THE SIGNAL RECONSTRUCTION OF SPEECH BY KPCA

Yan Hui, Zhang Xuegong, Li Yanda   Shen Liqin, Zhu Weibin***

*Institute of information processing, Department of automation, Key State Lab of Intelligent Technology and System, Tsinghua University, Beijing 100084, P.R. China.

Email: yanhui@simba.au.tsinghua.edu.cn

** Speech Group, IBM China Research Lab.

No 26, 6th Street, Shangdi, Haidian District Beijing, 100085, China

Topic J

ABSTRACT

A new method for speech signal reconstruction is proposed by performing a nonlinear Kernel Principal Component Analysis (KPCA). By the use of kernel functions, one can efficiently compute principal components in high-dimensional feature spaces, and reconstruct vectors mapping from input space by those dominant principal components. As the reconstructed vectors is expressed in high dimensional feature space and they could not exist pre-image in input space. For finding pre-image, we use iteration method to approximate the pre-image. The experimental results using KPCA in data reconstruction and de-noising in speech signal show that it had many potential advantages comparing with PCA.

1.PRINCIPLE

Principal Component Analysis (PCA) is an orthogonal basis transformation. The new basis is founded by diagonalizing the centered covariance matrix of a data set, The coordinates in the Eigenvector basis are called principal components. The size of an Eigenvalue corresponding to an Eigenvector v of covariance matrix equals the amount of variance in the direction of v. Furthermore, the directions of the first n Eigenvectors corresponding to the biggest n Eigenvalues cover as much variance as possible by n orthogonal directions. In many applications they contain the most interesting information: for instance, in data compression, where we project onto the directions with biggest variance to retain as much information as possible, or in de-noising, where we deliberately drop directions with small variance.

Assume that our data is mapped into feature space by nonlinear map

$$\Phi : x -> \Phi(x)\ , \Phi(x_1), \Phi(x_2), \Lambda\ \ , \Phi(x_l)\ , \tag{1}$$

And do PCA for the covariance matrix.

$$\bar{\mathbf{C}} = \frac{1}{l}\sum_{i=1}^{l} \Phi(\mathbf{x}_{\mathbf{i}})\Phi(\mathbf{x}_{\mathbf{i}})^T \tag{2}$$

We have to find Eigenvalues and Eigenvectors satisfying

$$\mathbf{V} \ \ = \bar{\mathbf{C}}\mathbf{V}\ . \tag{3}$$

This implies that we may consider the equivalent system.

$$< \Phi(\mathbf{x}_{\mathbf{j}}),\ l_i \mathbf{v}_{\mathbf{i}} >=< \Phi(\mathbf{x}_{\mathbf{j}}), \bar{C}\mathbf{v}_{\mathbf{i}} > \qquad j = 1, \Lambda\ \ , l\ . \tag{4}$$

By defining $K_{ij} =< \Phi(\mathbf{x}_{\mathbf{i}}), \Phi(\mathbf{x}_{\mathbf{j}}) >$, we get the expression

$$l\, l_i \ \ _{\mathbf{i}} = \mathbf{K} \ \ _{\mathbf{i}} \tag{5}$$

The same as PCA does. We solve the Eigenvalue problem for nonzero Eigenvalues. Clearly, all solutions belonging to nonzero Eigenvalues is principal component in high dimensional feature space. If n is large enough to take into account all directions belonging to Eigenvectors with non-zero Eigenvalue, we can reconstruct signal in feature space rather than in input space.

For getting the project of $\Phi(\mathbf{x})$ on the space F which are spanned by vectors $\mathbf{v}_1, \mathbf{v}_2, \Lambda\ , \mathbf{v}_n$, we can define following project operator **Pn,** representing the summary of project of $\Phi(\mathbf{x})$ on the Eigenvectors corresponding to the first n Eigenvalues.

$$P_n \Phi(\mathbf{x}) = \sum_{i=1}^{n} b_i \mathbf{v}_i \, . \qquad (6)$$

If the first n eigenvalues is big enough, the following equation is proved to be true.

$$P_n \Phi(\mathbf{x}_i) = \Phi(\mathbf{x}_i) \, . \qquad (7)$$

For we can not know if the pre-image of $P_n \Phi(\mathbf{x})$ existed, we try to approximate it by minimizing

$$E_Z = \| P_n \Phi(\mathbf{x}) - \Phi(\mathbf{z}) \|^2 \, . \qquad (8)$$

If we restrict our attention to kernels of the form

$$k(\mathbf{z}, \mathbf{x}) = \exp(-\frac{\| \mathbf{z} - \mathbf{x} \|^2}{c}) \, , \qquad (9)$$

We can devise an iteration scheme for z by

$$\mathbf{z}_{k+1} = \frac{\sum_{j=1}^{n} b_j \sum_{i=1}^{l} a_{ij} \exp(-\| \mathbf{z}_k - \mathbf{x}_i \|^2 / c) \mathbf{x}_i}{\sum_{j=1}^{n} b_j \sum_{i=1}^{l} a_{ij} \exp(-\| \mathbf{z}_k - \mathbf{x}_i \|^2 / c)} \, .$$

(10)

To test the feasibility of the algorithm, we run several toy and real world experiments. They were performed using Gaussian kernels of the form (9). We mainly focused on the application of *de-noising* in speech signal, which differs from *reconstruction* by the fact that we are allowed to make use of the original test data as starting points in the iteration. Using different principal components, we get reconstructed signal and most of noise been removed.

2.EXPERIMENT AND APPLICATION

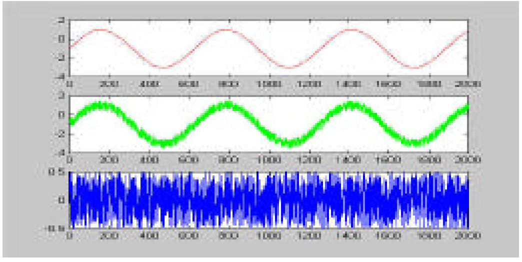

Figure 1: sin signal (From top to bottom in turn is: sin signal, sin signal with noise, white noise)

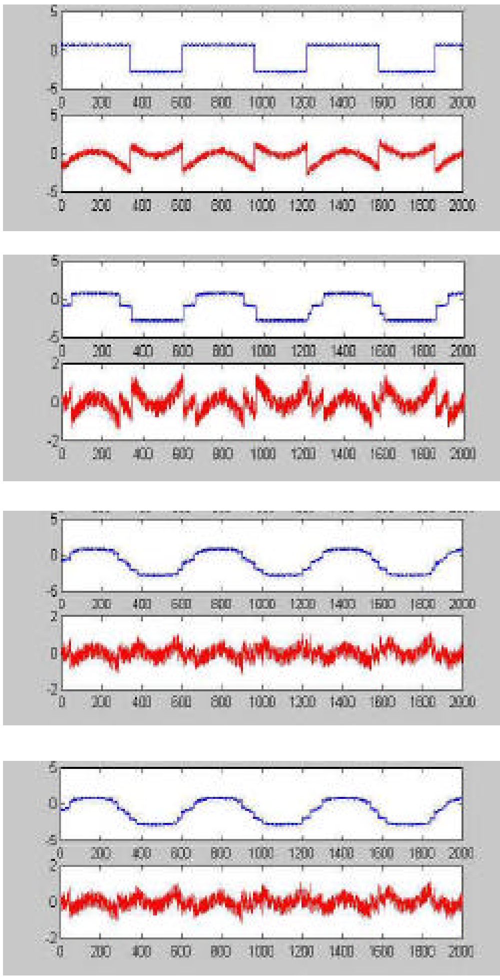

Figure 2: result of de-noising sin signal with different principle components (In each sub-figure, the upper plot is signal reconstructed by principal components)

Figure 2 illustrate four of result after reconstruction based on KPCA. The curve in top left corner is based on one principle component; the curve in top right corner is based on three principle components; the curve in third bottom left is based on five principle components; the curve in bottom right is based on seven principle components. You can see that the quality of signal is better with the increasing of number of principal components.

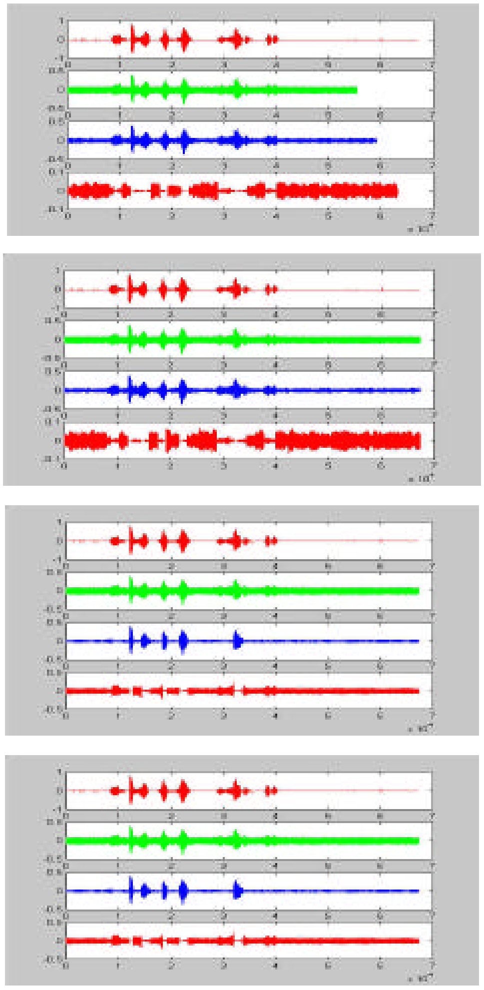

Figure 3: result of de-noising speech signal with different principal components (From top to bottom in turn is: initial speech signal, speech signal with noise, speech signal after de-noising, signal cut from speech signal)

Figure 3 illustrate four of result after reconstruction based on KPCA. The curve in top left corner is based on ten principle component; The curve in top right corner is based on eight principle components; The curve in bottom left corner is based on five principle components; The curve in bottom right corner is based on three principle components. You can also see that the quality of signal is better with the increasing of number of principal components.

3.CONCLUSION

The algorithm can be applied to both reconstruction and de-noising. In the former case, results were comparable to linear PCA, while in the latter case we obtained significantly better results. Our interpretation of this finding is as follows. Linear PCA can extract at most N components, where N is the dimensionality of the data. Being a basis transform, all N components together fully describe the data. If the data are noisy, this implies that a certain fraction of the components will be devoted to the extraction of noise. KPCA, on the other hand, allows the extraction of up to L features, where L is the number of training examples. Accordingly, KPCA can provide a larger number of features carrying information about the structure in the data. In addition, if the structure to be extracted is nonlinear, then linear PCA must necessarily fail.

4.REFERENCE

1. V. Vapnik. 1995. The Nature of Statistical Learning Theory. Springer-Verlag,New York.
2. Bernhard Scholkopf, Alex Smola, Klaus-Robert Muller. Nonlinear Component Analysis as a Kernel Eigenvalue Problem,Neural Computation,Vol.10,Issue 5,pp.1299-1319, The Mit Press ,1998.
3. Bernhard Scholkopf, Alex Smola, Klaus-Robert Muller. Kernel Principal Component Analysis, Technical Report, Max-Planck-Institut f. biol. Kybernetik,1998. GMD FIRST, Rudower Chaussee 5, 12489 Berlin, Germany.
4. Phil Knirsch, Alex Smola, and Chris Burges, Fast Approximation of Support Vector Kernel Expansions, and an Interpretation of Clustering as Approximation in Feature Spaces,Technical Report,GMD FIRST, Rudower Chaussee 5, 12489 Berlin, Bell Labs, Lucent Technologies, Crawfords Corner Rd., Holmdel NJ 07733, USA .