# Spectral Clustering Survey

HU, Pili[*]

May 14, 2012[†]

**Abstract**

Abstract.

[*]hupili [at] ie [dot] cuhk [dot] edu [dot] hk

[†]Last compile:May 14, 2012

# Contents

# 1 Introduction

Spectral Clustering(SC) was used in several disciplines long ago. For example, computer vision[11], load balancing [4], electronics design [3], etc. Spectral Embeding(SE) was also widely discussed in the community[2]. Outside spectral community, the machine learning community also developed many linear or non-linear Dimensionality Reduction(DR) methods, like Principal Component Anslysis (PCA), Kernel PCA (KPCA)[10], Locally Linear Embedding (LLE)[7], etc. Other technique like Multi-Dimensional Scaling(MDS) was successfully used in computational psychology for a very long time[1], which can be viewed as both ”embedding” or ”dimensionality reduction”.

According to our survey, although those methods target at different problem and are derived from different assumptions, they do share a lot in common. The most significant sign is that, the core procedure involves eigenvalue decomposition or singular value decomposition, aka ”spectral”. They all involve an intermidiate step of embedding high-dimensional / non-Euclidean / non-metric points into a low-dimensional Euclidean space (although some do not embed explicitly). In this case, we categorize all these algorithms as Spectral Embedding Technique(SET).

## 1.1 A Sample Spectral Clustering Algorithm

## 1.2 Linear Algebraic Properties

# 2 Spectral Clustering Framework

## 2.1 Metric Formulation

## 2.2 Spectral Embeding

## 2.3 Clustering

# 3 Spectral Clustering Justification

## 3.1 Random Walk

[von]

## 3.2 Normalized Cut

[shi]

## 3.3 Ratio Cut

[von]

### 3.4 Conductance

[von]

### 3.5 Matrix Perturbation

[andrew ng]

### 3.6 Low Rank Approximation

[matthew brand]

### 3.7 Density Estimation View

[mo chen, 2010]

### 3.8 Commute Time

[jihun ham, kernel]. view pseudo inverse of graph Laplacian by commute times on graphs.

## 4 Other Spectral Like Embedding

### 4.1 MDS

### 4.2 isomap

### 4.3 Laplacian Eigenmap

### 4.4 Hessian Eigenmap

### 4.5 PCA

### 4.6 LLE

### 4.7 Kernel PCA

### 4.8 Kernel Framework

### 4.9 Graph Framework

## 5 Conclusion

## Acknowledgements

## References

[1] I. Borg and P.J.F. Groenen. *Modern multidimensional scaling: Theory and applications*. Springer Verlag, 2005.

[2] M. Brand and K. Huang. A unifying theorem for spectral embedding and clustering. In *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics*, 2003.

[3] S.W. Hadley, B.L. Mark, and A. Vannelli. An efficient eigenvector approach for finding netlist partitions. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 11(7):885–892, 1992.

[4] B. Hendrickson and R. Leland. Multidimensional spectral load balancing. *Report SAND93-0074, Sandia National Laboratories, Albuquerque, NM*, ¡¡missing¿¿:¡¡missing¿¿, 1993.

[5] Pili Hu. Matrix calculus. GitHub, https://github.com/hupili/tutorial/tree/master/matrix-calculus, 3 2012. HU, Pili's tutorial collection.

[6] Pili Hu. Tutorial collection. GitHub, https://github.com/hupili/tutorial, 3 2012. HU, Pili's tutorial collection.

[7] S.T. Roweis and L.K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.

[8] L.K. Saul and S.T. Roweis. An introduction to locally linear embedding. Technical report, NYU, 2000.

[9] L.K. Saul and S.T. Roweis. Think globally, fit locally: unsupervised learning of low dimensional manifolds. *The Journal of Machine Learning Research*, 4:119–155, 2003.

[10] B. Schölkopf, A. Smola, and K.R. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation*, 10(5):1299–1319, 1998.

[11] J. Shi and J. Malik. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8):888–905, 2000.

# Appendix