KECE471 Computer Vision

Segmentation by Clustering

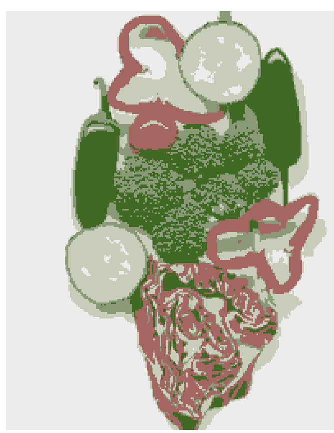
Chang-Su Kim

Chapter 14, Computer Vision by Forsyth and Ponce

Note: Dr. Forsyth's notes are partly used.
Jae-Kyun Ahn in Korea University made the first draft of these slides

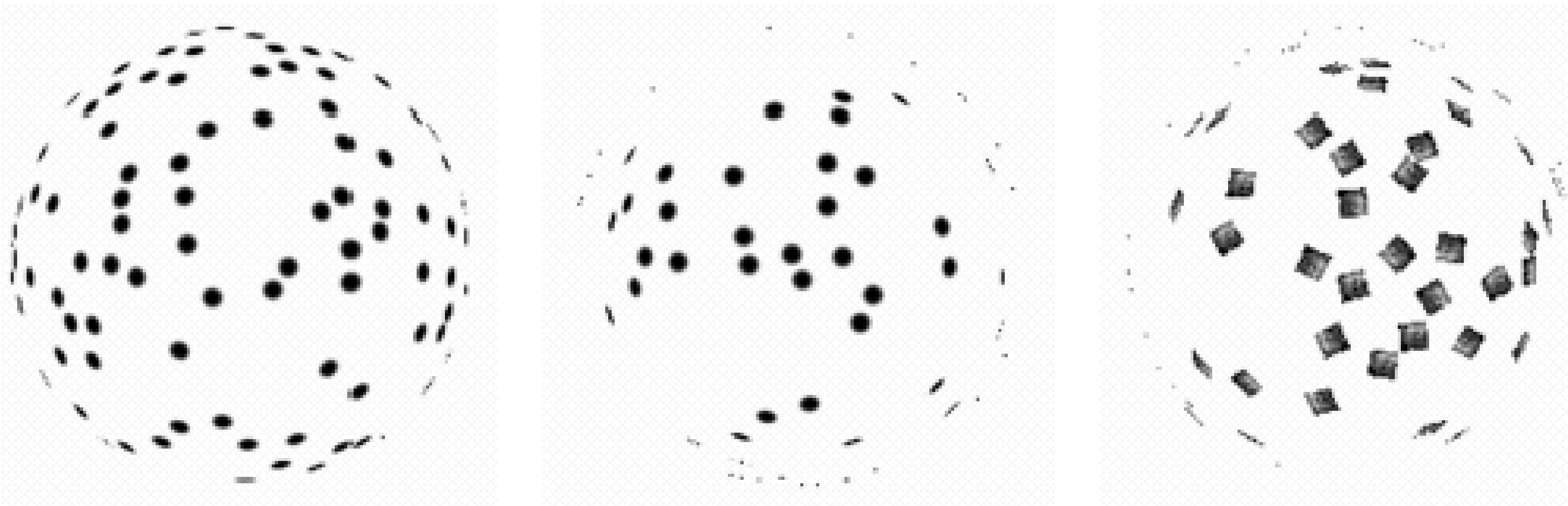
Segmentation and Grouping

- Segmentation
  - Obtain **compact representation** from an image/motion sequence/set of tokens
  - Grouping (or clustering)
    - collect together tokens that belong together
  - Fitting
    - associate a model with tokens

General ideas

- Tokens
  - whatever we need to group (pixels, points, surface elements, etc.)
- Top down segmentation
  - tokens belong together because they lie on the same object
- Bottom up segmentation
  - tokens belong together because they are locally coherent
- These two are not mutually exclusive



Why do these tokens seem to belong together?

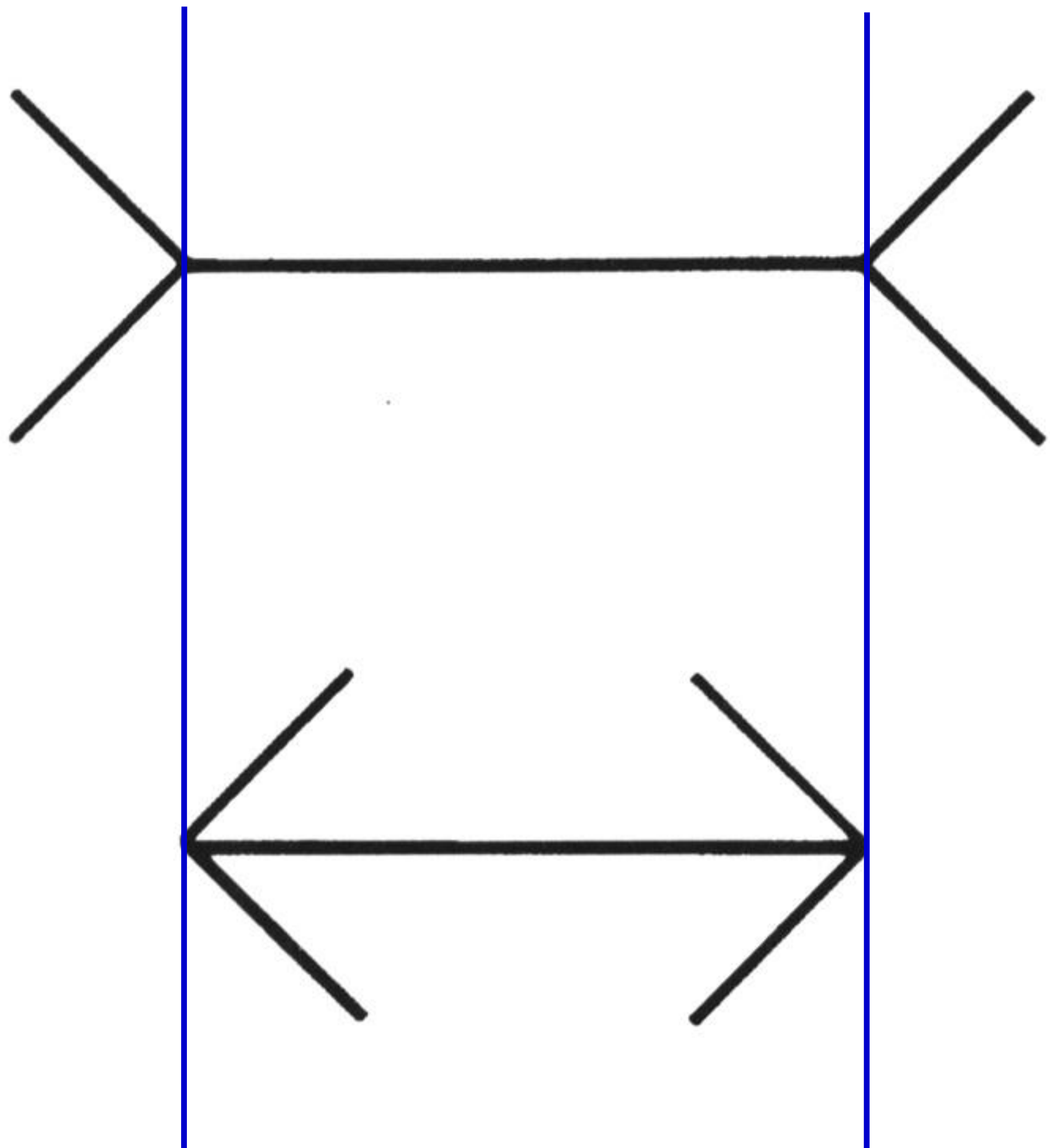

INTERSTELLAR

Basic ideas of grouping by humans

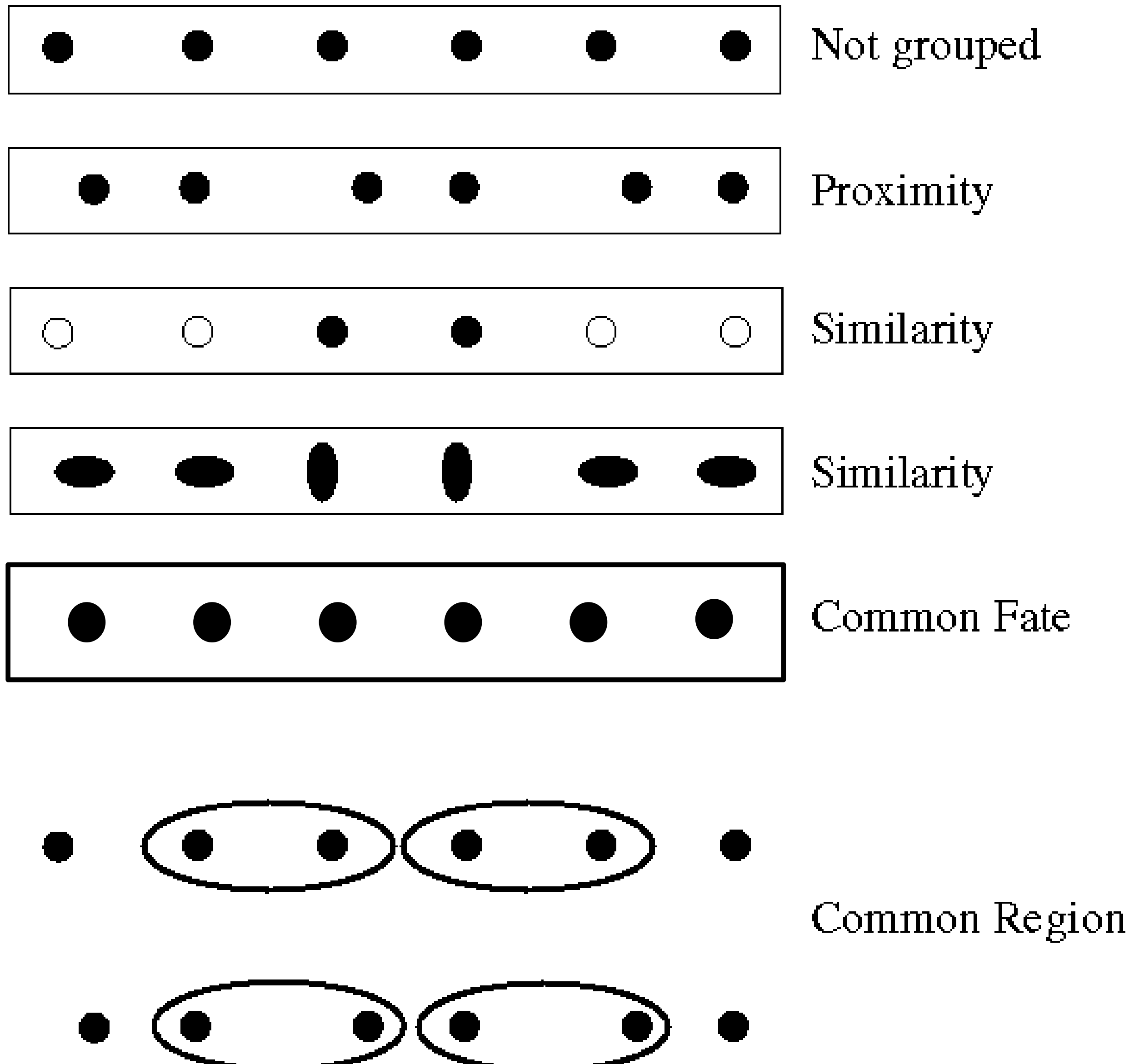
- Gestalt (group or whole)
  - Gestalt school (學派) of psychologist
  - The tendency of the visual system to assemble components of a picture together and perceive them together

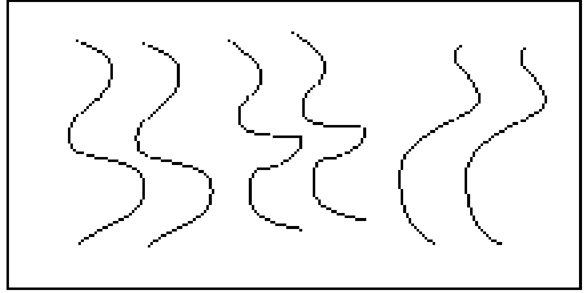
- Gestaltqualität (gestalt properties)
  - A set of factors that affect which elements should be grouped together

Perceiving objects as groups

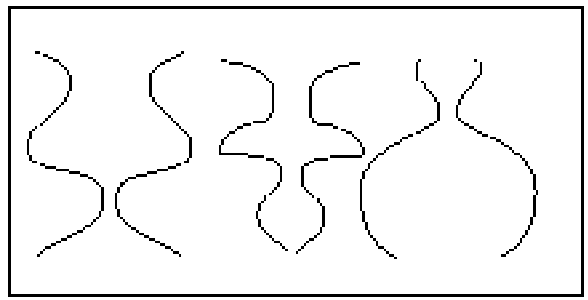


Gestaltqualität

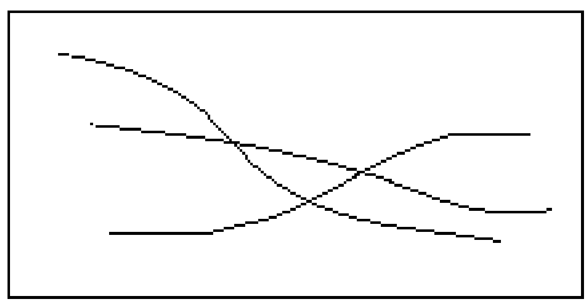




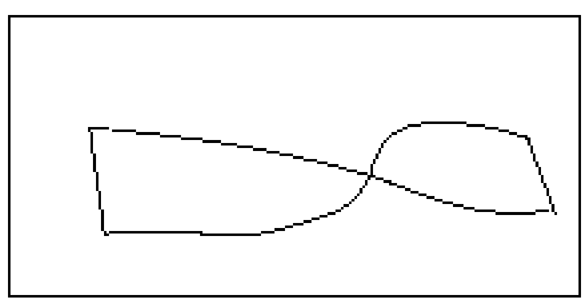
Parallelism



Symmetry



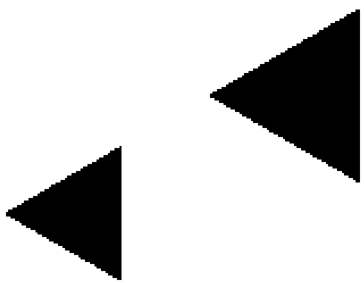
Continuity



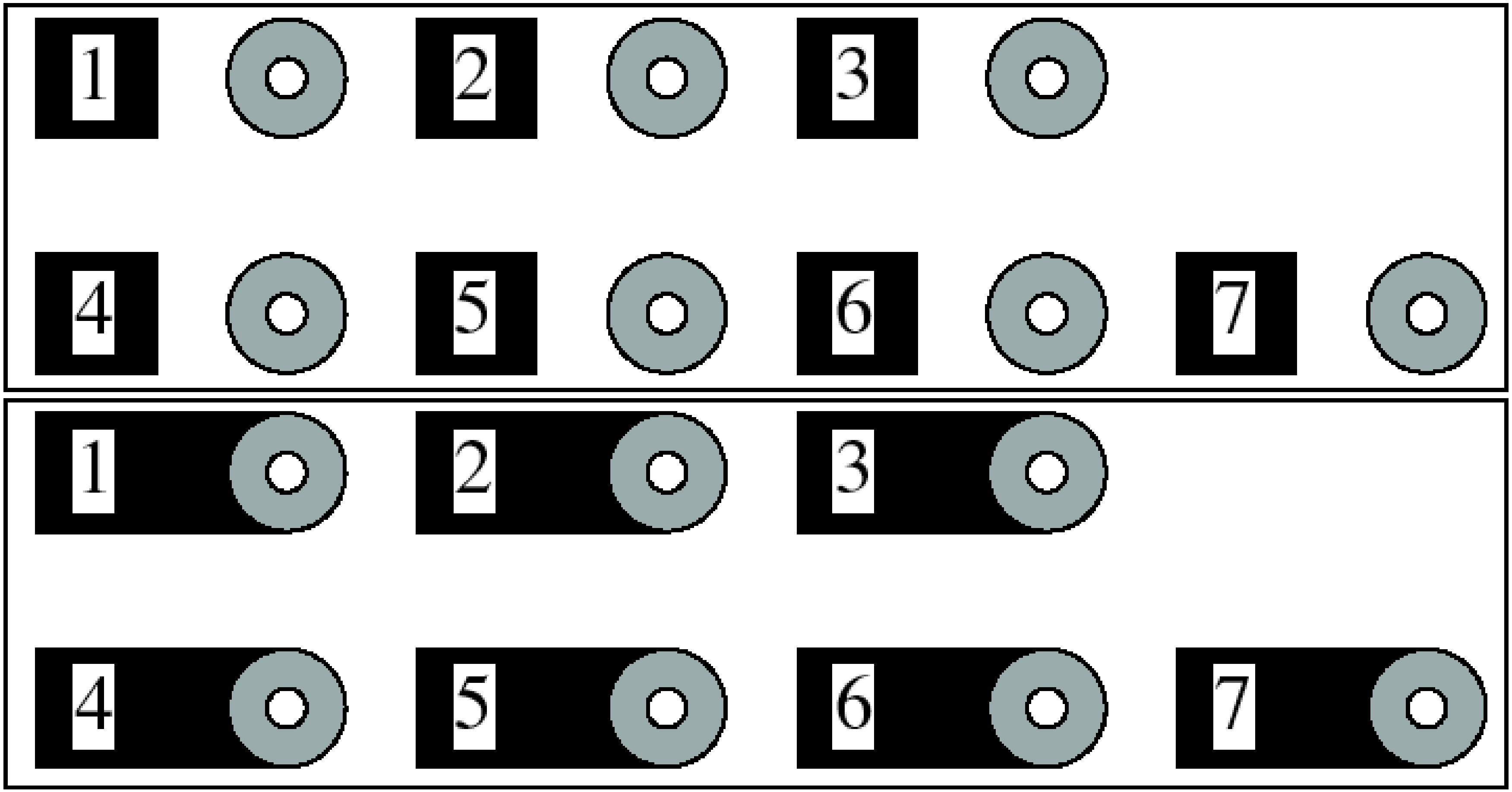
Closure

Grouping in case of occlusion

Illusory Contours



Common Region Cues



Example 1: Shot Boundary Detection

- Find the shots in a sequence of video
  - shot boundaries usually result in big differences between successive frames
- Strategy:
  - compute interframe distances
  - declare boundaries where these are big
- Possible distances
  - frame differences
  - histogram differences
  - edge differences
- Applications:
  - representation for movies, or video sequences
    - find shot boundaries
    - obtain "most representative" frame
  - supports search

cf. Video Tapestry

Video Tapestries with Continuous Temporal Zoom

Connelly Barnes[1] Dan B Goldman[2] Eli Shechtman[2,3] Adam Finkelstein[1]
[1]Princeton University [2]Adobe Systems [3]University of Washington

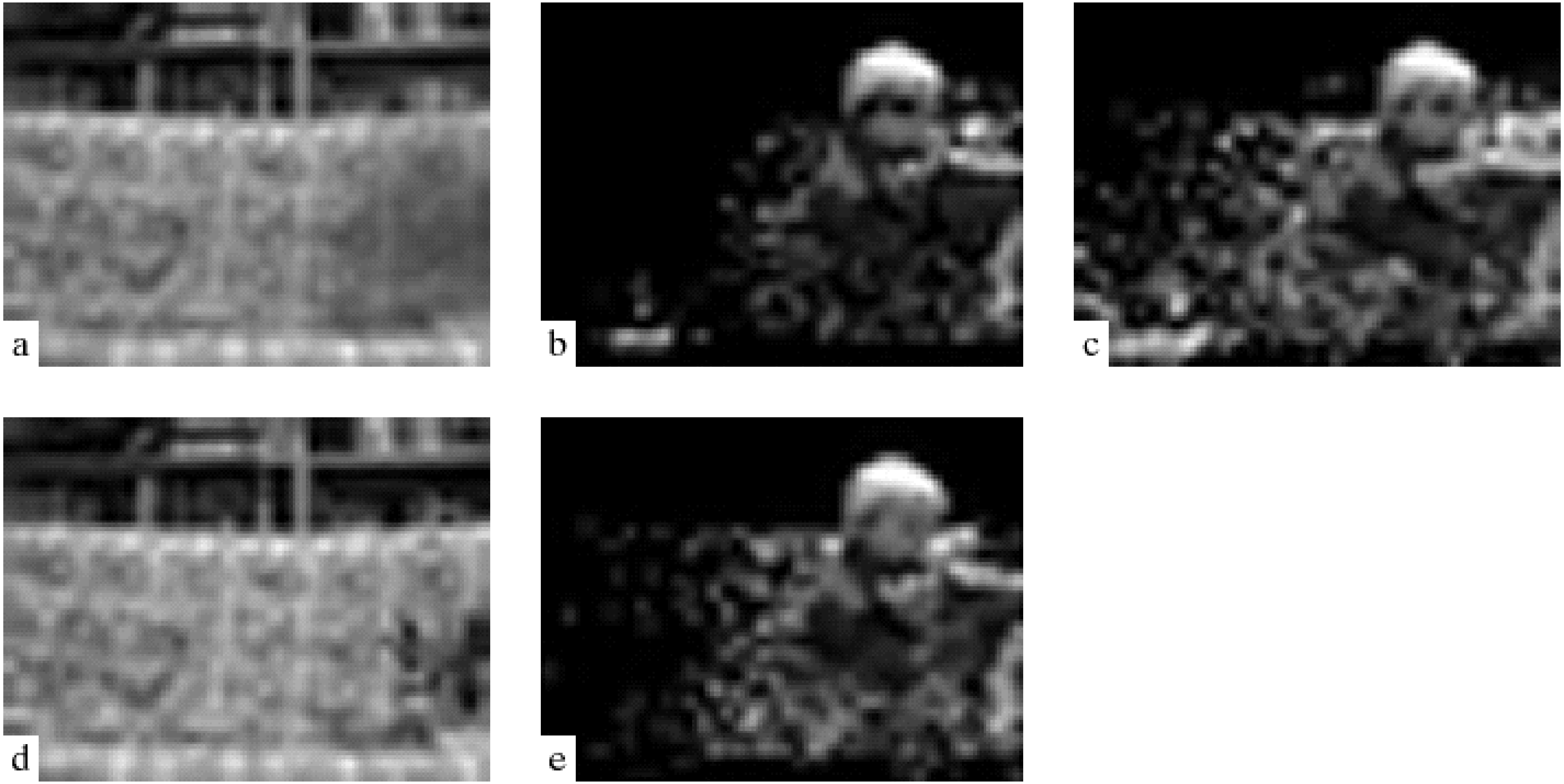
Figure 1: *A multiscale tapestry represents an input video as a seamless and zoomable summary image which can be used to navigate through the video. This visualization eliminates hard borders between frames, providing spatial continuity and also continuous zooms to finer temporal resolutions. This figure depicts three discrete scale levels for the film* Elephants Dream *(Courtesy of the Blender Foundation). The lines between each scale level indicate the corresponding domains between scales. See the video to view the continuous zoom animation between the scales. For Copyright reasons, the print and electronic versions of this paper contain different imagery in Figures 1, 4, 6, and 7.*

Example 2: Background Subtraction

- If we know what the background looks like, it is easy to identify interesting pixels

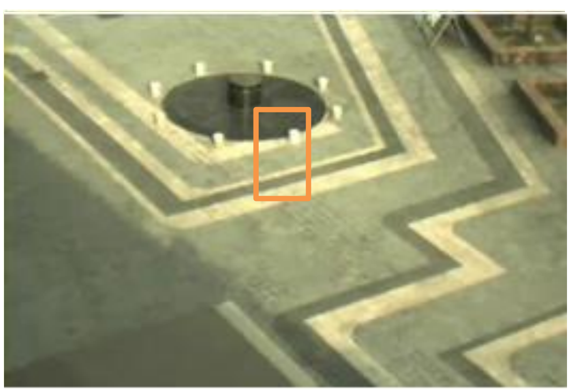
- Applications
  - Person in an office
  - Tracking cars on a road
  - surveillance

- Approach:
  - use a moving average to estimate background image
  - subtract from current frame
  - large absolute values are interesting pixels
    - trick: use morphological operations to clean up pixels

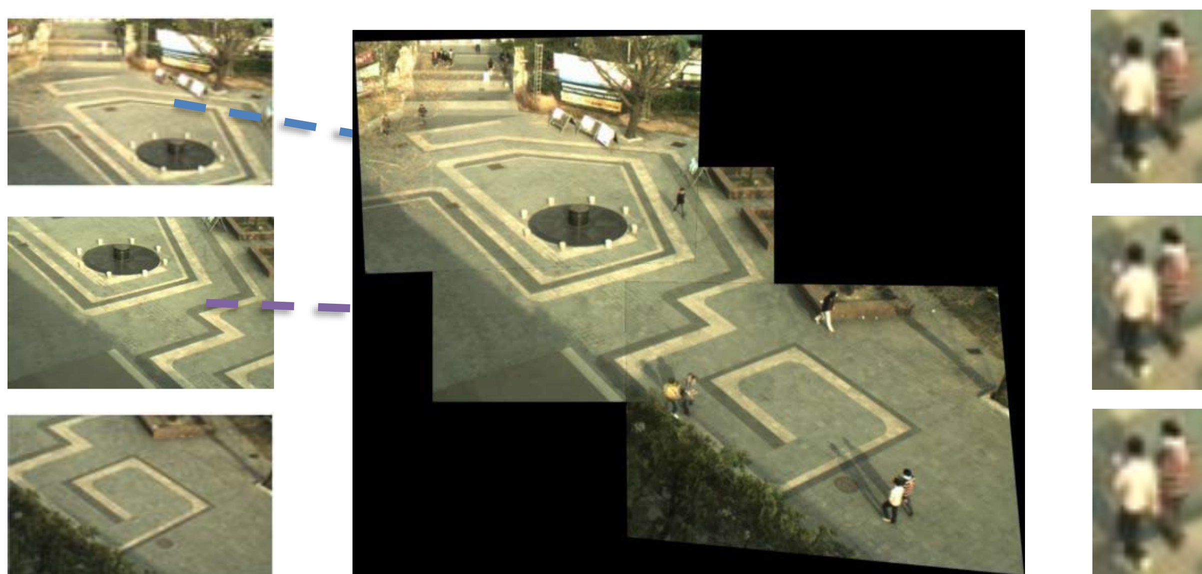

a
b
c
d
e

전역 시점 영상 합성

- 동영상에서의 전역 시점 영상 합성
  - 배경 영상 추정 기법
    - 확률 분포에 기반한 배경 영상 추정 및 획득
  - 전역 시점 영상 합성
    - 추정된 배경 영역들을 이용한 배경 합성 → 합성된 배경위에 객체를 합성



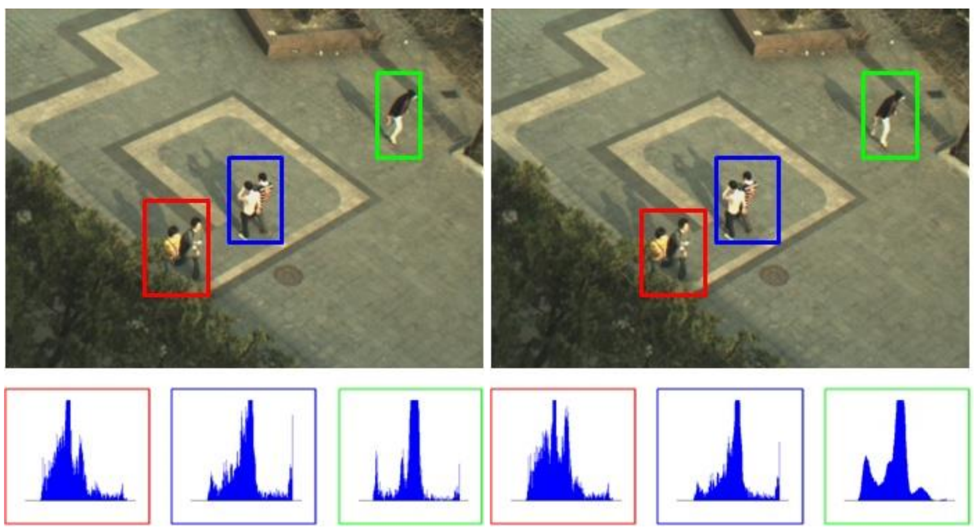
배경 영상 추정



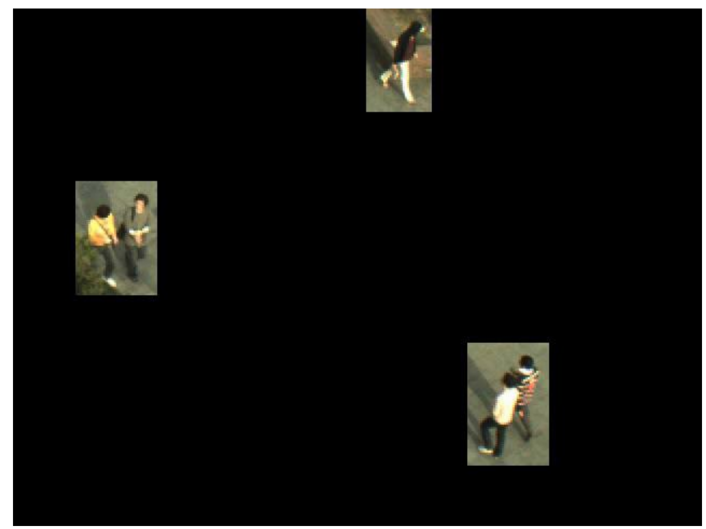
전역 시점 영상 합성: 배경 + 객체

개별 카메라의 이벤트 처리

- 객체 추적
  - 연속된 두 프레임에서 추출된 객체 추적 기법
    - 마스크 내의 히스토그램 유사성을 통한 객체 추적



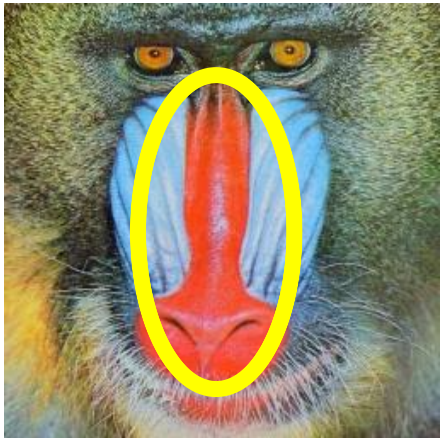
연속된 프레임에서 마스크내의 히스토그램 유사성

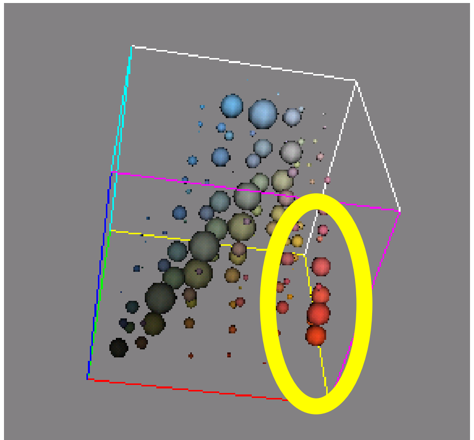


객체 추적 결과 영상

Segmentation as Clustering

- Clustering
  - Cluster together pixels, tokens, and etc that belong together
  - Agglomerative clustering
    - combine two close clusters to make one
  - Divisive clustering
    - split a cluster along best boundary





Agglomerative Clustering

- Each item is regarded as a cluster, and clusters are recursively merged to yield a good clustering
- Clustering by merging
- A bottom-up approach

```
Agglomerative Clustering


Make each point a separate cluster
Until the clustering is satisfactory
    Merge the two clusters with the smallest
    inter-cluster distance
End
```

Divisive Clustering

- The entire set is regarded as a cluster, and then clusters are recursively split to yield a good clustering
- Clustering by splitting
- A top-down approach

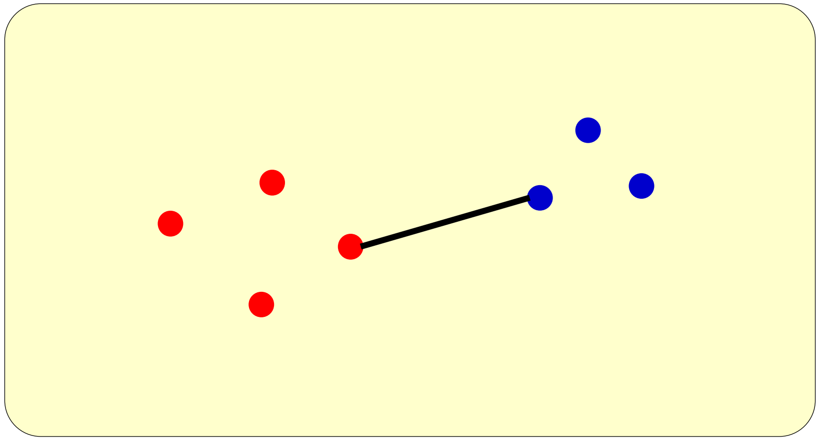
```
Divisive Clustering


Construct a single cluster containing all points
Until the clustering is satisfactory
     Split the cluster that yiels the two components
     with the largest inter-cluster distance
End
```

Inter-Cluster Distance

- Single-link clustering

$$\boxed{d(A,B) = \min_{a \in A, b \in B} d(a,b)}$$

  – It may yield elongated clusters

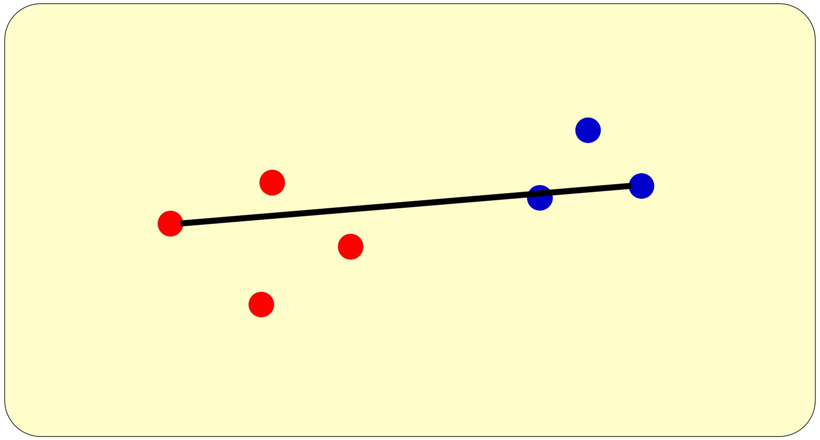


Inter-Cluster Distance

- Complete-link clustering

$$\boxed{d(A,B) = \max_{a \in A, b \in B} d(a,b)}$$

  - It usually yields round clusters

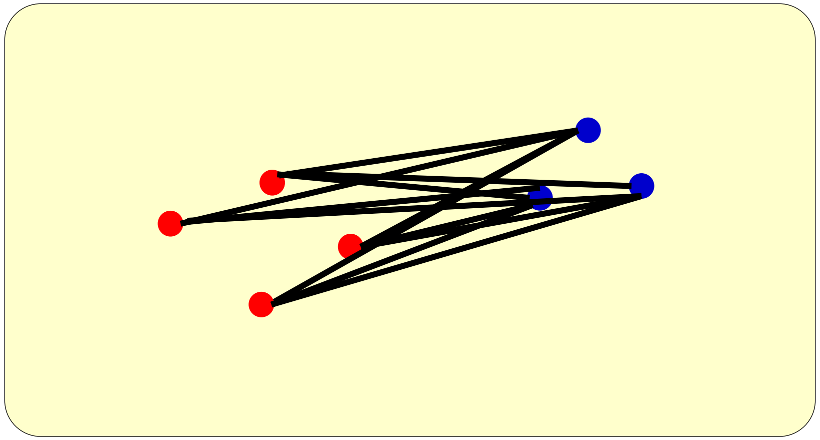


Inter-Cluster Distance

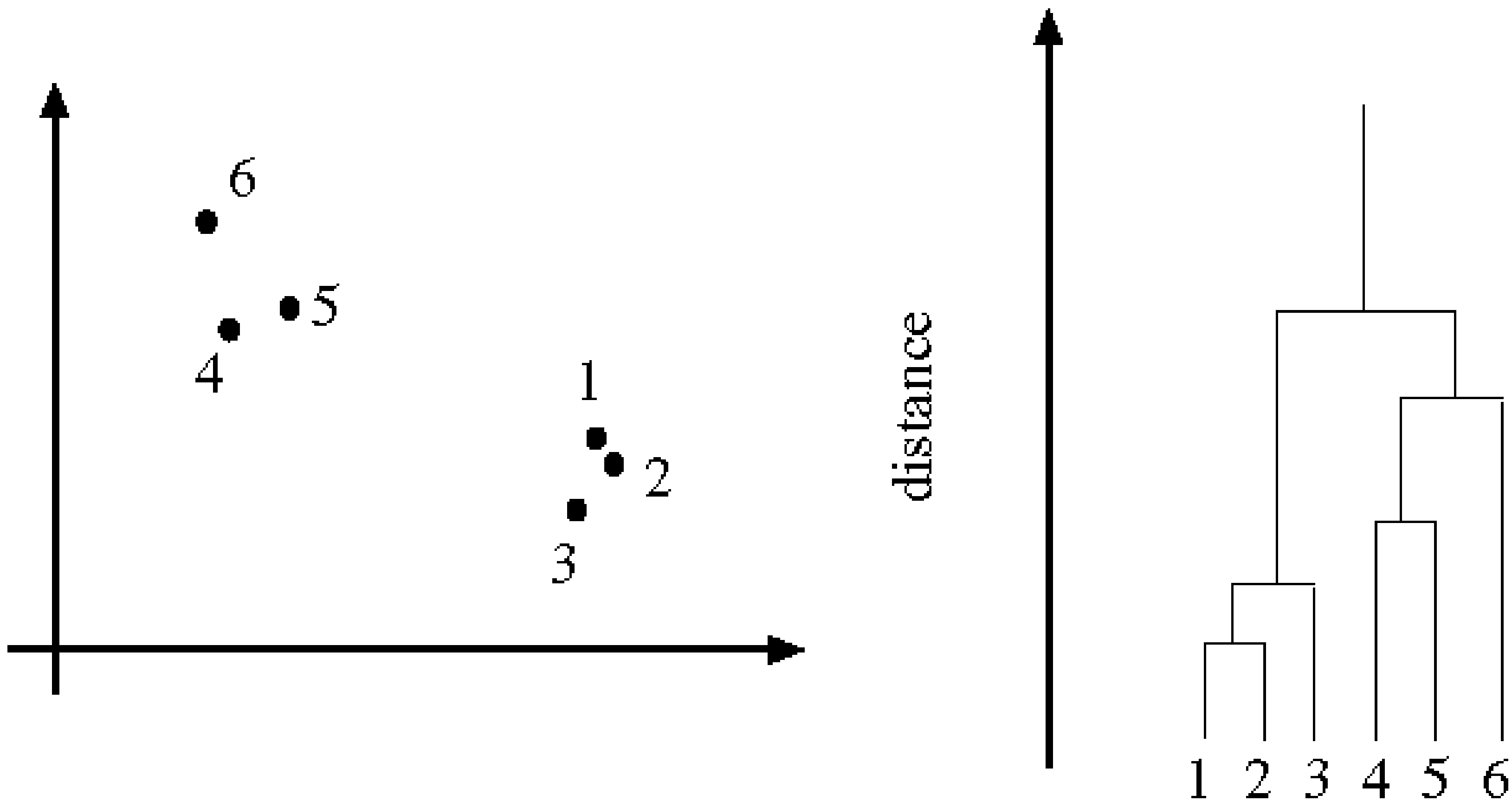
- Group-average clustering

$$d(A,B)=\frac{1}{|A|\times|B|}\sum_{a\in A,b\in B} d(a,b)$$

  - It usually yields round clusters

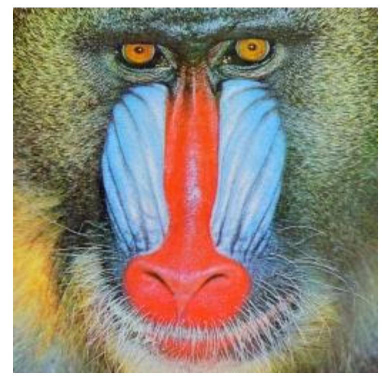


Dendrogram for Agglomerative Clustering



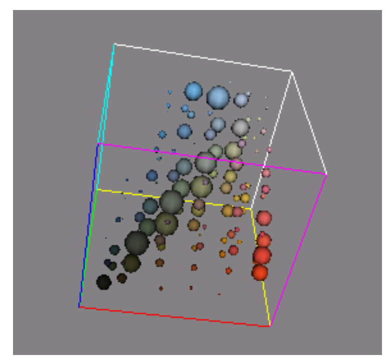
K - Means

- Application of vector quantization
  - Choose a fixed number of clusters
  - Choose cluster centers and point-cluster allocations to minimize error
  - Repeat until centers converge



- Error or cost function

$$\sum_{i \in \text{clusters}} \left\{ \sum_{j \in \text{elements of i'th cluster}} \left\| x_j - \mu_i \right\|^2 \right\}$$



K - Means

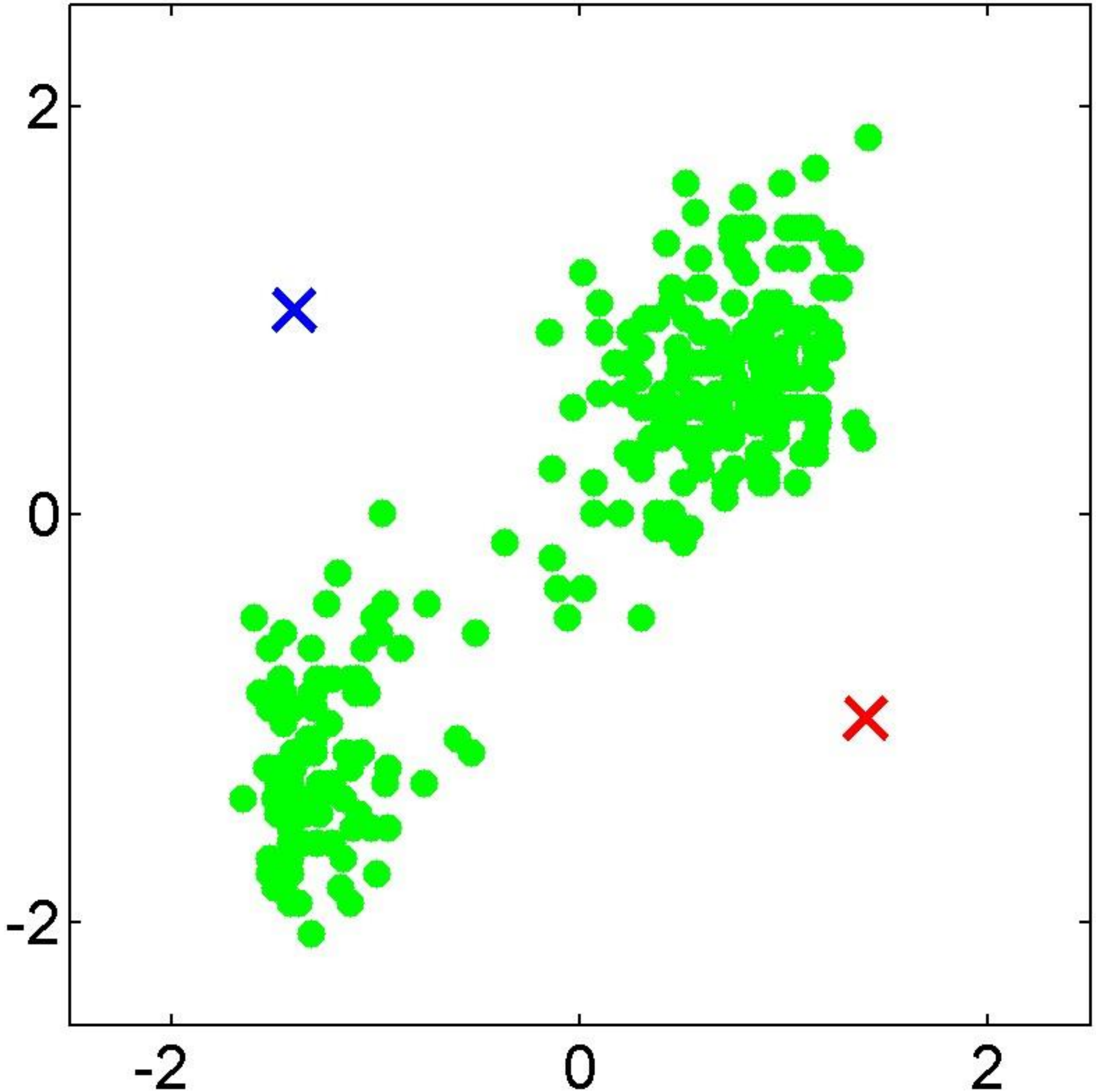
K-Means Algorithm

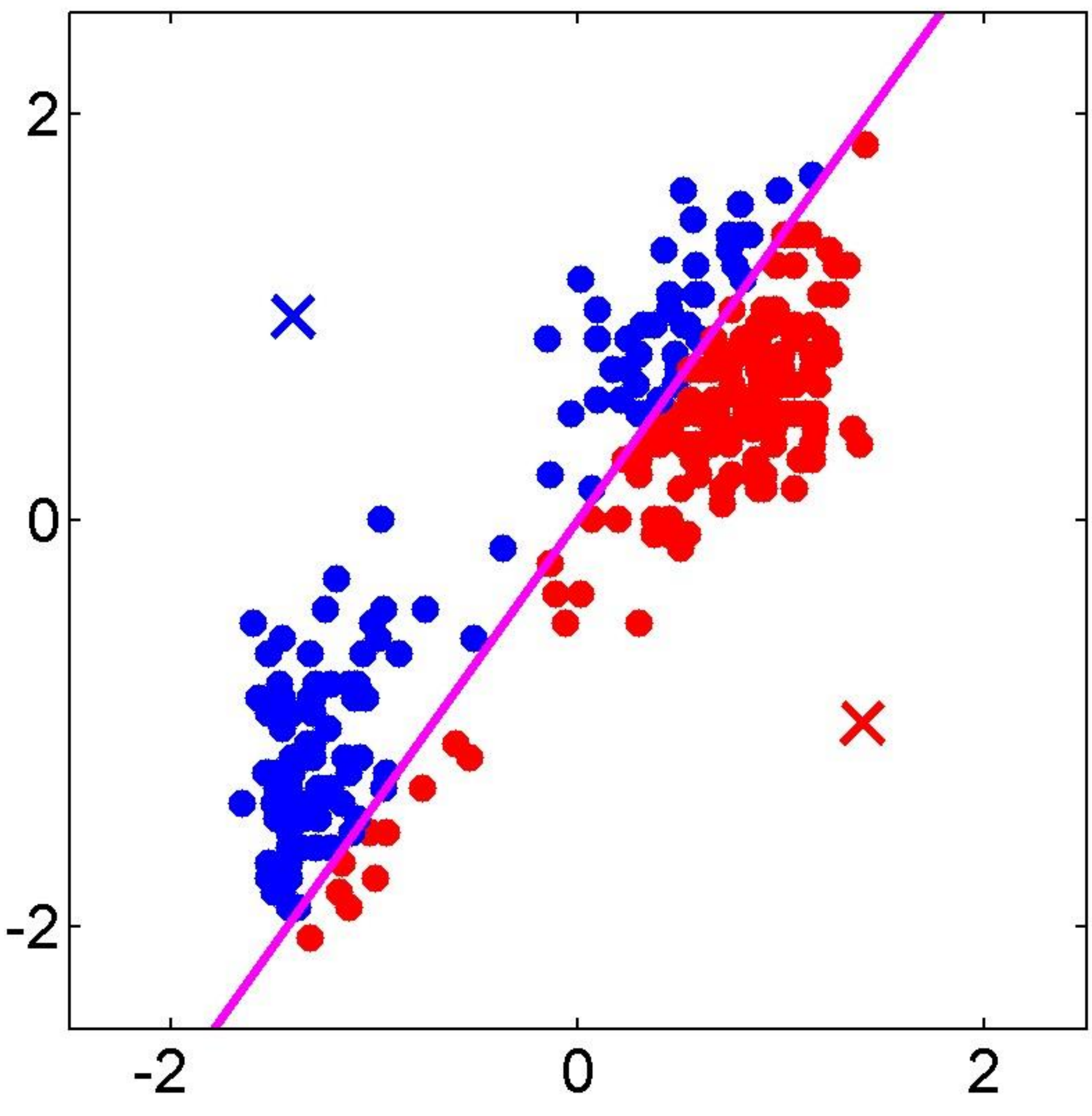
Choose k data points to act as cluster centers

Until the cluster centers are unchanged

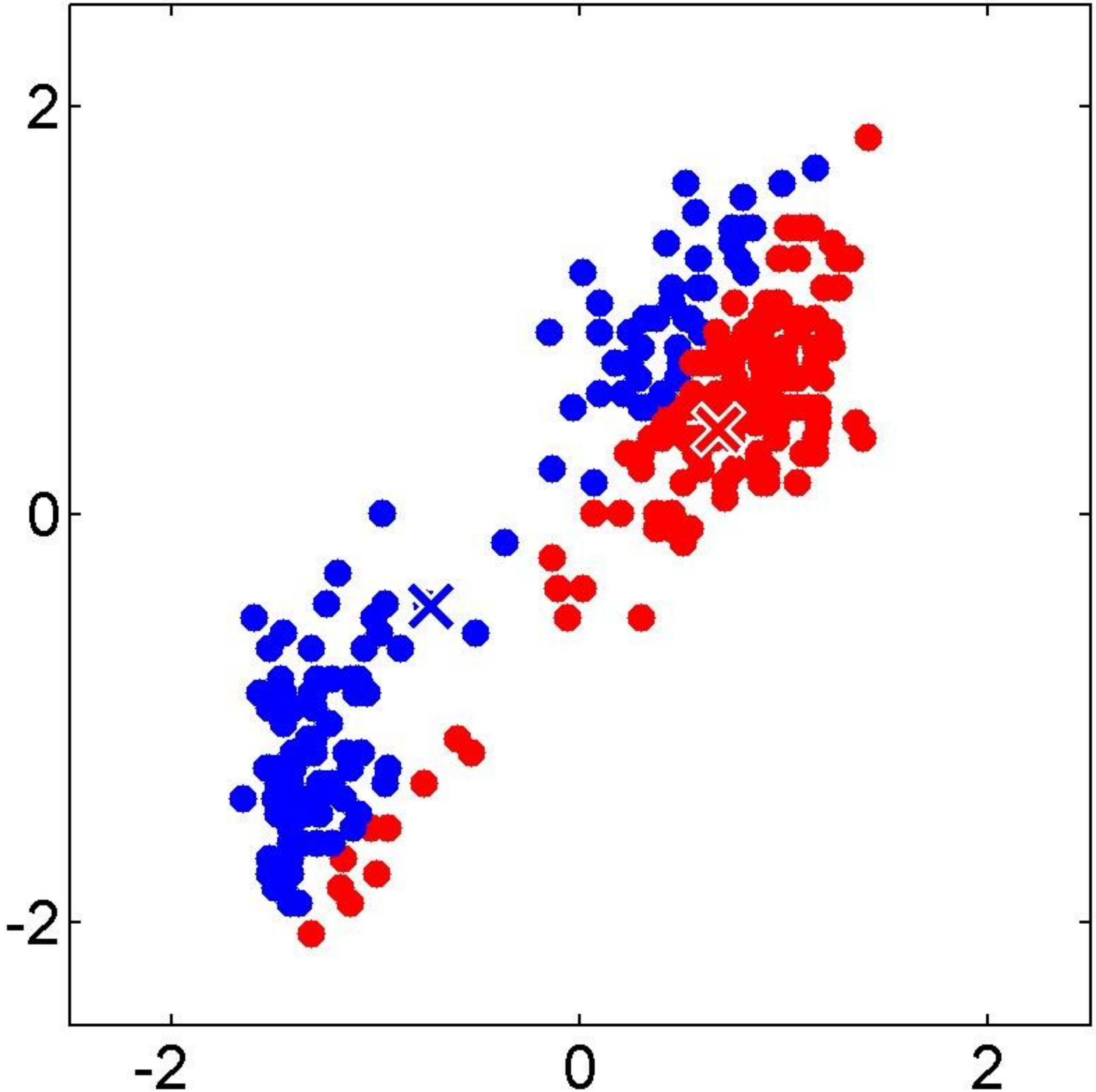
- Allocate each data point to cluster whose center is nearest (NN rule)
- Replace the cluster centers with the mean of the elements in their clusters (centroid rule)

End

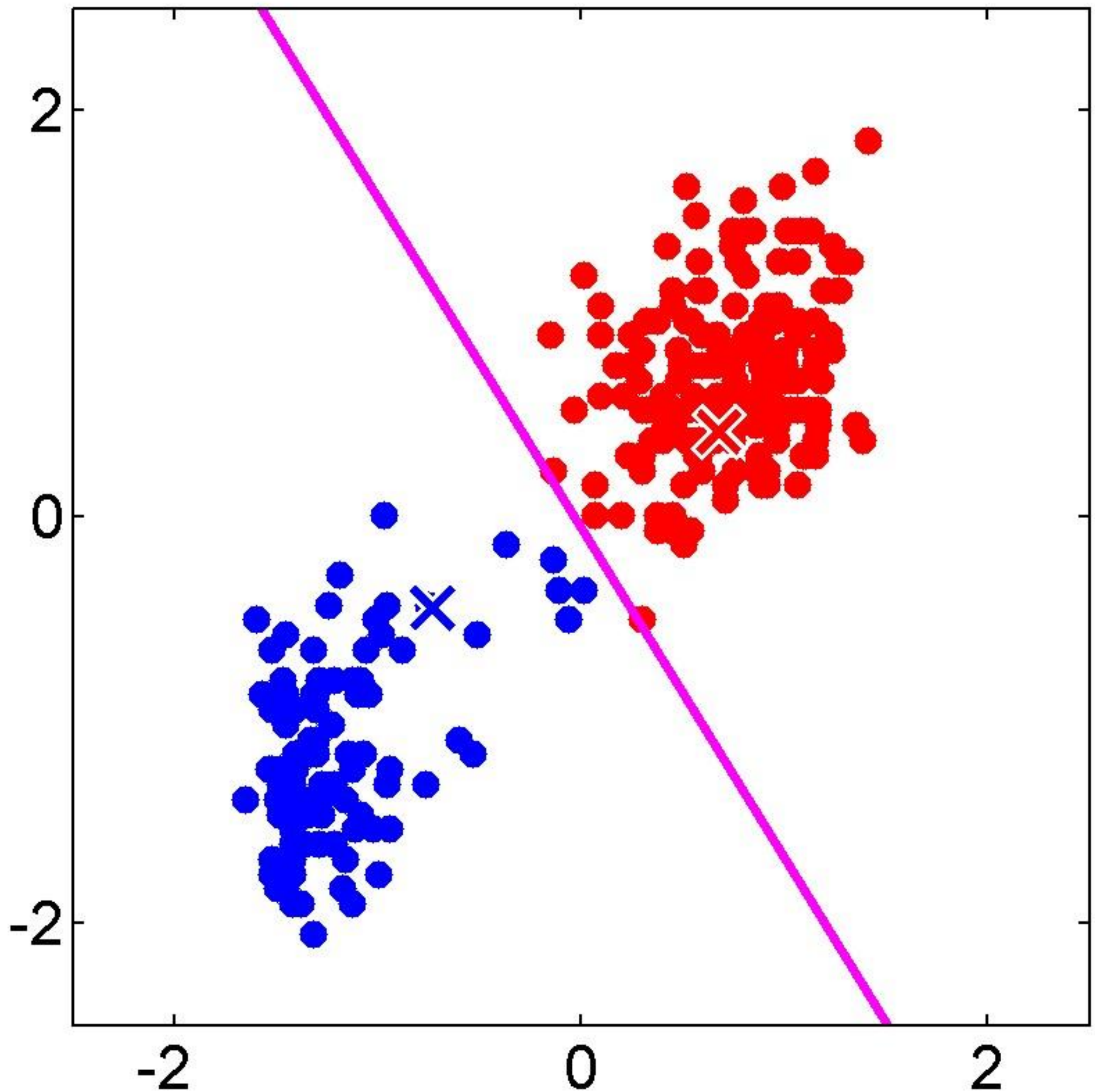

2
0
-2
-2
0
2



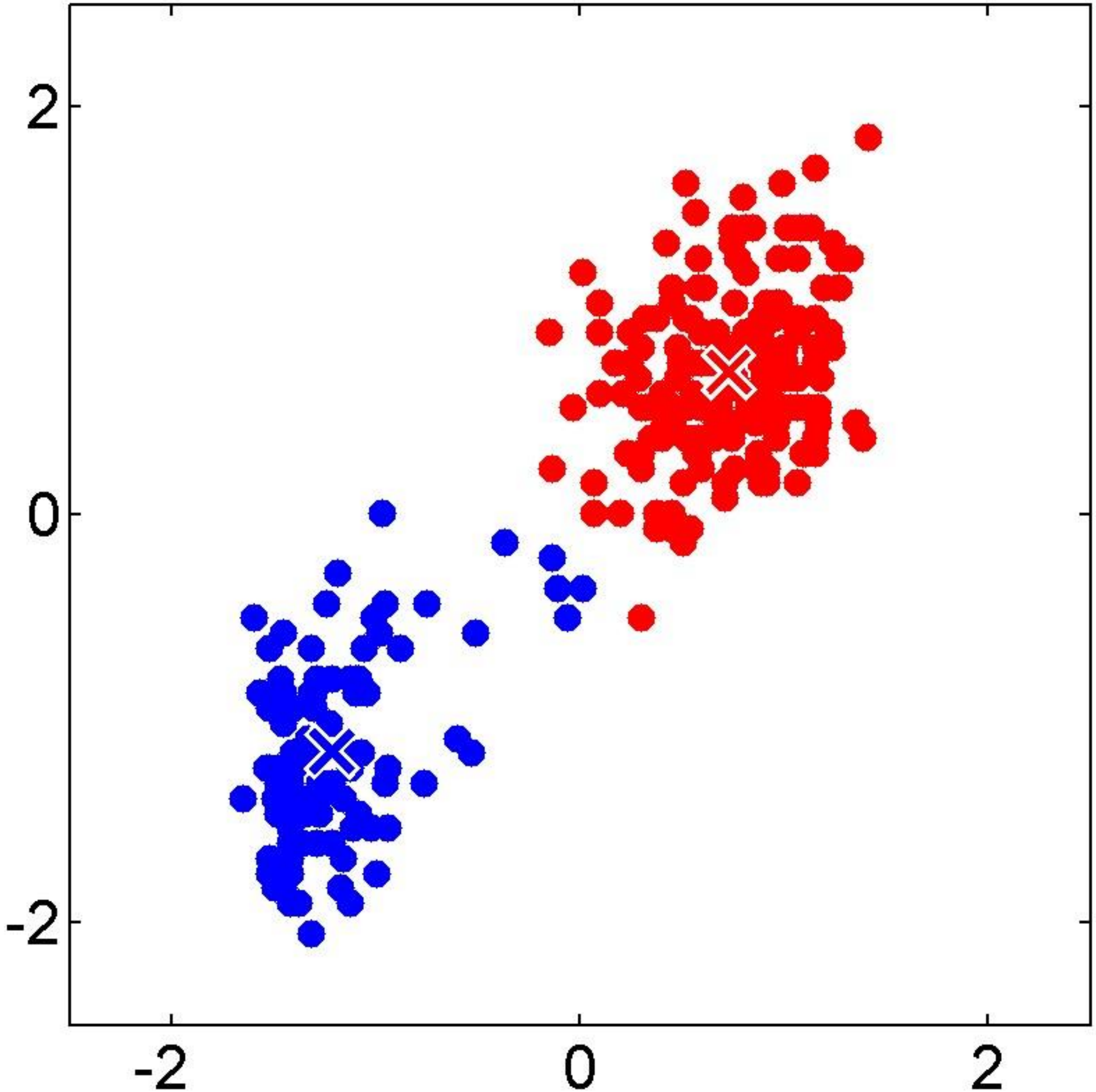
2
0
-2
-2
0
2



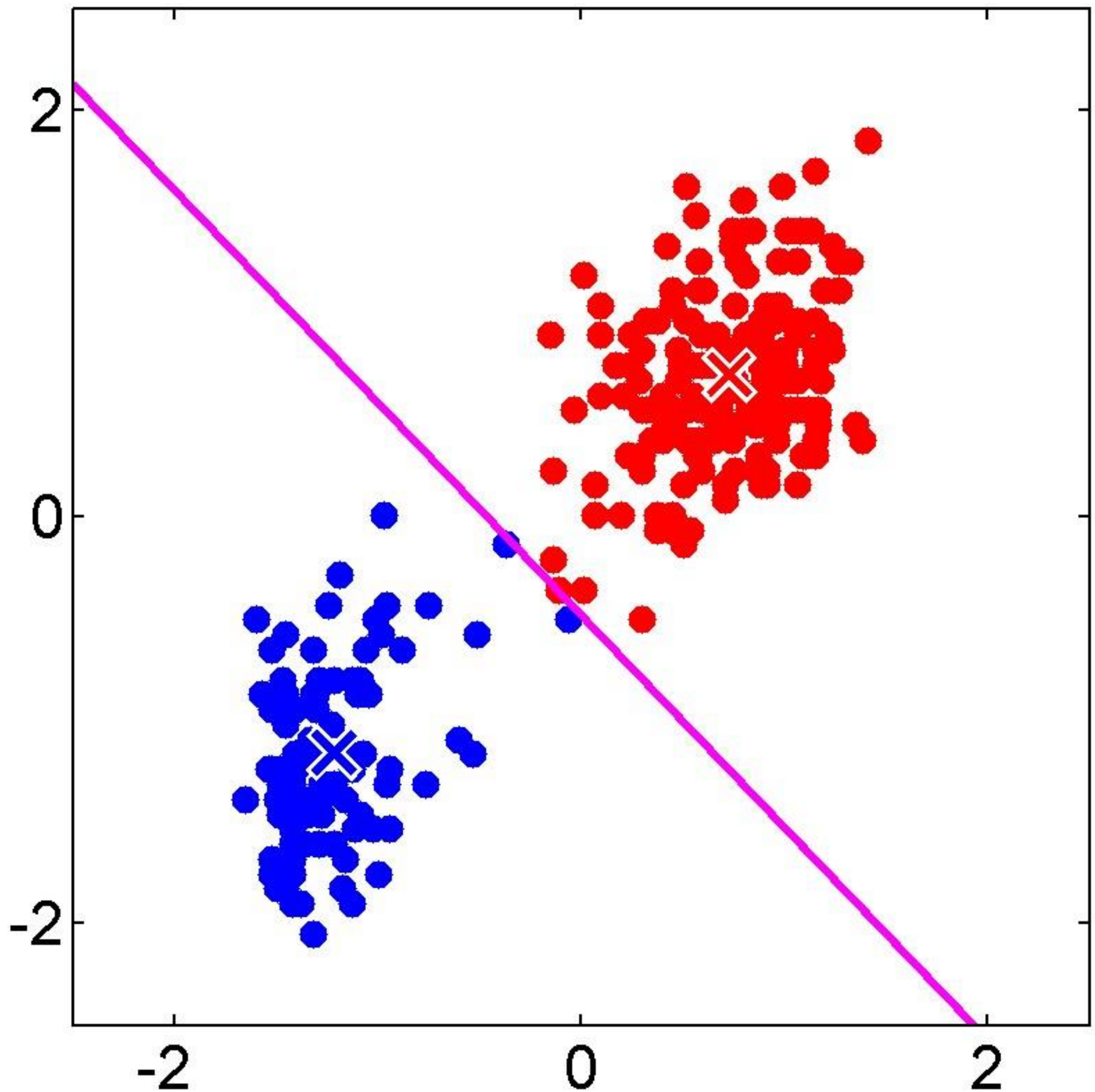
2
0
-2
-2
0
2



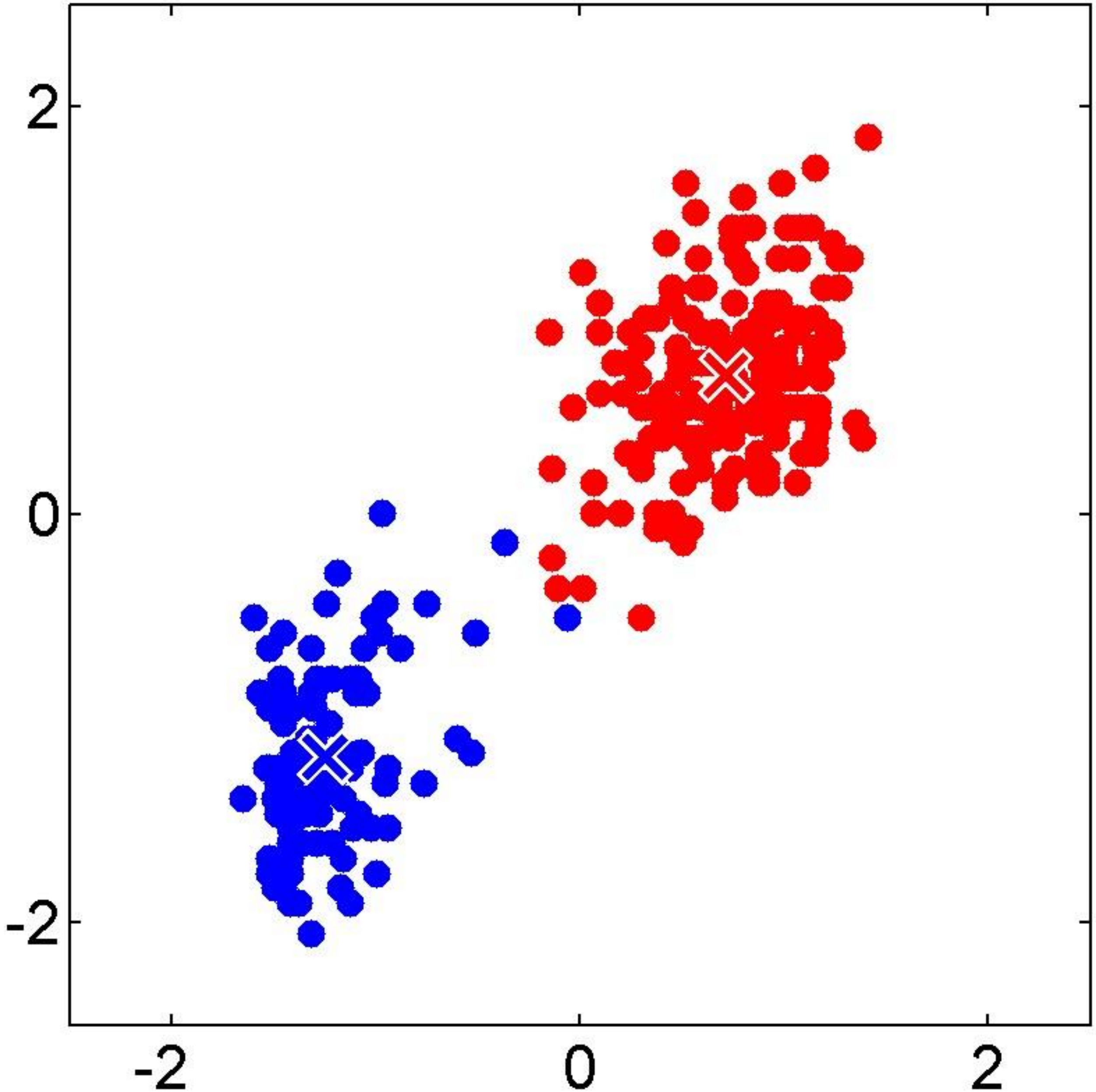
2
0
-2
-2
0
2



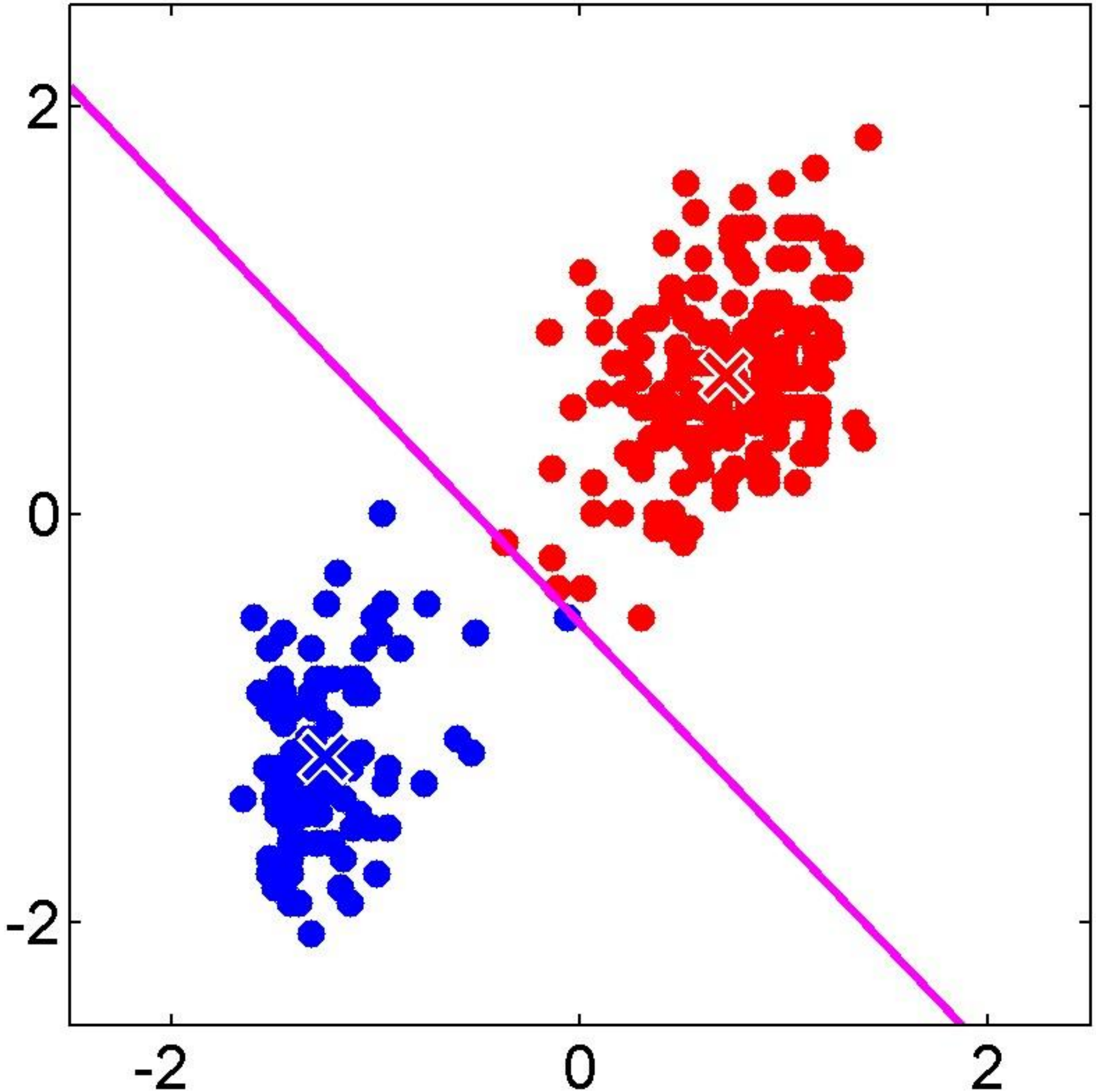
2
0
-2
-2
0
2



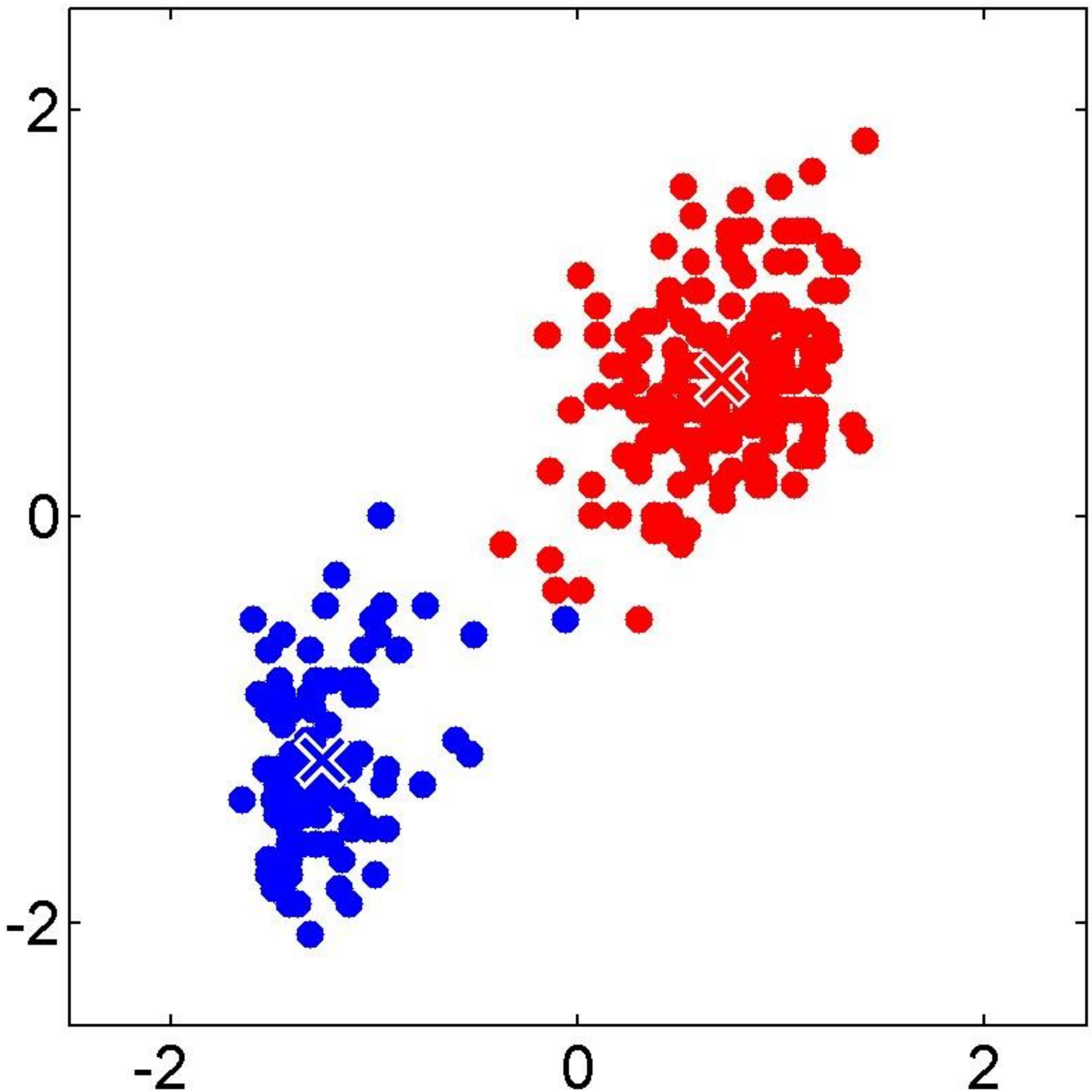
2
0
-2
-2
0
2



2
0
-2
-2
0
2



2
0
-2
-2
0
2


2
0
-2
-2
0
2

K - Means

Image

Clusters on intensity

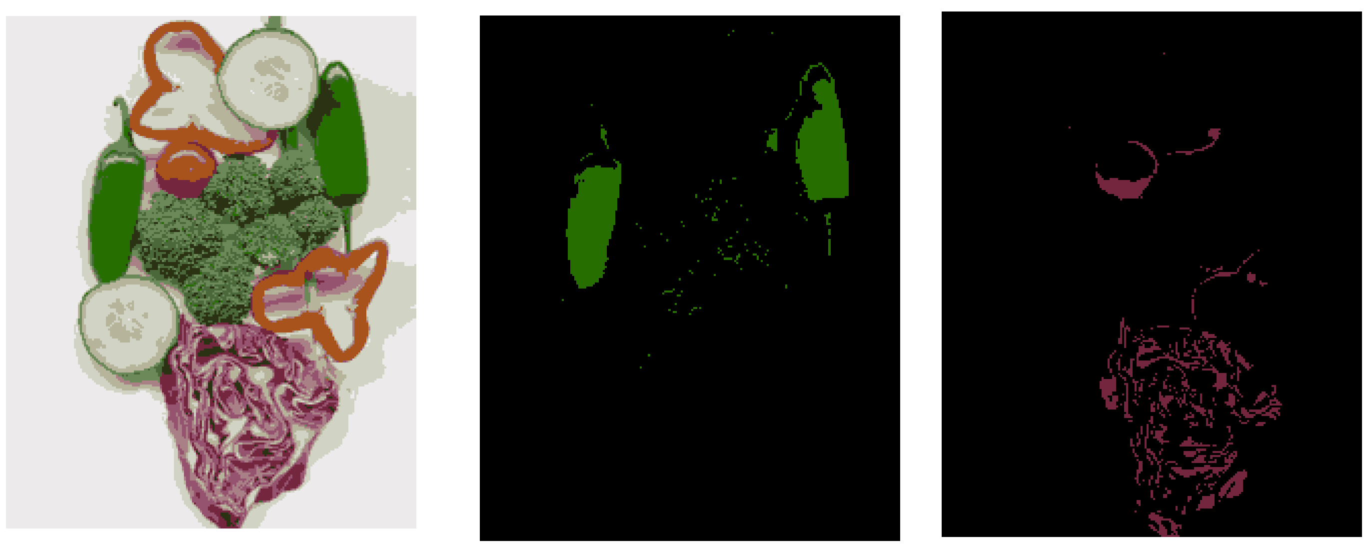
Clusters on color

K-means clustering using intensity alone and color alone
(5 clusters in each case)

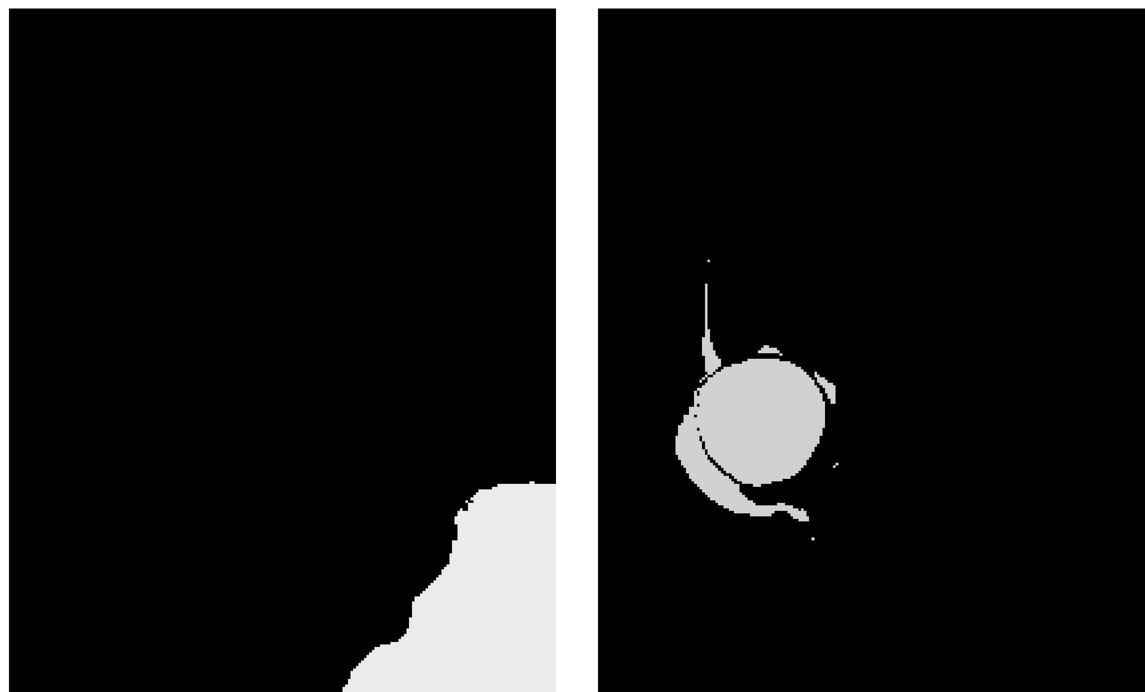
K - Means

Image

Clusters using color alone (11 clusters)

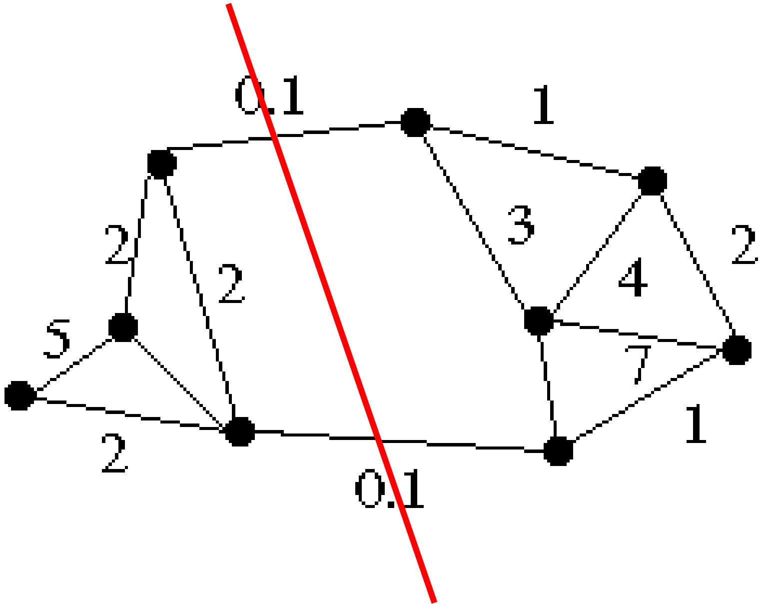


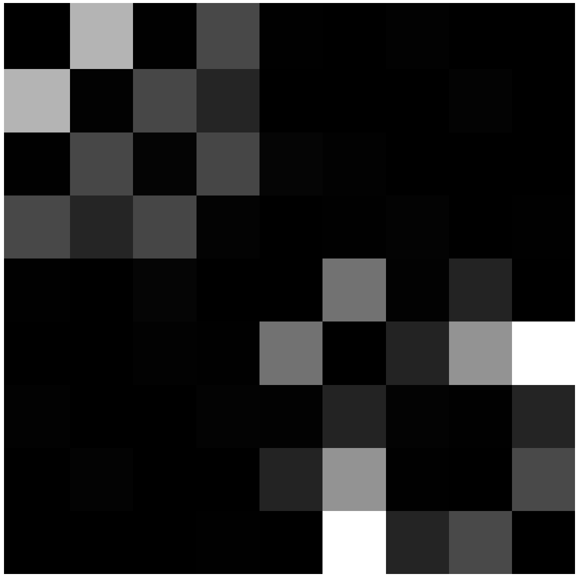
K-means using colour and
position, 20 segments



Graph Theoretic Clustering

- Graph cut
  - Represent an image using a weighted graph
    - Pixels become nodes
    - Affinity (similarity) becomes edge weights
  - Cut up this graph to get sub-graphs with strong interior links





Graph Theoretic Clustering

- Measuring affinity
  - Intensity

$$aff(x,y) = \exp\left\{-\left(\tfrac{1}{2\sigma_i^2}\right)\left(\|I(x) - I(y)\|^2\right)\right\}$$

  - Distance

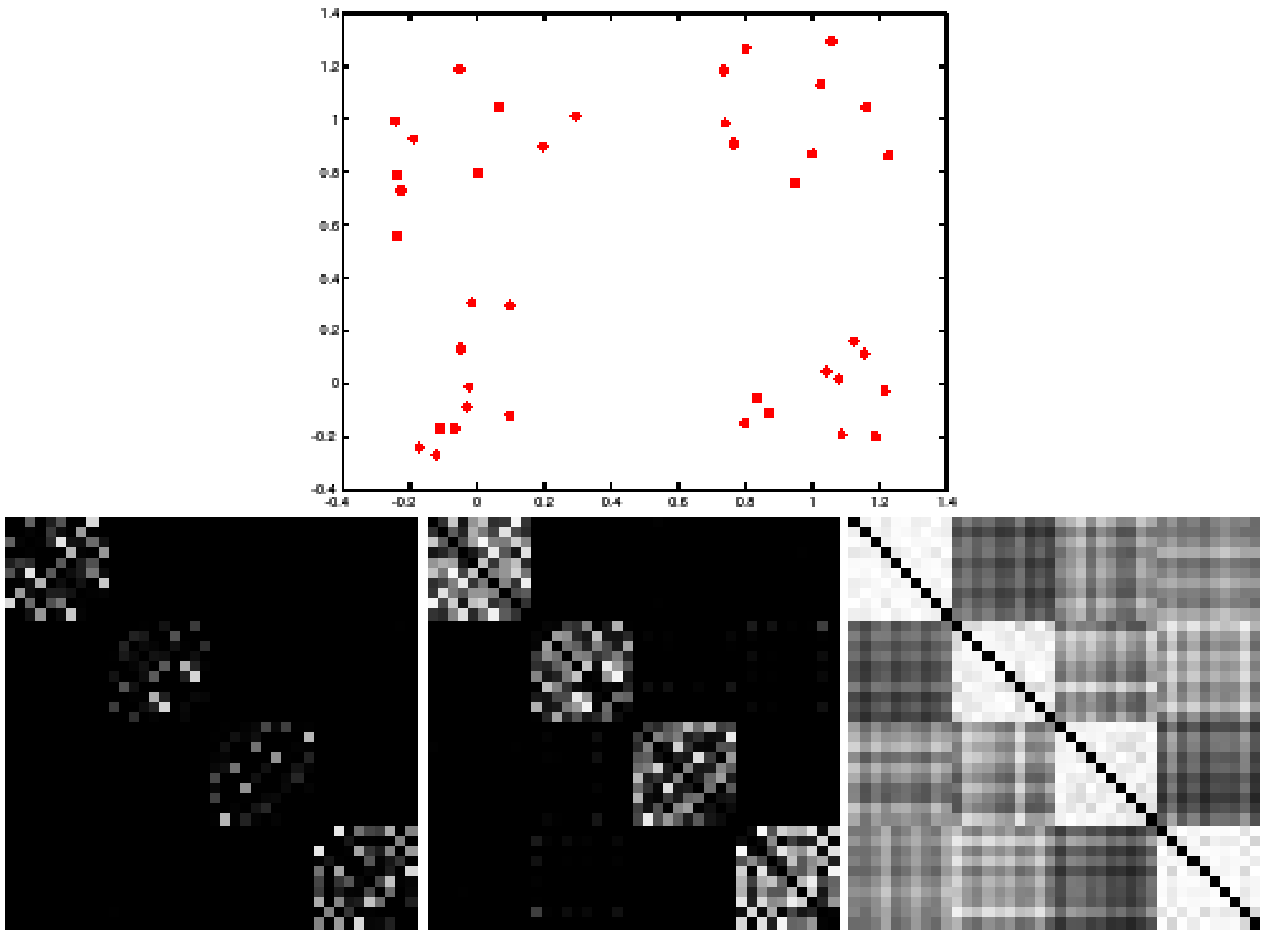
$$aff(x,y) = \exp\left\{-\left(\tfrac{1}{2\sigma_d^2}\right)\left(\|x - y\|^2\right)\right\}$$

  - Texture

$$aff(x,y) = \exp\left\{-\left(\tfrac{1}{2\sigma_t^2}\right)\left(\|c(x) - c(y)\|^2\right)\right\}$$

Graph Theoretic Clustering

- Scale(weight) affects affinity



Graph Theoretic Clustering

- How to shuffle the affinity matrix to obtain block diagonal structure?
  - Beyond the scope of the class

Normalized Cut

- Partitioning a graph $G = (V, E)$
- Cut

$$cut(A,B) = \sum_{u \in A, v \in B} w(u,v)$$

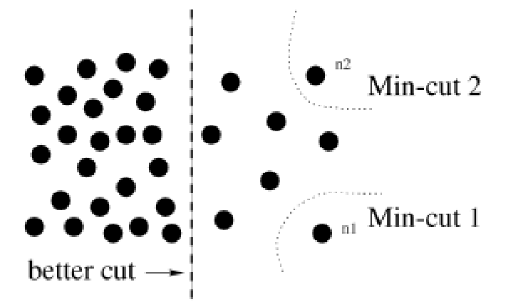


Fig. 1. A case where minimum cut gives a bad partition.

- Association

$$assoc(A,V) = \sum_{u \in A, v \in V} w(u,v)$$

- Normalized cut

$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)}$$

J. Shi and J. Malik, Normalized cuts and image segmentation, TPAMI 2000.

Normalized Cut

- Laplacian Matrix $D - W$
  - W: similarity matrix $w_{ij} = w(v_i, v_j)$
  - D: diagonal with $d_{ii} = \sum_j w_{ij}$

- The solution to the normalized cut problem

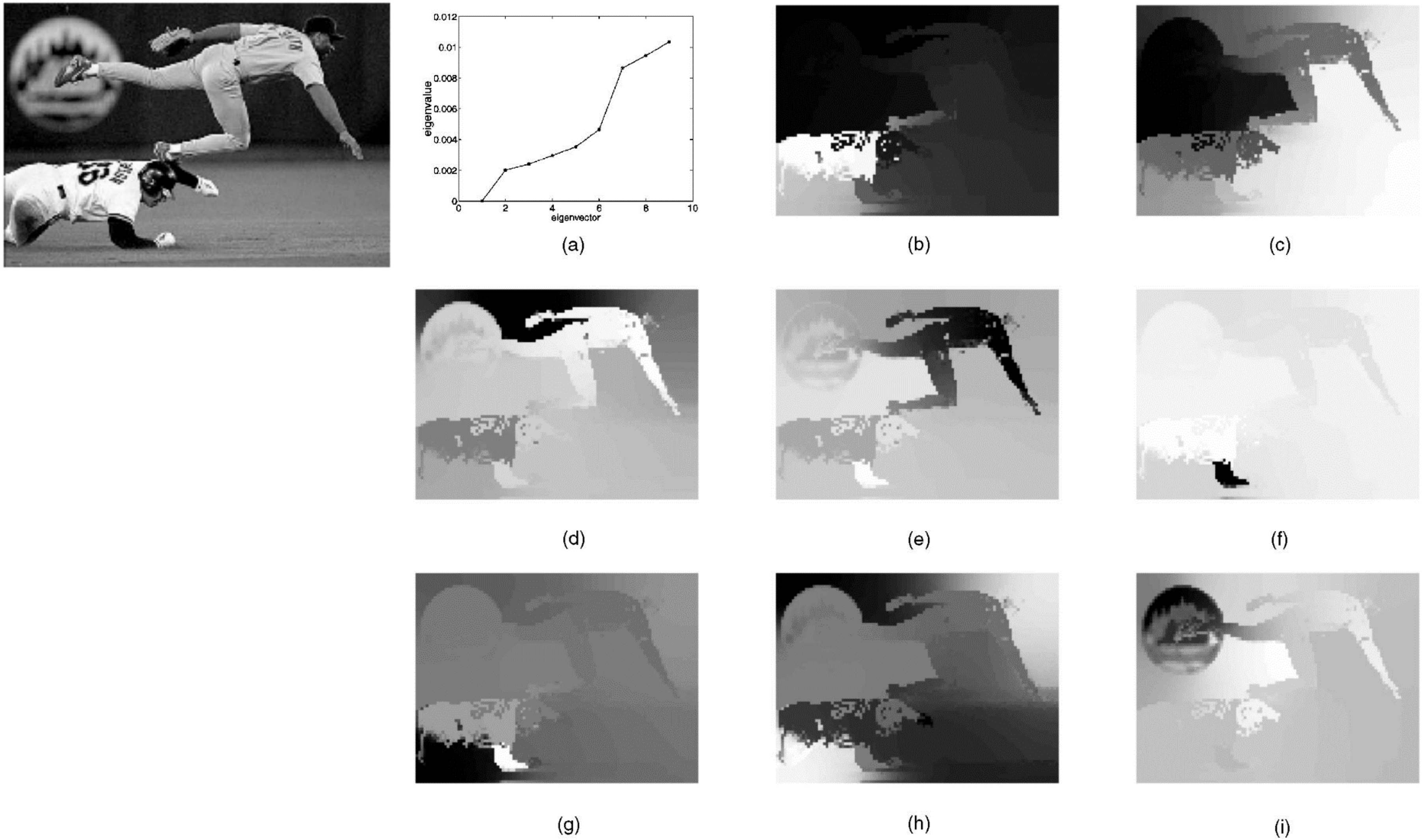
$$\min_{\mathrm{y}} \frac{y^T(D-W)y}{y^T D y}$$

subject to $y_i = 2$ or $-2b$ and $y^T D 1 = 0$

- It is a generalized eigenvalue problem
  - The eigenvector for the 2nd smallest eigenvalue gives the solution

J. Shi and J. Malik, Normalized cuts and image segmentation, TPAMI 2000.

Normalized Cut



J. Shi and J. Malik, Normalized cuts and image segmentation, TPAMI 2000.

ANOTHER EXAMPLE

Multiple Random Walkers

Multiple Random Walkers and Their Application to Image Cosegmentation

Chulwoo Lee Won-Dong Jang
School of Electrical Engineering
Korea University
{chulwoo, wdjang}@mcl.korea.ac.kr

Jae-Young Sim
School of ECE
UNIST
jysim@unist.ac.kr

Chang-Su Kim
School of Electrical Engineering
Korea University
changsukim@korea.ac.kr

Abstract

A graph-based system to simulate the movements and interactions of multiple random walkers (MRW) is proposed in this work. In the MRW system, multiple agents traverse a single graph simultaneously. To achieve desired interactions among those agents, a restart rule can be designed, which determines the restart distribution of each agent according to the probability distributions of all agents. In particular, we develop the repulsive rule for data clustering. We illustrate that the MRW clustering can segment real images reliably. Furthermore, we propose a novel image cosegmentation algorithm based on the MRW clustering. Specifically, the proposed algorithm consists of two steps: inter-image concurrence computation and intra-image MRW clustering. Experimental results demonstrate that the proposed algorithm provides promising cosegmentation performance.

one another and form their own dominant regions. Eventually, the power balance among the agents is achieved, and their distributions converge. By comparing the stationary distributions, clustering can be achieved. We demonstrate that this MRW process can cluster point data and segment real images reliably.

Moreover, we apply the proposed MRW system to the problem of segmenting similar images jointly. Recently, attempts to extract common foreground objects from a set of similar images have been made. This approach, called *cosegmentation*, was first addressed by Rother *et al.* [23] and has been researched actively [20, 12, 14, 3, 5, 30, 7, 24, 31]. Compared with segmenting each image independently, it is advantageous to delineate similar objects from multiple images. However, since repeating image features do not always imply the most important and informative parts of a scene, cosegmentation is still a challenging vision problem.

For cosegmentation, we introduce the notion of concur-

$\mathbf{r}_k^{(t)} = [r_{k,1}^{(t)}, \cdots, r_{k,N}^{(t)}]^T$.

We can make the agents interact with one another, by determining the restart distribution as

$$\mathbf{r}_k^{(t)} = (1-\delta^t)\mathbf{r}_k^{(t-1)} + \delta^t \phi_k(\mathcal{P}^{(t)}) \tag{6}$$

where the function ϕ_k is referred to as the restart rule. It determines a probability distribution ϕ_k from $\mathcal{P}^{(t)} = \{\mathbf{p}_k^{(t)}\}_{k=1}^{K}$, which is the set of the probability distributions of all agents at time t.

In (6), δ is a constant within $[0,1]$, called the cooling factor. In an extreme case of $\delta = 0$, the restart distribution $\mathbf{r}_k^{(t)}$ becomes time-invariant, and the MRW recursion of each agent in (5) is identical with the RWR recursion in (3). In the other extreme case of $\delta = 1$, $\mathbf{r}_k^{(t)} = \phi_k(\mathcal{P}^{(t)})$ does not directly depend on the previous distribution $\mathbf{r}_k^{(t-1)}$. Suppose that $0 < \delta < 1$. We have

$$\|\mathbf{r}_k^{(t)} - \mathbf{r}_k^{(t-1)}\|_\infty = \delta^t \|\phi_k(\mathcal{P}^{(t)}) - \mathbf{r}_k^{(t-1)}\|_\infty \leq \delta^t. \tag{7}$$

Thus, if $s \geq t \geq T$, $\|\mathbf{r}_k^{(s)} - \mathbf{r}_k^{(t)}\|_\infty \leq \delta^T/(1-\delta)$. So each element in the restart distribution $\mathbf{r}_k^{(t)}$ is a Cauchy sequence in terms of time t. Since a Cauchy sequence in $\mathbb{R}$ is convergent, the restart distribution $\mathbf{r}_k^{(t)}$ converges to a fixed distribution $\mathbf{r}_k^{(\infty)}$. Therefore, as t approaches infinity, the MRW recursion in (5) becomes the RWR recursion, and agent k has a stationary distribution eventually. To summarize we have the following convergence theorem.

In the MRW system,

$$\mathbf{p}_k = [p(\mathbf{x}_1|\omega_k), \cdots, p(\mathbf{x}_N|\omega_k)]^T \tag{9}$$

where $p(\mathbf{x}_i|\omega_k)$ is the probability that agent k is found at node i. According to the Bayes' rule, the posterior probability is given by

$$p(\omega_k|\mathbf{x}_i) = \frac{p(\mathbf{x}_i|\omega_k)p(\omega_k)}{\sum_l p(\mathbf{x}_i|\omega_l)p(\omega_l)}, \tag{10}$$

which represents the probability that node i is occupied by agent k. The repulsive restart rule sets the ith element of $\phi_k(\mathcal{P})$ as

$$\phi_{k,i} = \alpha \cdot p(\omega_k|\mathbf{x}_i) \cdot p(\mathbf{x}_i|\omega_k) \tag{11}$$

where α is a normalizing factor to make $\phi_k(\mathcal{P})$ a probability distribution. Suppose that agent k is dominant at node i, *i.e.*, it has a high posterior probability $p(\omega_k|\mathbf{x}_i)$ and a high likelihood $p(\mathbf{x}_i|\omega_k)$. Then, it restarts at that node with a high probability, and tends to become more dominant. This has the effect that a dominant agent at a node repels the other agents. The repulsive restart rule in (11) can be rewritten as

$$\phi_k(\mathcal{P}) = \alpha \mathbf{Q}_k \mathbf{p}_k \tag{12}$$

where $\mathbf{Q}_k$ is a diagonal matrix whose (i,i)th element is the posterior probability $p(\omega_k|\mathbf{x}_i)$.

For clustering, we perform the MRW simulation in (5) and (6), by employing the restart rule in (12), to obtain the

Acknowledgements

This work was supported partly by the Global Frontier R&D Program on Human-centered Interaction for Coexistence funded by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (MSIP) (2011-0031648), and partly by the NRF grant funded by the Korea government (MSIP) (No. 2012-011031).

References

[1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE TPAMI*, 34(11):2274–2282, Nov. 2012.

[2] N. Alon, C. Avin, M. Koucký, G. Kozma, Z. Lotker, and M. R. Tuttle. Many random walks are faster than one. In *Proc. ACM SPAA*, pages 110–128, 2008.

[3] D. Batra, A. Kowdle, D. Parikh, J. Luo, and T. Chen. iCoseg: Interactive co-segmentation with intelligent scribble guidance. In *Proc. IEEE CVPR*, pages 3169–3176, 2010.

[4] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1):107–117, 1998.

[5] K.-Y. Chang, T.-L. Liu, and S.-H. Lai. From co-saliency to co-segmentation: An efficient and fully unsupervised energy minimization model. In *Proc. IEEE CVPR*, pages 2129–2136, 2011.

[6] F. R. K. Chung. *Spectral Graph Theory*. American Mathe-

[17] C. Liu, J. Yuen, and A. Torralba. SIFT flow: Dense correspondence across scenes and its applications. *IEEE TPAMI*, 33(5):978–994, May 2011.

[18] M. Meila and J. Shi. Learning segmentation by random walks. In *Adv. Neural Inf. Process. Syst.*, pages 873–879, 2001.

[19] M. Meila and J. Shi. A random walks view of spectral segmentation. In *Proc. AISTATS*, 2001.

[20] L. Mukherjee, V. Singh, and C. Dyer. Half-integrality based algorithms for cosegmentation of images. In *Proc. IEEE CVPR*, pages 2028–2035, 2009.

[21] A. Y. Ng, M. I. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. In *Adv. Neural Inf. Process. Syst.*, volume 2, pages 849–856, 2002.

[22] J.-Y. Pan, H.-J. Yang, C. Faloutsos, and P. Duygulu. Automatic multimedia cross-modal correlation discovery. In *Proc. ACM SIGKDD*, pages 653–658, 2004.

[23] C. Rother, T. Minka, A. Blake, and V. Kolmogorov. Cosegmentation of image pairs by histogram matching - incorporating a global constraint into MRFs. In *Proc. IEEE CVPR*, pages 993–1000, 2006.

[24] J. Rubio, J. Serrat, A. Lopez, and N. Paragios. Unsupervised co-segmentation through region matching. In *Proc. IEEE CVPR*, pages 749–756, 2012.

[25] S. E. Schaeffer. Graph clustering. *Computer Science Review*, 1(1):27–64, 2007.

[26] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE TPAMI*, 22(8):888–905, 2000.

Change photo

Chang-Su Kim

Edit | Follow

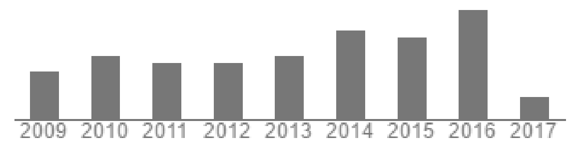
Professor of Korea University, Korea University
Image Processing, Computer Vision
Verified email at korea.ac.kr - Homepage
My profile is public

Title | Add | More | 1–20 | Cited by | Year

Title	Cited by	Year
Technologies for 3D mesh compression: A survey J Peng, CS Kim, CCJ Kuo Journal of Visual Communication and Image Representation 16 (6), 688-733	275	2005
Motion-compensated frame interpolation using bilateral motion estimation and adaptive overlapped block motion compensation BD Choi, JW Han, CS Kim, SJ Ko IEEE Transactions on Circuits and Systems for Video Technology 17 (4), 407-416	272	2007
Spatial and temporal error concealment techniques for video transmission over noisy channels WY Kung, CS Kim, CCJ Kuo IEEE transactions on circuits and systems for video technology 16 (7), 789-803	149	2006
Optimized contrast enhancement for real-time image and video dehazing JH Kim, WD Jang, JY Sim, CS Kim	94	2013

Google Scholar

Citation indices	All	Since 2012
Citations	3270	1756
h-index	30	19
i10-index	87	47



Co-authors Edit...

- Sang-Uk Lee
- C.-C. Jay Kuo
- Jae-Young Sim
- Chul Lee
- Won-Dong Jang
- Tae-Young Chung
- Jin-Hwan Kim
- Chulwoo Lee
- Jae-Kyun Ahn
- Dae-Youn Lee

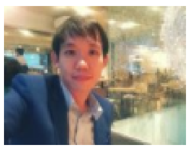
...

Large-scale 3D point cloud compression using adaptive radial distance prediction in hybrid coordinate domains JK Ahn, KY Lee, JY Sim, CS Kim IEEE Journal of Selected Topics in Signal Processing 9 (3), 422-434	7	2015
SOWP: Spatially Ordered and Weighted Patch Descriptor for Visual Tracking HU Kim, DY Lee, JY Sim, CS Kim Proceedings of the IEEE International Conference on Computer Vision, 3011-3019	7	2015
Multihypothesis Trajectory Analysis for Robust Visual Tracking DY Lee, JY Sim, CS Kim Proceedings of the IEEE Conference on Computer Vision and Pattern ...	11	2015
Multiple Random Walkers and Their Application to Image Cosegmentation C Lee, WD Jang, JY Sim, CS Kim Proceedings of the IEEE Conference on Computer Vision and Pattern ...	19	2015
On DCT coefficient distribution in video coding using quad-tree structured partition JW Kang, CS Kim Asia-Pacific Signal and Information Processing Association, 2014 Annual ...	2	2014
VGEF: Contrast enhancement of dark images using value gap expansion force and sorted histogram equalization Y Kim, JH Kim, CS Kim Asia-Pacific Signal and Information Processing Association, 2014 Annual ...		2014
Hybrid representation and rendering of indoor environments using meshes and point clouds HU Kim, JH Kim, CS Kim Ubiquitous Robots and Ambient Intelligence (URAI), 2014 11th International ...		2014
Stereo video deraining and desnowing based on spatiotemporal frame warping JH Kim, JY Sim, CS Kim 2014 IEEE International Conference on Image Processing (ICIP), 5432-5436	1	2014

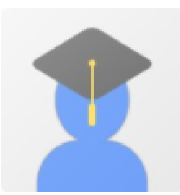
Depth-guided adaptive contrast enhancement using 2D

Chul Lee
Assistant Professor of Computer Engineering, Pukyong National University
Verified email at pknu.ac.kr
Cited by 349
Image restoration   image enhancement   computational imaging

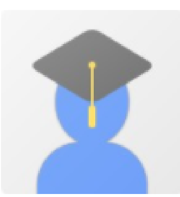
Won-Dong Jang
Media Communications Lab, Korea University
Verified email at mcl.korea.ac.kr
Cited by 153
Image Processing   Computer Vision



Tae-Young Chung
System S/W platform(Multimedia), Software R&D Center,Samsung Electronics
Verified email at mcl.korea.ac.kr
Cited by 212
3D video   coding   error concelment   h.264   multi-view plus depth



Jin-Hwan Kim
Korea University
Verified email at mcl.korea.ac.kr
Cited by 281
Image Enhancement   Computer Vision



Chulwoo Lee
Northumbria University
Verified email at northumbria.ac.uk
Cited by 229
Image processing   Contrast enhancement   High dynamic range imaging   Computer vision

Jae-Kyun Ahn
Korea University
Verified email at korea.ac.kr
Cited by 76

RECENT PUBLICATIONS

See all publications

Online video object segmentation via convolutional trident network
Won-Dong Jang and Chang-Su Kim, ***CVPR***, Honolulu, HI, Jul. 2017.

Contour-constrained superpixels for image and video processing
Se-Ho Lee, Won-Dong Jang, and Chang-Su Kim, ***CVPR***, Honolulu, HI, Jul. 2017.

Semi-supervised video object segmentation using multiple random walkers [pdf] [project page]
Won-Dong Jang and Chang-Su Kim, ***BMVC***, York, UK, Sep. 2016.

Streaming video segmentation via short-term hierarchical segmentation and frame-by-frame Markov random field optimization [pdf] [project page]
Won-Dong Jang and Chang-Su Kim, ***ECCV***, Amsterdam, Netherlands, Oct. 2016.

RGB-D image segmentation based on multiple random walkers [pdf]
Se-Ho Lee, Won-Dong Jang, Byung Kwan Park, and Chang-Su Kim, *IEEE* ***ICIP***, Phoenix, AZ, Sep. 2016.

Primary object segmentation in videos via alternate convex optimization of foreground and background distributions [pdf] [project page]
Won-Dong Jang, Chulwoo Lee, and Chang-Su Kim, *IEEE Conference on* ***CVPR***, Las Vegas, NV, Jun. 2016.

POD: Discovering primary objects in videos based on evolutionary refinement of object recurrence, background, and primary object models [pdf]
Yeong Jun Koh, Won-Dong Jang, and Chang-Su Kim, *IEEE Conference on* ***CVPR***, Las Vegas, NV, Jun. 2016.

Multiple random walkers and their application to image cosegmentation [pdf] [project page]
Chulwoo Lee, Won-Dong Jang, Jae-Young Sim, and Chang-Su Kim, *IEEE Conference on* ***CVPR***, Boston, MA, Jun. 2015.