given image. Images are ordered by gender (firstly) and age (secondly). The color range is normalized between the interval $[0, 1]$

Aiming to prove the first of above statement, two subsequent analyses are performed on the obtained representation space. Firstly, a quadratic regression is computed, so that the trend along the second component is emphasized. Secondly, since the new representation enhances each cluster properties, making the natural groups easily detectable, a simple *k*-means clustering algorithm is used to find, with less difficulties, the natural groups on the new space. As expected, resulting regression (colored in red line) and clustering results (Figure 8(a)) show a trend along the second axis. Besides, the cluster age distribution is provided in Figure 8(b) showing age clustering, proving that the subject age tends to increase along the computed trend curve.

Lastly, for proving the second statement, a cluster measure

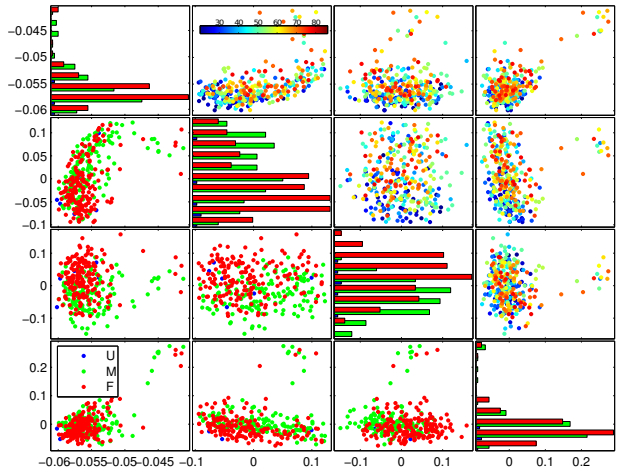


Fig. 6. Scatter matrix for the four largest decomposition eigenvectors using the voxel-wise approach. Upper scatters are grouped by age. Lower scatters are grouped by gender.

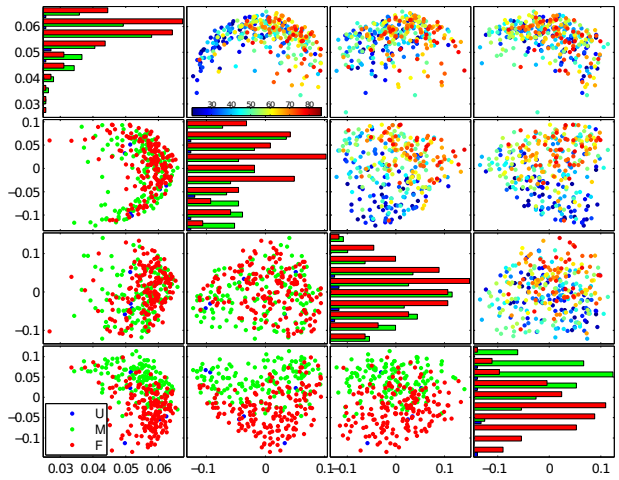
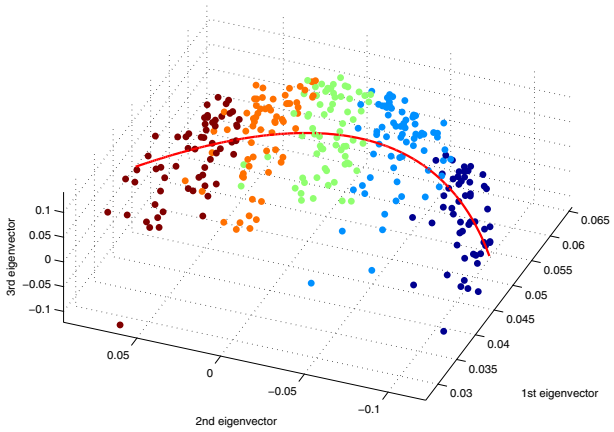


Fig. 7. Scatter matrix for the four largest decomposition eigenvectors using the ISK along the axial axis. Upper scatters are grouped by age. Lower scatters are grouped by gender.

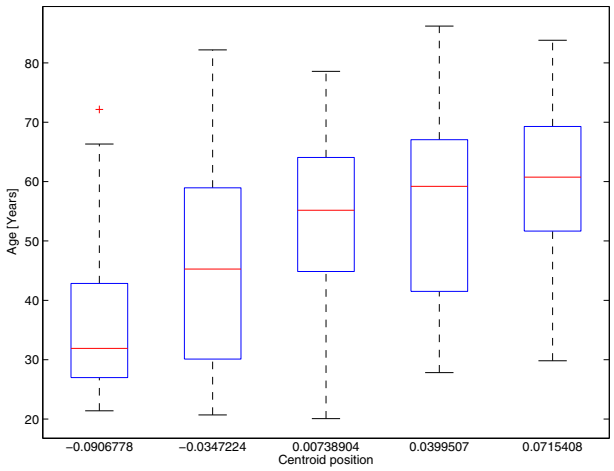
is employed to quantify the separability of male and female clusters. The measure is computed from the inter-cluster over the intra-cluster variance, so that the larger the value, the farther the cluster distributions. The test is carried out along each considered decomposition eigenvectors and over the four dimensional representation space. Obtained separability measures (see Table I) show a larger separation through the fourth component for most of the ISK representations.

Approach	1st	2nd	3rd	4th	All
Baseline	2.33	5.23	8.33	2.62	12.22
ISK Axial	2.74	2.25	2.06	14.66	15.98
ISK Sagittal	3.03	4.61	8.48	11.20	18.74
ISK Coronal	2.57	3.05	12.79	2.72	14.50

TABLE I. CLUSTER SEPARABILITY MEASURE FOR THE FIRST FOUR COMPONENTS OF CONSIDERED REPRESENTATION APPROACHES.



(a) Quadratic regression and clustering



(b) Age distribution for each cluster

Fig. 8. Trend estimation along the second decomposition axis results using the ISK axial image representation.

IV. Discussion and Concluding Remarks

In the current work, a kernel-based image representation is introduced that is specifically devoted to 3D MR image unsupervised clustering. Besides, a low-dimensional, compact space is build by means of spectral clustering, so that the demographic categories are easier distinguished. From the obtained results, the following comments arise:

As seen in Figure 2, the estimated parameters for the Gaussian kernel are close to each other. Therefore, since the inter slice difference is taken from an image for the three axes, the difference dynamic range and the parameter search space are the same for all axes. Moreover, given that the confidence interval, produced by the 30-trials-computed standard deviation, is narrow, the proposed parameter tuning proved to be a stable criterion. In this sense, obtained ISK using estimated parameters maximally enhances the inter-slice relationship, as can be seen in Figure 3.

In Figure 4, obtained parameters for ISK tuning (see Figure 4(b)) lie around the same value. Such a fact may be explained by a couple reasons: i) the feature values range from 0 to 1 for all axes, so the search space is the same. ii) Although the kernel shape is different for the three axes (Figure 3), the latent phenomenon is the same for all axes. Therefore, if above considerations are meet, tuned parameters for different views tend to converge to the same value.

Regarding the age as a demographic category, by visual inspection of the first and second components depicted in Figure 7, it can be seen that the proposed methodology is able to unfold the age better than any other component pair, even on the baseline decomposition results (see Figure 6). Moreover, a quadratic dependance between second and first eigenvectors can be inferred. Additionally, a larger dispersion is shown on older subjects than on younger ones. This finding can be due to a larger head shape dispersion on older humans, which is according to anatomical head knowledge. It is known that brain anatomy is steady on middle age humans, while change (gray matter volume diminish) faster on older humans.

Aiming to prove above statements, a quadratic regression is performed (see Figure 8), which proves to fit adequately the relation between the first two ISK-Axial eigenvectors. Moreover, since the average age on each cluster tends to increase as the centroid position increases, it can be said that subject age is directly described by the relationship between the first two eigenvectors.

Regarding the gender, on a component-wise analysis, the fourth ISK-Axial and third voxel-wise eigenvectors seem more suitable to distinguish gender than remaining components in Figures 6 and 7. Therefore, a two-sample hypothesis test is employed to quantify the separability between gender clusters for the first four eigenvectors of all considered representations. A component-wise and a multivariate test are performed. Results on Table I show that the largest separability is found when using the first four eigenvectors and proved that our proposal is more suitable to distinguish gender than the voxel-wise baseline.

Taking into account the aforementioned results, the proposed kernel-based representation methodology is proved to find the natural inherent distributions of MR images, namely, age and gender categories. In this sense, our proposal is suitable to support MR image clustering and similarity measurement tasks required on template-based image segmentation.

As future work, three main research lines are proposed. i) Given that obtained decomposition vectors do not follow Gaussian distributions, other relaxed embedding techniques, such as local linear embedding and laplacian eigenmaps, will be tested aiming to improve the representation quality. ii) Supervised decomposition techniques will be proved to find representations suitable to distinguish other categories such as ethnicity or pathology subclasses. iii) The most straightforward research direction is to test methodology as a template subset selector on MR image segmentation tasks, so that the structure classification results are improved.

Acknowledgment

This work was supported by *Programa Nacional de Formación de Investigadores "Generación del Bicentenario"*, 2011/2012, *Programa Nacional de Jóvenes investigadores e innovadores*, 2012, and the research project *Desarrollo de un sistema efectivo y apropiado de estimulación del volumen de*

tejido activo cerebral para el mejoramiento de los resultados terapéuticos en pacientes con enfermedad de Parkinson intervenidos quirúrgicamente code 111056934461, all funded by COLCIENCIAS.

REFERENCES

[1] Z. Lao, D. Shen, Z. Xue, B. Karacali, S. M. Resnick, and C. Davatzikos, "Morphological classification of brains via high-dimensional shape transformations and machine learning methods," *NeuroImage*, vol. 21, no. 1, pp. 46–57, Jan. 2004.

[2] P. a. Valdés-Hernández, N. von Ellenrieder, A. Ojeda-Gonzalez, S. Kochen, Y. Alemán-Gómez, C. Muravchik, and P. a. Valdés-Sosa, "Approximate average head models for EEG source imaging." *Journal of neuroscience methods*, vol. 185, no. 1, pp. 125–32, Dec. 2009.

[3] G. Strobbe, D. Cárdenas-Peña, V. Montes-Restrepo, P. Van Mierlo, G. Castellanos-Dominguez, and S. Vandenberghe, "Selecting volume conductor models for EEG source localization of epileptic spikes: preliminary results based on 4 operated epileptic patients," in *First International Conference on Basic and Clinical multimodal Imaging (BaCI)*, vol. c, 2013.

[4] A. Ericsson, P. Aljabar, and D. Rueckert, "Construction of a patient-specific atlas of the brain: Application to normal aging," in *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. Ieee, May 2008, pp. 480–483.

[5] U. Vovk, F. Pernus, and B. Likar, "A review of methods for correction of intensity inhomogeneity in MRI." *IEEE transactions on medical imaging*, vol. 26, no. 3, pp. 405–21, Mar. 2007.

[6] D. Cardenas-Pena, J. D. Martinez-Vargas, and G. Castellanos-Dominguez, "Local binary fitting energy solution by graph cuts for MRI segmentation." *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, vol. 2013, no. 2, pp. 5131–4, Jan. 2013.

[7] D. J. Blezek and J. V. Miller, "Atlas stratification." *Medical image analysis*, vol. 11, no. 5, pp. 443–57, Oct. 2007.

[8] P. Aljabar, R. a. Heckemann, a. Hammers, J. V. Hajnal, and D. Rueckert, "Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy." *NeuroImage*, vol. 46, no. 3, pp. 726–38, Jul. 2009.

[9] J. Ye, G. Alexander, E. Reiman, K. Chen, T. Wu, J. Li, Z. Zhao, R. Patel, M. Bae, R. Janardan, and H. Liu, "Heterogeneous data fusion for alzheimer's disease study," in *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD 08*. New York, New York, USA: ACM Press, 2008, pp. 1025–1033.

[10] B. Magnin, L. Mesrob, S. Kinkingnéhun, M. Pélégrini-Issac, O. Colliot, M. Sarazin, B. Dubois, S. Lehéricy, and H. Benali, "Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI." *Neuroradiology*, vol. 51, no. 2, pp. 73–83, Feb. 2009.

[11] O. Colliot, G. Chetelat, M. Chupin, B. Desgranges, B. Magnin, H. Benali, B. Dubois, L. Garnero, F. Eustache, and S. Lehericy, "Discrimination between Normal Aging by Using Automated Segmentation of the Purpose : Methods : Results : Conclusion :," *Neuroradiology*, vol. 248, no. 1, pp. 194–201, 2008.

[12] E. Gerardin, G. Chételatd, M. Chupin, R. Cuingnet, B. Desgranges, H.-S. Kime, M. Niethammerf, B. Dubois, S. Lehéricy, L. Garnero, F. Eustache, and O. Colliot, "Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging," *NeuroImage*, vol. 47, no. 4, pp. 1476–1486, 2009.

[13] B. Schölkopf and A. J. Smola, *Learning with kernels*. The MIT Press, 2002.

[14] A. Y. Ng, M. I. Jordan, Y. Weiss *et al.*, "On spectral clustering: Analysis and an algorithm," *Advances in neural information processing systems*, vol. 2, pp. 849–856, 2002.

[15] M. Filippone, F. Camastra, F. Masulli, and S. Rovetta, "A survey of kernel and spectral methods for clustering," *Pattern recognition*, vol. 41, no. 1, pp. 176–190, 2008.

[16] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, no. 4, pp. 395–416, 2007.

[17] A. Evans, D. Collins, S. R. Mills, E. D. Brown, R. L. Kelly, and T. Peters, "3D statistical neuroanatomical models from 305 MRI volumes," in *Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record.*, 1993, pp. 1813–1817.