




Inspire…Educate…Transform.

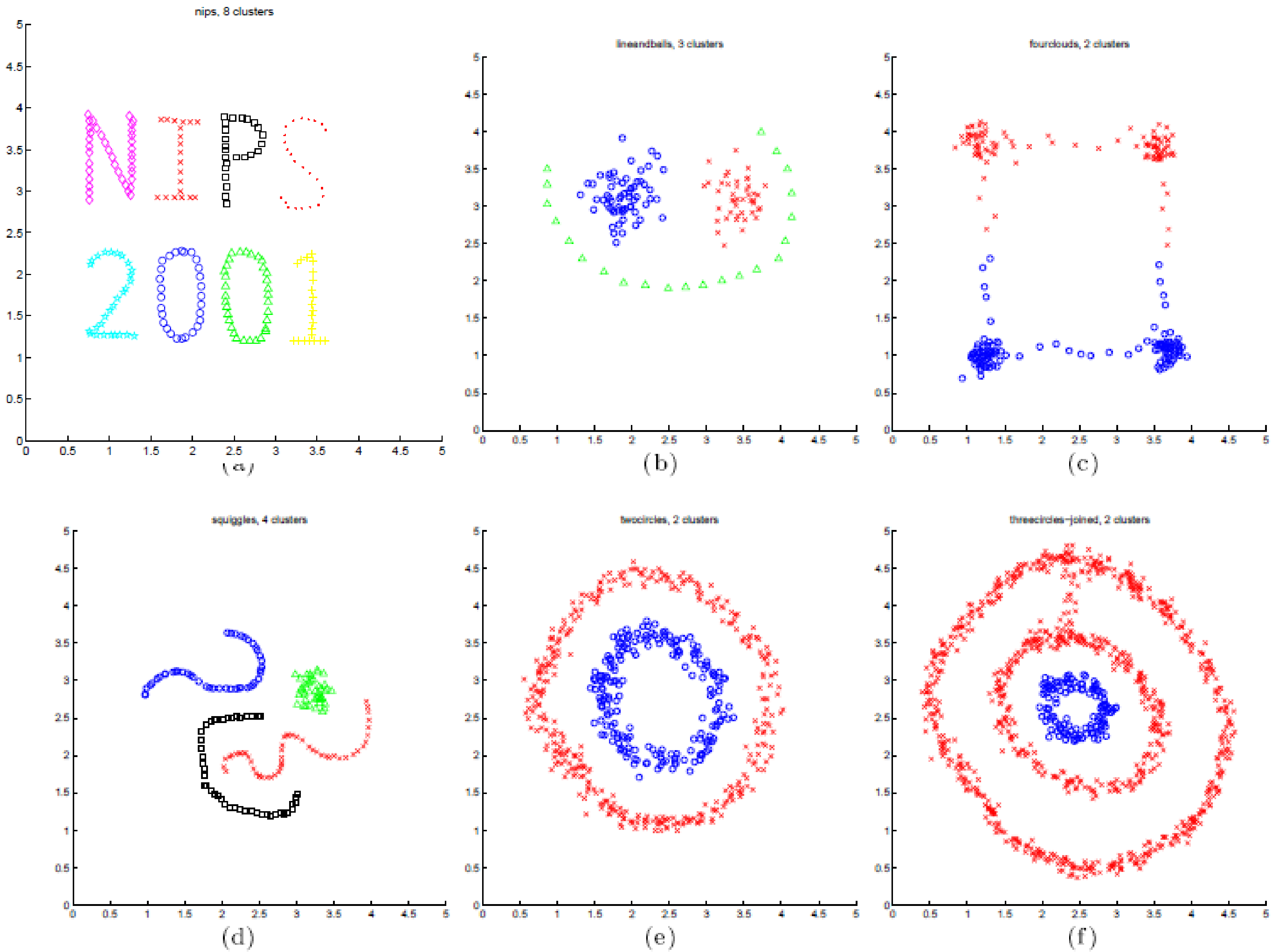
Spectral Clustering

Dr. K. V Dakshinamurthy
President, INSOFE

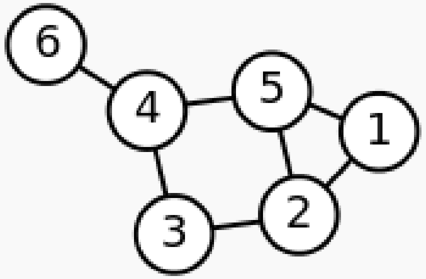
https://charlesmartin14.wordpress.com/2012/10/09/spectral-clustering/

http://ai.stanford.edu/~ang/papers/nips01-spectral.pdf

SPECTRAL CLUSTERING


nips, 8 clusters
(a)
lineandballs, 3 clusters
(b)
fourclouds, 2 clusters
(c)
squiggles, 4 clusters
(d)
twocircles, 2 clusters
(e)
threecircles-joined, 2 clusters
(f)

WHEN AFFINITY IS MORE IMPORTANT

Labeled graph	Degree matrix	Adjacency matrix	Laplacian matrix
	$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 2 & -1 & 0 & 0 & -1 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ -1 & -1 & 0 & -1 & 3 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}$

The elements of L are given by

$$L_{i,j} := \begin{cases} \deg(v_i) & \text{if } i = j \\ -1 & \text{if } i \neq j \text{ and } v_i \text{ is adjacent to } v_j \\ 0 & \text{otherwise} \end{cases}$$

where $\deg(v_i)$ is degree of the vertex i.

- The number of times 0 appears as an eigenvalue in the Laplacian is the number of connected components in the graph.
- The smallest non-zero eigenvalue of L is called the spectral gap.
- The second smallest eigenvalue of L is the algebraic connectivity (or Fiedler value) of G.

Affinity/Adjacency in continuous sense

$$W(i,j) = e^{-d(x_i, x_j)/2\sigma^2}$$

$$d(x_i, x_j) = ||x_i - x_j||^2$$ One possible definition

W(i,i) = 0

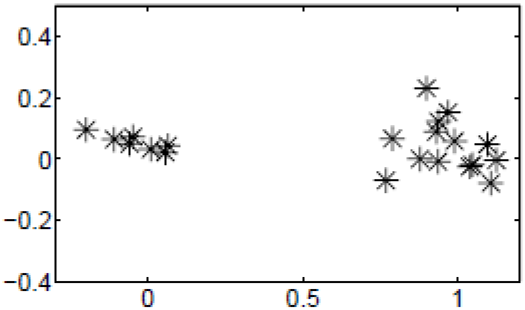
$$d(x_i, x_j) = d(x_j, x_i)$$ Mandatory condition

The symmetric normalized Laplacian matrix is defined as:[1]

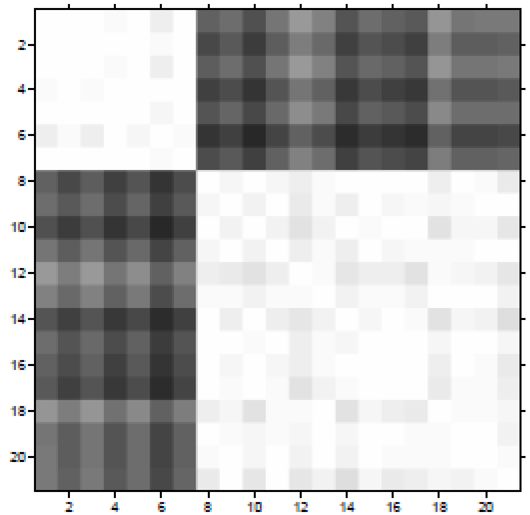
$$L^{\text{sym}} := D^{-1/2} L D^{-1/2} = I - D^{-1/2} A D^{-1/2},$$

The elements of L^{sym} are given by

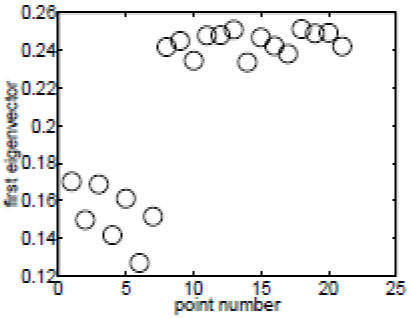
$$L^{\text{sym}}_{i,j} := \begin{cases} 1 & \text{if } i = j \text{ and } \deg(v_i) \neq 0 \\ -\frac{1}{\sqrt{\deg(v_i)\deg(v_j)}} & \text{if } i \neq j \text{ and } v_i \text{ is adjacent to } v_j \\ 0 & \text{otherwise.} \end{cases}$$

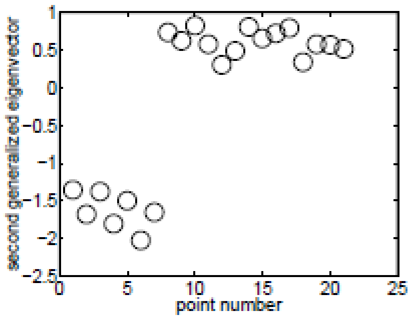

0.4
0.2
0
−0.2
−0.4
0
0.5
1

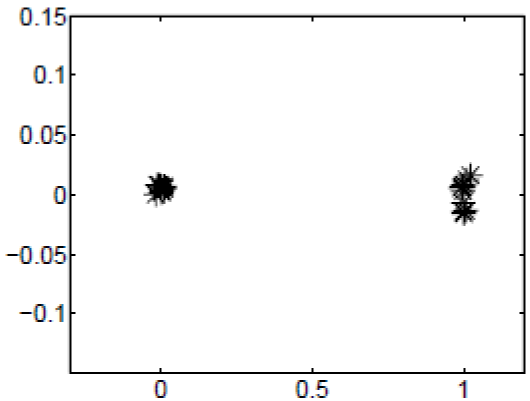
a

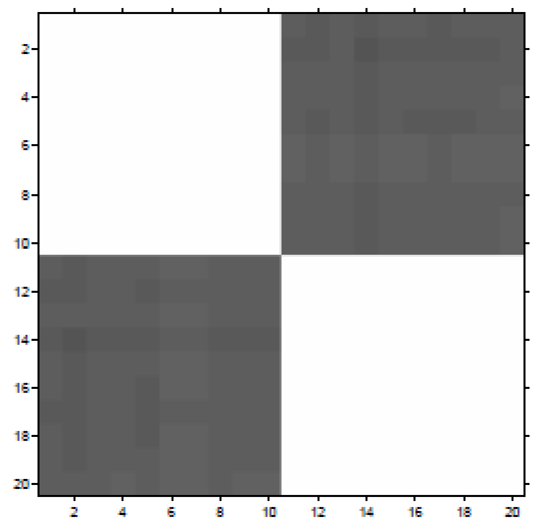

2
4
6
8
10
12
14
16
18
20

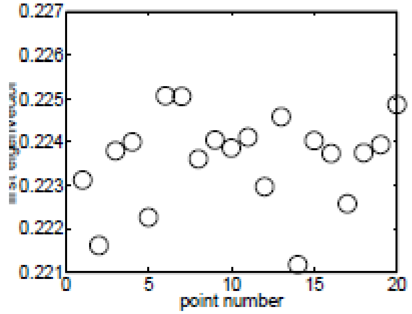
b

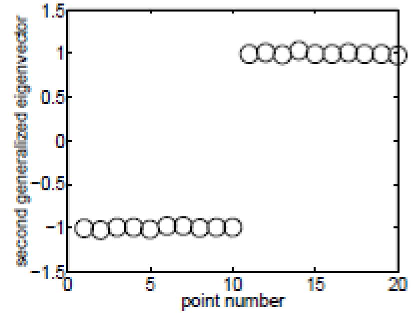

first eigenvector
point number
0.26
0.24
0.22
0.2
0.18
0.16
0.14
0.12
0
5
10
15
20
25

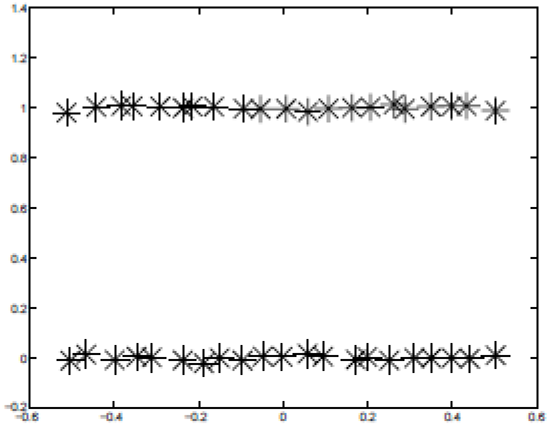

second generalized eigenvector
point number
1
0.5
0
−0.5
−1
−1.5
−2
−2.5
0
5
10
15
20
25

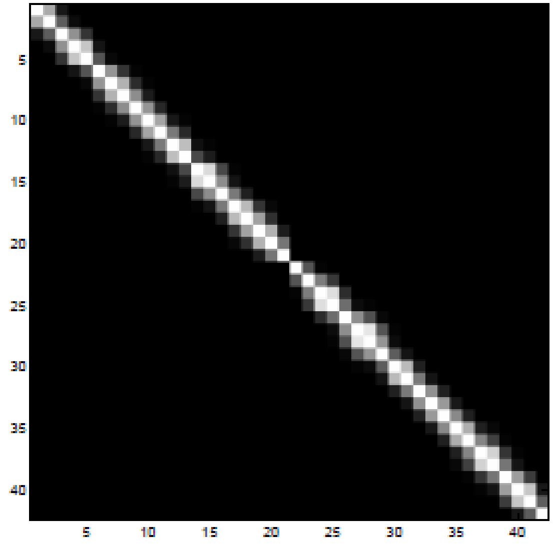

0.15
0.1
0.05
0
−0.05
−0.1
0
0.5
1

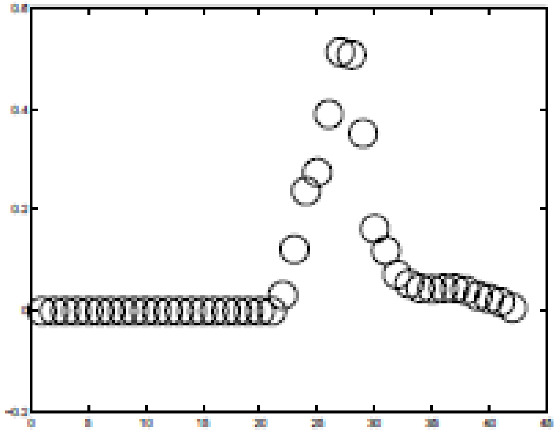

2
4
6
8
10
12
14
16
18
20

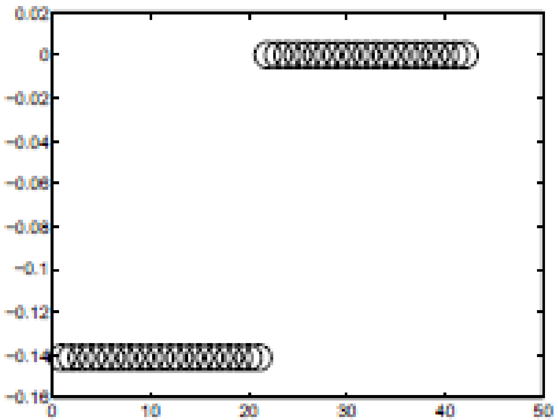

0.227
0.226
0.225
0.224
0.223
0.222
0.221
0
5
10
15
20
point number

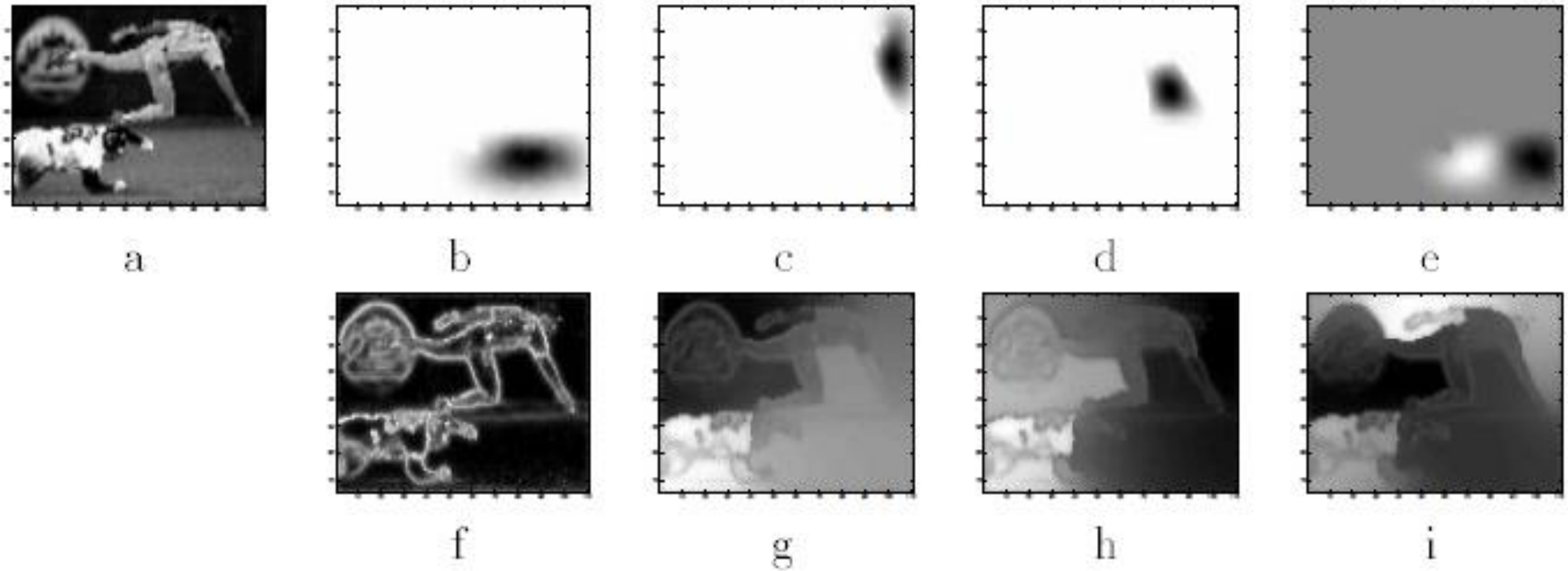

second generalized eigenvector
1.5
1
0.5
0
−0.5
−1
−1.5
0
5
10
15
20
point number

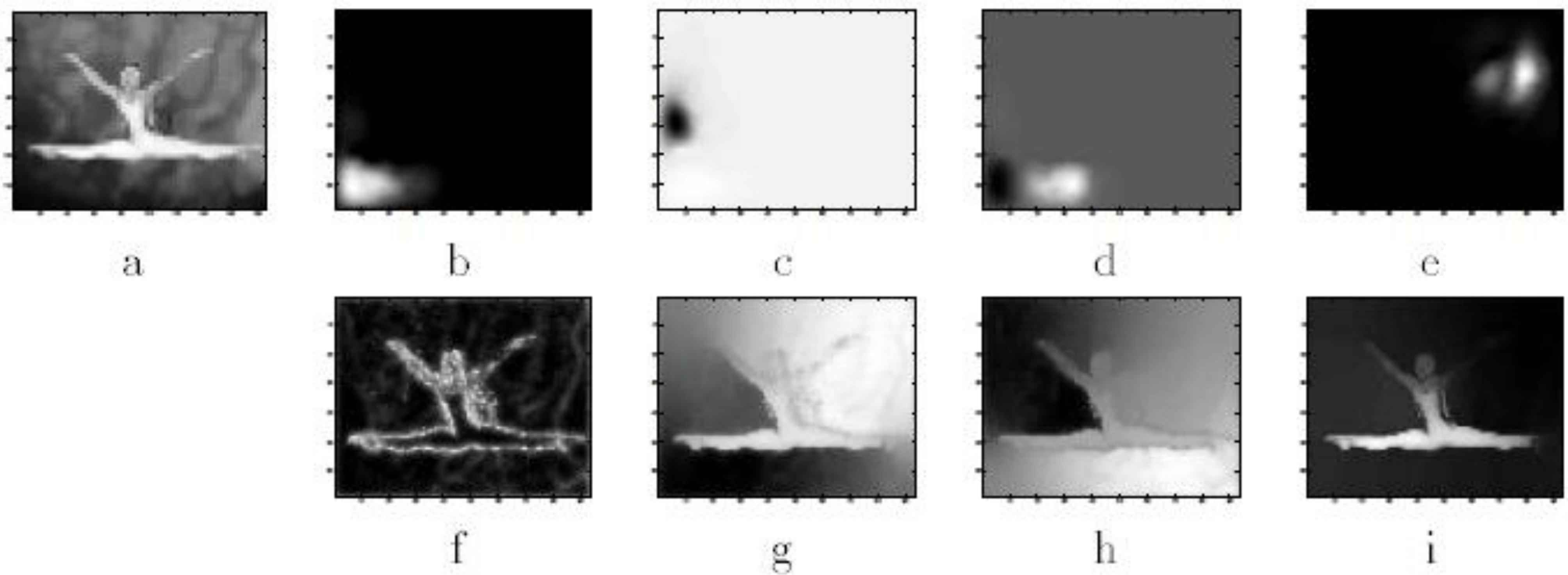

1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.6
-0.4
-0.2
0
0.2
0.4
0.6


5
10
15
20
25
30
35
40
5
10
15
20
25
30
35
40




0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1
-0.12
-0.14
-0.16
0
10
20
30
40
50

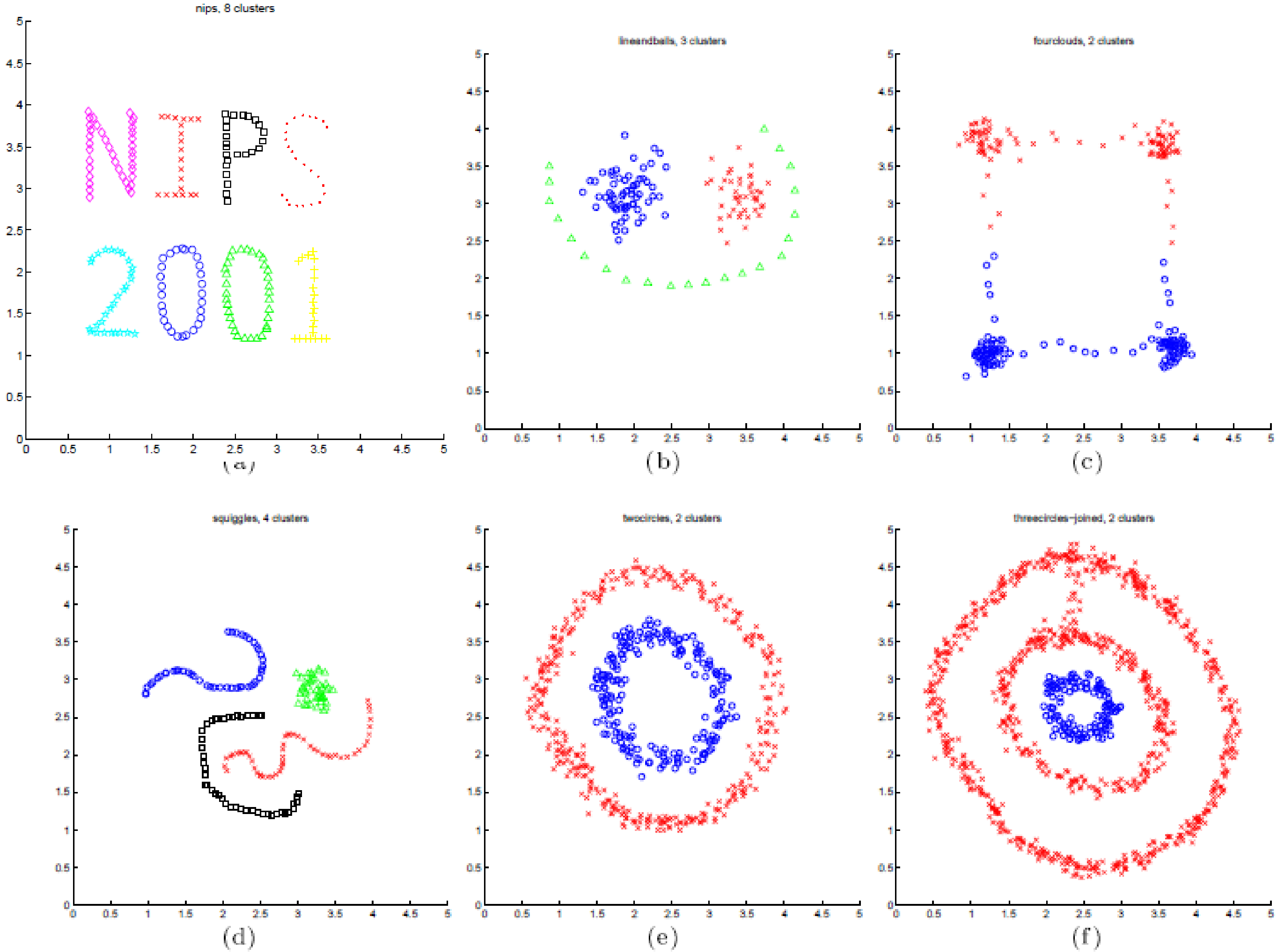

a
b
c
d
e
f
g
h
i


a
b
c
d
e
f
g
h
i

A spectral algorithm

Given a set of points $S = \{s_1, \ldots, s_n\}$ in $\mathbb{R}^l$ that we want to cluster into k subsets:

1. Form the affinity matrix $A \in \mathbb{R}^{n \times n}$ defined by $A_{ij} = \exp(-||s_i - s_j||^2/2\sigma^2)$ if $i \neq j$, and $A_{ii} = 0$.
2. Define D to be the diagonal matrix whose (i, i)-element is the sum of A's i-th row, and construct the matrix $L = D^{-1/2}AD^{-1/2}$.[1]
3. Find $x_1, x_2, \ldots, x_k$, the k largest eigenvectors of L (chosen to be orthogonal to each other in the case of repeated eigenvalues), and form the matrix $X = [x_1 x_2 \ldots x_k] \in \mathbb{R}^{n \times k}$ by stacking the eigenvectors in columns.
4. Form the matrix Y from X by renormalizing each of X's rows to have unit length (i.e. $Y_{ij} = X_{ij}/(\sum_j X_{ij}^2)^{1/2}$).
5. Treating each row of Y as a point in $\mathbb{R}^k$, cluster them into k clusters via K-means or any other algorithm (that attempts to minimize distortion).
6. Finally, assign the original point s_i to cluster j if and only if row i of the matrix Y was assigned to cluster j.


nips, 8 clusters
lineandballs, 3 clusters
fourclouds, 2 clusters
squiggles, 4 clusters
twocircles, 2 clusters
threecircles-joined, 2 clusters
(a)
(b)
(c)
(d)
(e)
(f)

http://www.cs.stanford.edu/~acoates/papers/coatesng_nntot2012.pdf

http://www.cs.stanford.edu/~acoates/papers/CoatesLeeNg_nips2010_dlwkshp_singlelayer.pdf

http://fastml.com/the-secret-of-the-big-guys/

AN ADVANCED SPECTRAL FEATURE ENGINEERING

This is hot!

- Do K-means clustering of the data
- For every point, create a hard cluster space representation

	Cluster 1	Cluster 2	…	Cluster 100	Target
Record 1	1	0	0	0	A
Record 2	0	0	..1..	0	B
…					A

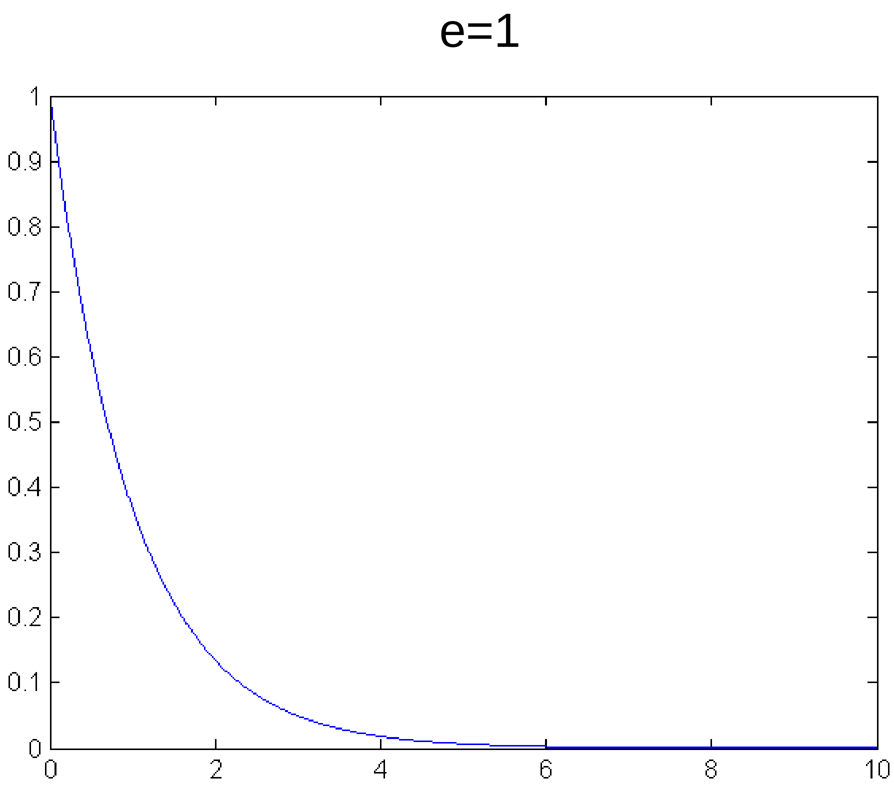
- Now, build a regression

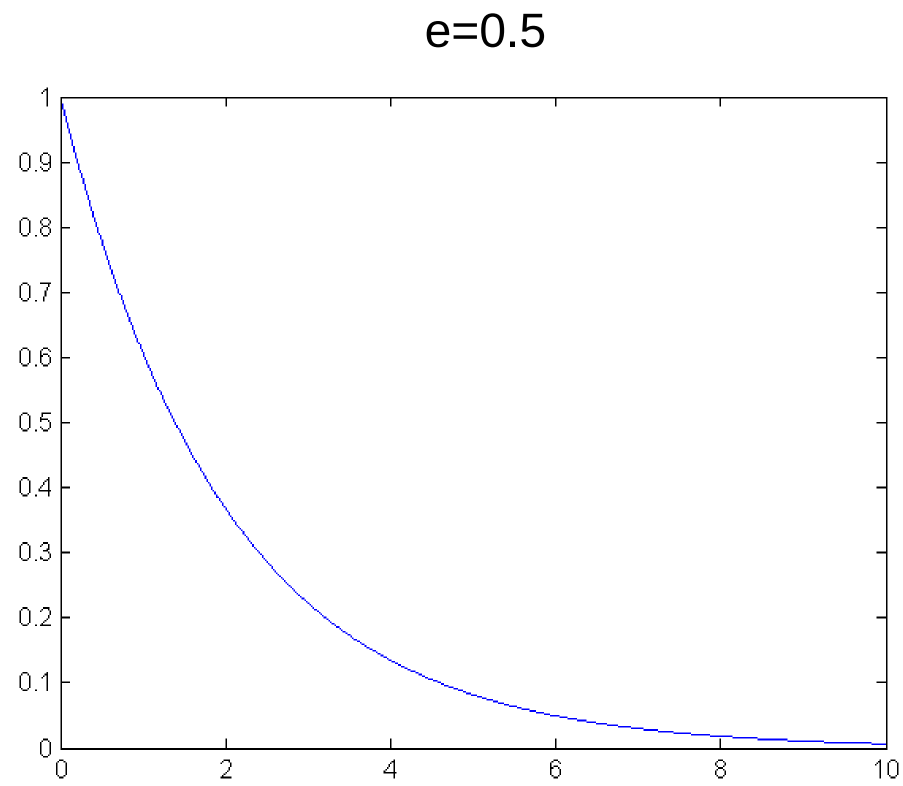
This is hotter!

- Do K-means clustering of the data
- For every point, create a soft cluster space representation ($\phi(r) = e^{-(\varepsilon r)^2}$)

	Cluster 1	Cluster 2	…	Cluster 100	Target
Record 1	1	0	0, 0, 0, 1, 0, 0	0	A
Record 2	0	1	0, 0, 1, 0, 0, 0	0	B

- Now, build a regression


e=1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
2
4
6
8
10


e=0.5
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
2
4
6
8
10

A powerful method

- run rSofia K-means to find cluster centers (hundreds or thousands)
- map the data to these centers using RBF
- learn a linear model on the mapped data

Will it get better with spectral k-means?

- R&D question

International School of Engineering

Plot 63/A, 1st Floor, Road # 13, Film Nagar, Jubilee Hills, Hyderabad - 500 033

For Individuals: +91-9502334561/63 or 040-65743991

For Corporates: +91-9618483483

Web: http://www.insofe.edu.in

Facebook: https://www.facebook.com/insofe

Twitter: https://twitter.com/Insofeedu

YouTube: http://www.youtube.com/InsofeVideos

SlideShare: http://www.slideshare.net/INSOFE

LinkedIn: http://www.linkedin.com/company/international-school-of-engineering