Computerized Medical Imaging and Graphics 41 (2015) 93–107

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

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

A framework for optimal kernel-based manifold embedding of medical image data

Veronika A. Zimmer[a,*], Karim Lekadir[a], Corné Hoogendoorn[a],
Alejandro F. Frangi[a,b], Gemma Piella[a]

[a] *Information and Communication Technologies Department, Universitat Pompeu Fabra, Barcelona, Spain*
[b] *Department of Mechanical Engineering, The University of Sheffield, United Kingdom*

ARTICLE INFO

Article history:
Received 23 December 2013
Received in revised form 16 April 2014
Accepted 1 June 2014

Keywords:
Nonlinear dimensionality reduction
Kernel principal component analysis
Manifold embedding quality
Multilevel kernel combinations

ABSTRACT

Kernel-based dimensionality reduction is a widely used technique in medical image analysis. To fully unravel the underlying nonlinear manifold the selection of an adequate kernel function and of its free parameters is critical. In practice, however, the kernel function is generally chosen as Gaussian or polynomial and such standard kernels might not always be optimal for a given image dataset or application. In this paper, we present a study on the effect of the kernel functions in nonlinear manifold embedding of medical image data. To this end, we first carry out a literature review on existing advanced kernels developed in the statistics, machine learning, and signal processing communities. In addition, we implement kernel-based formulations of well-known nonlinear dimensional reduction techniques such as Isomap and Locally Linear Embedding, thus obtaining a unified framework for manifold embedding using kernels. Subsequently, we present a method to automatically choose a kernel function and its associated parameters from a pool of kernel candidates, with the aim to generate the most optimal manifold embeddings. Furthermore, we show how the calculated selection measures can be extended to take into account the spatial relationships in images, or used to combine several kernels to further improve the embedding results. Experiments are then carried out on various synthetic and phantom datasets for numerical assessment of the methods. Furthermore, the workflow is applied to real data that include brain manifolds and multispectral images to demonstrate the importance of the kernel selection in the analysis of high-dimensional medical images.

1. Introduction

Manifold embedding is a widely used technique for the analysis of high-dimensional medical image data. In particular, kernel-based methods have received significant attention due to the distinctive approach used for the nonlinear dimensionality reduction problem. More specifically, kernel methods apply a kernel mapping onto a higher dimensional feature space where the original nonlinear data becomes linear or near-linear [1]. As a result, more complex high-dimensional data can be identified and unfolded. More recently, it has been shown that another major advantage of kernel-based manifold embedding is that it provides a unified framework, i.e., many well-known techniques such as Isomap [2], Locally Linear Embedding (LLE) [3], Multidimensional Scaling (MDS) [4], can also be formulated as a kernel problem [5,6].

A key challenge in working with kernel-based manifold embedding, however, is using a kernel transformation that is appropriate to the application and the associated type of nonlinearity. For example, radial basis functions enable to pick out circles (or hyperspheres) in the high-dimensional space, while a linear kernel can better detect lines (or hyperplanes). Despite the importance of this problem [7], the choice of the kernel transformation has received limited attention in the medical image analysis community thus far. In practice, the application of kernel-based embedding is generally achieved using the Gaussian or Polynomial kernels (e.g., [8–13]). However, there is no guarantee that these standard kernels are the best choices for a given dataset or application. The first goal of this paper is therefore to revisit existing kernels in order to promote their use in the medical imaging community for more optimal results.

While the introduction of new functions into the pool of kernel candidates can potentially improve the manifold embedding, given their number and the associated free parameters, the subsequent empirical definition of the optimal kernel transformation

---

* Corresponding author. Tel.: +34 93 542 1348.
E-mail address: veronika.zimmer@upf.edu (V.A. Zimmer).

http://dx.doi.org/10.1016/j.compmedimag.2014.06.001

for each image data type and application becomes infeasible in practice. The second objective of this work is therefore to present a new method to measure the quality of the manifold embedding, and subsequently to automatically select the optimal kernel and associated parameters.

Some of the existing works on the subject are application-specific, such as in Ref. [14] where the labels of the data are used to evaluate the preservation of the intra- and inter-class relationships. However, the technique is only relevant for supervised applications. Other existing quality measures are in nature either local [15–17] or global [18], which could respectively leave global and local inconsistencies undetected. A solution would be to combine local and global measures as proposed by [19], but the optimal weighting of each term in the analysis also depends on the shape of the manifold and associated nonlinearity. Alternatively, the rank coefficient proposed in Ref. [20] has both a local and global nature but it is based on the Euclidean inter-point distances of the embedding space, which does not model well the nonlinear nature of manifolds.

The third goal of this paper is to study the effect of different kernels in manifold embedding on two typical high dimensional problems in medical imaging, namely brain manifold embedding and multispectral image dimensionality reduction. Brain manifold embedding considers a population of brain volumes and generates a lower dimensional representation where the relationships between subjects is easier to visualize and interpret [8,21–23]. In this case, each image is considered as a point in the high dimensional space, where the dimensionality in the space is defined by the number of voxels.

Multispectral imaging is the other typical high dimensional problem considered in this paper. Multispectral images can be very useful for tissue characterization, since the different tissue constituents often require different acquisition parameters and protocol settings in order to be adequately highlighted. This generates a high-dimensional image stack. With the recent advances in multispectral imaging hardware and sequence design, multispectral imaging has gained significant popularity in both research and clinical studies for various modalities [24–26] and applications [27–30]. However, the large number of image spectra that can be obtained (dozens) means nonlinear dimensionality reduction is necessary to summarize all the available information in a lower dimensional space, with minimal loss of information.

This manuscript is an extension of our previous work [31]. Differences include a more detailed presentation of the methodology, a kernel derivation of Isomap, LLE and MDS, a multiscale extension of the quality assessment measure, a numerical evaluation with synthetic manifolds, an application to brain manifold learning, and a validation with a new multispectral dataset of the pancreas. This paper is organized as follows: Sections 2.1 and 2.2 present a review of the kernel principal component analysis (*k*PCA) method, as well as of various advanced kernel transformations used in this study, including kernel derivations of other known manifold embedding techniques. Section 2.3 presents the quality measure and its multiscale extension for the assessment of manifold embedding, while Section 2.4 describes a method to combine multiple kernels based on this measure. Finally, a study of the effect of the kernels in manifold embedding is presented in Section 3, based on various synthetic, phantom and real image datasets.

2. Methods

2.1. Kernel PCA

*K*PCA [1] is a nonlinear dimensionality reduction technique, which maps an input manifold X^p into a lower dimensional representation Y^q, $q < p$. With this method, the data $\{\mathbf{x}_i\}_{i=1}^N$ is transformed nonlinearly onto a feature space F by the mapping $\Phi : X^p \rightarrow F$. This mapping is described by a kernel function, which represents the inner product between the points in the feature space. In an ideal scenario, this step allows the removal of the underlying nonlinearity, which means the standard linear PCA can be applied directly in the feature space. The covariance matrix of the transformed data in the feature space is

$$\mathbf{C} = \frac{1}{N}\sum_{i=1}^{N}\Phi(\mathbf{x}_i)\Phi(\mathbf{x}_i)^T. \tag{1}$$

The eigenvectors of **C** are computed by using the so-called *kernel trick*. The advantage of this approach is that the mapping Φ is never computed explicitly. Only transformations that enable the dot product to be calculated in the original space by a kernel function k are considered. The associated kernel matrix **K** is given by

$$\mathbf{K}_{ij} = k(\mathbf{x}_i, \mathbf{x}_j) = \langle\Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j)\rangle = \Phi(\mathbf{x}_i)\cdot\Phi(\mathbf{x}_j)^T.$$

To compute the eigenvectors of Eq. (1), the covariance matrix has to be diagonalized by solving

$$\mathbf{C}\mathbf{v} = \lambda\mathbf{v}.$$

Since

$$\mathbf{C}\mathbf{v} = \frac{1}{N}\sum_{i=1}^{N}\Phi(\mathbf{x}_i)\Phi(\mathbf{x}_i)^T\mathbf{v} = \lambda\mathbf{v},$$

the eigenvector **v** is

$$\mathbf{v} = \frac{1}{N\lambda}\sum_{i=1}^{N}\Phi(\mathbf{x}_i)\Phi(\mathbf{x}_i)^T\mathbf{v} = \frac{1}{N\lambda}\sum_{i=1}^{N}(\Phi(\mathbf{x}_i)\cdot\mathbf{v})\,\Phi(\mathbf{x}_i)^T.$$

Because $\Phi(\mathbf{x}_i)\cdot\mathbf{v}$ is a scalar, all solutions **v** with $\lambda \neq 0$ lie in the span of $\mathbf{x}_1, \ldots, \mathbf{x}_N$, i.e.,

$$\mathbf{v} = \sum_{i=1}^{N}\alpha_i\Phi(\mathbf{x}_i),$$

thereby reducing the problem to finding the coefficients α_i. This is done by solving the eigenvalue equation

$$N\lambda\boldsymbol{\alpha} = \mathbf{K}\boldsymbol{\alpha},$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)^T$ and $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$. Finally, to obtain the low dimensional data representation, the data is projected onto the eigenvector **v**.

Note that we assumed in Eq. (1) that the data in the feature space has zero mean. If this were not the case, the kernel matrix would be $\widetilde{\mathrm{K}} = \mathbf{HKH}$ with the centering matrix $\mathbf{H} = \mathbf{I}_N - \frac{1}{N}\mathbf{e}\mathbf{e}^T$, where $\mathbf{I}_N$ is the $N \times N$ identity matrix and $\mathbf{e} = (1, \ldots, 1)^T \in \mathbb{R}^N$.

2.2. Kernel functions

An important step in *k*PCA is the choice of the kernel function and its free parameters (e.g., bandwidth for the Gaussian kernel). In this section, we review some alternatives to the kernels more commonly used in the medical imaging community (e.g., Gaussian). Our hypothesis is that by using a more comprehensive list of kernels, the manifold embedding results can be improved by adapting the kernel function to the data manifold and to the application. However, in choosing the kernel functions, it is important that they conform to some conditions so that they can be applied within the *k*PCA framework. In particular, they must be symmetric, continuous and