

## ORIGINAL RESEARCH ARTICLE

# Performance evaluation of KPCA pre-imaging methods for speech signal denoising

Shashikant L. Sahare\*, Ashok R. Khedkar, Sandeep S. Musale

Department of Electronics and Telecommunication Cummins, College of Engineering for Women, Pune 411052, India

\* Corresponding author: Shashikant L. Sahare, Shashikant.sahare@cumminscollege.in

### ABSTRACT

Kernel principal component analysis (KPCA) has gained wider interest amongst the researchers in nonlinear dimensionality reduction, data compression, feature extraction, and denoising applications. In KPCA, data from low dimensional input space implicitly mapped to higher dimensional feature space where linear PCA is enforced. However, for denoising application data need to invert back to input space, which is impossible and known as pre-imaging problems. Over recent years, several pre-imaging methods have proposed each with its benefits and disadvantages. In this paper, we evaluated the performance for selected pre-imaging methods for denoising speech signal, whose intelligibility and quality degraded by background noises. Further, we extend our work by comparing the performance of these pre-imaging methods by objective evaluation of denoised speech signal on generated toy examples and NOIZUES database.

**Keywords:** denoising; projective subspace; delay embedding; kernel principal component analysis (KPCA); pre-imaging problem

### ARTICLE INFO

Received: 1 August 2023  
Accepted: 4 September 2023  
Available online: 22 November 2023

### COPYRIGHT

Copyright © 2023 by author(s).  
Journal of Autonomous Intelligence is published by Frontier Scientific Publishing. This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).  
<https://creativecommons.org/licenses/by-nc/4.0/>

## 1. Introduction

Principal component analysis (PCA) is a subspace based unsupervised technique<sup>[1,2]</sup> which projects data from  $d$ -dimensional input space onto a subspace spanned by predefined principal components, thus PCA also synonymously refers as projective subspace technique. However, PCA shows a limitation if the input data is nonlinear. There exist a solution using kernel-PCA (KPCA)<sup>[3,4]</sup> which had gained wide interest over the past decades for nonlinear feature extraction, higher dimensional reduction, and denoising which outperform PCA. In KPCA,  $d$ -dimensional data from input space  $\mathcal{X}$  is implicitly maps to  $M$  (or even infinite) dimension feature space  $\mathcal{F}$  i.e.,  $\varphi : \mathcal{X} \mapsto \mathcal{F}$ , where a traditional PCA is performed. However, these projections are still in  $\mathcal{F}$ , and needs to reverse map back into  $\mathcal{X}$  in order to get denoised data  $\mathbf{z}$  in  $\mathcal{X}$ . Unfortunately, reversing  $\mathcal{F}$  to  $\mathcal{X}$  does not exist, if it exists, then there will be no one to one corresponding mapping. This is called an ill posed problem<sup>[5]</sup> and known as pre-image problem.

Over few decades numbers of pre-imaging methods has been proposed by Mika et al.<sup>[4]</sup>, Kwok and Tsang<sup>[5]</sup>, Rathi et al.<sup>[6]</sup>, Honeine and Richard<sup>[7]</sup>, Abrahamsen and Hansen<sup>[8]</sup>, etc. to approximate pre-images  $\mathbf{z}$  by minimizing the distance between feature space point  $\phi(\mathbf{z})$  and the projection  $P_q\phi(\mathbf{x})$ . However, experimentation and performance evaluation of these methods was performed on generated

toy example and by image denoising. In this paper, we evaluated the performance for selected<sup>[3-8]</sup> pre-imaging methods used for denoising speech signal, whose intelligibility and quality degraded by background noises.

This paper is organized as follows: section II covers brief overview of KPCA<sup>[3]</sup> and its formulation. Section III discusses various existing KPCA de-noising methods. In section IV, first we discussed the experimental results performed on the toy example and second by denoising speech signal from NOIZEUS database followed by objective tests. Section V draws conclusions on KPCA de-noising methods.

## 2. Review of KPCA

In KPCA polynomial, RBF Gaussian, etc., kernels used to map non-linear data from  $d$ -dimensional input space  $\mathcal{X}$  to  $M$ -dimensional such that  $d \ll M$  to feature space  $\mathcal{F}$  as  $\Phi : \mathcal{X} \mapsto \mathcal{F}$ , where linear PCA is enforced. Let  $X = (x_1, x_2, \dots, x_d)'$  be multivariate data in  $\mathcal{X}$  having  $d$ -dimension and  $N$  number of data samples and  $\varphi(x) = [\varphi_1(x), \varphi_2(x), \dots, \varphi_M(x)]^T$  be its corresponding feature vectors in  $\mathcal{F}$ . Before formulating linear PCA in  $\mathcal{F}$ , we must first find centered feature vector as  $\tilde{\varphi}(x) = \varphi(x) - \tilde{\varphi}$  where  $\tilde{\varphi} = N^{-1} \sum_{i=1}^N \varphi(x_i)$ . Centering of feature vectors  $\varphi(x)$  leads to  $N \times M$  centered feature vectors  $\tilde{\Phi} = [\tilde{\varphi}(x_1) \tilde{\varphi}(x_2), \dots, \tilde{\varphi}(x_N)]'$  and the corresponding  $M \times M$ .

Covariance matrix given as

$$C_{\Phi} = \frac{1}{N} \tilde{\Phi} \tilde{\Phi}^T = \frac{1}{N} \sum_{n=1}^N \tilde{\varphi}(x_n) \tilde{\varphi}(x_n)^T \quad (1)$$

Next step is to find nonzero eigenvalues  $\lambda_i \geq 0$  and its corresponding eigenvectors  $v_i \in \mathcal{F} \setminus \{0\}$ , which satisfy eigenvalues equation

$$C_{\Phi} v_i = \lambda_i v_i \quad (2)$$

where  $i = 1, 2, \dots, M$ , and substituting Equation (1) in Equation (2), we get

$$v_i = \frac{1}{N \lambda_i} \sum_{n=1}^N \tilde{\varphi}(x_n) \tilde{\varphi}(x_n)^T \cdot v_i \quad (3)$$

Equation (3) shows that solution of eigenvector must lie in the span of centered feature vectors of input training data i.e.,  $\in \text{span}\{\tilde{\varphi}(x_1), \tilde{\varphi}(x_2), \dots, \tilde{\varphi}(x_N)\}$ , such that  $v_i$  expressed as

$$v_i = \sum_{n=1}^N \alpha_{in} (\tilde{\varphi}(x_n)) \quad (4)$$

where  $\{\alpha_{in}\}_{n=1}^N$  are the coefficients of the expansion. By substituting Equation (4) in Equation (3) and multiplying both side with  $\tilde{\varphi}(x_n)^T$  we get

$$\tilde{K} \alpha_i = v_i \alpha_i \quad (5)$$

where  $\tilde{K} = \tilde{k}(x, x) = \tilde{\Phi}^T \tilde{\Phi}$  is  $N \times N$  a positive semi-defined centered kernel matrix.  $v_i = N \lambda_i$  are ordered eigenvalues and  $\alpha_i$  are corresponding eigenvector of  $\tilde{K}$ . Further eigenvector  $\alpha_i$  can be normalized by  $\alpha_i \leftarrow \alpha_i / \sqrt{v_i}$ . This is ensured by requiring that the eigenvectors  $v_i$  of the covariance matrix are normalized in  $\mathcal{F}$ . The projection  $\tilde{\varphi}(x)$  onto  $i$ -th component given by

$$\beta_i = \tilde{\varphi}(x_n)' v_i = \sum_{n=1}^N \alpha_{in} \tilde{\varphi}(x_n)' \tilde{\varphi}(x_n) = \sum_{n=1}^N \alpha_{in} \tilde{k}(x, x_n) \quad (6)$$

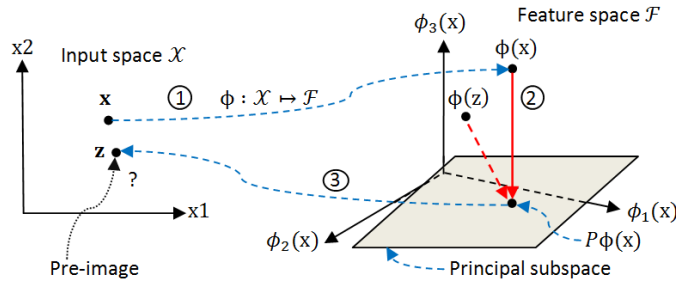
Finally, projecting  $\tilde{\varphi}(x)$  onto subspace spanned by the first  $q$  eigenvectors is given by

$$P_q \phi(x) = \sum_{i=1}^q \beta_i v_i + \tilde{\varphi} = \sum_{i=1}^q \beta_i \sum_{n=1}^N \alpha_{in} \tilde{\varphi}(x_n) + \tilde{\varphi} = \sum_{n=1}^N \tilde{\gamma}_n \tilde{\varphi}(x_n) + \tilde{\varphi} \quad (7)$$

where  $\tilde{\gamma}_n = \sum_{i=1}^q \beta_i \alpha_{in}$ .

### 3. Overview of KPCA denoising methods

In KPCA, denoising in input space  $\mathcal{X}$  requires three steps as shown in **Figure 1**. In first step, test point  $\mathbf{x}$  is explicitly mapped into feature space  $\mathcal{F}$  using polynomial or Gaussian kernel function to get corresponding image  $\phi(\mathbf{x})$ . Second step consist of projecting  $\phi(\mathbf{x})$  onto principal subspace spanned by first  $q$  principal eigenvectors with the largest eigenvalues, yielding projection point  $P\phi(\mathbf{x})$ . Here, the hope is that first  $\lambda_1, \lambda_2, \dots, \lambda_q$  eigenvalues corresponding to the signal of interest and discarding  $\lambda_{q+1}, \lambda_{q+2}, \dots, \lambda_M$  eigenvalues corresponding to noisy components<sup>[3,4]</sup>. In third and last step mapping  $P\phi(\mathbf{x})$  back into  $\mathcal{X}$  gives denoised data  $\mathbf{z}$ . Since, test point  $\mathbf{x}$  explicitly mapped to  $\mathcal{F}$ , reverse mapping is impossible and exact pre-image of  $P\phi(\mathbf{x})$  never exist, if exist there will no one to one mapping. This is ill-posed problem and known as pre-image problem, which consist of finding point  $\mathbf{z} \in \mathcal{X}$  such that minimizing the distance between  $\phi(\mathbf{z})$  and  $P\phi(\mathbf{x})$  or to approximate  $\phi(\mathbf{z}) \simeq P\phi(\mathbf{x})$ .



**Figure 1.** Denoising of data  $\mathbf{x}$  consists of three steps.

#### 3.1. Mika's et al.<sup>[4]</sup> fixed point iteration method

Mika et al.<sup>[4]</sup> pioneered fixed-point iterative method, similar to gradient descent method to minimize the Euclidian distance between  $\phi(\mathbf{z})$  and  $P\phi(\mathbf{x})$  i.e.,

$$\mathbf{z} = \arg \min_{\mathbf{z}} \|\phi(\mathbf{z}) - P\phi(\mathbf{x})\|^2 \quad (8)$$

$\|\phi(\mathbf{z}) - P\phi(\mathbf{x})\|^2 = \|\phi(\mathbf{z})\|^2 - 2\phi(\mathbf{z})^T P\phi(\mathbf{x}) + \|P\phi(\mathbf{x})\|^2$  solving above and neglecting terms independent of  $\mathbf{z}$  gives

$$J(\mathbf{z}) = k(\mathbf{z}, \mathbf{z}) - 2 \sum_{n=1}^N \gamma_n k(\mathbf{z}, \mathbf{x}_n) \quad (9)$$

where  $J(\mathbf{z})$  is called objective function and  $\gamma_n = \tilde{\gamma}_n + N^{-1}(1 - \sum_j^N \tilde{\gamma}_n)$ . Denoised point  $\mathbf{z}$  obtained by differentiating objective function  $J(\mathbf{z})$  w.r.t to  $\mathbf{z}$  and setting it to zero, for RBF Gaussian kernel iteratively updating of  $\mathbf{z}$  given by

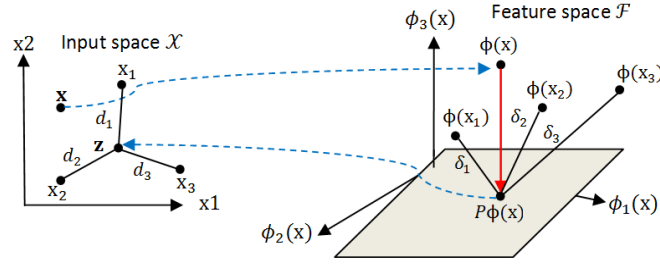
$$\mathbf{z}_{t+1} = \frac{\sum_{n=1}^N \gamma_n \exp(-\|\mathbf{z}_t - \mathbf{x}_n\|^2 / 2\sigma^2) \mathbf{x}_n}{\sum_{n=1}^N \gamma_n \exp(-\|\mathbf{z}_t - \mathbf{x}_n\|^2 / 2\sigma^2)} \quad (10)$$

However, the performance of this method solely relies on initial guesses of  $\mathbf{z}$ . As initialization is undefined, this method suffers from numerical instability and often converges to local minima<sup>[5]</sup>. Moreover, this method is limited to the use of Radial Basis Function (RBF) kernel only. Whereas in polynomial kernel this iterative method fails to converge even after repeated restart<sup>[5]</sup> and restarting with different initial guesses of  $\mathbf{z}$ .

#### 3.2. Kwok and Tsang's<sup>[5]</sup> distance constraint

Let  $d_i = \|\mathbf{z} - \mathbf{x}_i\|^2$  and  $\delta_i = \|P\phi(\mathbf{z}) - \phi(\mathbf{x}_i)\|^2$  be Euclidean distance measure in the input space and feature space as shown in **Figure 2**. For every  $i = 1, 2, \dots, k$ , the input space distance  $\mathbf{d} = \{d_1, \dots, d_k\}$  and the

feature space distance  $\delta = \{\delta_i, \dots, \delta_k\}$  are preserved, using multi-dimensional scaling (MDS) technique such that  $\mathbf{d} = \delta$ .



**Figure 2.** Pre-imaging method proposed by Kwok and Tsang's<sup>[5]</sup>.

This method assumes that pre-images will lie in the span of  $k$  nearest neighbors as compare to all training point in input space in Mika et al.<sup>[4]</sup>. After selecting  $k$  nearest neighbors and then they are centering to get  $\tilde{X} = [\tilde{x}_1, \dots, \tilde{x}_k]$ . Decomposition of  $\tilde{X}$  using SVD yields.

$$\tilde{X} = LDU^T = LS \quad (11)$$

where,  $L = [l_1, \dots, l_k]$  is orthonormal basis matrix of the subspace and  $S = [s_1, \dots, s_k]$  is projection matrix of  $\tilde{X}$  onto  $L$ .

Define  $\mathbf{d}^2 = [d_1^2, \dots, d_k^2]^T$  input space distances and  $\mathbf{d}_c^2 = [\|s_1\|^2, \dots, \|s_k\|^2]^T$  distances from  $x_i$  to centroid  $\tilde{x} = k^{-1} \sum_{j=1}^k x_j$ . Then, from above Kwok and Tsang<sup>[5]</sup> formulate the approximate pre-image to

$$\mathbf{z} = Ls + \tilde{x} \quad (12)$$

where  $s = -0.5 \cdot (SS^T)^{-1} \cdot S \cdot (\mathbf{d}^2 - \mathbf{d}_c^2)$ . This method is non-iterative method, which does not suffer from any local minima problem and simply require algebraic manipulation (12). However, this method relies on the assumption that distances in the input space and feature space are approximately same i.e.,  $\mathbf{d}^2 \simeq \mathbf{d}_c^2$  which is not always possible. If  $\mathbf{d}^2 = \mathbf{d}_c^2$  then  $s = 0$  resulting pre-image  $\mathbf{z}$  equals to nearest neighbor's centroid  $\tilde{x}$ . It is observed that less the number of  $k$ , variation between  $\mathbf{x}$  and  $\mathbf{z}$  will be smaller. Similarly large the number of  $k$ , pre-image  $\mathbf{z}$  will be closer to overall mean of data  $X$ . Thus, performance of this method heavily relies on selection of nearest neighbor  $k$ 's

### 3.3. Rathi's et al.'s<sup>[6]</sup> direct method

Rathi's et al.'s<sup>[6]</sup> modify Mika et al.<sup>[4]</sup> iterative method and does not use MDS (Kwok and Tsang's<sup>[5]</sup>) but use feature space distances to approximate pre-image  $\mathbf{z}$ . Using the relationship between squared input space distance and squared feature space distance and invertible kernel as RBF kernel the approximate pre-image given by Equation (13),

$$\mathbf{z} \approx \frac{\sum_{n=1}^N \gamma_n \left( 1 - \frac{1}{2} \delta^2(P_q \phi(x), \phi(x_n)) \right) x_n}{\sum_{n=1}^N \gamma_n \left( 1 - \frac{1}{2} \delta^2(P_q \phi(x), \phi(x_n)) \right)} \quad (13)$$

Equation (13) shows it is modification of Equation (10). Contrary to [Kwok] this method does not compute SVD thus reduces computational time for higher dimension input data  $X$ .

### 3.4. Honeine's and Richard conformal map approach

Honeine and Richard<sup>[7]</sup> proposed pre-imaging method based on conformal mapping approach which preserve angle between  $x_i$  and  $x_j$  in the input space by preserving inner product distances.

This method is developed in two stages, in first stage basis in reproducing kernel Hilbert space (RKHS)  $H$  is obtained having isometry with the input space basis  $\mathcal{X}$ . Further this isometry is given with respect to the training data sets. In the second method projections into the future space  $\tilde{\varphi}(\cdot)$  is represented in this basis, which

gives the same inner product in the  $\mathcal{X}$ . Let the  $l$  number of basis obtained in  $\mathbf{H}$  be denoted by  $\Psi_x$  by so that we can write in below equation as

$$\psi_k(x) = \sum_{i=1}^n \alpha_{in} k(x_i, x)$$

where  $\Psi_x = [\psi_1(x) \ \psi_2(x) \ \psi_3(x), \dots, \psi_k(x)]^T$ , now  $\Psi_x$  is represented by  $\Psi_x = Ak_x$ . where  $k_x = [k_x(x_1, x), k_x(x_2, x), \dots, k_x(x_n, x)]^T$  and  $A$  is matrix of order a  $l \times n$  of unknowns whose  $(k, i)$ -th entry is  $\alpha_{kn}$ . This matrix  $A$ , needs to be optimization and solution was obtained as

$$\hat{A} = \operatorname{argmin} \frac{1}{2} \|P - KA^T AK\|^2 + \lambda \operatorname{tr}(A^T AK) \quad (14)$$

Derivation of above equation with respect to  $A^T A$  and letting it to zero we get

$$\hat{A}^T \hat{A} = K^{-1}(P - \lambda K^{-1})K^{-1} \quad (15)$$

where  $P = \langle x_i, x_j \rangle$  and  $K = k(x_i, x_j)$  are Gramm matrices, respectively. This looks like solving inner product of  $A^T A$  in RKHS use to obtained preimages  $z$ . Thus, in this method optimized preimage is given by

$$z = XP^{-1}(P - \lambda K^{-1})\gamma \quad (16)$$

where  $X = (x_1, x_2, \dots, x_d)^T$  and  $\gamma$  from Equation (7).

### 3.5. Abrahamsen's<sup>[8]</sup> input space distance regularization

Abrahamsen and Hansen<sup>[8]</sup> proposed method which overcome the stability issues in Mika's et al.<sup>[4]</sup> by penalizing cost function given Equation (8).

$$z = \operatorname{argmin}_z \|\phi(z) - P\phi(x)\|^2 + \lambda \|z - x_0\|^2 \quad (17)$$

where  $\lambda$  non negative regularization parameter and  $x_0$  is noisy data in space  $\mathcal{X}$ , solving Equation (9) gives objective function as

$$J_\lambda(z) = k(z, z) - 2 \sum_{n=1}^N \gamma_n k(z, x_n) - \lambda \|z - x_0\|^2 \quad (18)$$

Differentiating  $J_\lambda(z)$  w.r.t to  $z$  and setting it to zero, for RBF Gaussian kernel iteratively updating of  $z$  given by

$$z_{t+1} = \frac{\frac{1}{\sigma^2} \sum_{n=1}^N \gamma_n \exp(-\|z_t - x_n\|^2 / 2\sigma^2) x_n + \lambda x_0}{\frac{1}{\sigma^2} \sum_{n=1}^N \gamma_n \exp(-\|z_t - x_n\|^2 / 2\sigma^2) + \lambda} \quad (19)$$

It is clear from equation (19) that denominator never reached to zero because  $\lambda$  is non negative and non-zero. Hence stability in this method is guaranteed.

## 4. Experimentation

To evaluate the performance of above mentioned KPCA de-noising methods on 1-D time series signal (e.g., speech signal), one needs to transform it to multi-dimensional signal using delay embedding technique<sup>[9,10]</sup>.

Let 1-D time series sequence  $x = (x[0], x[1], \dots, x[K-1])$  is transformed into multi-dimensional lagged vector by using time delay embedding technique<sup>[11]</sup> (Taken's Theorem). Transformed 1-D sequence to multidimensional sequences produces  $N = K - d + 1$  lagged vectors in input space  $\mathcal{X}$  as

$$x_k = [x[k-1+d-1], \dots, x[k-1]]^T, \quad k = 1, \dots, N \quad (20)$$

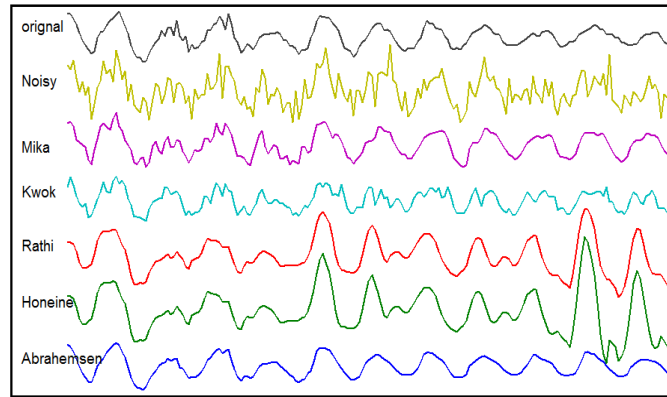
The lagged vectors  $x_k$  lie in an input space  $\mathcal{X}$  of dimension  $d$  and constitute the columns of the trajectory matrix  $X = [x_1 \ x_2 \ x_3, \dots, x_N]$ , ( $K > d$ ) i.e.,

$$X = \begin{bmatrix} x[d-1] & \cdots & x[K-1] \\ \vdots & \ddots & \vdots \\ x[0] & \cdots & x[K-d] \end{bmatrix}$$

$$X = \begin{bmatrix} x[d-1] & x[d] & \dots & x[K-1] \\ x[d-2] & x[d-1] & \dots & x[K-2] \\ x[d-3] & x[d-2] & \dots & x[K-3] \\ \vdots & \vdots & \ddots & \vdots \\ x[1] & x[2] & \dots & x[K-d+1] \\ x[0] & x[1] & \dots & x[K-d] \end{bmatrix} \quad (21)$$

#### 4.1. Performance evaluation for the toy example

As shown in **Figure 3**, the toy example is generated by clipping a 50ms duration small segment of speech signal known as original signal. Then 10dB of AWGN noise is added with it to produce noisy signal. This noisy signal is given to all the above-mentioned pre-imaging methods, which gives denoised signals. Performance evaluation is obtained by computing signal to noise ratio (SNR) between original and denoised signals. In this experimentation number of principal components  $q = 8$  and RBF kernel with  $\sigma^2 = 0.02$  are used respectively.



**Figure 3.** Original, noisy, and denoised signal obtained<sup>[4-8]</sup>.

SNR obtained for Mika et al.<sup>[3]</sup> is 22.33db, Kwok is 28.26db, Rathi is 30.25db, Honeine is 32.10db and for Abrahamsen is 38.23db. Here Abrahamsen show better performance in terms of SNR and obtained denoised signal is closer to original signal.

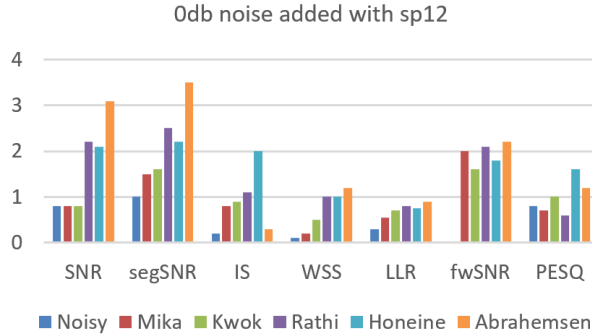
#### 4.2. Performance evaluation for speech signals

In this experiment, we use above mentioned KPCA denoising methods for denoising speech signal, which are corrupted by additive noise. Since, speech signals are semi random signal and cannot be analysed as a whole, we split the signal into segments of 20ms duration each using windowing and overlapping method. These segmented signals are one-dimensional time series signal and projective subspace technique such as PCA, KPCA, etc., cannot be implemented. Delay embedding method given by Equations (20) and (21) is the only solution to map one-dimensional time series signal to multidimensional signals.

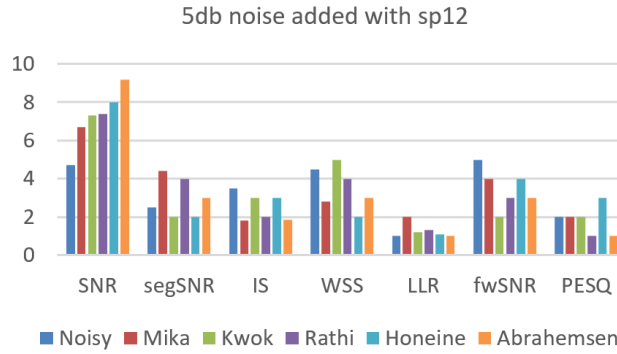
The denoising method performance solely relies on selecting number of principal components  $q$ , kernel type (e.g., polynomial, RBF, etc.) and its parameter  $\sigma^2$ , if RBF kernel is chosen. All the aforementioned methods for approximating preimages rely solely on selection of model parameters. Selection of these model parameters plays very important role in denoising for best results. To achieve best denoising results, these model parameters are heuristically selected for the performance evaluation of above mentioned KPCA denoising methods. Objective evaluation of speech quality parameters such as SNR, segmental SNG (segSNR), Itakura Saito ratio (IS), weighted spectral slope (WSS), log likelihood ratio (LLR), frequency SNR (fwSNR) and perceptual evaluation of speech quality (PESQ) are used for evaluating the performance<sup>[1,10]</sup>. Speech corpus (speech signal data) is used from NOIZEUS: A noisy speech corpus for evaluation of speech enhancement algorithms<sup>[10]</sup>. Here, we used 'sp12.wav' i.e., 'The drip of the rain made a pleasant sound', which is female voice sampled at 8 KHz and added with street noise of 5dB and 0dB noise. Delay embedding method is used to map 1-D signal to multi-

dimensional signal with the delay factor  $\tau = 1$  and number of delay vectors  $d = 60$ . Denoising of noisy signal is performed using Mika's at el.<sup>[3]</sup>, where 1000 iterations were used, Kwok and Tasang<sup>[5]</sup> by choosing nearest neighbored  $k = 15$ , Honeine and Richard<sup>[7]</sup> by setting regularize parameter  $\lambda = 10^{-7}$ , and Abrahamsen  $\lambda = 0.02$ .

**Figures 4 and 5** shows performance of pre-imaging methods when 0dB and 5dB of street back ground noise is added. These results show that for 0dB noise Abrahamsen method is better in all evaluation parameters except PESQ. For 5dB of noise Abrahamsen method is better in all evaluation parameters except IS, WSS, and PESQ, where, Honeine method is better. This above experiment is conducted by heuristically fixing  $q = 32$  and  $\sigma^2 = 0.01$  respectively.



**Figure 4.** Performance evaluation with 0dB of street noise.



**Figure 5.** Performance evaluation with 5dB of street noise added.

## 5. Conclusion

We proposed the effectiveness of KPCA for denoising speech signals, which is used extensively for image denoising. Here we have shown that how one-dimensional speech signal can be denoised using KPCA and pre-imaging methods. We compared and evaluated pre-imaging methods to state-of-the-art techniques using NOIZUES data base. By using toy example, we observed that Abrhamsen method is better than other method in terms of SNR. Also, performance of these methods is evaluated using speech signal, where Abrahamsen and Honeine techniques outperform others in terms of SNR and PESQ. Further better denoising can be achieved if noise and energy-based method<sup>[11]</sup> is used to estimate number of  $q$  and  $\sigma^2$ .

## Author contributions

Conceptualization, SLS; methodology, SLS; software, SLS and ARK; validation, SLS, ARK and SSM; formal analysis, ARK; investigation, ARK; resources, SLS; data curation, SLS and ARK; writing—original draft preparation, ARK and SSM; writing—review and editing, ARK; visualization, SLS and SSM; supervision, SSM; project administration, SLS. All authors have read and agreed to the published version of the manuscript.



## Conflict of interest

The authors declare no conflict of interest.

## References

1. Loizou PC. *Speech Enhancement: Theory and Practice*, 2nd ed. CRC Press; 2013.
2. Jolliffe IT. *Principal Component Analysis*. Springer Verlag; 1986.
3. Scholkopf B, Smola A, Müller KR. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation* 1998; 10(5): 1299–1319.
4. Mika S, Schölkopf B, Smola A, et al. Kernel PCA and de-noising in feature spaces. In: Kearns M, Solla S, Cohn D (editors). *Advances in Neural Information Processing Systems*. MIT Press; 1999. pp. 536–542.
5. Kwok JTY, Tsang IWH. The pre-image problem in kernel methods. *IEEE Transactions on Neural Network* 2004; 15(6): 1517–1525. doi: 10.1109/TNN.2004.837781
6. Rathi Y, Dambreville S, Tannenbaum A. Statistical shape analysis using kernel PCA. *Processing of SPIE-IS&T Electronic Imaging* 2006; 6064: 425–432. doi: 10.1117/12.641417
7. Honeine P, Richard C. Solving the pre-image problem in kernel machines: A direct method. In: Proceedings of 19th IEEE Workshop on Machine Learning for Signal Processing (MLSP); 1–4 September 2009; Grenoble, France. pp. 1–6.
8. Abrahamsen TJ, Hansen LK. Input space regularization stabilizes pre-images for kernel PCA de-noising. In: Proceedings of the 2009 IEEE International Workshop on Machine Learning for Signal Processing; 1–4 September 2009; Grenoble, France. pp. 1–6.
9. Teixeira AR, Tome AM, Lang EW, et al. On the use of KPCA to extract artifact in one-dimensional biomedical signals. 16th IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing 2006. doi: 10.1109/MLSP.2006.275580
10. Loizou PC, Kim G. Reasons why current speech-enhancement algorithms do not improve speech intelligibility and suggested solutions. *IEEE Transaction on Audio, Speech, and Language Processing* 2011; 19(1): 47–56. doi: 10.1109/TASL.2010.2045180.
11. Jorgensen KW, Hansen LK. Model selection for Gaussian kernel PCA denoising. *IEEE Transactions on Neural Networks and Learning Systems* 2012; 23(1): 163–168. doi: 10.1109/TNNLS.2011.2178325