

Characterizing Pathological Deviations from Normality Using Constrained Manifold-Learning

Nicolas Duchateau, Mathieu De Craene,
Gemma Piella, and Alejandro F. Frangi

Center for Computational Imaging & Simulation Technologies in Biomedicine
(CISTIB) – Universitat Pompeu Fabra and CIBER-BBN, Barcelona, Spain

Abstract. We propose a technique to represent a pathological pattern as a deviation from normality along a manifold structure. Each subject is represented by a map of local motion abnormalities, obtained from a statistical atlas of motion built from a healthy population. The algorithm learns a manifold from a set of patients with varying degrees of the same pathology. The approach extends recent manifold-learning techniques by constraining the manifold to pass by a physiologically meaningful origin representing a normal motion pattern. Individuals are compared to the manifold population through a distance that combines a mapping to the manifold and the path along the manifold to reach its origin. The method is applied in the context of cardiac resynchronization therapy (CRT), focusing on a specific motion pattern of intra-ventricular dyssynchrony called septal flash (SF). We estimate the manifold from 50 CRT candidates with SF and test it on 38 CRT candidates and 21 healthy volunteers. Experiments highlight the need of nonlinear techniques to learn the studied data, and the relevance of the computed distance for comparing individuals to a specific pathological pattern.

1 Introduction

By definition, a disease is an impairment of the normal condition of an organism. Considering different grades of the same disease as progressive deviations from normality addresses therefore the understanding of this disease and facilitates its diagnosis in a given patient. This approach is particularly of interest for cardiac resynchronization therapy (CRT), where the definition of relevant criteria for selecting candidates likely to respond to the therapy is still a topic under active debate [5]. The advantages of considering specific groups of mechanical dyssynchrony in the selection process were recently discussed in [11]. Each of these groups corresponded to one pathological pattern of myocardial motion and deformation with different grades of abnormality with respect to a healthy cardiac function. However, the approach lacks of reproducible tools for the grading of a given pathological pattern within a population and for the quantitative comparison of individuals to each of these specific populations. The aim of this paper is to demonstrate the relevance of describing each pathological pattern as a deviation from normality along a manifold structure, allowing the computation of an appropriate distance between individuals and each pathological pattern.

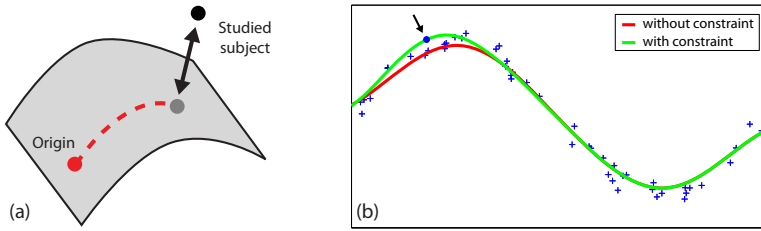


Fig. 1. (a) Distance proposed in this paper, combining a mapping to the manifold and the path along the manifold to reach its origin. (b) Interpolation of a 1D synthetic dataset using inexact matching, before and after the addition of a constraint forcing the curve to pass by the point indicated by the black arrow.

The definition of an optimal space for the comparison of different populations was already addressed by dimensionality reduction techniques, such as principal component analysis (PCA), Kernel-PCA [10], principal geodesic analysis (PGA) [4], linear discriminant analysis [9] or multivariate statistics [14]. Nonetheless, the flexibility of these techniques is limited when a new subject or a new population is added to the existing dataset, as dimensionality reduction is applied to the whole set of studied subjects. In addition, the dimensionality reduction could be biased towards certain populations if they show higher variability. An alternative for moving beyond these limitations consists in separating the analysis for each coherent group of subjects. Using PCA, Kernel-PCA or PGA within a given population is of limited interest for our application, as we target the comparison of individuals to the whole population and not just to its mean or centroid. The comparison of an individual to its k -nearest neighbors (k -NN) does not take into account the local topology of the dataset and assimilates all distances to Euclidean distances [13] [7]. In contrast, manifold learning techniques intrinsically take into account this geometry, and allow relevant comparison of individuals to the studied population through the use of a mapping distance. This mapping results from the “pre-image problem,” used in the literature for denoising [10] [8], segmentation [3], face recognition [15] and regression [1]. A distance based on this mapping mechanism was introduced in [6], but its use was limited to the estimation of reconstruction errors inherent to a reduction of dimensionality.

In this paper, we extend manifold-learning techniques to embed the definition of a relevant origin within the manifold. We propose a distance for comparing individuals to the manifold population, which combines a mapping to the manifold and the path along the manifold to reach its origin (Fig. 1a). The originality of our method resides in the use of motion abnormality maps as input, as introduced in [2], which allows the gradation of the disease and the definition of a physiologically meaningful origin within the manifold, representing a normal motion pattern. Each pathological pattern is therefore considered a deviation from normality along a manifold structure. We present the application of the proposed method to septal flash (SF), a specific motion pattern of intra-ventricular dyssynchrony associated to a high response rate to CRT [11].

2 Methods

The computation of a distance between individuals and a given population considered as a pathologic deviation from normality consists of two steps: the quantification of abnormality for each subject in the dataset, and the estimation of a manifold for this population, constrained to pass by an origin representing normality.

2.1 Atlas-Based Maps of Motion Abnormalities

The inputs for our method consist of 2D spatiotemporal maps of myocardial motion abnormalities, obtained from a statistical atlas of motion built from healthy volunteers [2]. Each map corresponds to one subject in the dataset, and is used as a 2D image in which the horizontal dimension is time (systole) and the vertical one is the position along the septum. Each pixel value corresponds to a p -value index used to locally encode abnormality, in a logarithmic scale, multiplied by the sign of the radial velocity. This choice was made to highlight the inward and outward events of SF, when present (Fig.2a). The color-code associates blue and red color to highly abnormal inward and outward motion of the septum, respectively. According to these conventions, the origin used to constrain the manifold (Sec. 2.2) is defined as an image having 0 value at each pixel, representing a normal motion pattern.

2.2 Manifold-Based Distances to a Population

All the images considered in this paper belong to an ambient space \mathcal{A} . Let's denote $\mathcal{I} = \{\mathbf{I}_0, \dots, \mathbf{I}_N\} \subset \mathcal{A}$ the dataset of $N + 1$ images used for the manifold estimation. The image \mathbf{I}_0 corresponds to the manifold origin for normality. This image is added to the original dataset $\{\mathbf{I}_1, \dots, \mathbf{I}_N\}$ before any computation, so that every image \mathbf{I}_i , $i > 0$ is connected to \mathbf{I}_0 through the isomap graph resulting from the computations described below. This amounts to considering every element of \mathcal{I} as a deviation from the origin along a specific path on the manifold structure.

The space of manifold coordinates is denoted $\mathcal{C} \subset \mathbb{R}^m$, m being the dimensionality of the manifold, while $f : \mathcal{A} \rightarrow \mathcal{C}$ and $g : \mathcal{C} \rightarrow \mathcal{A}$ stand for the correspondence functions between \mathcal{A} and \mathcal{C} . The computation of these functions is based on interpolation techniques adapted from [1] and explained in the following paragraphs. We denote $d : \mathcal{A} \rightarrow \mathbb{R}$ the metric used to compare elements of \mathcal{A} .

Manifold estimation. The isomap algorithm [13] is used to estimate the manifold. First, a graph is built for the dataset \mathcal{I} , based on the k -NN algorithm, connecting all the images among themselves according to the metric d . Then, Euclidean embedding of the manifold data provides a set of coordinates $\mathcal{X} = \{\mathbf{x}_0, \dots, \mathbf{x}_N\} \subset \mathcal{C}$.

From ambient space to manifold coordinates. The estimation of $f : \mathcal{A} \rightarrow \mathcal{C}$ can be formulated as an exact matching problem on a reproducible kernel Hilbert space \mathcal{V} [12] of functions $\mathcal{A} \rightarrow \mathcal{C}$, namely:

$$\operatorname{argmin}_{f \in \mathcal{V}} \left(\frac{1}{2} \|f\|_{\mathcal{V}}^2 \right) \quad \text{under the constraint } f(\mathbf{I}_i) = \mathbf{x}_i, \forall i \in [0, N] \quad (1)$$

with solution: $f(\mathbf{I}) = \sum_{i=0}^N K_f(\mathbf{I}, \mathbf{I}_i) \cdot \mathbf{a}_i$ with $\mathbf{a}_{\mathcal{I}} := \mathbf{K}_f^{-1} \cdot \mathbf{x}_{\mathcal{I}}$,

where \mathbf{K}_f is the matrix $(K_f(\mathbf{I}_i, \mathbf{I}_j))_{(i,j) \in [0,N]^2}$, K_f being chosen of the exponential form $K_f(\mathbf{I}, \mathbf{J}) := \exp(-d(\mathbf{I}, \mathbf{J})^2/\sigma_f^2)$, $(\mathbf{I}, \mathbf{J}) \in \mathcal{A}^2$, σ_f being its bandwidth, and $\mathbf{a}_{\mathcal{I}}$ and $\mathbf{x}_{\mathcal{I}}$ the vectors $(\mathbf{a}_i)_{i \in [0,N]}$ and $(\mathbf{x}_i)_{i \in [0,N]}$, respectively.

Back from manifold coordinates to the ambient space. The estimation of $g: \mathcal{C} \rightarrow \mathcal{A}$ is a variant of the previous computation, formulated as an inexact matching problem on a reproducible kernel Hilbert space \mathcal{W} [12] of functions $\mathcal{C} \rightarrow \mathcal{A}$, with a constraint to force the manifold to pass by the origin image \mathbf{I}_0 :

$$\operatorname{argmin}_{g \in \mathcal{W}} \left(\frac{1}{2} \|g\|_{\mathcal{W}}^2 + \frac{\gamma}{2} \sum_{i=1}^N d(g(\mathbf{x}_i), \mathbf{I}_i) \right) \quad \text{under the constraint } g(\mathbf{x}_0) = \mathbf{I}_0 \quad (2)$$

with solution: $g(\mathbf{x}) = \sum_{i=0}^N K_g(\mathbf{x}, \mathbf{x}_i) \cdot \mathbf{b}_i$ with $\mathbf{b}_{\mathcal{I}} := (\mathbf{K}_g + \frac{1}{\gamma} \mathbf{M})^{-1} \cdot \mathbf{I}_{\mathcal{I}}$,

where \mathbf{K}_g is the matrix $(K_g(\mathbf{x}_i, \mathbf{x}_j))_{(i,j) \in [0,N]^2}$, K_g being chosen of the exponential form $K_g(\mathbf{x}, \mathbf{y}) := \exp(-\|\mathbf{x} - \mathbf{y}\|^2/\sigma_g^2)$, $(\mathbf{x}, \mathbf{y}) \in \mathcal{C}^2$, σ_g being its bandwidth, \mathbf{M} is the matrix $(M_{i,j})_{(i,j) \in [0,N]^2}$, with $M_{i,i} = 1 \forall i \neq 0$ and 0 otherwise, and $\mathbf{b}_{\mathcal{I}}$ and $\mathbf{I}_{\mathcal{I}}$ the vectors $(\mathbf{b}_i)_{i \in [0,N]}$ and $(\mathbf{I}_i)_{i \in [0,N]}$, respectively.

The addition of such a constraint is illustrated in Fig. 1b, which displays the interpolated curve obtained from a 1D synthetic dataset using inexact matching before and after forcing the curve to pass by one point, as described in Eq. 2.

Mapping to the manifold and induced distance. Any image $\mathbf{I} \in \mathcal{A}$ can be associated to an element of the manifold $\hat{\mathbf{I}}$ by means of the composition of the above-defined functions, using $\hat{\mathbf{I}} = g(f(\mathbf{I}))$. This composition allows defining a distance between any image $\mathbf{I} \in \mathcal{A}$ and the manifold [6], namely: $d_{\text{mapping}}(\mathbf{I}) = d(\hat{\mathbf{I}}, \mathbf{I})$. This distance is complemented by a second one, which compares individuals to normality along the manifold structure: $d_{\text{manifold}}(\mathbf{I}) = \|f(\mathbf{I}) - \mathbf{x}_0\|^2$. Total distance to normality is then written as $\sqrt{(d_{\text{mapping}})^2 + (d_{\text{manifold}})^2}$.

3 Experiments and Results

Patient population and processed data. Using the method presented in Sec. 2.2, a manifold was estimated from a population of 50 CRT candidates with SF. This manifold is expected to represent pathologic deviations from normality, each point of the manifold being a SF pattern. A second dataset was used for testing the distances proposed in Sec. 2.2. This population was made of 38 CRT candidates (7 having SF and 31 without SF) and 21 healthy volunteers. All patient data was acquired before the implantation of the CRT device. The presence of SF was assessed by two experienced cardiologists, from the visual inspection of echocardiographic M-mode images, as described in [11].

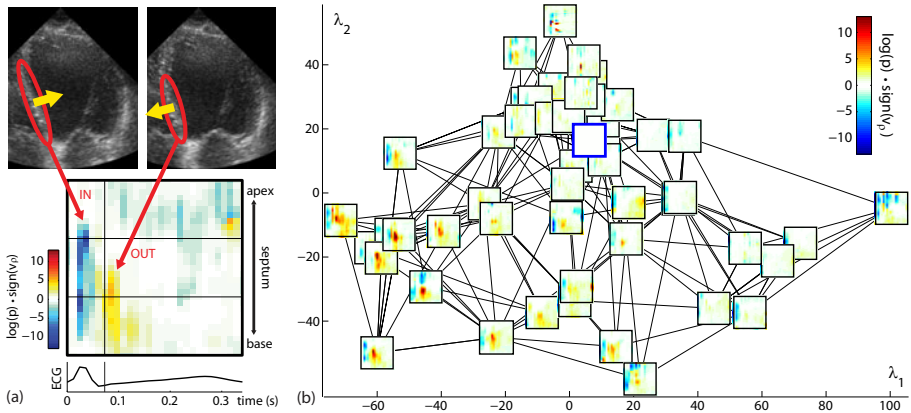


Fig. 2. (a) Map of septal motion abnormalities for one CRT candidate with SF. The color-scale encodes abnormality (p -value) in a logarithmic scale, multiplied by the sign of the radial velocity \mathbf{v}_ρ to highlight SF (Sec. 2.1). (b) Two-dimensional embedding of the manifold of SF p -value maps (output of isomap) according to its two first dimensions. The blue-framed map corresponds to the image used to constrain the manifold, representing a normal motion pattern.

A 2D spatiotemporal map of myocardial motion abnormalities obtained from a statistical atlas of motion [2] was associated to each subject, as explained in Sec. 2.1. The atlas was built from the set of 21 healthy volunteers. The abnormality maps had a size of 31×20 pixels, corresponding to the sampling of the systolic period (horizontal dimension) and the septum along its medial line (vertical dimension), respectively.

The sum of squared differences was used for the metric $d : \mathcal{A} \rightarrow \mathbb{R}$. The number of neighbors for the k -NN computations was set to $k = 5$, which guaranteed that all the images from the manifold dataset were connected among themselves, as tested experimentally. A two-dimensional embedding of the computed manifold (output of isomap) is represented in Fig. 2b, showing the link between each image and its k -NN. The bandwidths for the kernels K_f and K_g introduced in Sec. 2.2 were set to the average k -NN distances over the manifold population and its corresponding set of coordinates \mathcal{X} , respectively. The value of γ involved in the inexact matching problem (Eq. 2) was set to 10, representing a compromise between the smoothness of the manifold and its closeness to the data.

Manifold accuracy. The influence of the manifold dimensionality on the reconstruction error was estimated by computing the average and standard deviation of $d_{mapping}$ over the manifold population, as shown in Fig. 3a. The curve reaches 95% of its final value when the manifold dimensionality is higher than 10, which is the value we chose for the rest of the paper. The reconstruction error obtained with a linear approximation of the manifold dataset (using PCA) is higher, as visible in this figure, justifying the choice of nonlinear techniques to characterize the population of CRT candidates with SF.

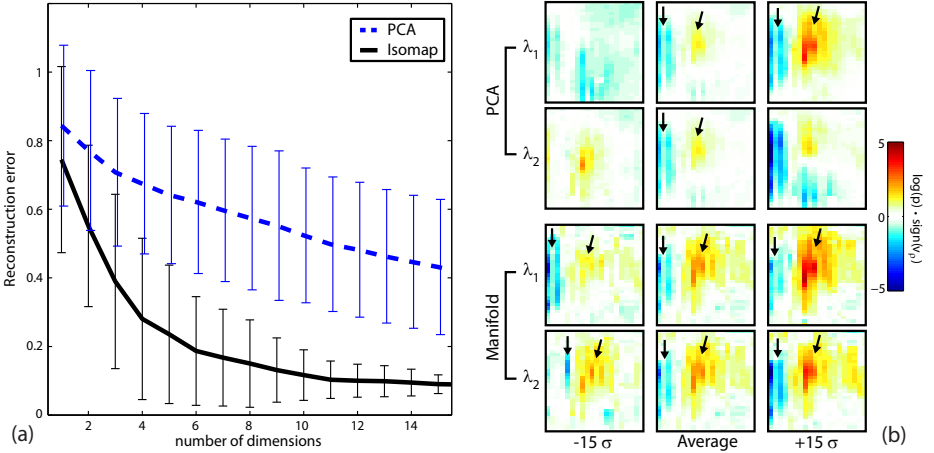


Fig. 3. (a) Evolution of the reconstruction error against the number of dimensions used. Comparison between PCA (dashed blue) and isomap (black). Values correspond to the average \pm standard deviation over the manifold dataset \mathcal{I} . (b) Variations around the average map along the two first principal directions of the manifold dataset \mathcal{I} , obtained using either PCA or isomap. Arrows indicate the inward and outward events of SF, when this pattern is present on the map.

The limitations of PCA on the studied dataset, compared to isomap, are also visible on Fig. 3b. This figure represents the variations around the average map along the two first principal directions of the manifold dataset \mathcal{I} , obtained using either PCA or isomap. As indicated by the black arrows, PCA does not guarantee that the computed maps still correspond to a SF, while this pattern is preserved by the use of manifold-learning.

Distance to the manifold. Figure 4 represents the distance between all the subjects involved in this study and the manifold. We separated the analysis between $d_{mapping}$ and $d_{manifold}$ for interpretation purposes. The patients from the manifold dataset have low $d_{mapping}$, which corresponds to the reconstruction error plotted in Fig. 3a. As the manifold is built from this population, they largely span the space associated to $d_{manifold}$. There is no SF patient from the manifold dataset close to the origin according to $d_{manifold}$, while almost all the healthy volunteers have lower values (vertical lines indicate the median and 1st/3rd quartiles of $d_{manifold}$ for the healthy subjects). This provides an estimation of the threshold above which SF can be detected, as being a deterioration from normality. This threshold may come from the accuracy of the patient selection process using M-mode images [11], and from the minimum accuracy of the abnormality maps [2]. Among the testing subjects, patients having SF are closer to the manifold than patients without SF, according to $d_{mapping}$. Higher values of $d_{manifold}$ are observed in the subjects having higher SF abnormalities on the maps. A larger bandwidth

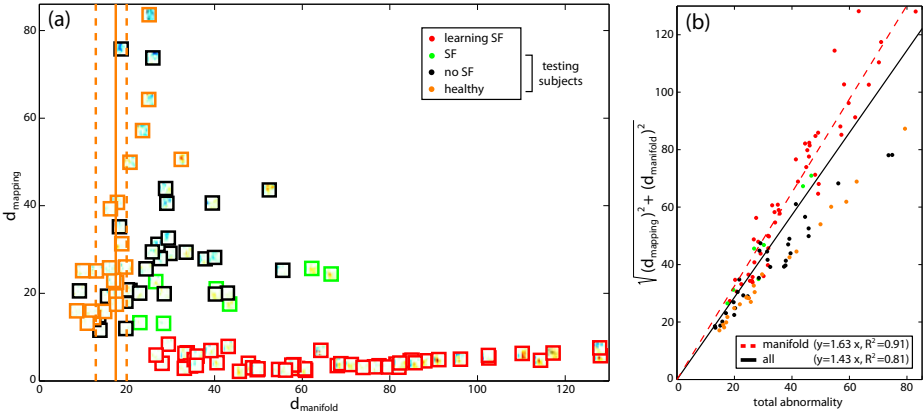


Fig. 4. (a) Subject ordering according to $d_{manifold}$ and $d_{mapping}$, used as horizontal and vertical axis, respectively. Vertical orange lines indicate the median and $1^{st}/3^{rd}$ quartiles of $d_{manifold}$ for the healthy subjects. (b) Patient ordering according to the amount of total abnormality and the proposed distance $\sqrt{(d_{mapping})^2 + (d_{manifold})^2}$.

for the kernels K_f and K_g would bring the testing patients with SF closer to the manifold, but would also increase the reconstruction error.

As the 2D maps processed in this study locally contain a measure of abnormality, it is also of interest to compare the total abnormality of each map against the distances introduced in Sec. 2.2. This comparison is shown in Fig. 4b. Total abnormality was computed for each subject using the L^2 norm of its abnormality map. Linear regression over the plotted data led to R^2 coefficients of 0.91 (manifold data only, dashed red line) and 0.81 (whole data, black line). This suggests that the dimensionality reduction inherent to the manifold estimation preserves the concept of abnormality embedded in the processed maps.

4 Conclusion

We have proposed a method for modeling a specific pathological motion pattern as a manifold. This manifold represents pathological motion as a deviation from normality, being by construction the manifold origin. The method was used to compute a distance between individuals and a given pathological pattern. Experiments demonstrate the need of nonlinear embedding of the learning data, and the relevance of the proposed method for grading different stages of motion abnormality. In the context of CRT, the method can improve the selection of responders to the therapy, allowing reproducible comparison of a new candidate to specific patterns of mechanical dyssynchrony that condition CRT outcome.

Acknowledgments. This work was supported by the Spanish Industrial and Technological Development Center (cvREMOD CEN-20091044), the Spanish

Ministry of Science and Innovation, Plan E and ERDF (STIMATH TIN2009-14536-C02-01), and the European Commission's Seventh Framework Program (EuHeart FP7-ICT-224495). AFF is partially funded by the ICREA-Academia program.

References

1. Davis, B.C., Fletcher, P.T., Bullitt, E., Joshi, S.: Population shape regression from random design data. *Int. J. Comput. Vis.* 90, 255–266 (2010)
2. Duchateau, N., De Craene, M., Piella, G., Silva, E., Doltra, A., Sitges, M., Bijmens, B.H., Frangi, A.F.: A spatiotemporal statistical atlas of motion for the quantification of abnormal myocardial tissue velocities. *Med. Image Anal.* 15(3), 316–328 (2011)
3. Etyngier, P., Keriven, R., Segonne, F.: Projection onto a shape manifold for image segmentation with prior. In: 14th IEEE International Conference on Image Processing, pp. IV361–IV364. IEEE Press, New York (2007)
4. Fletcher, P.T., Lu, C., Pizer, S.M., Joshi, S.: Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Trans. Med. Imaging* 23(8), 995–1005 (2004)
5. Fornwalt, B.K.: The dyssynchrony in predicting response to cardiac resynchronization therapy: A call for change. *J. Am. Soc. Echocardiogr.* 24(2), 180–184 (2011)
6. Gerber, S., Tasdizen, T., Fletcher, P.T., Joshi, S., Whitaker, R.: Manifold modeling for brain population analysis. *Med. Image Anal.* 14(5), 643–653 (2010)
7. Kim, K.-H., Choi, S.: Neighbor search with global geometry: a minimax message passing algorithm. In: 24th International Conference on Machine Learning, pp. 401–408 (2007)
8. Kwok, J.T.-Y., Tsang, I.W.-H.: The pre-image problem in kernel methods. *IEEE Trans. Neural Netw.* 15(6), 1517–1525 (2004)
9. Martinez, A.M., Kak, A.C.: PCA versus LDA. *IEEE Trans. Pattern Anal. Mach. Intell.* 23(2), 228–233 (2001)
10. Mika, S., Schölkopf, B., Smola, A., Müller, K.R., Scholz, M., Rätsch, G.: Kernel PCA and de-noising in feature spaces. *Adv. Neural Inf. Process Syst.* 11, 536–542 (1999)
11. Parsai, C., Bijmens, B.H., Sutherland, G.R., Baltabaeva, A., Claus, P., Marciniak, M., Paul, V., Scheffer, M., Donal, E., Derumeaux, G., Anderson, L.: Toward understanding response to cardiac resynchronization therapy: left ventricular dyssynchrony is only one of multiple mechanisms. *Eur. Heart J.* 30(8), 940–949 (2009)
12. Saitoh, S.: *Theory of Reproducing Kernels and its Applications*. Pitman Res. Notes Math. Ser., p. 189. Wiley, Chichester (1988)
13. Tenenbaum, J.B., De Silva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. *Science* 290(5500), 2319–2323 (2000)
14. Worsley, K.J., Taylor, J.E., Tomaiuolo, F., Lerch, J.: Unified univariate and multivariate random field theory. *NeuroImage* 23(suppl. 1), S189–S195 (2004)
15. Zheng, W.S., Lai, J.-H., Yuen, P.C.: Penalized preimage learning in kernel principal component analysis. *IEEE Trans. Neural Netw.* 21(4), 551–570 (2010)