

## Dark Background Image-denoising Based on KPCA Method

XIAO Yiran<sup>1, a</sup>, TIAN Xiaolin<sup>1,2, b</sup>

<sup>1</sup>Faculty of Information Technology, Macau University of Science and Technology, Macau, China

<sup>2</sup>Space Science Institute, Macau University of Science and Technology, Macau, China

<sup>a</sup>1753071742@qq.com, <sup>b</sup>xltian@must.edu.mo

**Keywords:** Image-denoising, Principle Component Analysis, Kernel Function, Dark background image

**Abstract.** In this paper, a dark background image-denoising method based on KPCA is discussed. First of all, the analysis of KPCA features is used to extract features from the training samples, and then discards the features which had small variance to form the feature space. Second, for reducing the noise, the principal components analysis is used to restructure the pattern for the smallest error in the feature space. Also, the value of parameter in kernel function is adjusted to fit the dark-background images. According to the testing results, this method is effective and operational.

### Introduction

Image noise-removal is a very common technology and has various methods, actually. The most classic two are histogram equalization method and median filtering method. However, these two simple methods can not fit for images with complex information well. Numerous algorithms are proposed for better enhancement during these years, and KPCA is one of them.

KPCA's predecessor is PCA (Principle Component Analysis). PCA is a linear and de-correlation technique that creates new images from the uncorrelated values of different images. The main intensity of PCA is to identify the principle components, in this image denoising can be done by applying moving window by calculating the local characteristics of an image for estimating the PCA transformation matrix.

However, this method has a strict demand for the original space, must subject to Gaussian distribution. Only if the space is linear, the PCA will work. So gradually, many people use the principle of KPCA to improve the method in image denoising recently. KPCA is a statistical de-correlation technique and also a nonlinear model in the image processing technology [1]. It is used for pattern recognition and the dimensionality reduction. KPCA is a non-linear form of PCA used for selecting the non-linear principle components by applying a kernel trick.

### Proposed KPCA Method

In the conventional PCA [2], the original space is linear, it must subject to Gaussian distribution. Only if the space is linear, the PCA will work. However, in the KPCA, the original space can be nonlinear, and we can use non-linear mapping  $\Phi$  to map the original space into a high dimensional feature space  $F$ . And the  $F$  space subject to Gaussian distribution approximately, not completely, not standard Gaussian distribution, because we just extract the principle components in the original space. And then we work on the new space  $F$  to reconstruct the mapping of the original space. To calculate the space  $F$ , we have to work out the mapping  $\Phi$  first, however, the non-linear mapping  $\Phi$  is very difficult to calculate. But, with the kernel trick [3], we don't need to calculate the  $\Phi$ , just have to perform dot product in the original space.

A set of data in the known input space are  $x_k, k=1, \dots, M, x_k \in \mathbb{R}^N$ . Usually, this set of samples is not necessarily subject  $n$ -dimensional Gauss distribution. The samples can be mapped into feature space by a nonlinear mapping function  $\Phi$  according to the idea of feature space.

$$\Phi: \mathbb{R}^N \rightarrow F \quad (1)$$

$$x \rightarrow \Phi(x) \quad (2)$$

Firstly, assume that the mapping samples in  $F$  are centered, that is  $\sum_{k=1}^M \Phi(x_k) = 0$ , and then compute the covariance matrix.

$$C = \frac{1}{M} \sum_{j=1}^M \Phi(x_j) \Phi(x_j)^T \quad (3)$$

Then we can get the nonzero eigenvalues ( $\lambda > 0$ ) of  $C$  and the corresponding feature vector  $V$  by doing eigenvalue decomposition of  $C$ , and satisfied with:

$$CV = \lambda V \quad (4)$$

Because  $\Phi$  is unknown, so  $C$  cannot be directly calculated. However, all nonzero feature vectors of  $C$  must be mapped to the data set. That is:

$$V = \sum_{i=1}^M \alpha_i \Phi(x_i) \quad (5)$$

Here,  $\alpha_i (i=1, \dots, M)$  represents the influence on characteristic direction  $V$  by  $\Phi(x_i)$ . Also, we have:

$$\lambda(\Phi(x_k) \cdot CV) \quad k = 1, \dots, M \quad (6)$$

So we can get the formula below by combining them:

$$\lambda \sum_{i=1}^M \alpha_i (\Phi(x_k) \cdot \Phi(x_j)) = \frac{1}{M} \sum_{i=1}^M \alpha_i (\Phi(x_k) \cdot \sum_{j=1}^M \Phi(x_j)) (\Phi(x_j) \cdot \Phi(x_i)) \quad (7)$$

and  $k=1, \dots, M$ .

By defining a  $M \times M$  kernel matrix  $K$  [4]:

$$K_{i,j} = (\Phi(x_i) \cdot \Phi(x_j)) = k(x_i, x_j) \quad (8)$$

And then we can calculate the eigenvalues and eigenvectors of  $K$  matrix instead of the  $\Phi$  calculation:

$$K\alpha = M\lambda\alpha \quad (\alpha = (\alpha_1, \dots, \alpha_M)^T) \quad (9)$$

As we know,  $K$  is a positive semi definite matrix. So, for all the  $X \in F$ :

$$(X \cdot KX) = \|(\Phi(x_1), \dots, \Phi(x_M))X\|^2 \geq 0 \quad (10)$$

Obviously, the characteristic values of the positive semidefinite matrix are non- negative. So after characteristic decomposition, the characteristic values can be shown in descending order  $\lambda_1 \geq \dots \geq \lambda_p \dots \geq \lambda_M$ ,  $\lambda_p$  represents the first nonzero eigenvalue.

Second, the feature vector  $V$  is a unit vector that must be satisfied with:

$$V_k^T V_k = 1 \quad k = 1, \dots, p \quad (11)$$

So we should normalize  $\alpha^1, \dots, \alpha^p$ , and to meet this, we should make  $\lambda_k(\alpha^k \cdot \alpha^k) = 1$ .

For a new sample  $x$ , it is necessary to extract its principal component by mapping the corresponding samples  $\Phi(x)$  in the  $F$  space to the  $V_k$ .

$$(V^k \cdot \Phi(x)) = \sum_{i=1}^M \alpha_i^k (\Phi(x_i) \cdot \Phi(x)) = \sum_{i=1}^M \alpha_i^k k(x_i, x) \quad (12)$$

There are nearly ten types of kernels till now. In this research, the Radial Basis Function Kernel was used:

$$k(X, Y) = \exp(-q\|X - Y\|^2) \quad (13)$$

Now we can get  $p$  principal components, In order to reconstruct  $\Phi(x)$ , we let the  $P_n \Phi(x)$  be the reconstruct-vectors:

$$P_n \Phi(x) = \sum_{k=1}^p \beta_k V^k \quad (14)$$

Muller and Scholkop [5] proposed a Kernel PCA denoising algorithm based on the principle of minimum reconstruction error in feature space. The target function is the distance of the point in the feature space:

$$\rho(z) = \|\Phi(z) - P_n \Phi(x)\|^2 = k(z, z) - 2 \sum_k \beta_k \sum_i \alpha_i^k k(z, x_i) + \omega \quad (15)$$

We can now simplify the formula by (13):

$$\rho(z) = \Phi(z) \cdot P_n \Phi(x) + \omega \quad (16)$$

$$z_{t+1} = \frac{\sum_{i=1}^M \gamma_i \exp(-\|z_t - x_i\|^2 / c) x_i}{\sum_{i=1}^M \gamma_i \exp(-\|z_t - x_i\|^2 / c)} \quad (17)$$

here,  $\gamma_i = \sum_{k=1}^p \beta_k \alpha_i^k$ .

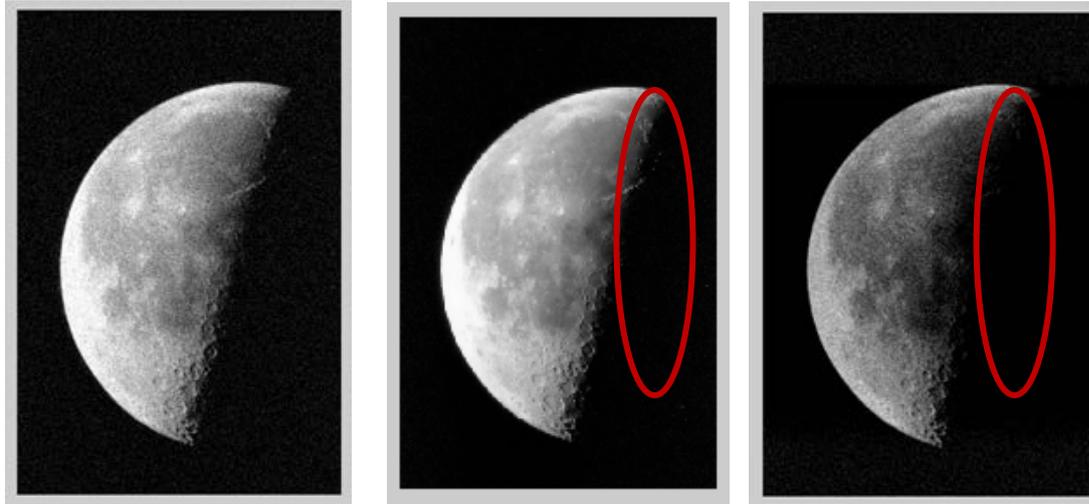
## Experimental Results

Since KPCA fits dark background images well, a picture of the far view of the moon was chosen to run my KPCA algorithm. In the Gaussian Radial Basis Function kernel, there is a parameter  $q$ . The value of  $q$  is selectable for different kinds of images. For most black background pictures, such as remote sensing images of lunar, the most suitable range is [10, 15]. Here the value is 15 for the testing image after many attempts.

Two figures are compared here. One is the result run by median filter, which is the classic one. Another is the result of KPCA with the combination of polynomial kernel and Gauss kernel. We can find the denosing result by KPCA method is much more obvious, especially in the red circle part. Also, the method can be used in nonlinear image denoising, it has broader applicability than conventional PCA method because the distribution of the image data is not limited.

As the results shown below, the KPCA method with correct  $q$  value in kernel function, has much better noise-removal ability in the dark-background. In the red circle part shown in the figure 1 below, the removal of background noise is much clear in KPCA result. However, in the Median Filter result, the noise is not completely denoised. Furthermore, this improved method

has a great possibility to use in remote sensing image of Lunar or Mars, which background is dark but important for lots of analysis.



i). Original Picture with Noise

ii). Result of Median Filter

iii). Result of KPCA Method

Fig 1.Experimental Results

### Acknowledgements

This work was financially supported by the Science and Technology Development Fund of Macao (No. 059/2013/A2).

### References

- [1] Liu Jingjing, Jin Hui, Gan Ling, *A Method for Face Denoising Based on Analysis of KPCA*, School of Information and Electrical Engineering CUMT, Xuzhou, (221008).
- [2] L. Zhang, W. Dong, D. Zhang, S. Guangming, *Two-stage imagedenoising by principal component analysis with local pixel grouping*, Pattern Recognition, vol. 43, issue 4, pp. 1531-1549, Apr 2010.
- [3] Yao Min. Digital image processing [M]. Beijing: Mechanical Industry Press.
- [4] J. L. Starck, E. J. Candes, and D. L. Donoho, *The curvelet transform for image denoising*, IEEE Transactions on Image Processing, vol. 11, no.6, pp. 670–684, 2002.
- [5] Venkatrama Phani Kumar, S.; Kishore, K.V.K.; Kumar, K.H., *HybridFace Recognition with Locally Discriminating Projection*, Signal Acquisition and Processing, 2010. ICSAP '10. International Conference on , vol., no., pp.327,331, 9-10 Feb. 2010.