COLLEGE OF ENGINEERING
ELECTRICAL ENGINEERING & COMPUTER SCIENCE
UNIVERSITY OF MICHIGAN

Segmentation ~~and Clustering~~

EECS 598-08 Fall 2014
Foundations of Computer Vision

Instructor: Jason Corso (jjcorso)
web.eecs.umich.edu/~jjcorso/t/598F14

Readings: FP 6.2, 9; SZ 5.2-5.4

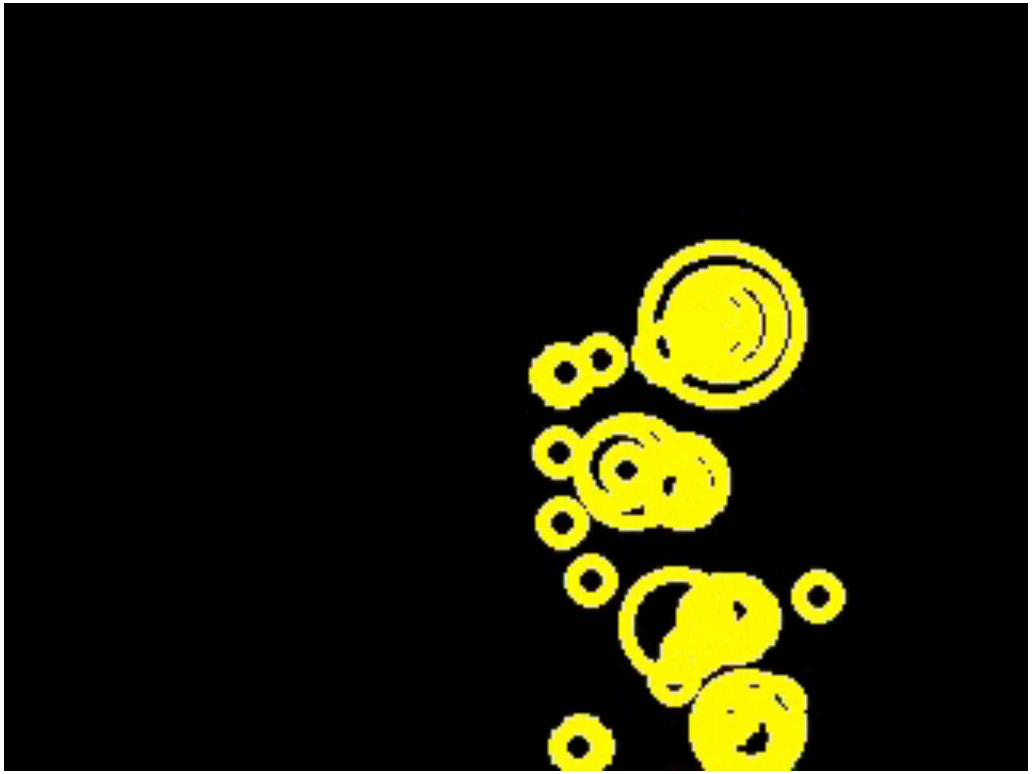
Date: 10/1/14

Materials on these slides have come from many sources in addition to myself; individual slides reference specific sources.

Plan

- Motivation for segmentation
- Gestalt Psychology / human perception for segmentation
- Piecewise Constant/Smooth Models

Some motivation; what do you see?

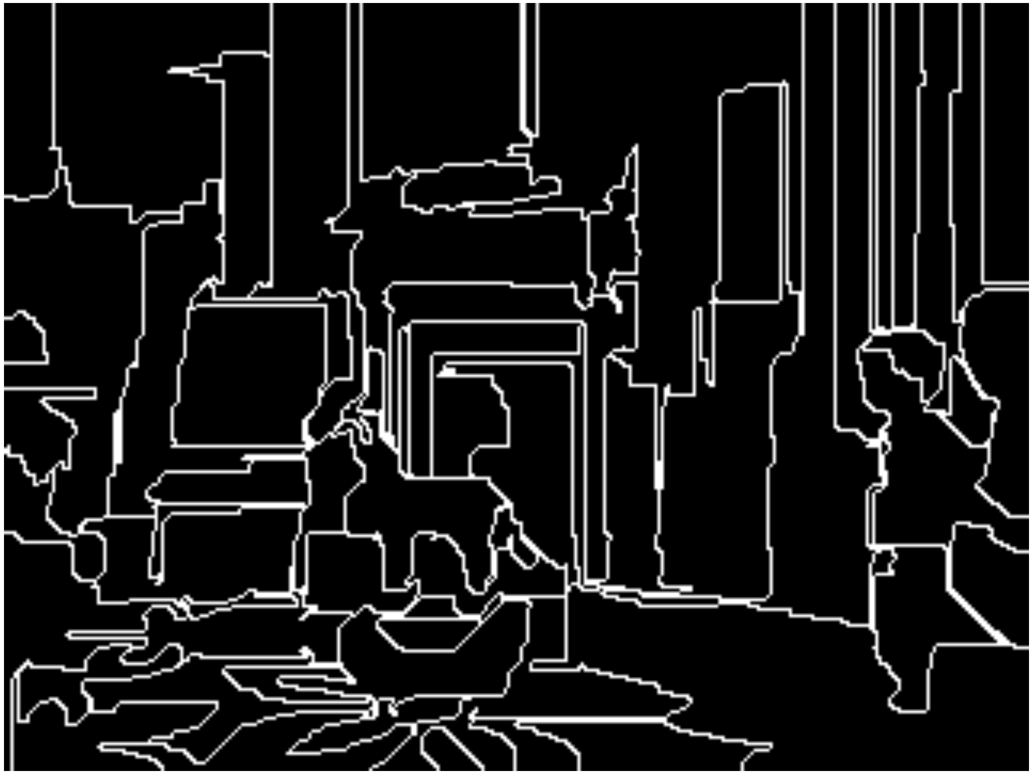


Method: Laptev. “On Space-Time Interest Points.” IJCV 64(2/3):107-123. 2005.

Some motivation; what do you see?

Sources: Maas 1971 with Johansson; downloaded from Youtube.

Some motivation; what do you see?

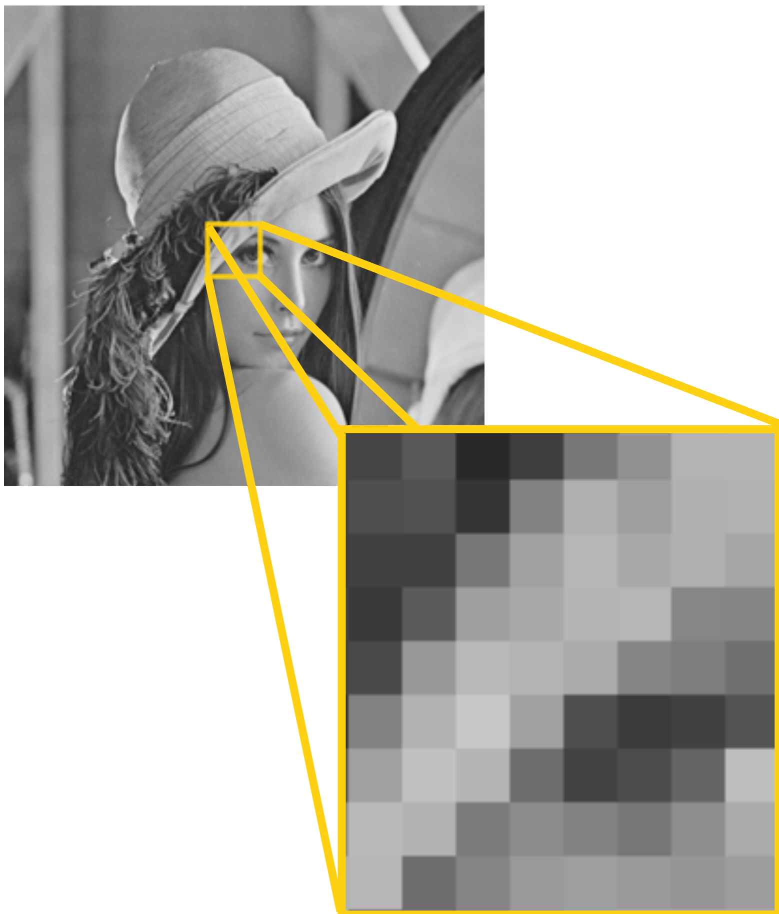


Method: Supervoxel segment boundaries. Xu and Corso CVPR 2012.

Segmentation: Toward a Representation with Rich Semantics?

Background

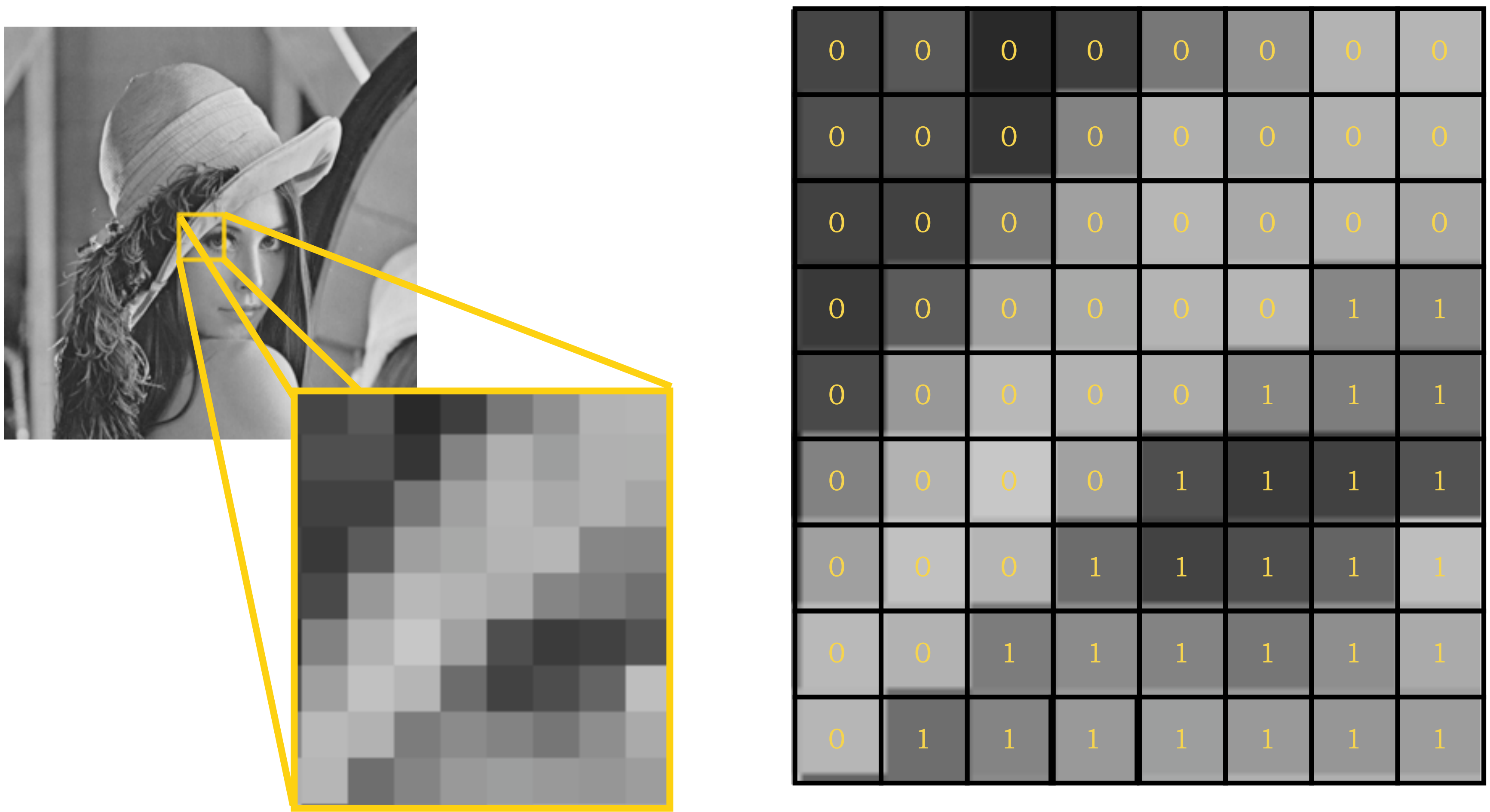
Images



62	70	31	47	100	125	164	166
62	63	40	112	159	140	160	161
50	50	100	143	167	153	150	148
43	73	142	152	165	167	115	114
57	134	170	164	155	114	106	93
111	163	187	144	61	45	50	62
143	180	166	89	51	60	81	176
141	163	105	120	112	99	123	154
167	91	113	135	140	135	135	139

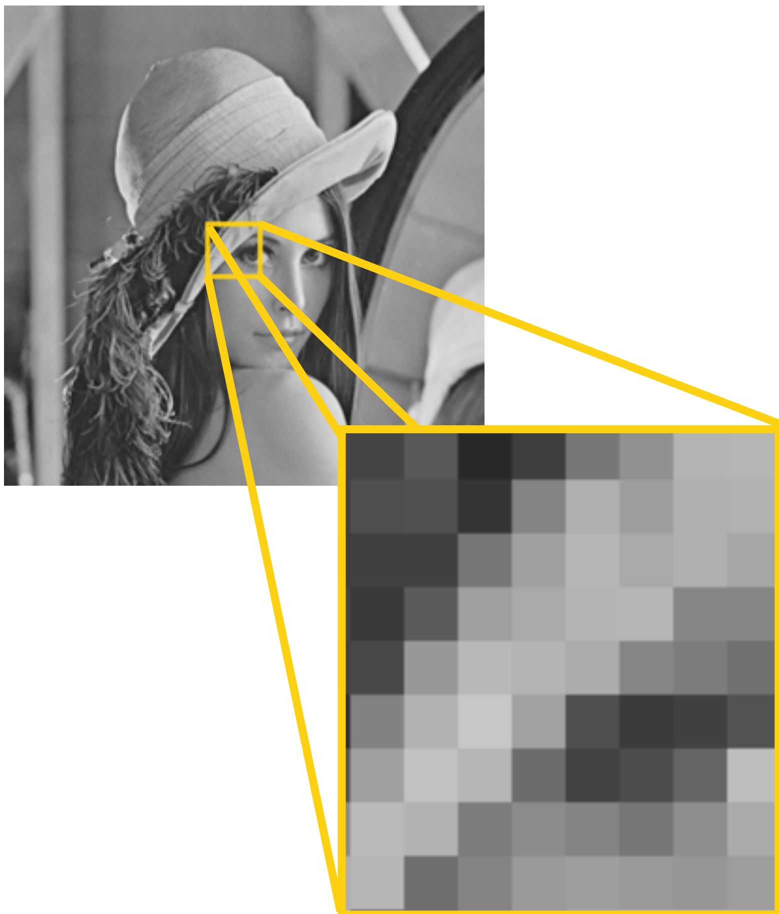
Background

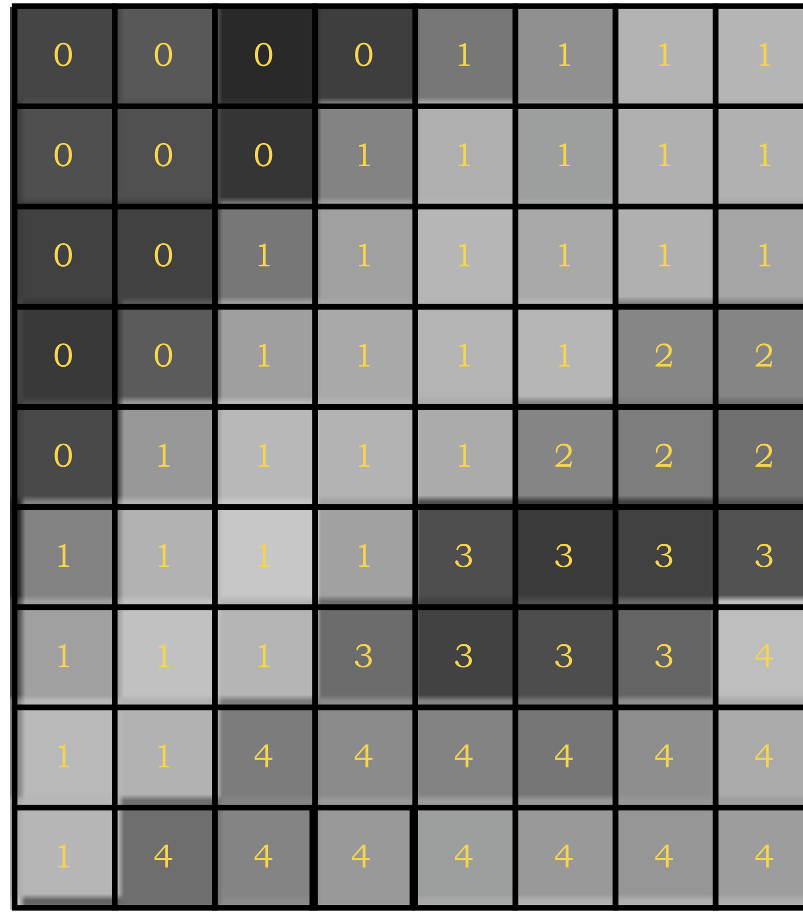
Segmentation



Background

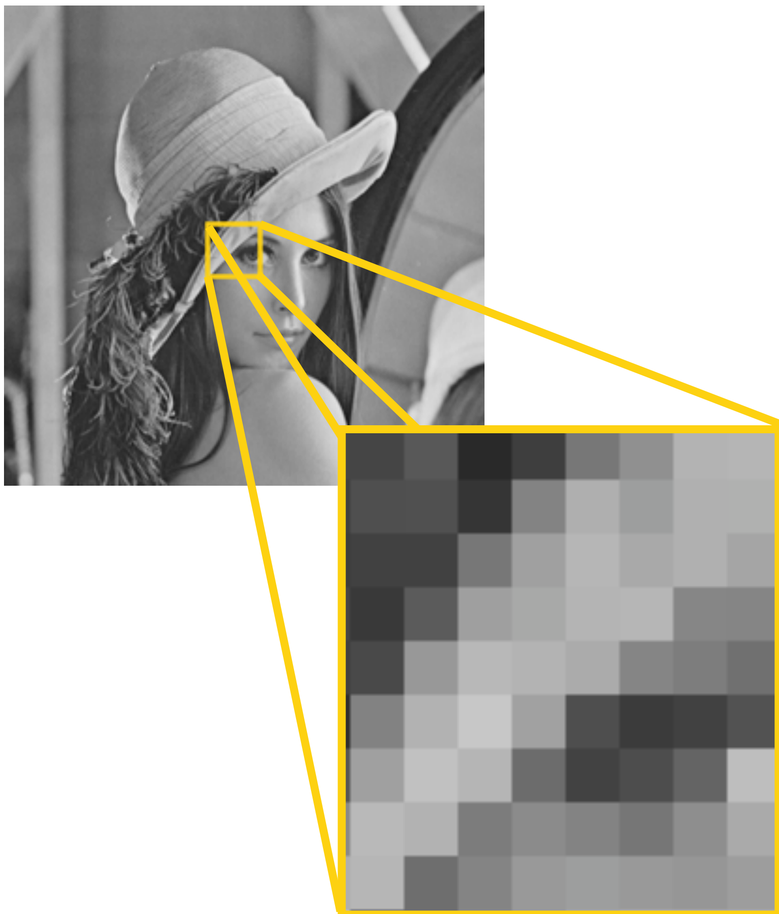
Segmentation

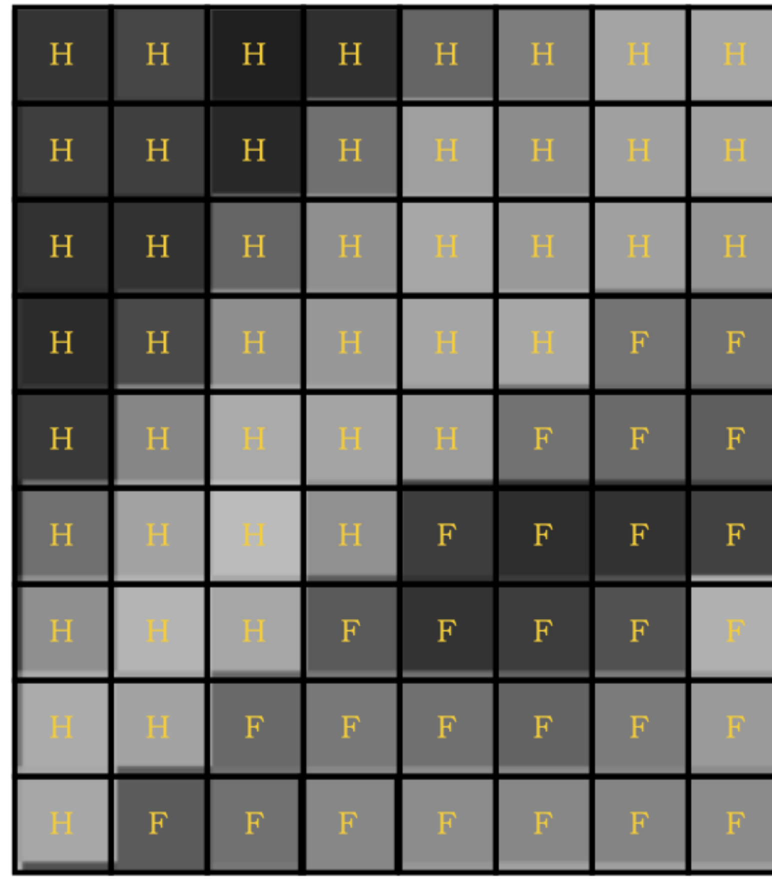




Background

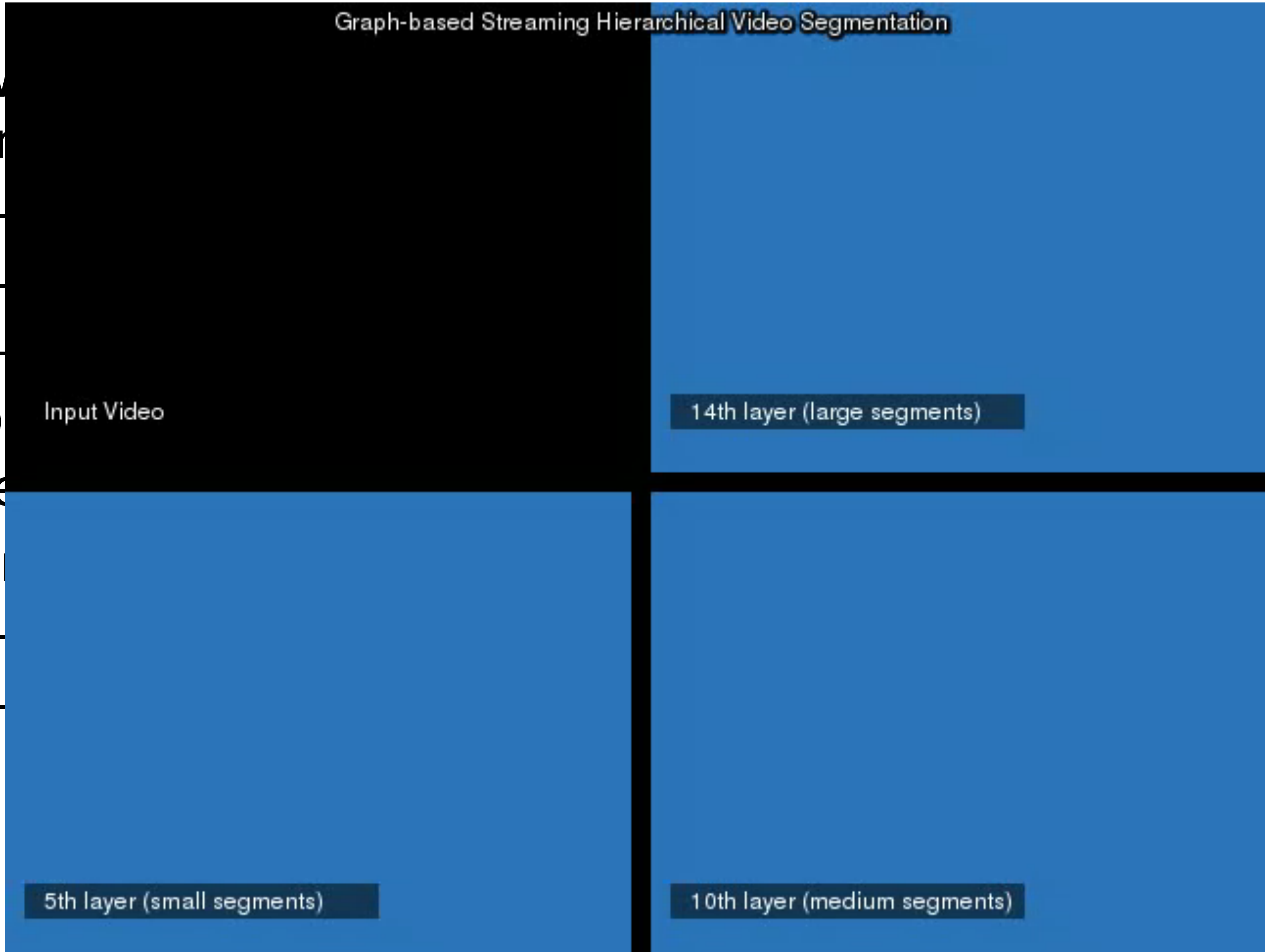
Segmentation and Classification





Segmentation: A Complementary “Feature”?

- W
  ur
- D
  se
- P



General ideas

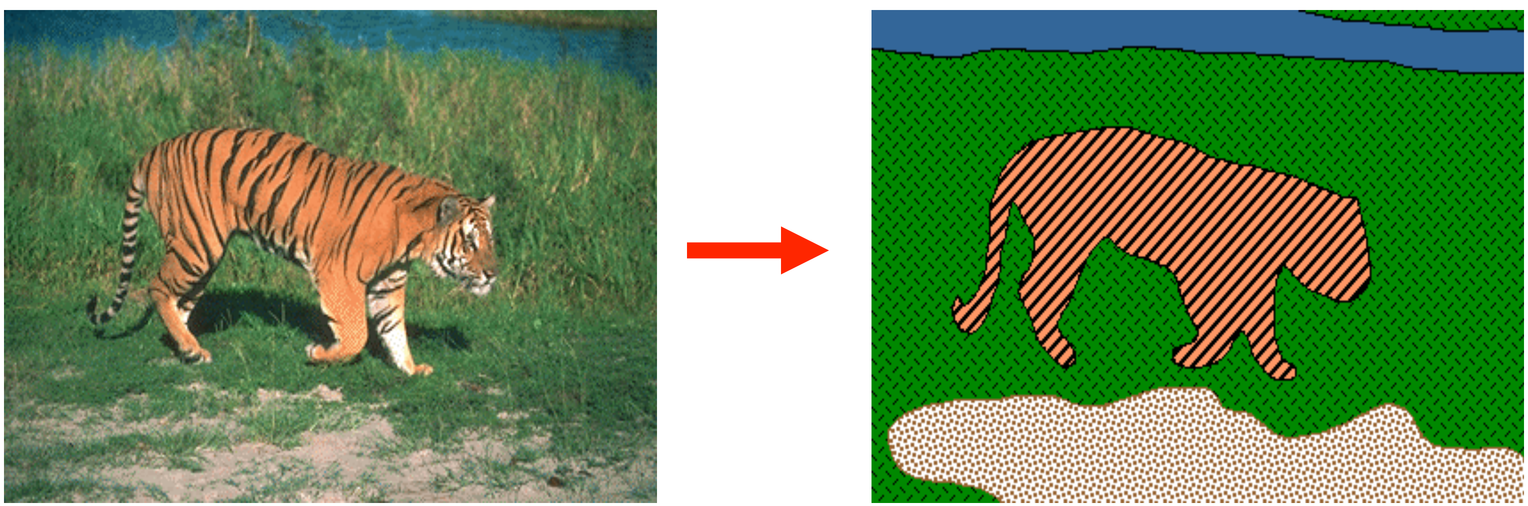
- Tokens
  - whatever we need to group (pixels, points, surface elements, etc., etc.)
- Bottom up segmentation
  - tokens belong together because they are locally coherent
- Top down segmentation
  - tokens belong together because they lie on the same visual entity (object, scene…)

> These two are not mutually exclusive

Source: S. Savarese slides.

What is Segmentation?

- **Grouping image elements that "belong together"**
  - **Partitioning**
    - Divide into regions/sequences with coherent internal properties
  - **Grouping**
    - Identify sets of coherent tokens in image

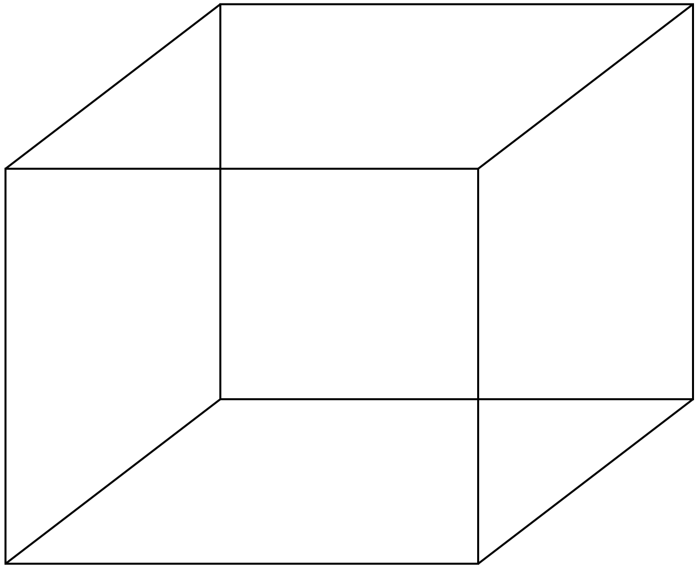


Source: S. Savarese, C. Rasmussen, S. Seitz slides.

What makes a good spatial segmentation method?

- Rationale for oversegmentation
  - Pixels are not natural elements in images.
  - The number of pixels is very high.
- **Spatial uniformity** – prefers compact and uniformly shaped superpixels.
  - Embeds basic Gestalt principles of continuity, closure, etc.
- **Spatial boundary preservation –** as superpixel boundaries should align with perceptual boundaries when present and should be stable when they are not.
- **Computation –** the overall computational cost for a particular application should be reduced via superpixels.
- **Performance –** the overall performance of a method should be increased.
- **Parsimony –** The above properties should be maintained with as few superpixels as possible.

Gestalt Principles of Visual Perception



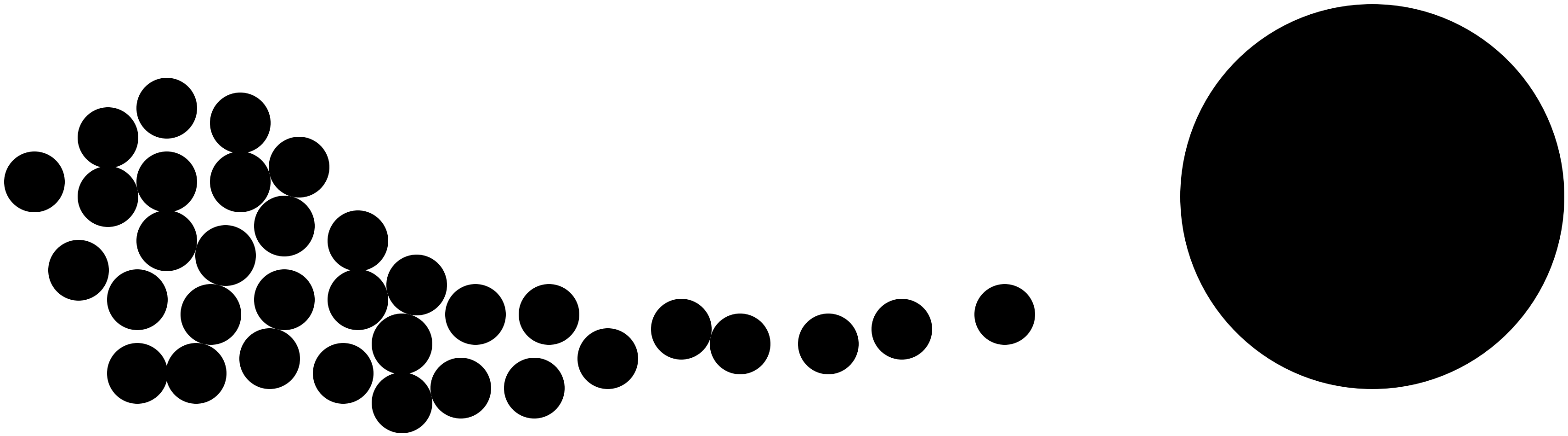
We organize pieces into patterns,

construct wholes out of parts,

and find meaning where there was none before...

Source for this section: C. Cumbie-Jones (http://webspace.ringling.edu/~ccjones/curricula/07-08/seqdesign/Gestalt.ppt) and
B. Schrank (http://lmc.gatech.edu/~bschrank/2720/)
In this section, specific sources are not given per slide since this may impact the visual gestalt of the slide!

What is a Gestalt?



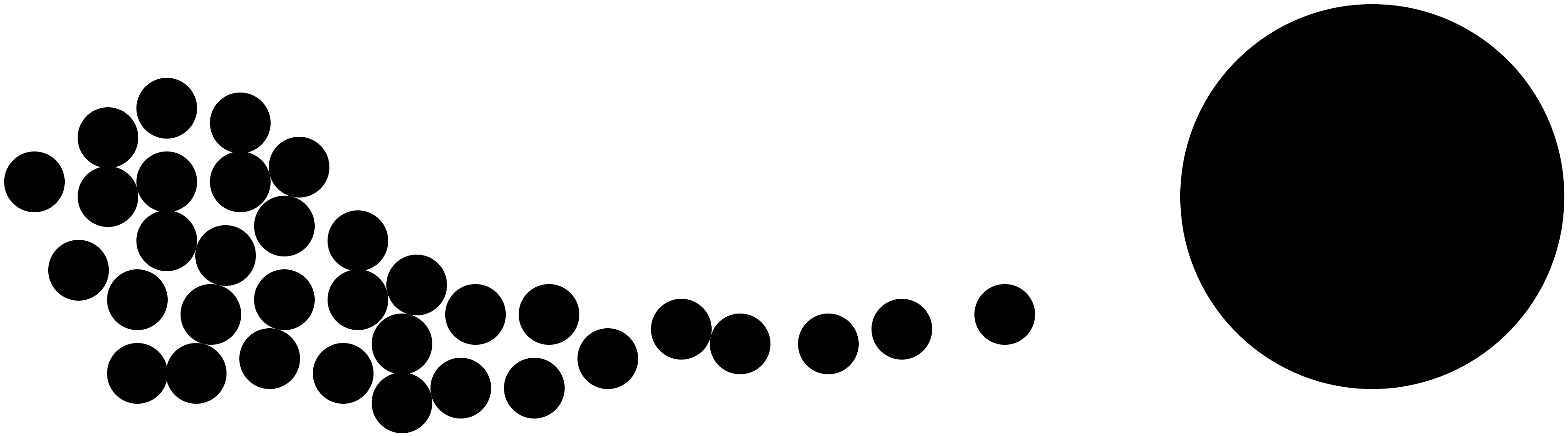
'Gestalt' means 'pattern' in German.

A gestalt is a configuration, pattern, or organized field having specific properties that cannot be derived from the summation of its component parts.

A gestalt is a unified whole.

Source: B. Schrank slides.

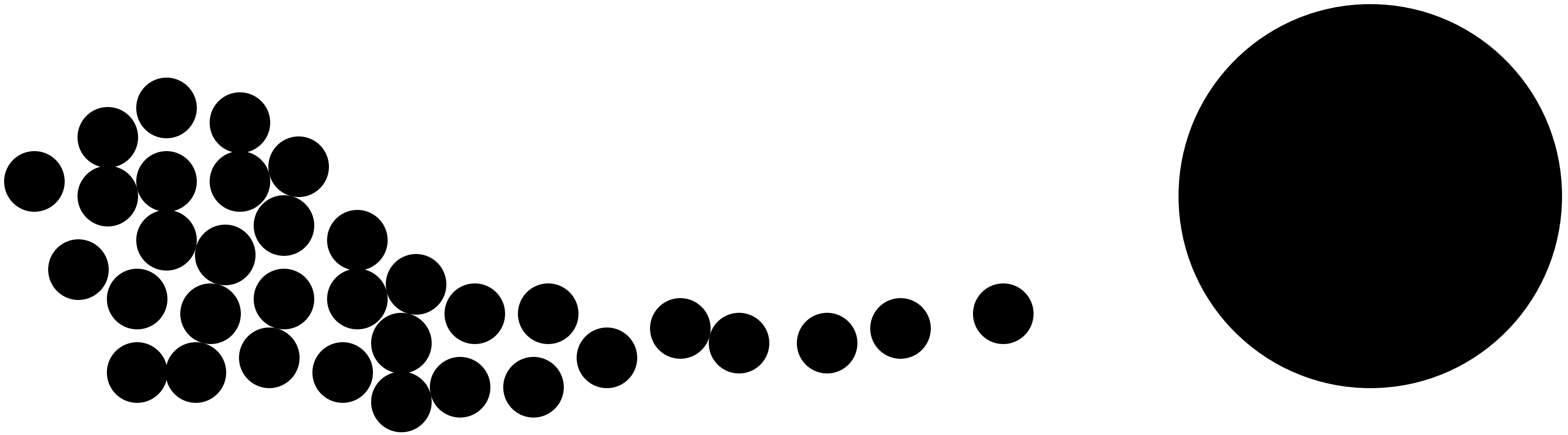
What is Gestalt Psychology?



Gestalt Psychology is the theory or doctrine that physiological or psychological phenomena do not occur through the summation of individual elements, as reflexes or sensations, but through gestalts functioning separately or interrelatedly.

Source: B. Schrank slides.

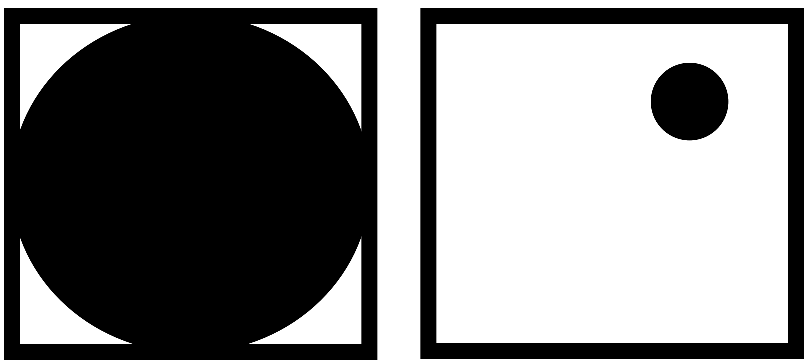
What is Gestalt Psychology?

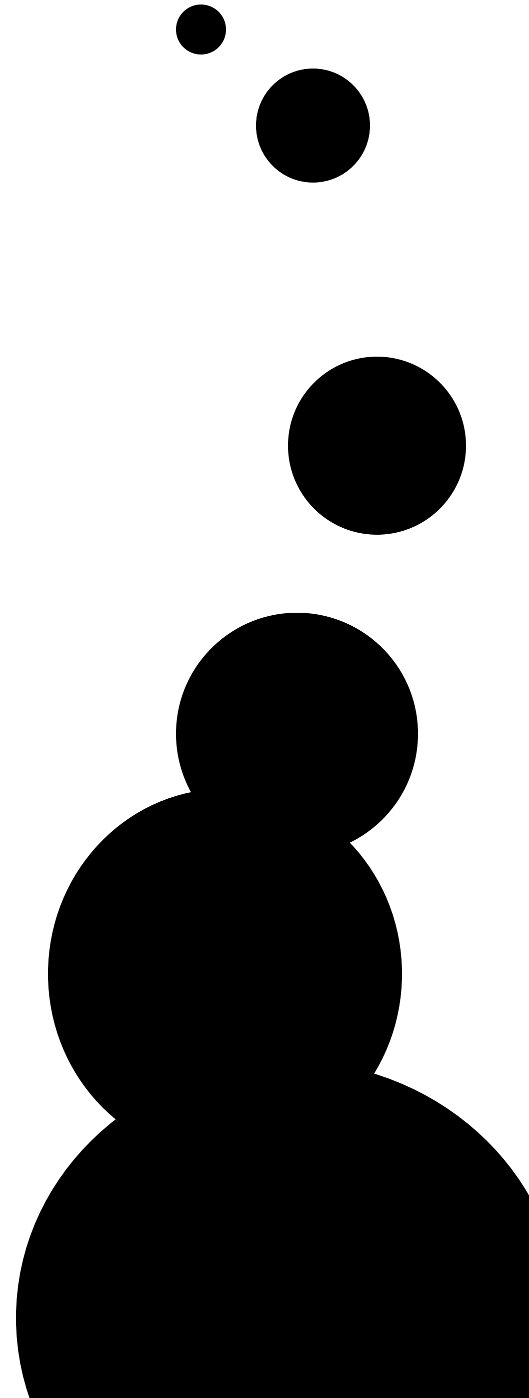


What is your gestalt of the images above? What is the meaning beyond random circles?

Source: B. Schrank slides.

What is Gestalt Psychology?





Although we may not be aware of it consciously, because we tend to relate what we see to our own bodily reactions to situations in space, shapes appear to fall or be pulled by gravitational forces, appear to lean over, to fly, to move fast or slow, to be trapped or be free.

-Sausmarez

Source: B. Schrank slides.

Gestalt Principles of Visual Perception

We impose visual organization on stimuli

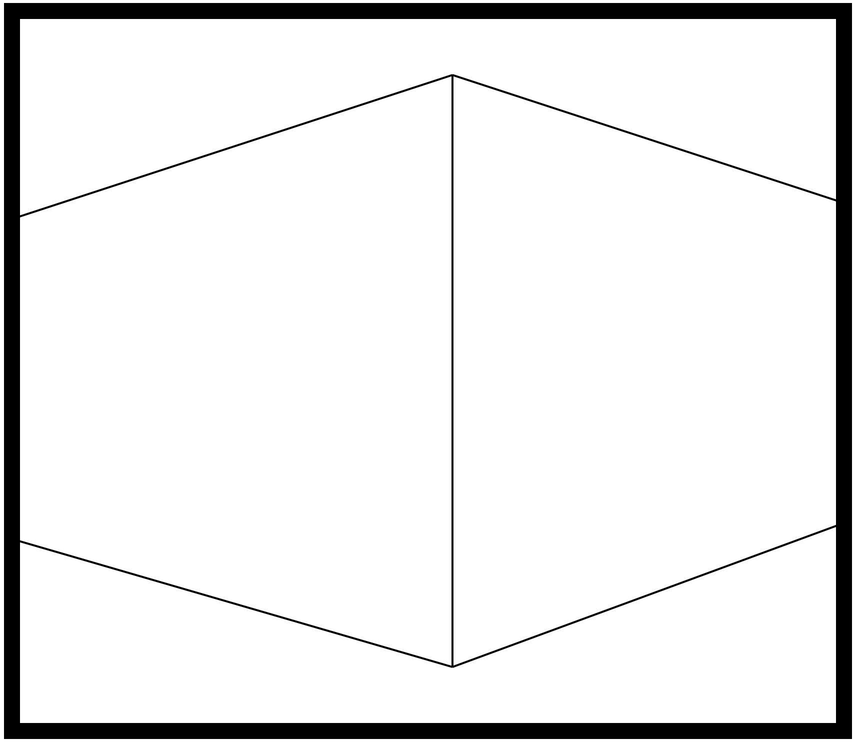
W.E. Hill, 1915

German postcard, 1880

Source: C Cumbie-Jones slides.

Gestalts are Constructed from Nature and Nurture

Architecture and our rectangular world has had a dramatic Influence on our Interpretation of Lines.



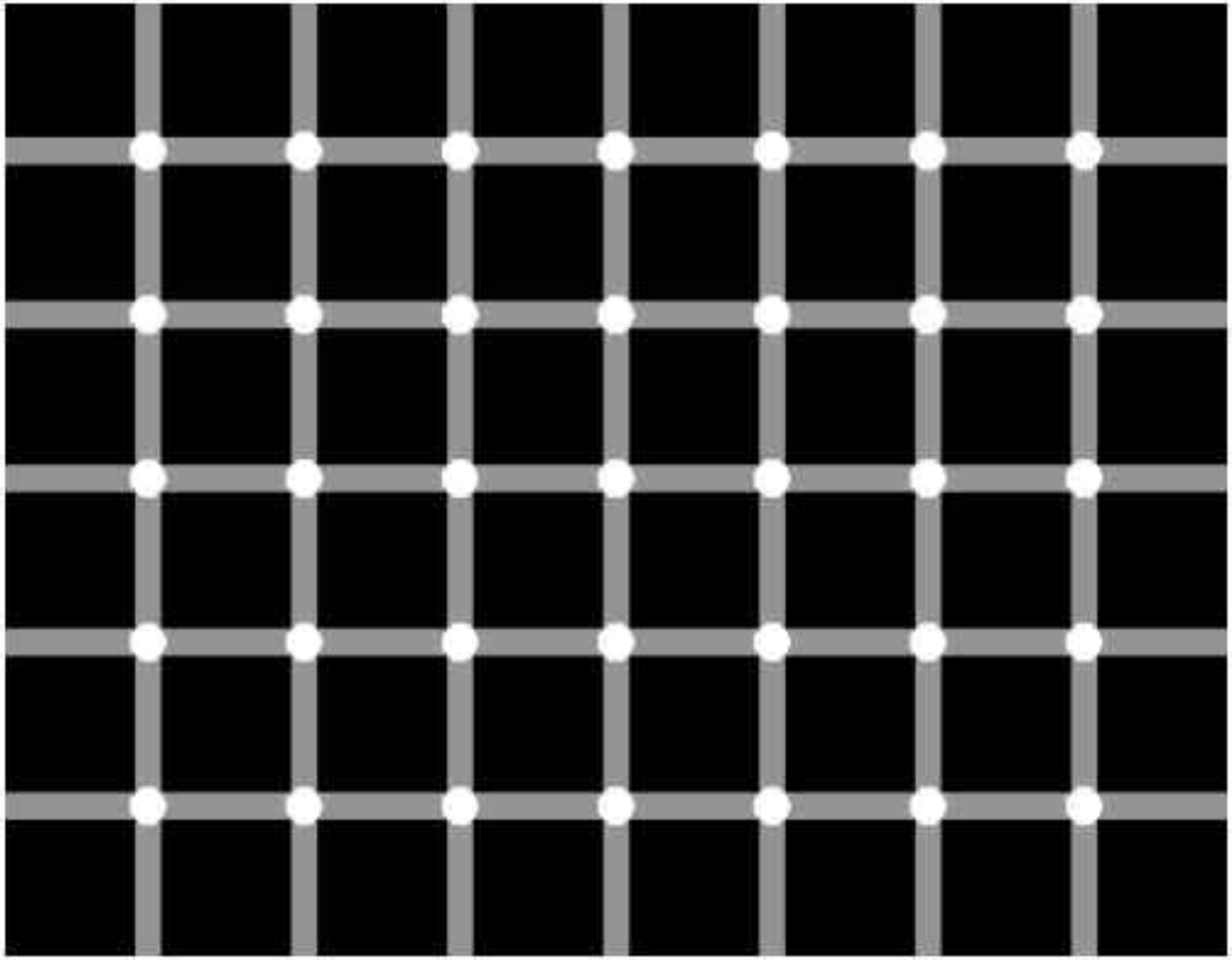
Source: B. Schrank slides.

Gestalts are Constructed from Nature and Nurture

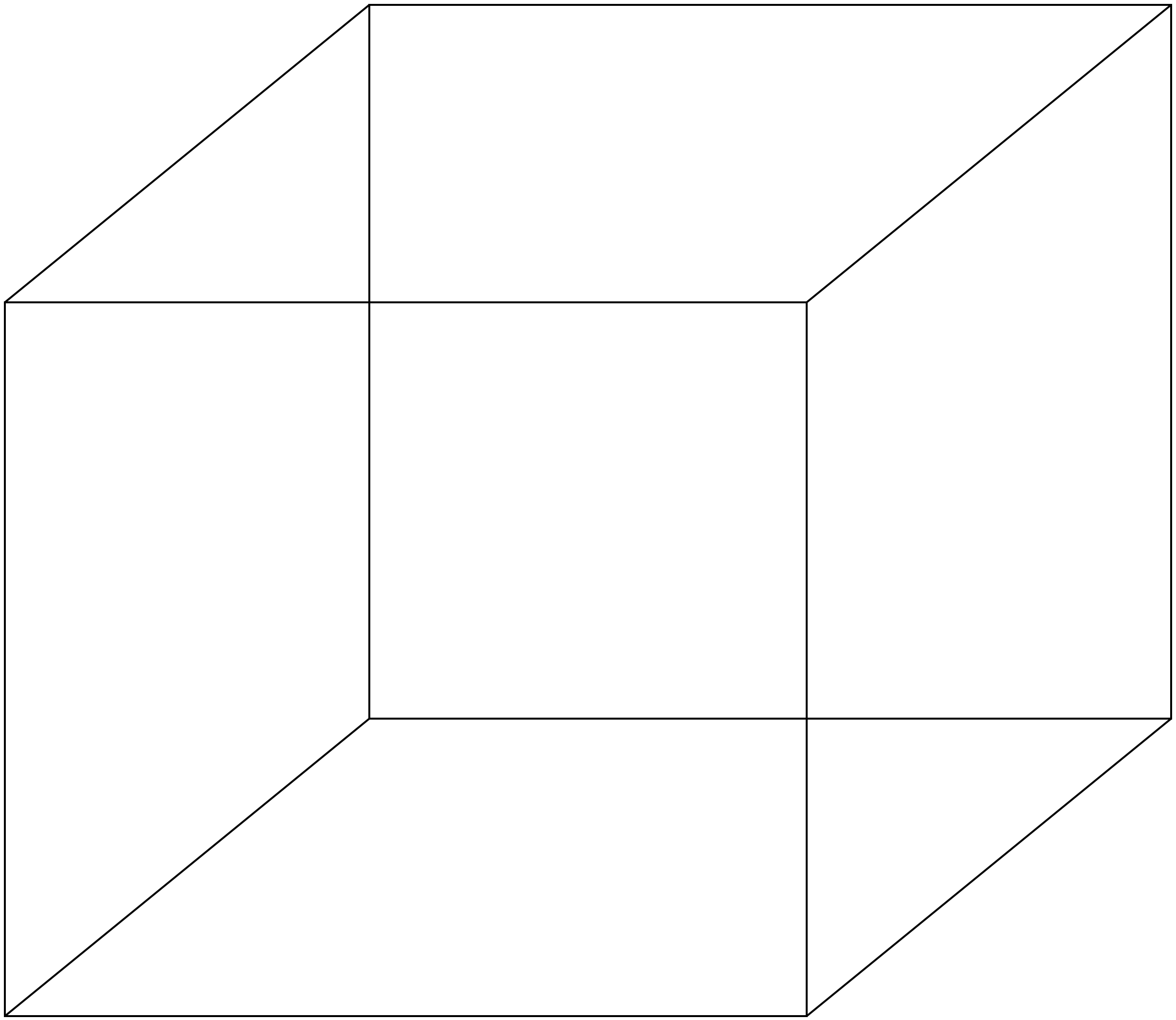
Even more physically wired Gestalts are prevalent, such as how we tend to naturally 'fill in' lacunas...

Source: B. Schrank slides.

Gestalts are Constructed from Nature and Nurture

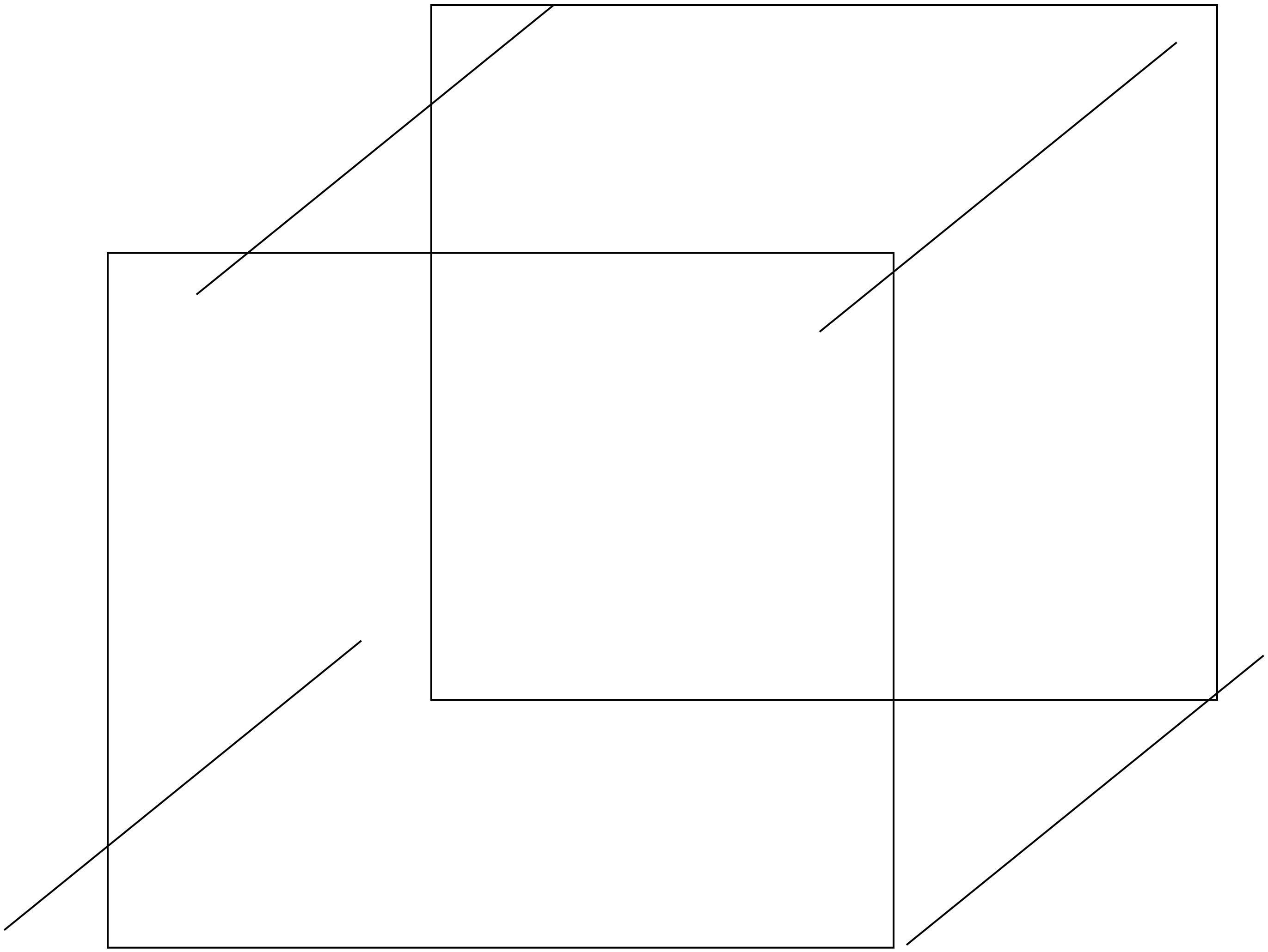


Source: B. Schrank slides.



Gestalt is also subtle...

Source: B. Schrank slides.



Do you feel the quiet desire for the cube to be complete and neat?

Source: B. Schrank slides.

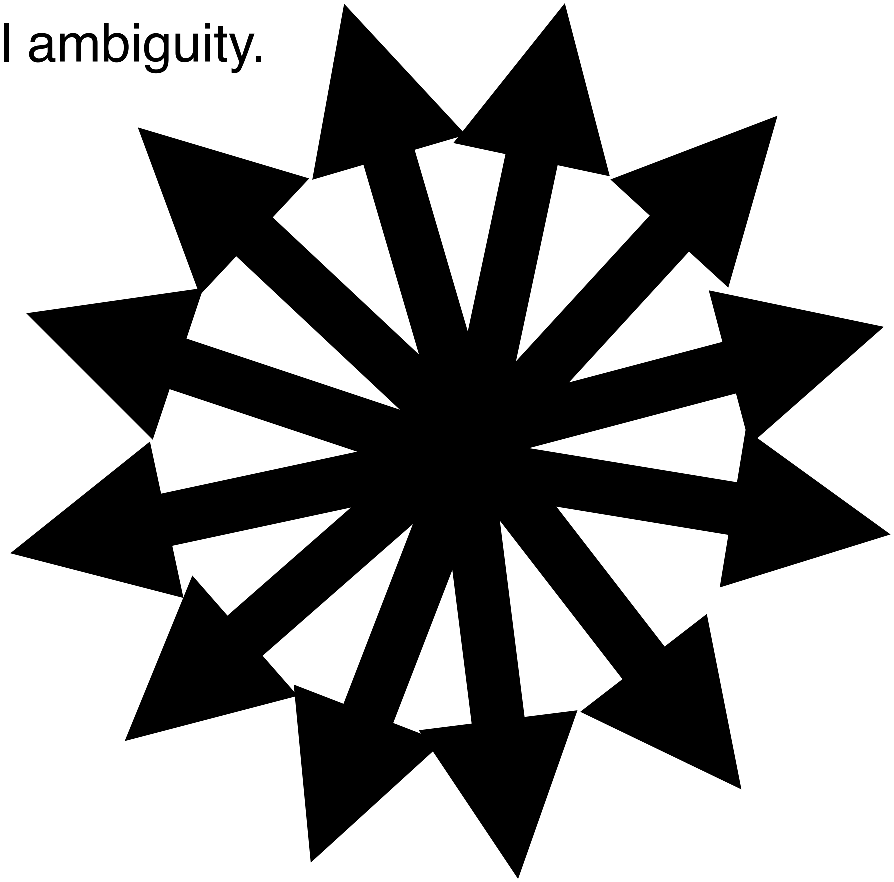
Some examples of Visual Gestalt

- Equivocation
- Continuance
- Closure
- Common Fate
- Constancy
- Similarity
- Proximity

Source: B. Schrank slides.

Equivocation

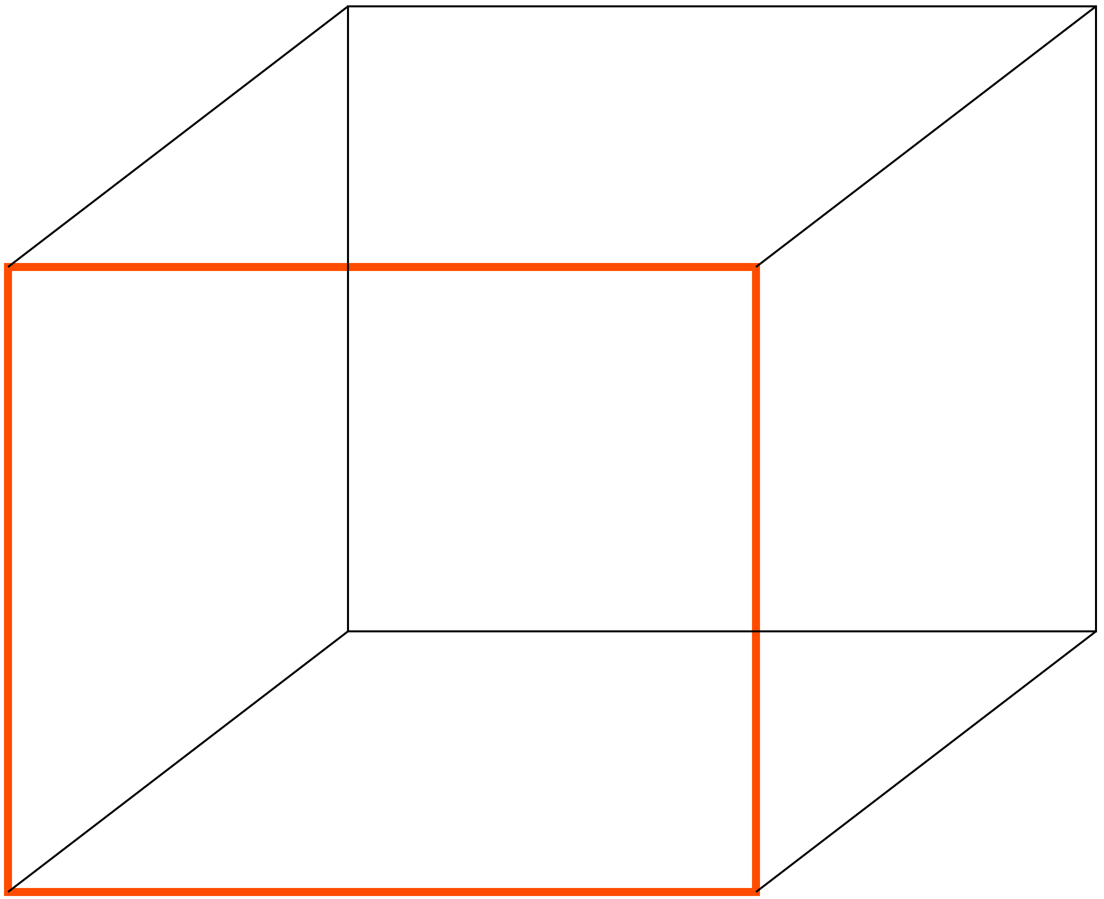
Equivocation is perceptual ambiguity.



For example,
Do you see the parts?
(The radial of arrows)

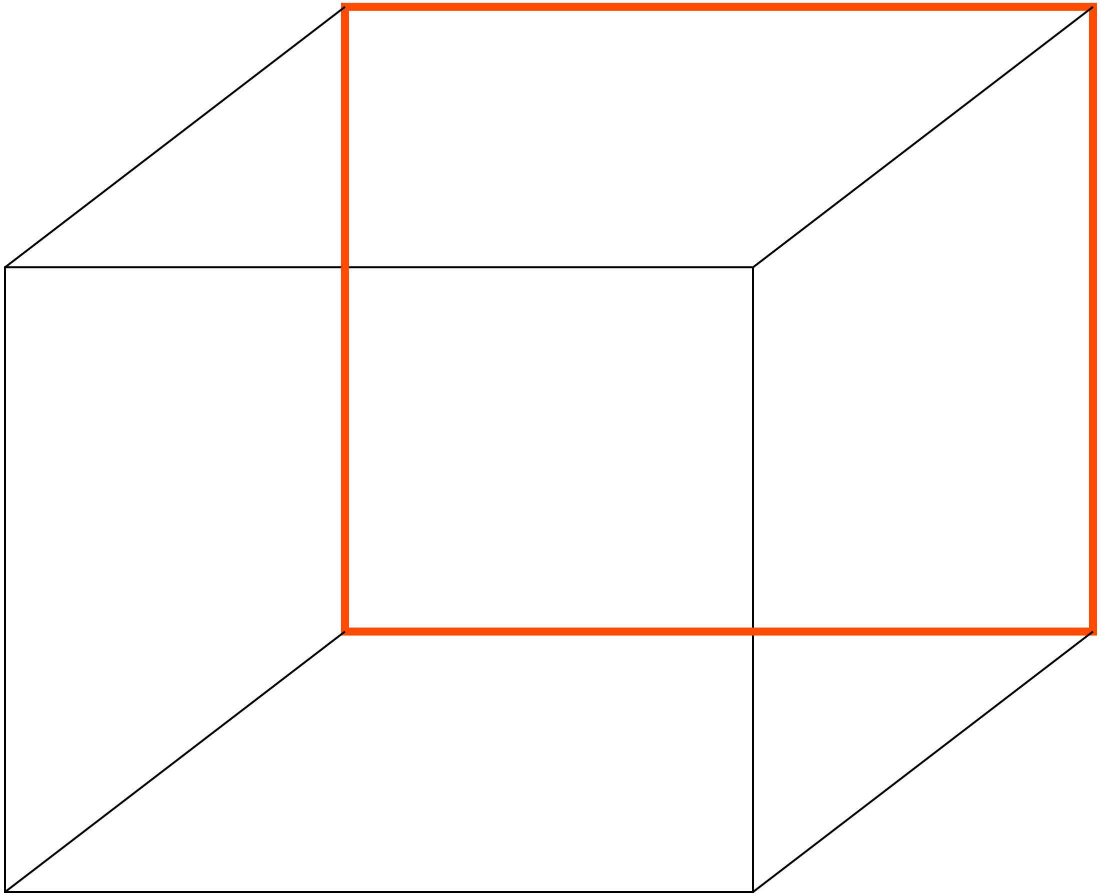
Or the whole?
(The spiked wheel or sun)

Source: B. Schrank slides.



Equivocation in the Necker Cube oscillates the closest plane between the two planes facing the viewer.

Source: B. Schrank slides.



Equivocation in the Necker Cube oscillates the closest plane between the two planes facing the viewer.

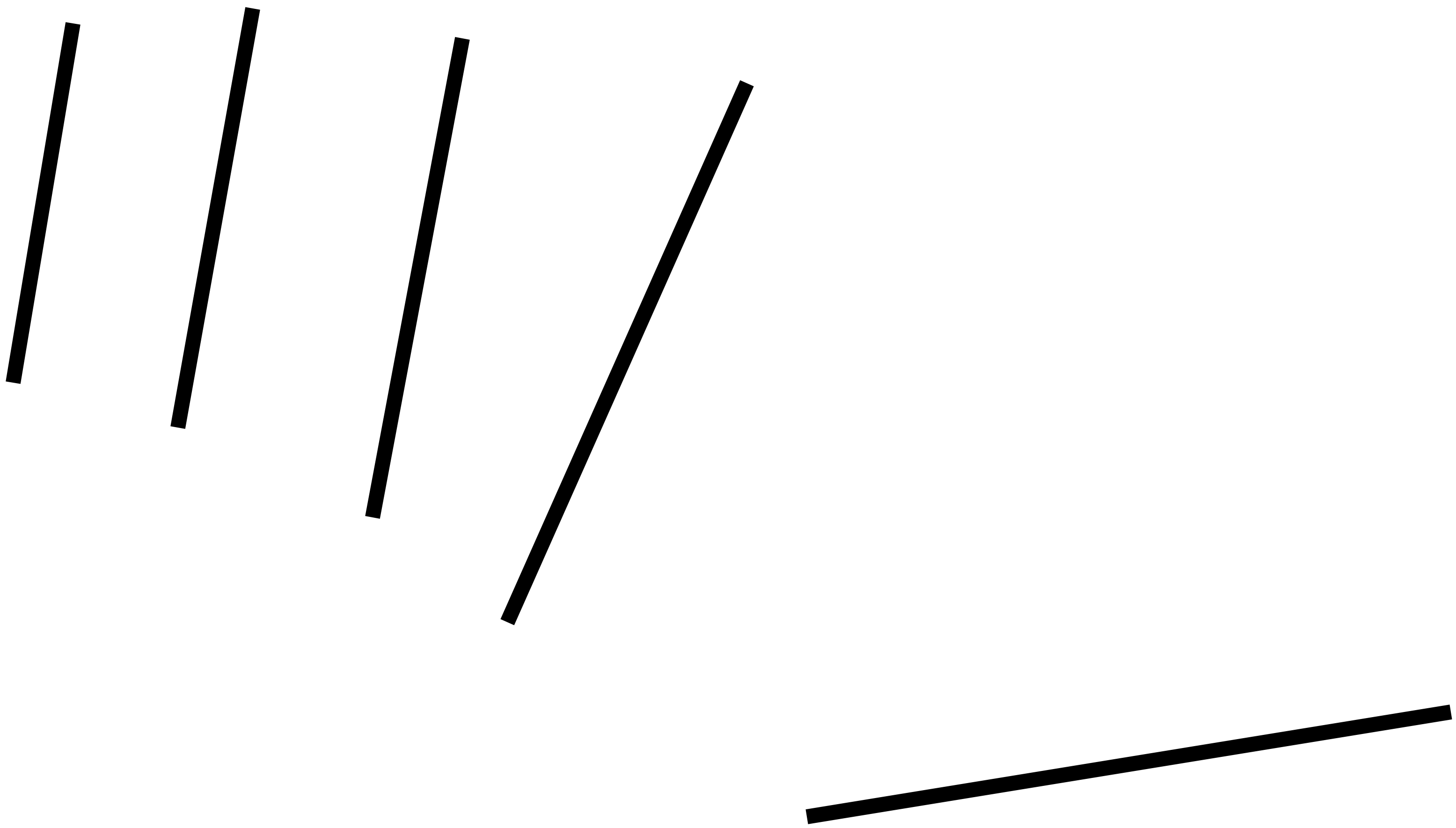
Source: B. Schrank slides.

Continuance

We tend to connect similar phenomena, psychologically constructing a timeline through them as a sequence...

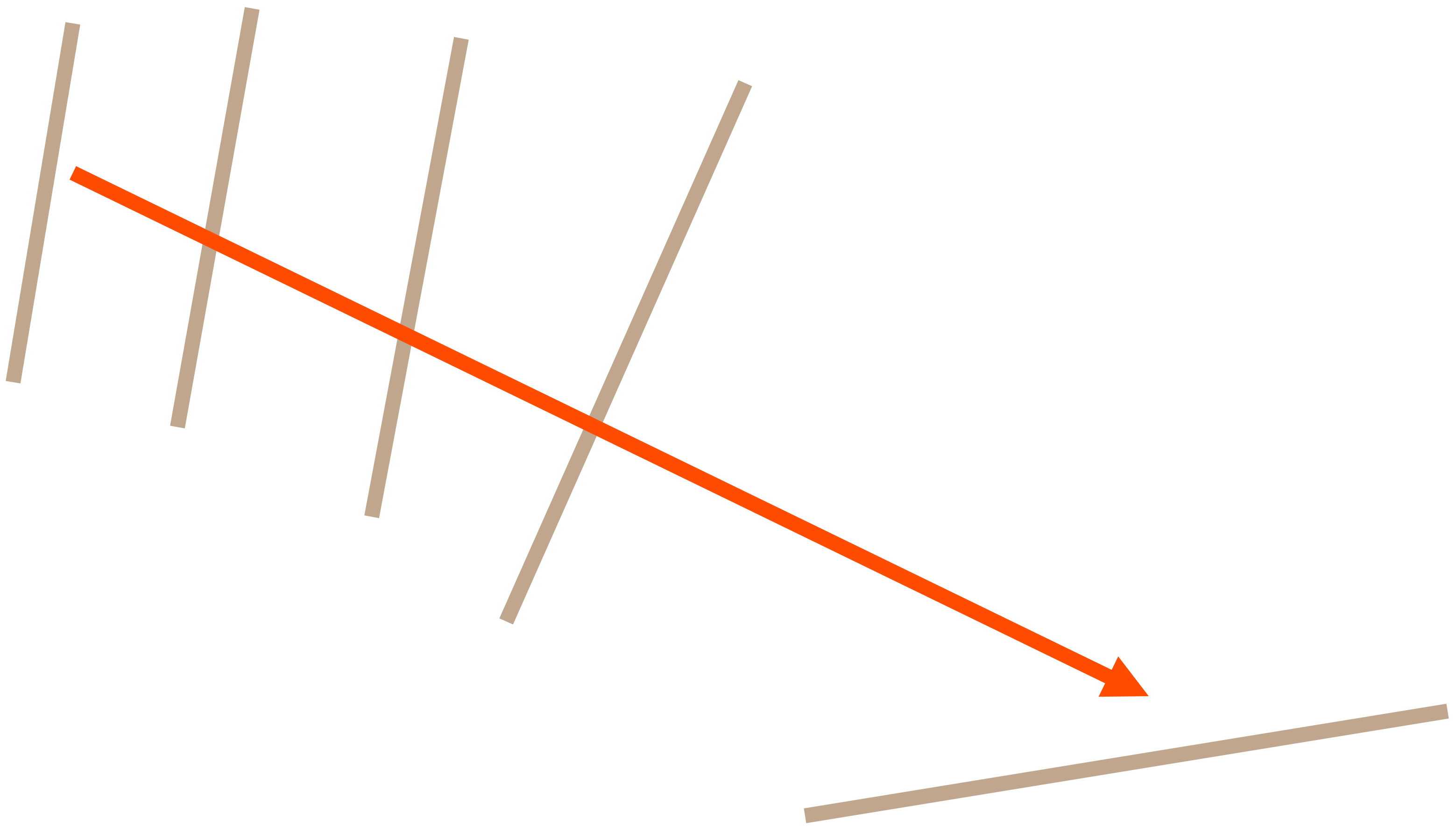
Source: B. Schrank slides.

Continuance



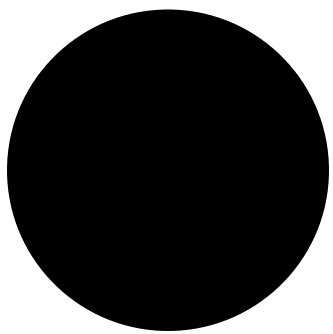
Source: B. Schrank slides.

Continuance



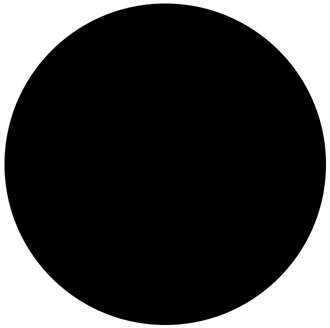
Source: B. Schrank slides.

Continuance... Is it the same circle?



Source: B. Schrank slides.

Continuance... Is it the same circle?



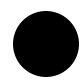
Source: B. Schrank slides.

Continuance…

●

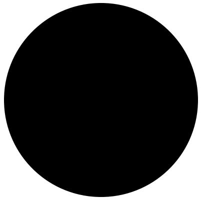
Source: B. Schrank slides.

Continuance… Is that circle approaching us?



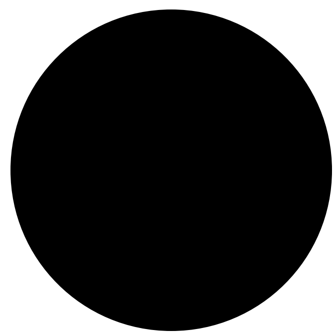
Source: B. Schrank slides.

Continuance… Is that circle approaching us?



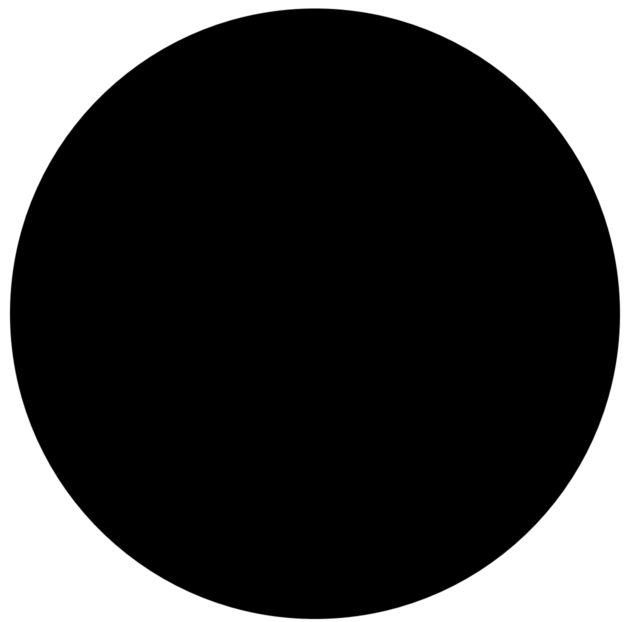
Source: B. Schrank slides.

Continuance… Is that circle approaching us?



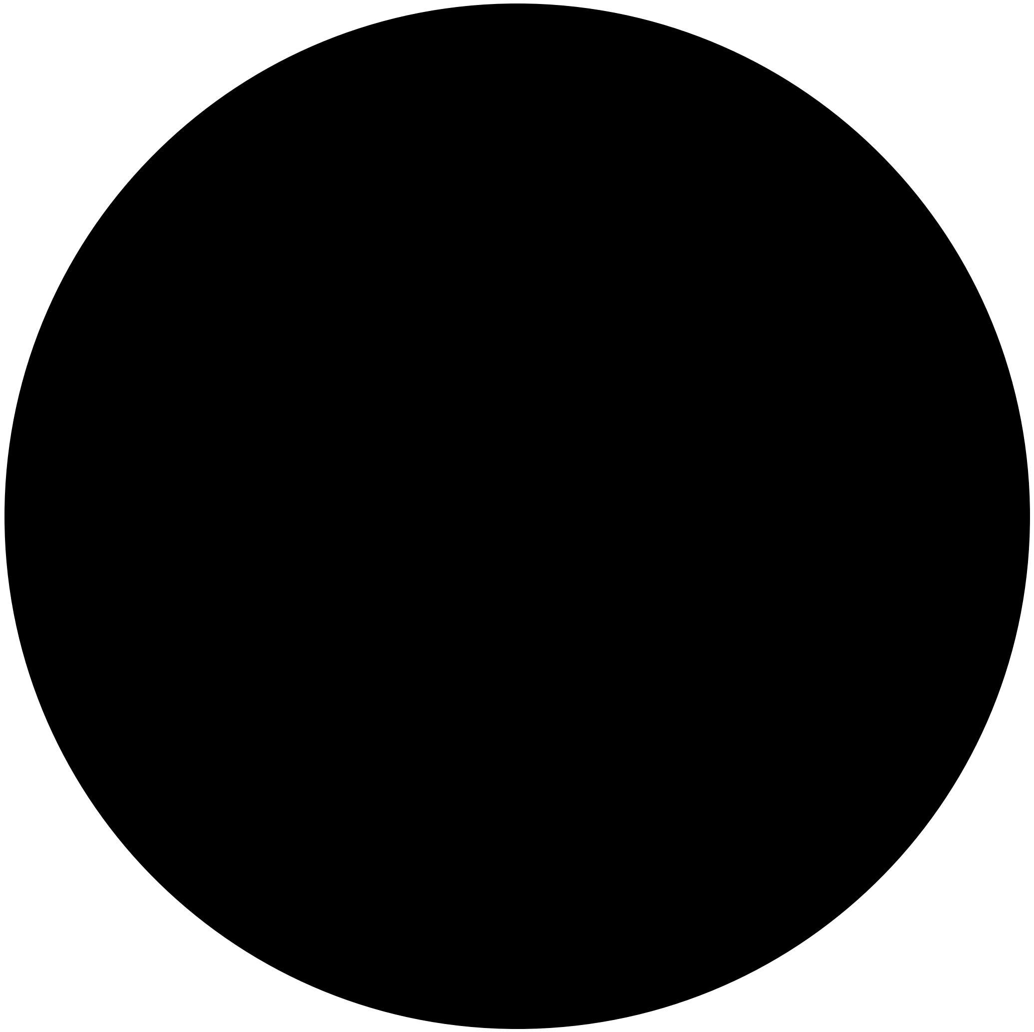
Source: B. Schrank slides.

Continuance… Is that circle approaching us?



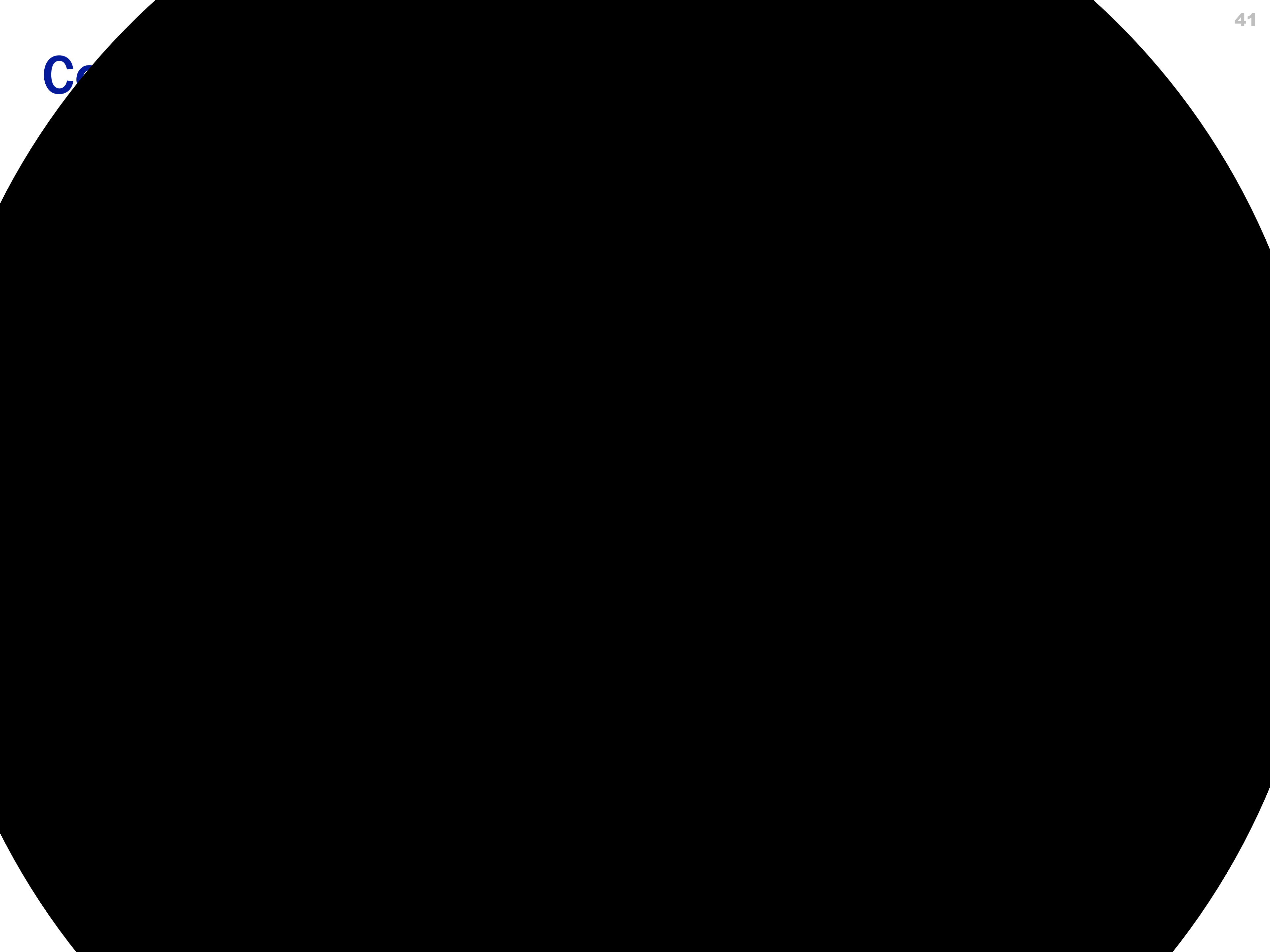
Source: B. Schrank slides.

Continuance… Is that circle approaching us?

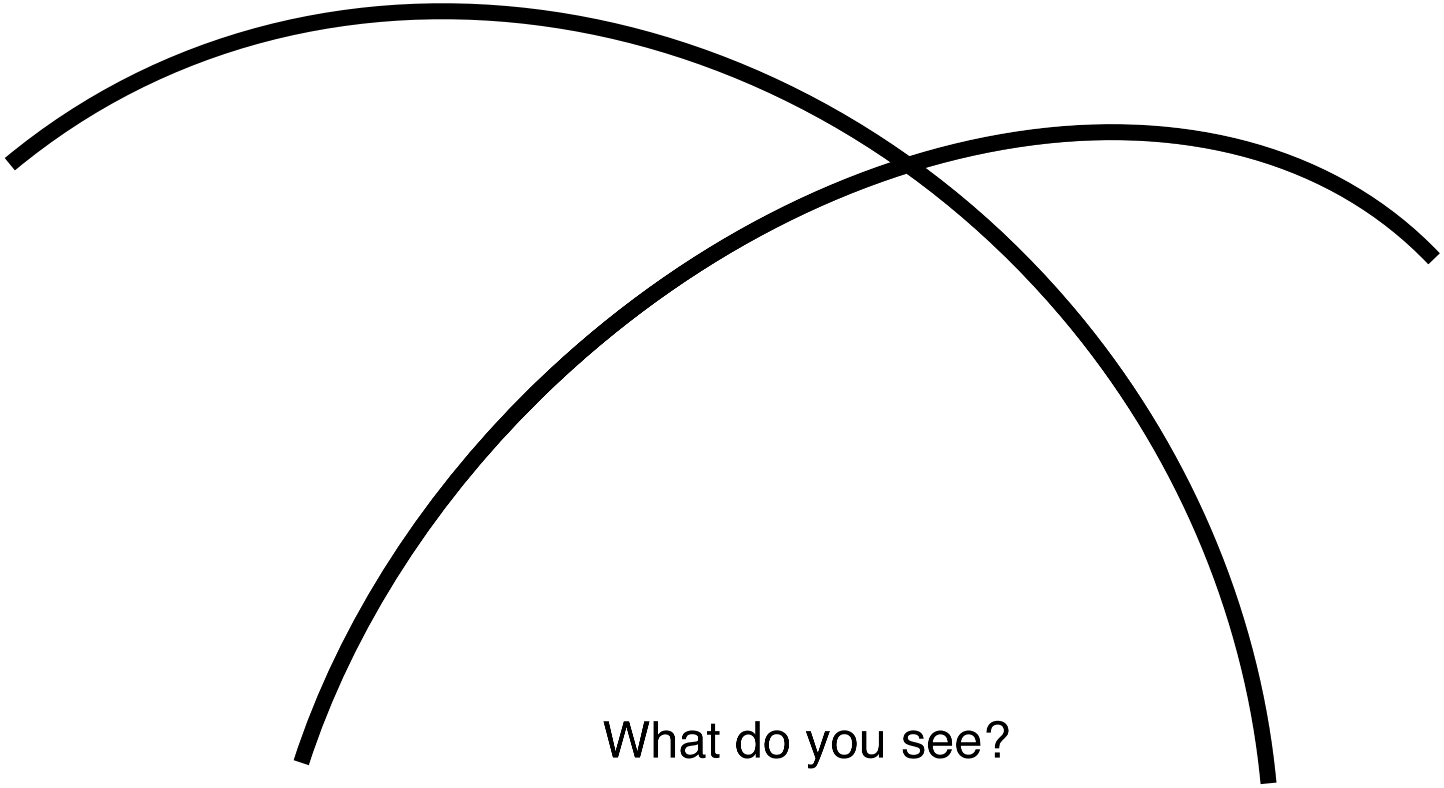


Source: B. Schrank slides.

Co



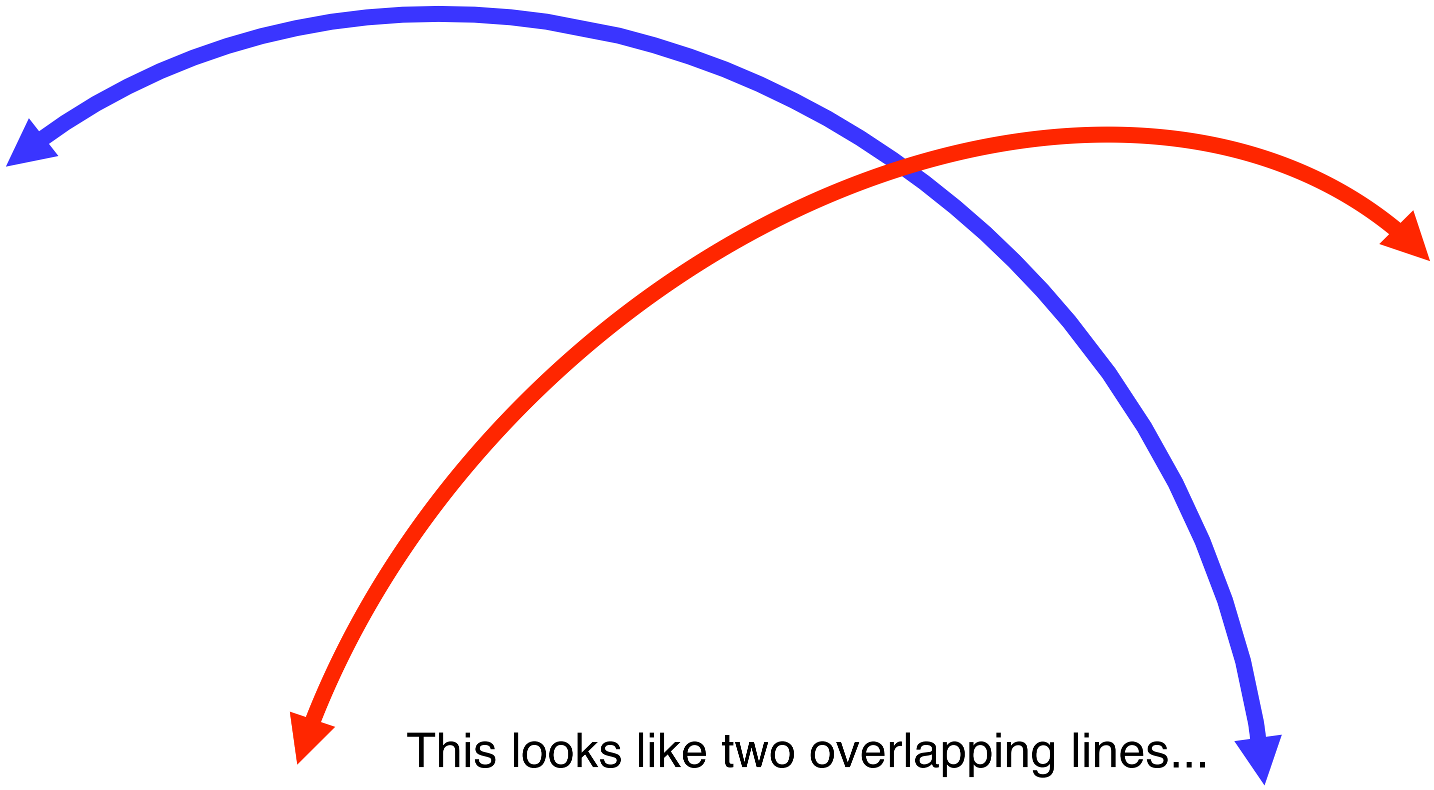
Continuance (cont'd)



What do you see?

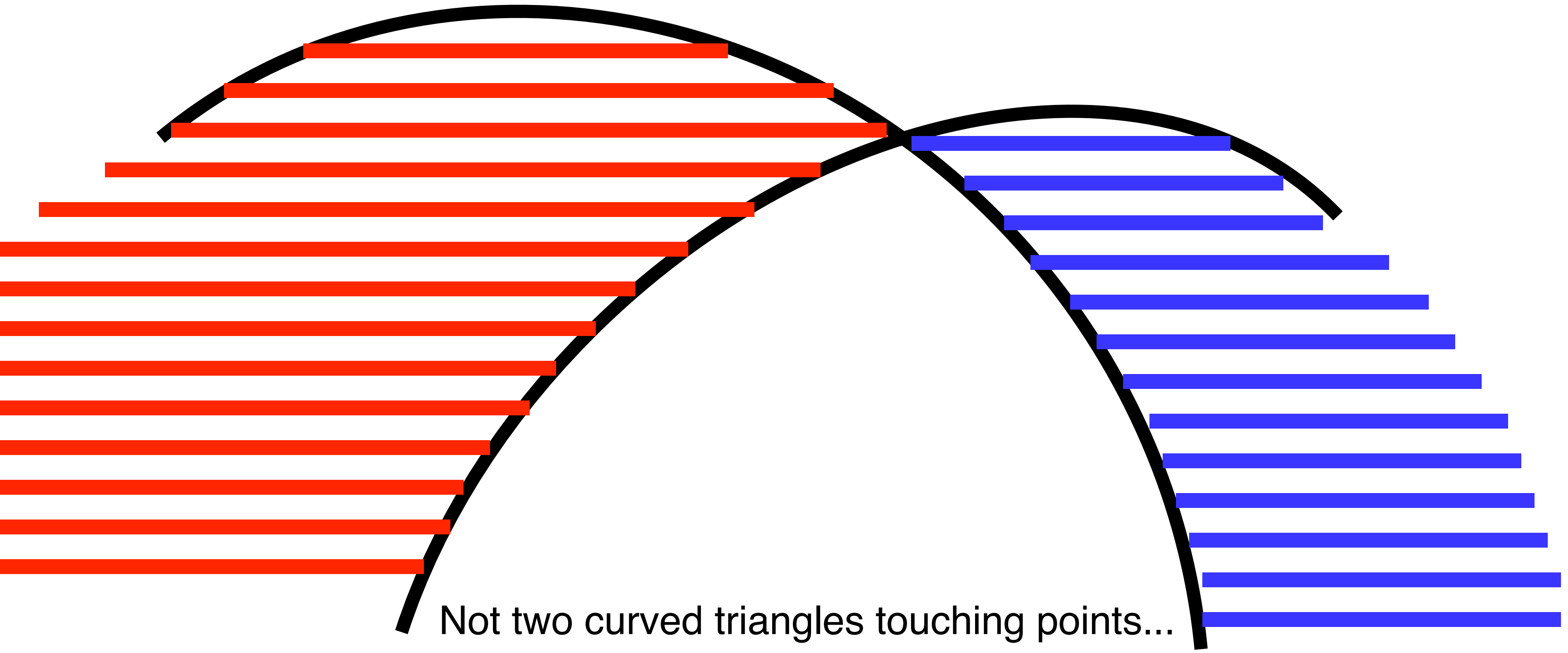
Source: B. Schrank slides.

Continuance of Line



This looks like two overlapping lines...

Source: B. Schrank slides.

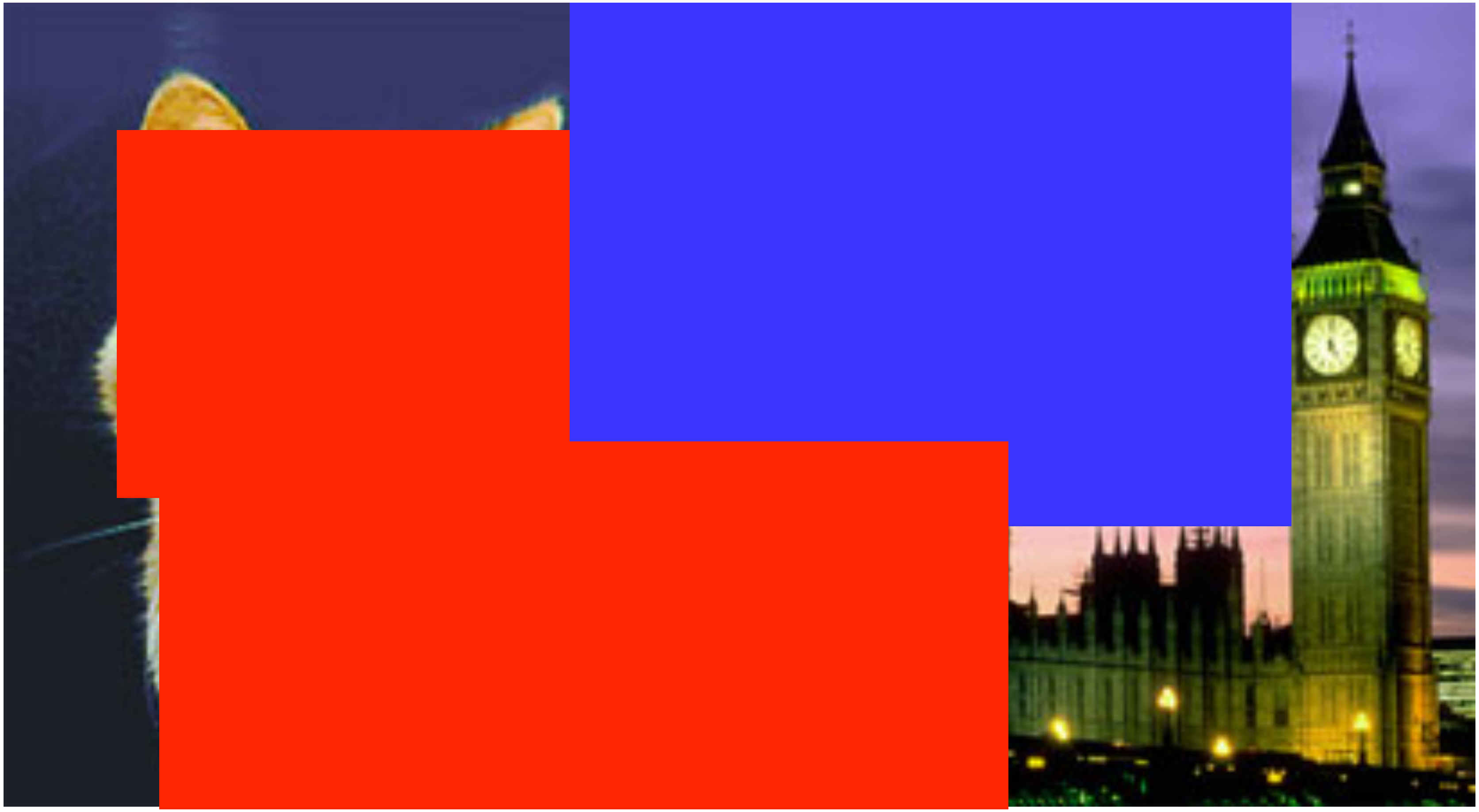


Source: B. Schrank slides.

Beware of Unintended Continuance

Source: B. Schrank slides.

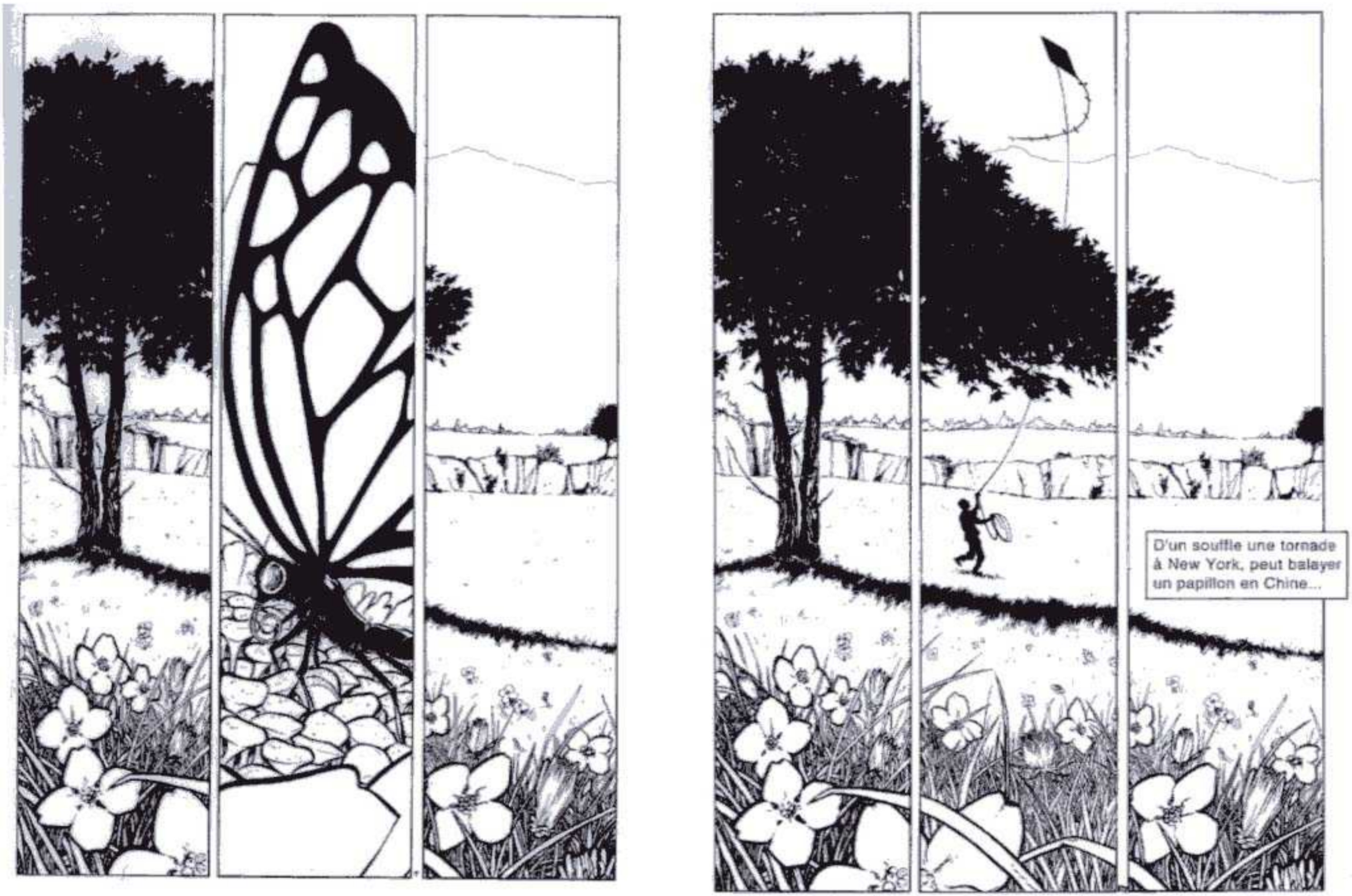
Beware of Unintended Continuance



Source: B. Schrank slides.

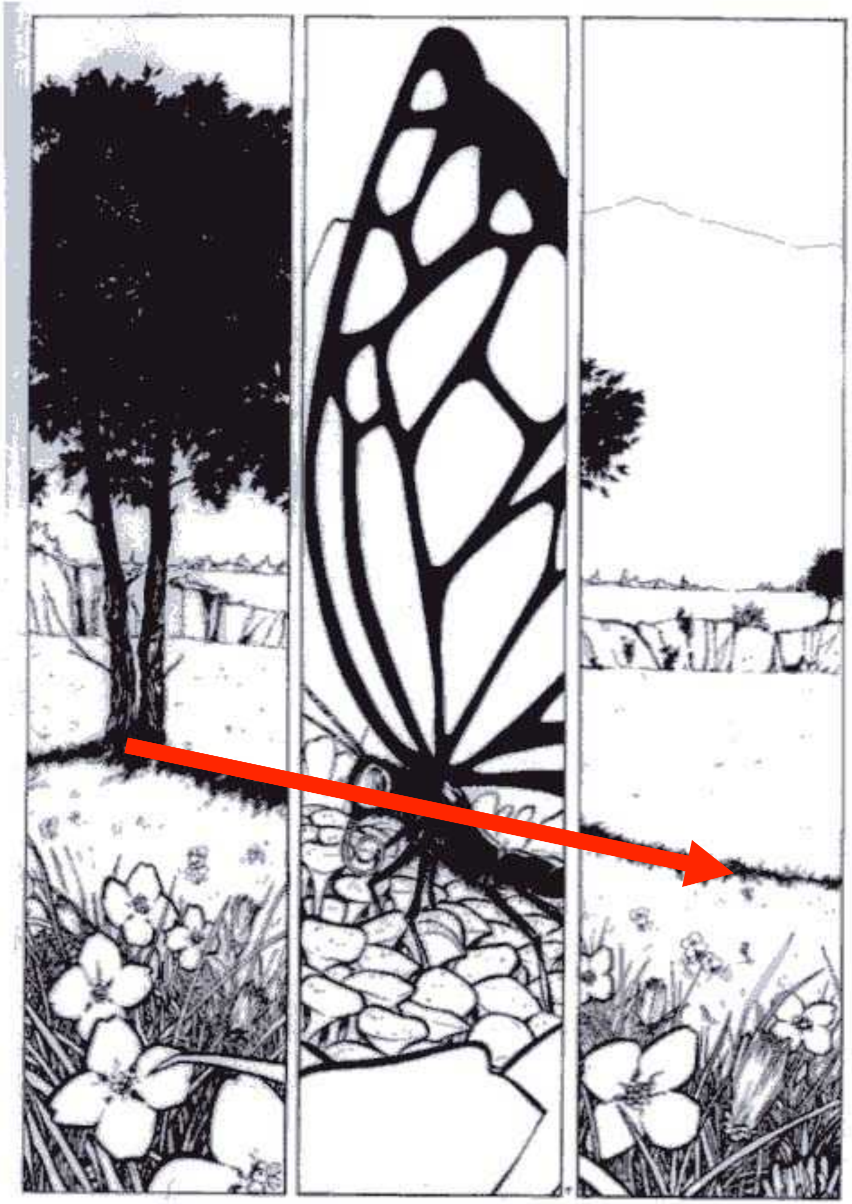
However, blending images through continuance can be beautiful...

Source: B. Schrank slides.



Schelle, *Théorie du chaos*, pp. 5 and 112

Source: B. Schrank slides.



Schelle, *Théorie du chaos*, pp. 5 and 112

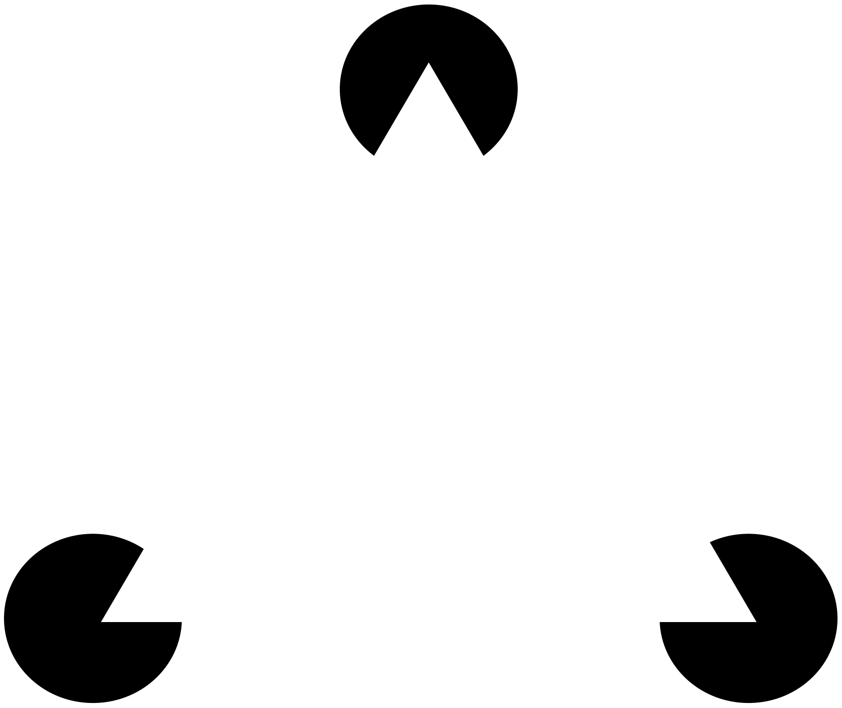
Source: B. Schrank slides.

Closure

Closure is the tendency to psychologically complete an incomplete picture or element.

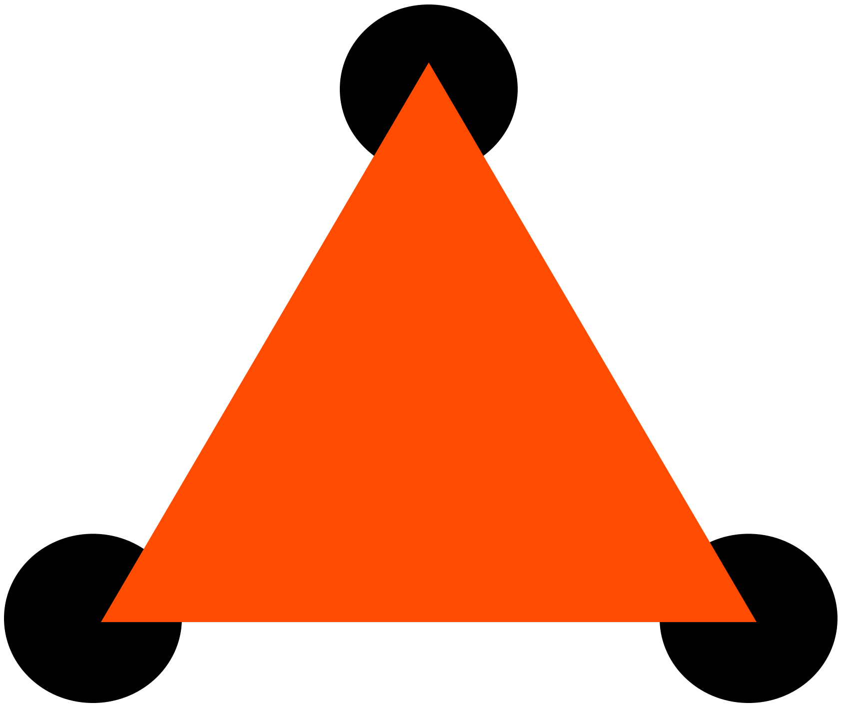
Source: B. Schrank slides.

Closure is most effective with recognizable shapes and images.



Source: B. Schrank slides.

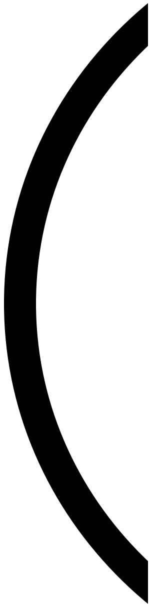
Closure is most effective with recognizable shapes and images.



The Kanisza triangle as figure-ground illusory contours

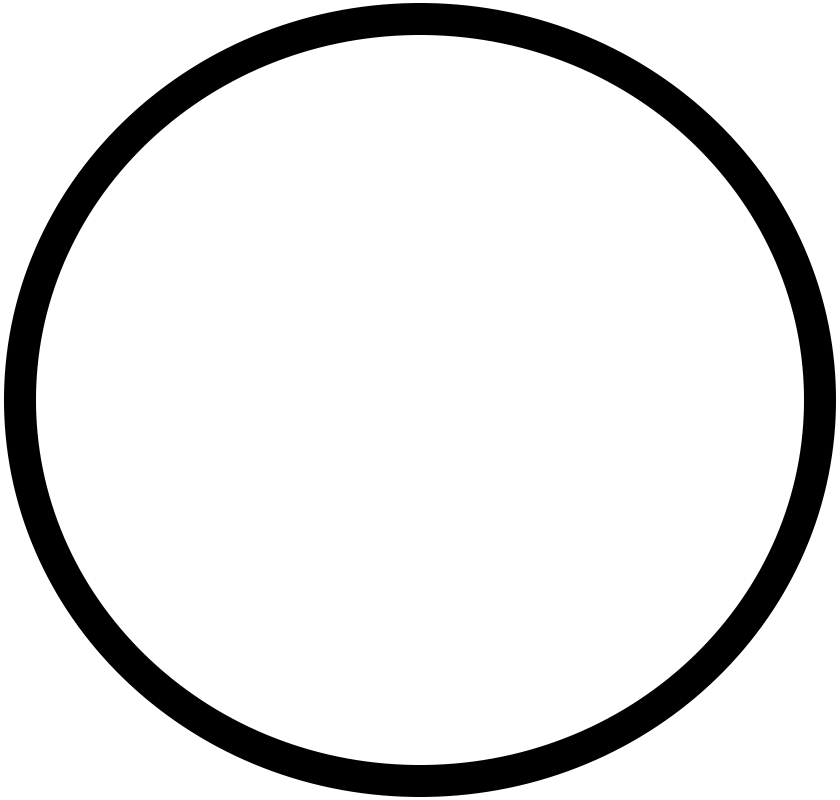
Source: B. Schrank slides.

Closure is most effective with recognizable shapes and images.



Source: B. Schrank slides.

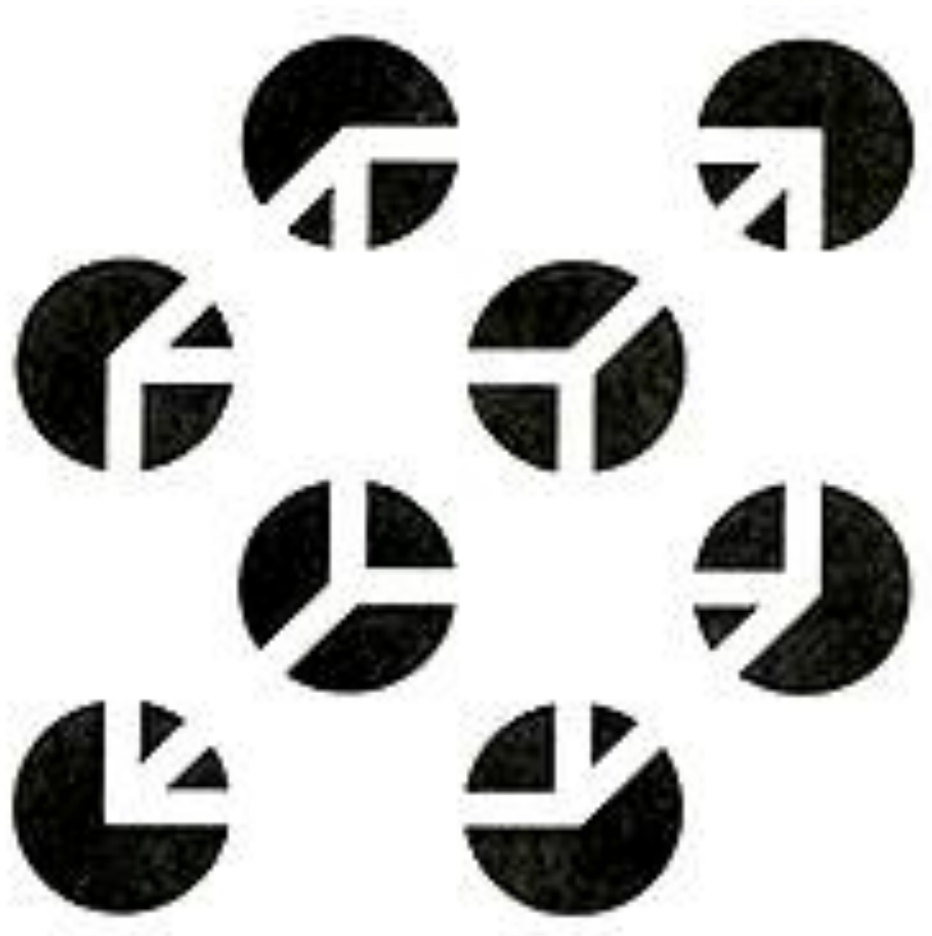
Closure is most effective with recognizable shapes and images.



Source: B. Schrank slides.

Gestalt Principles of Visual Perception

Law of Closure



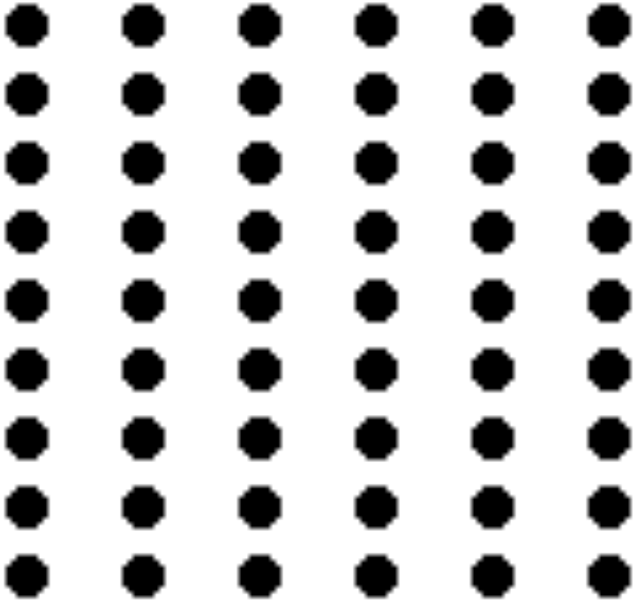
Source: C Cumbie-Jones slides.

Grouping

We ascribe a group relationship to elements in a visual field based on various attributes they have in common.

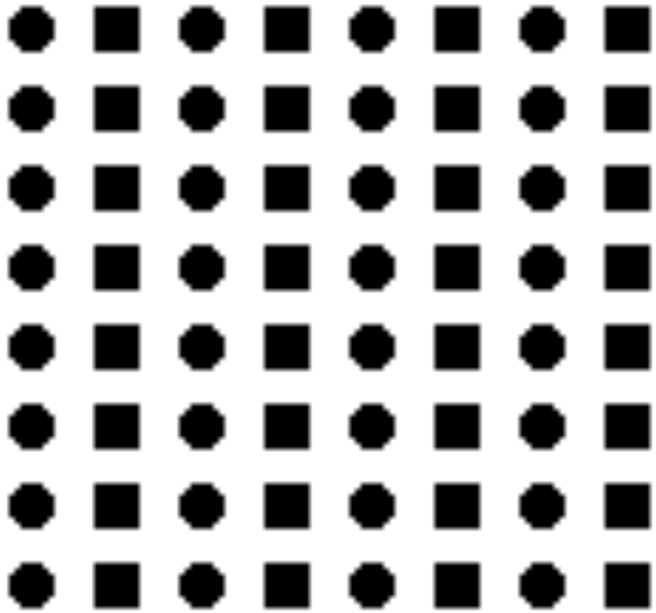
Source: B. Schrank slides.

Grouping through Proximity



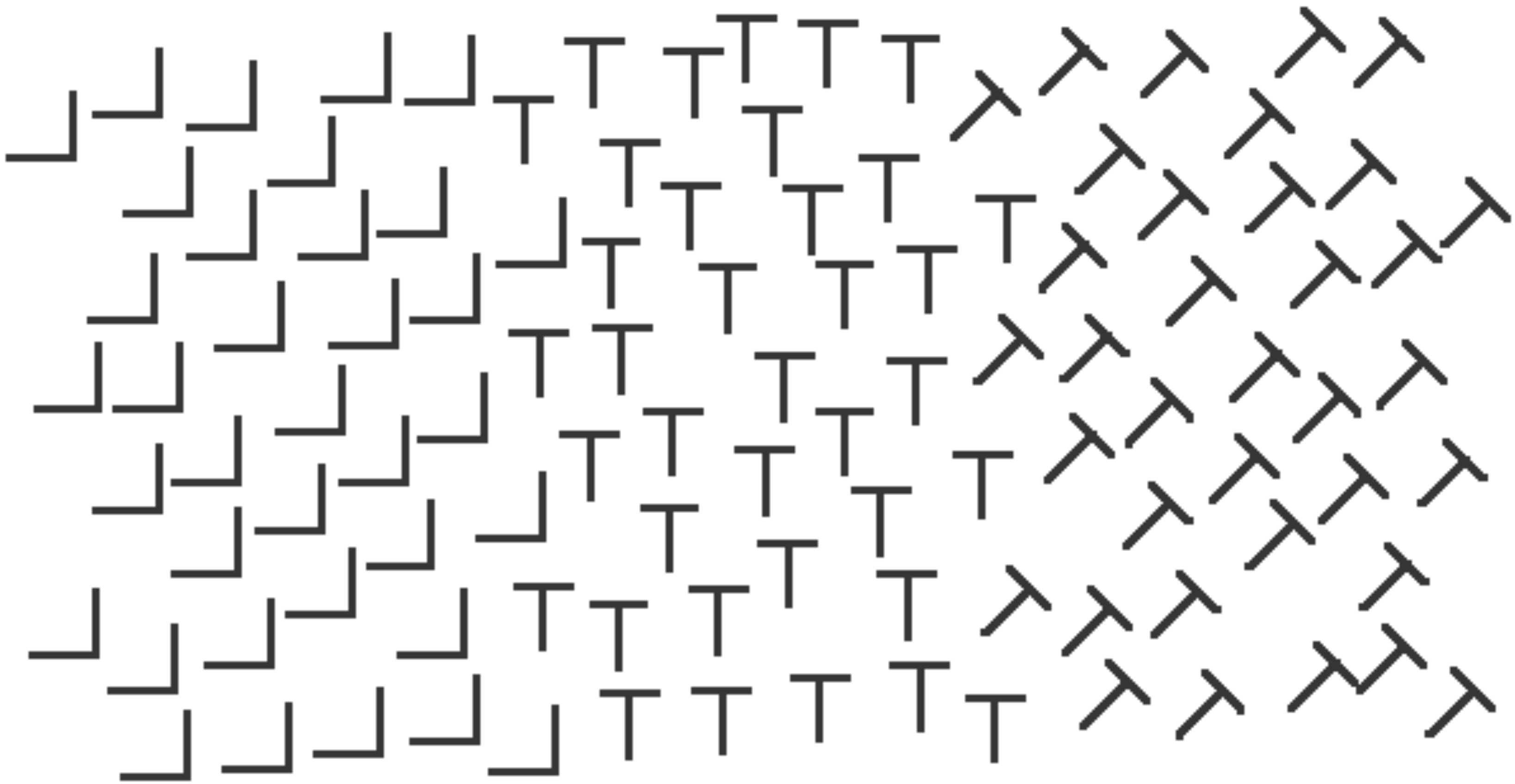
Source: C Cumbie-Jones slides.

Grouping through Similarity



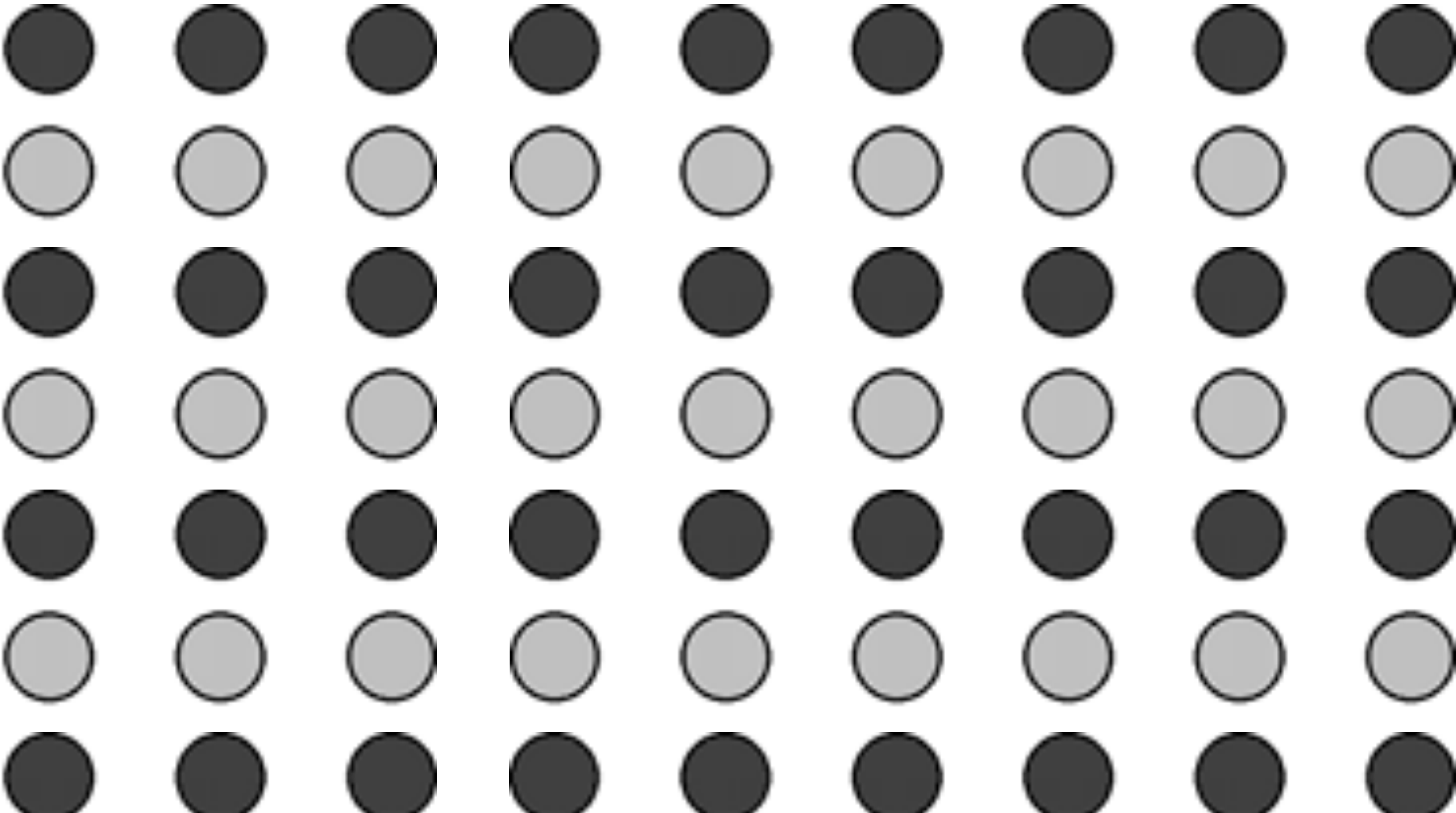
Source: C Cumbie-Jones slides.

Grouping through Orientation



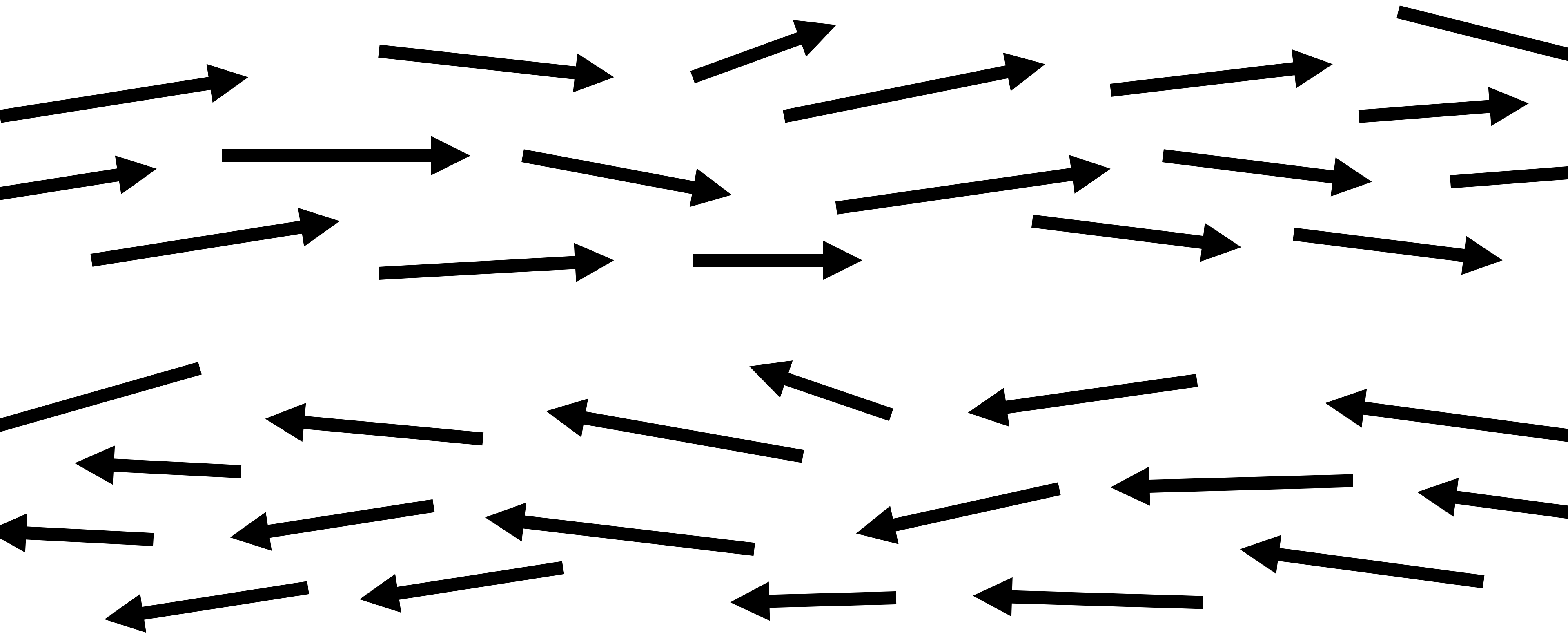
Source: B. Schrank slides.

Grouping by color can outweigh alignment:



Source: B. Schrank slides.

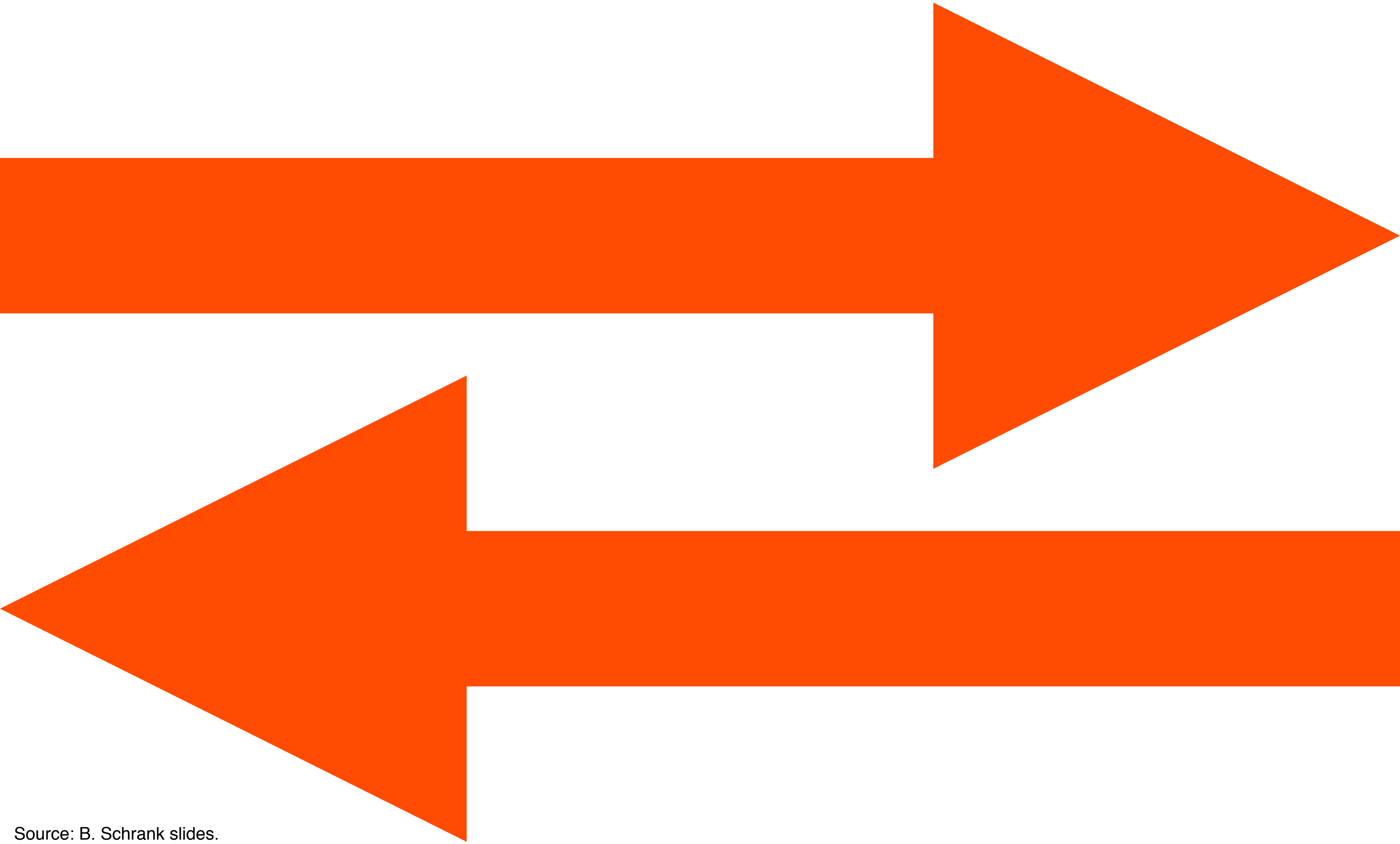
Common Fate



Parts of the visual field exhibiting the same motion are grouped together.

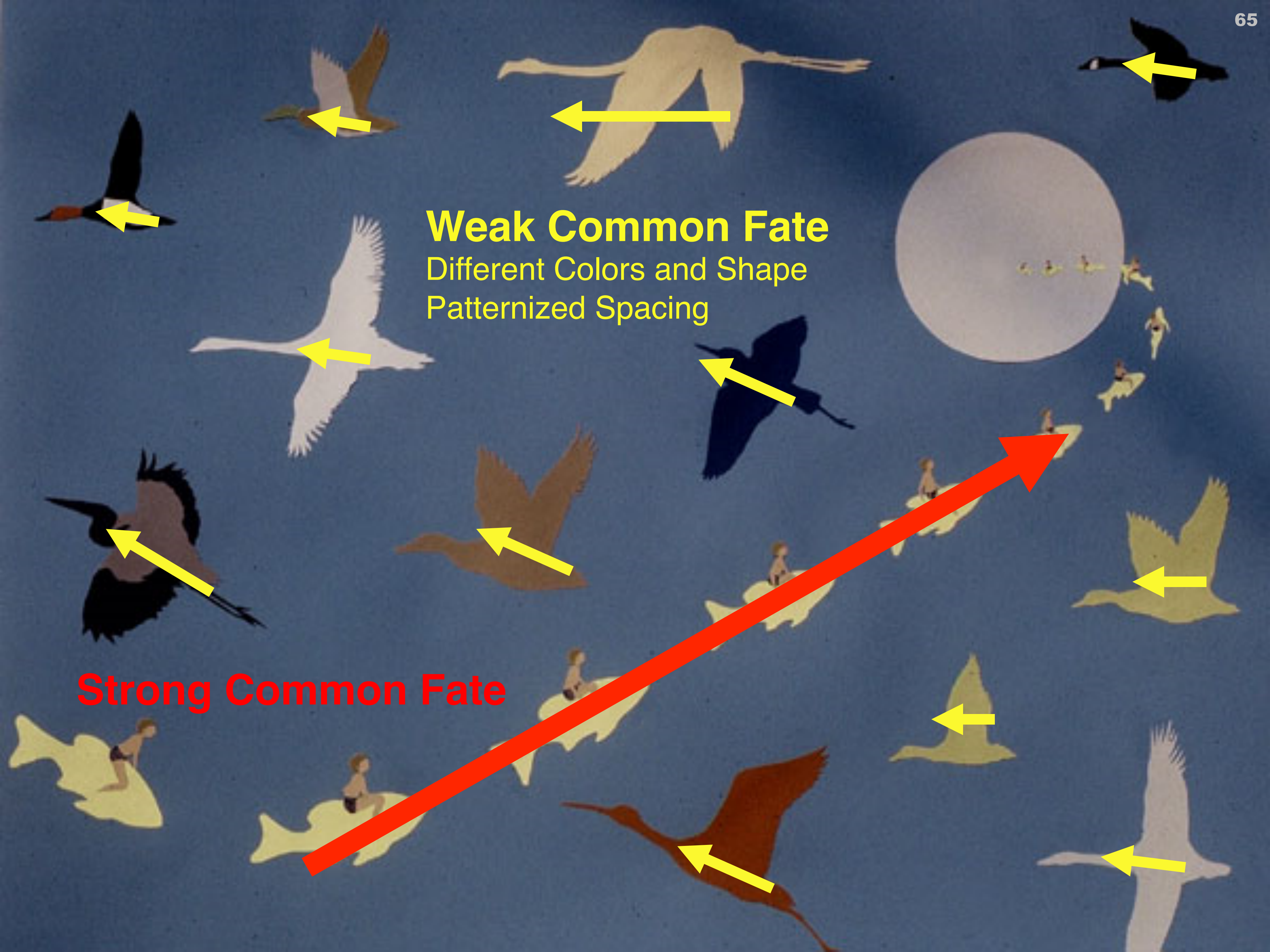
Source: B. Schrank slides.

Common Fate



Source: B. Schrank slides.

Source: B. Schrank slides.


Weak Common Fate
Different Colors and Shape
Patternized Spacing
Strong Common Fate

Scale Constancy

A form tends to preserve its proper shape, size and color... An object is perceived correctly as to the size or intensity within a wide range of actual stimulus variations. An automobile seen at a distance of 100 yards does not appear smaller than one seen at 20 yards even though there is a greater disparity in the size of the retinal image.

-Fryer

Source: B. Schrank slides.

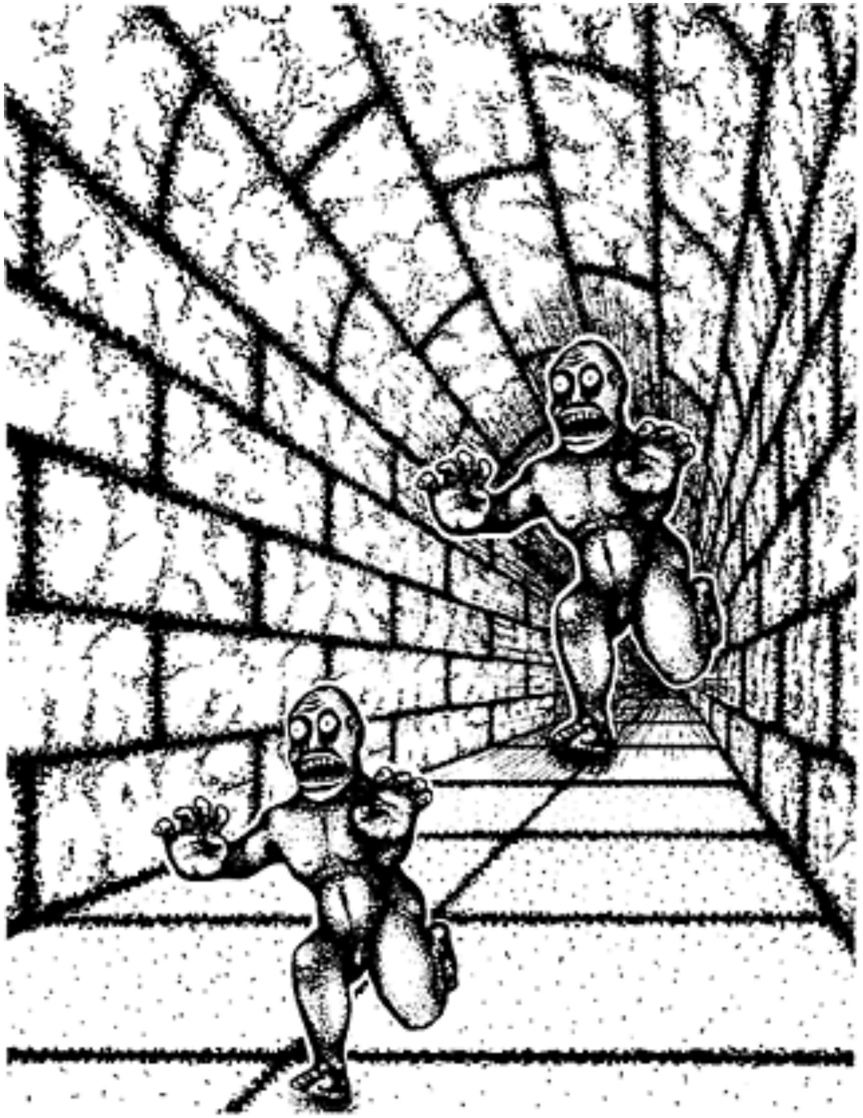
Scale Constancy

A form tends to preserve its proper shape, size and color... An object is perceived correctly as to the size or intensity within a wide range of actual stimulus variations. An automobile seen at a distance of 100 yards does not appear smaller than one seen at 20 yards even though there is a greater disparity in the size of the retinal image.

-Fryer

Source: B. Schrank slides.

Recall...



Reversing scale constancy to retain context. The figures are actually the same measurement.

Source: B. Schrank slides.

Color Constancy

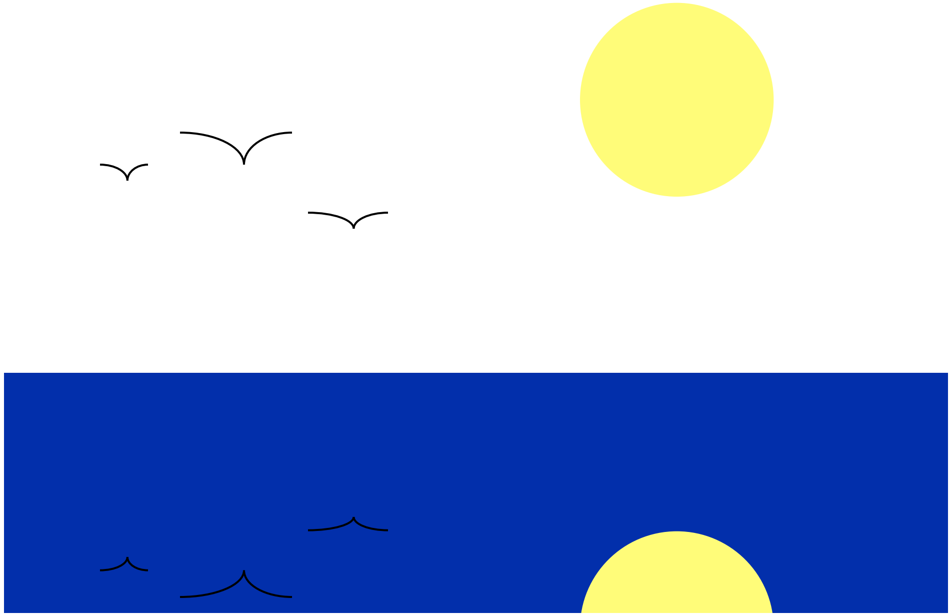
A lawn is seen as the same shade of green, even though part of it lies in bright sunshine and part in shadow.

-Fryer

Source: B. Schrank slides.

The impressionists tried to reverse this gestalt and paint what they see before their mind makes sense of it, stripping away the richness of reality (of course, squinting helps).

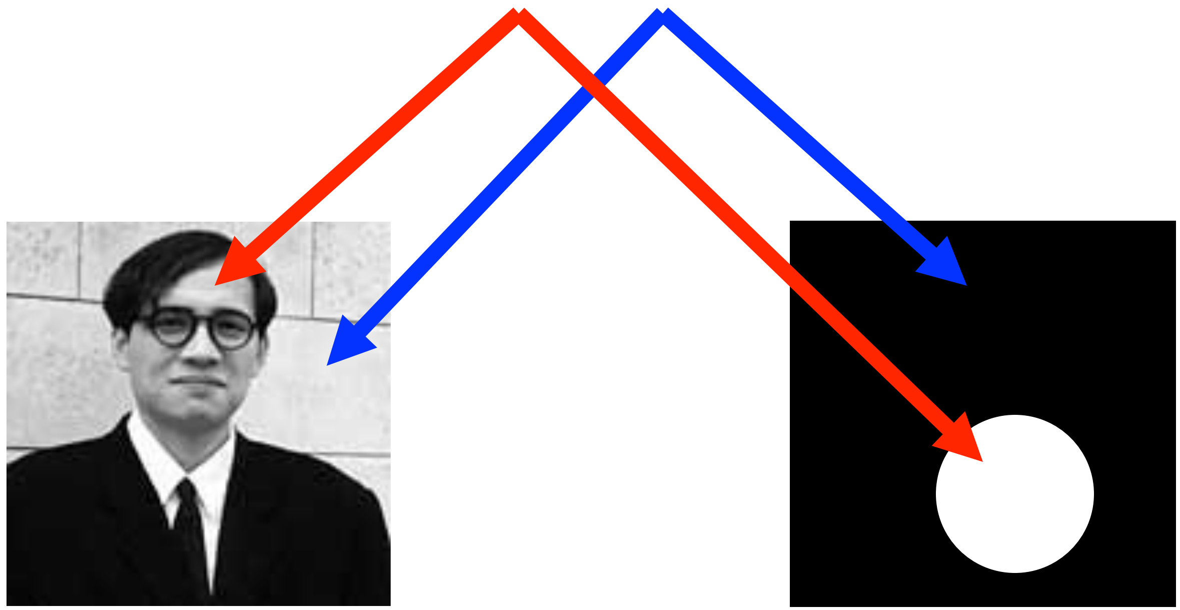
Source: B. Schrank slides.



a Composition is a combination of elements to make a unified whole.

Source: B. Schrank slides.

Figure/Ground



A form tends to be a figure set upon the ground, and a figure-ground dichotomy is fundamental to all perception.

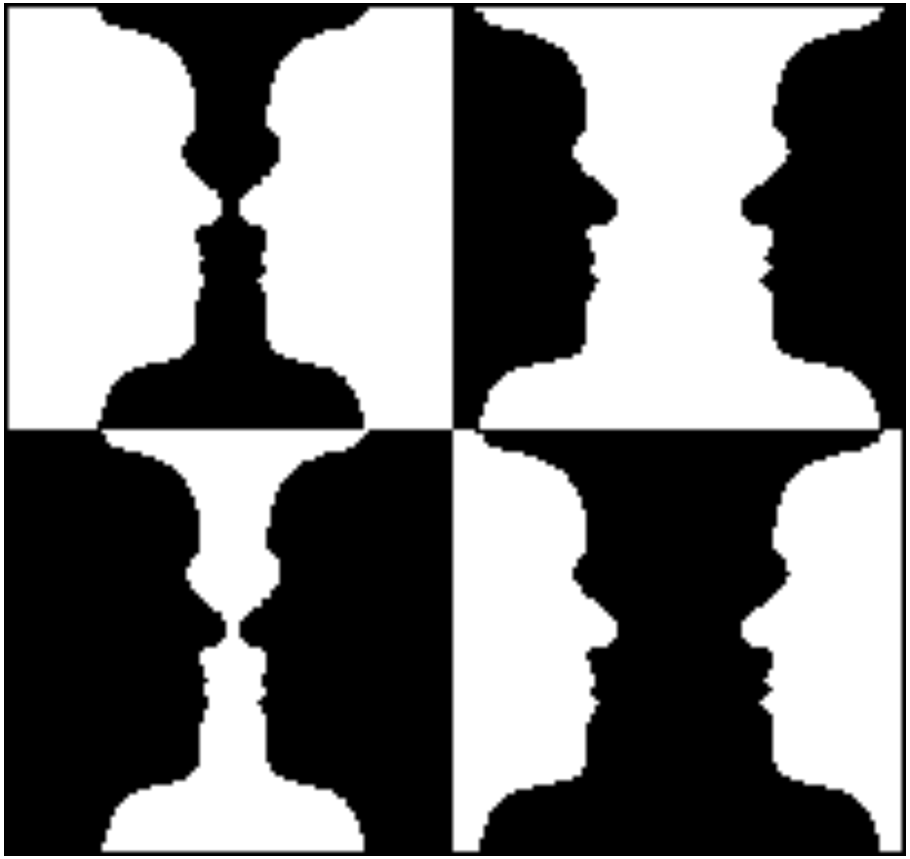
-Fryer

Source: B. Schrank slides.

Reversible Figure/Ground Relationship

Source: C Cumbie-Jones slides.

Reversible Figure/Ground Relationship



Can be affected by the principle of smallness:

Smaller areas tend to be seen as figures against a larger background.

Source: C Cumbie-Jones slides.

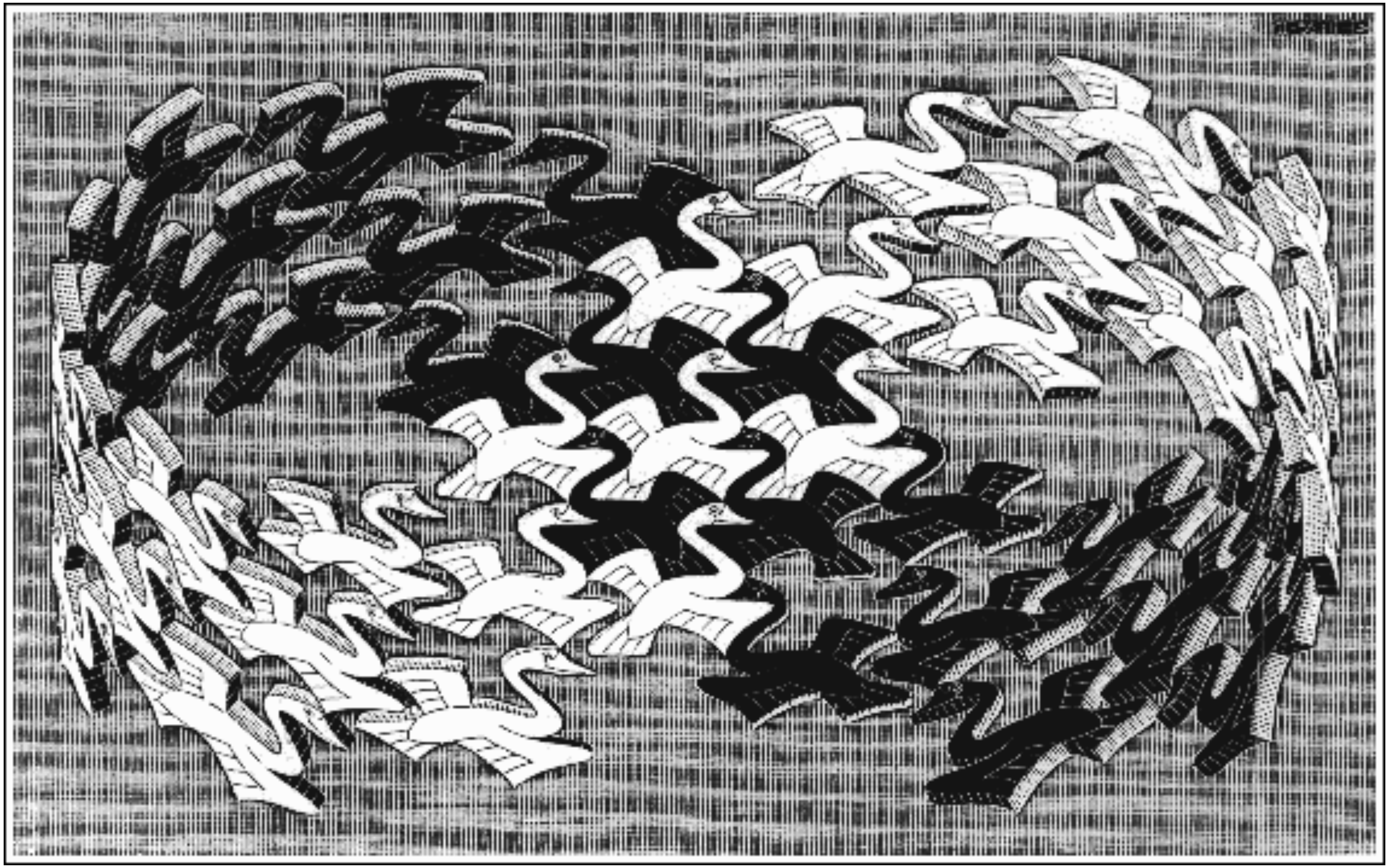
Figure-Ground Equivocation

Tessellation – interlocking figure/ground

M.C. Escher

Source: C Cumbie-Jones slides.

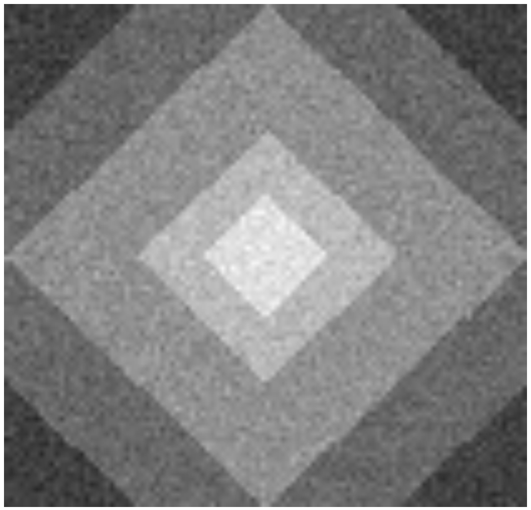
Figure-Ground Equivocation



Source: B. Schrank slides.

Early Segmentation Models

Piece-Wise Constant Models (image restoration)



observed noisy image I

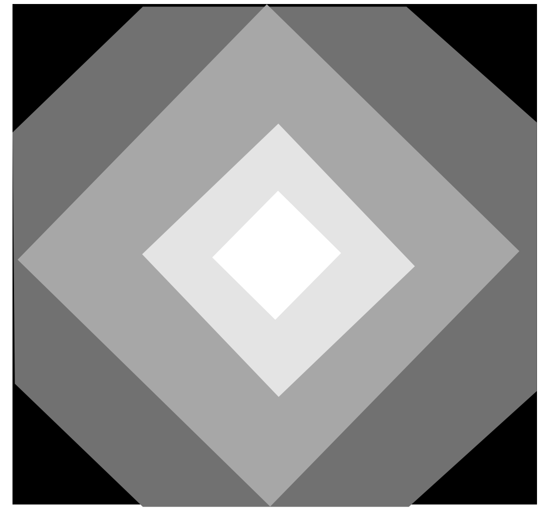
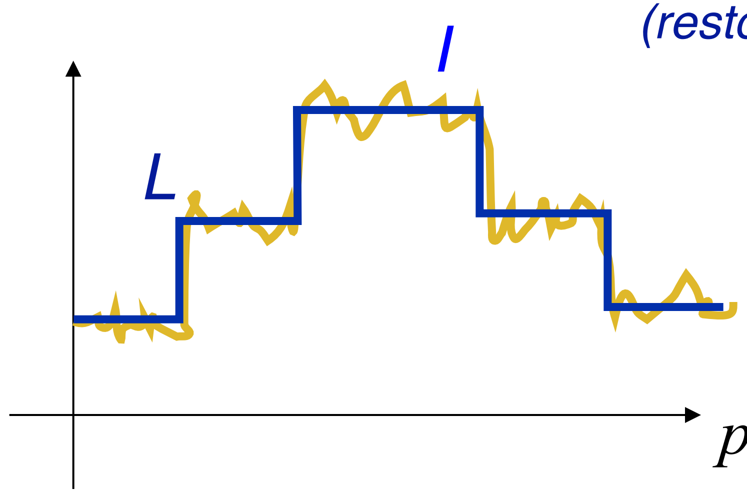
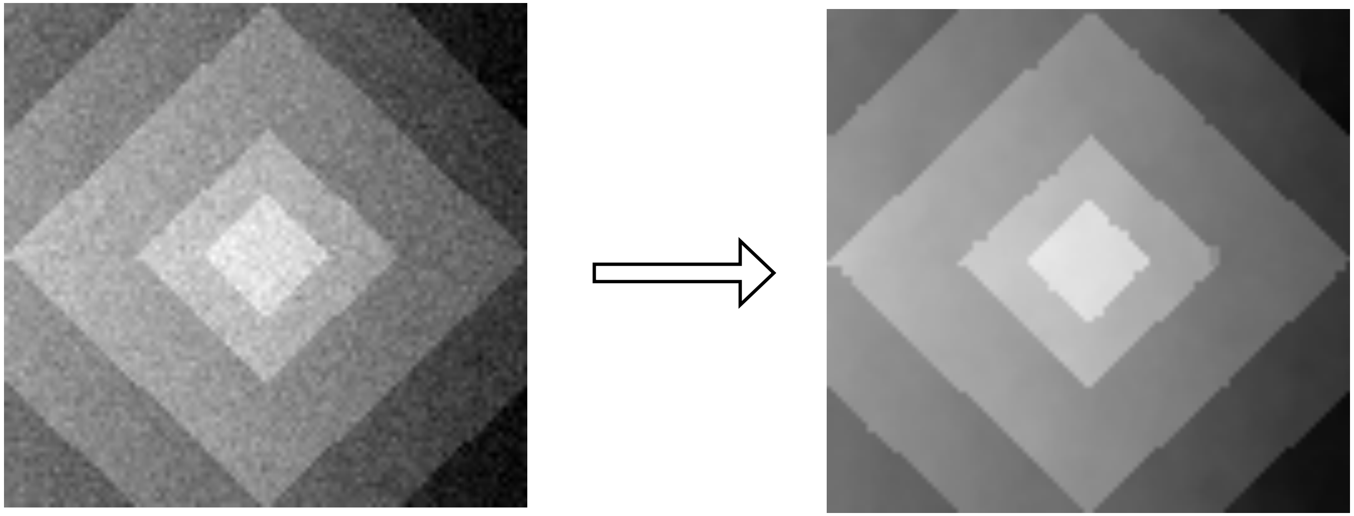


image labeling L
(restored intensities)



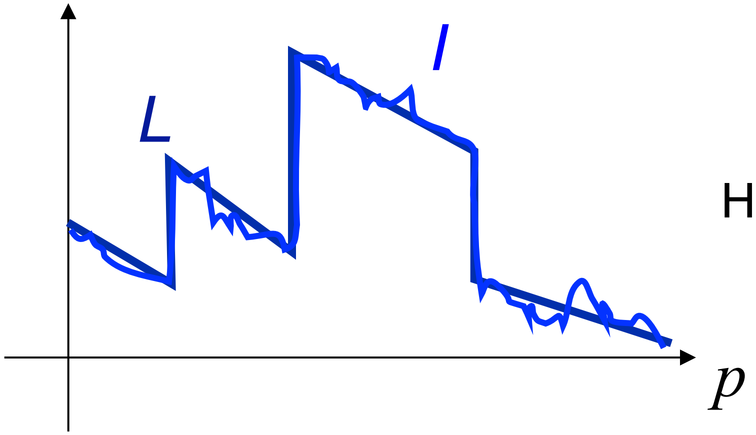
Source: Y. Boykov slides.

Piecewise Smooth Models



observed noisy image I

image labeling L
(restored intensities)



How to compute *L* from *I* ?

Source: Y. Boykov slides.

Piecewise Smooth Models

- Mumford-Shah Model (1989)
- Recall the functional view of an image: $f : \mathbb{R}^2 \rightarrow \mathbb{R}$
- Consider a decomposition of the domain of the image into a set of regions $\{R_1, R_2, \ldots, R_n\}$ such that $R = \bigcup_i R_i$
- Let ∂ represent the boundary between regions in R
- Assumptions:
  - f varies smoothly within each R_i
  - f changes rapidly and is discontinuous across boundaries ∂

- Seek some approximation $\hat{f}$ of f that observes our assumptions and yet best matches the original image f

Mumford-Shah

$$E(\hat{f}, \partial) \doteq \alpha \int\int_{R} (\hat{f} - f)^2 dx dy + \beta \int\int_{R-\partial} \|\nabla \hat{f}\|^2 dx dy + \gamma |\partial|$$

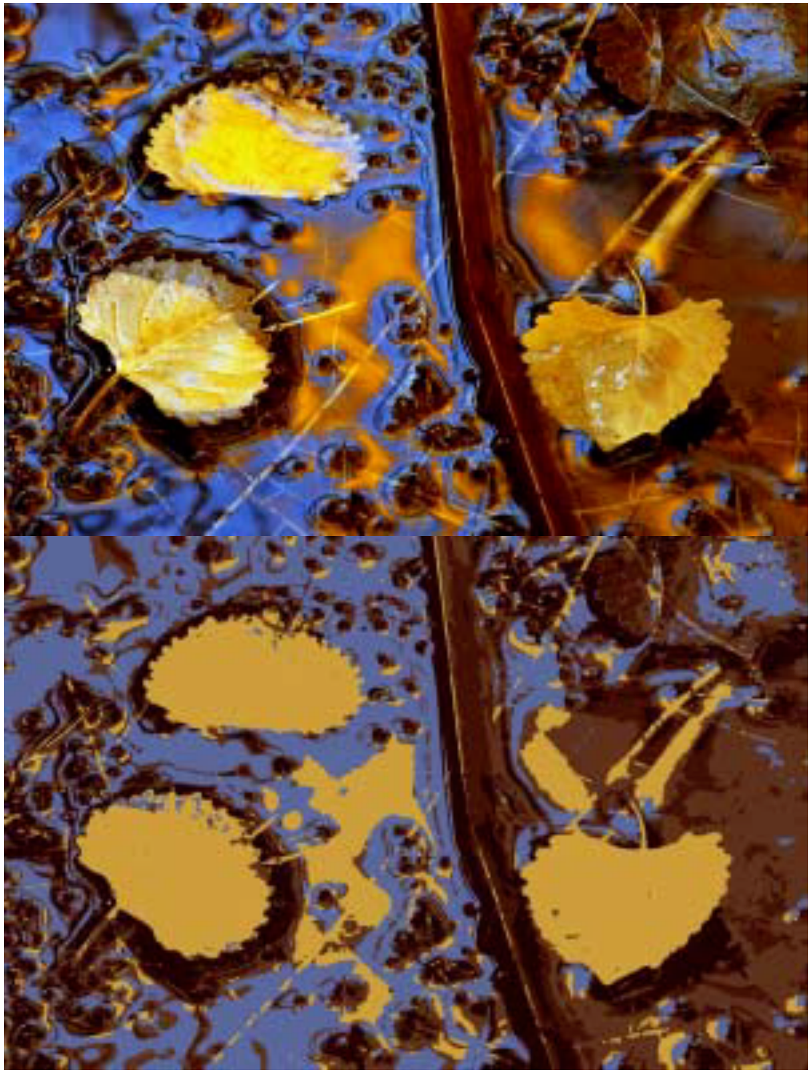
Goodness of Fit

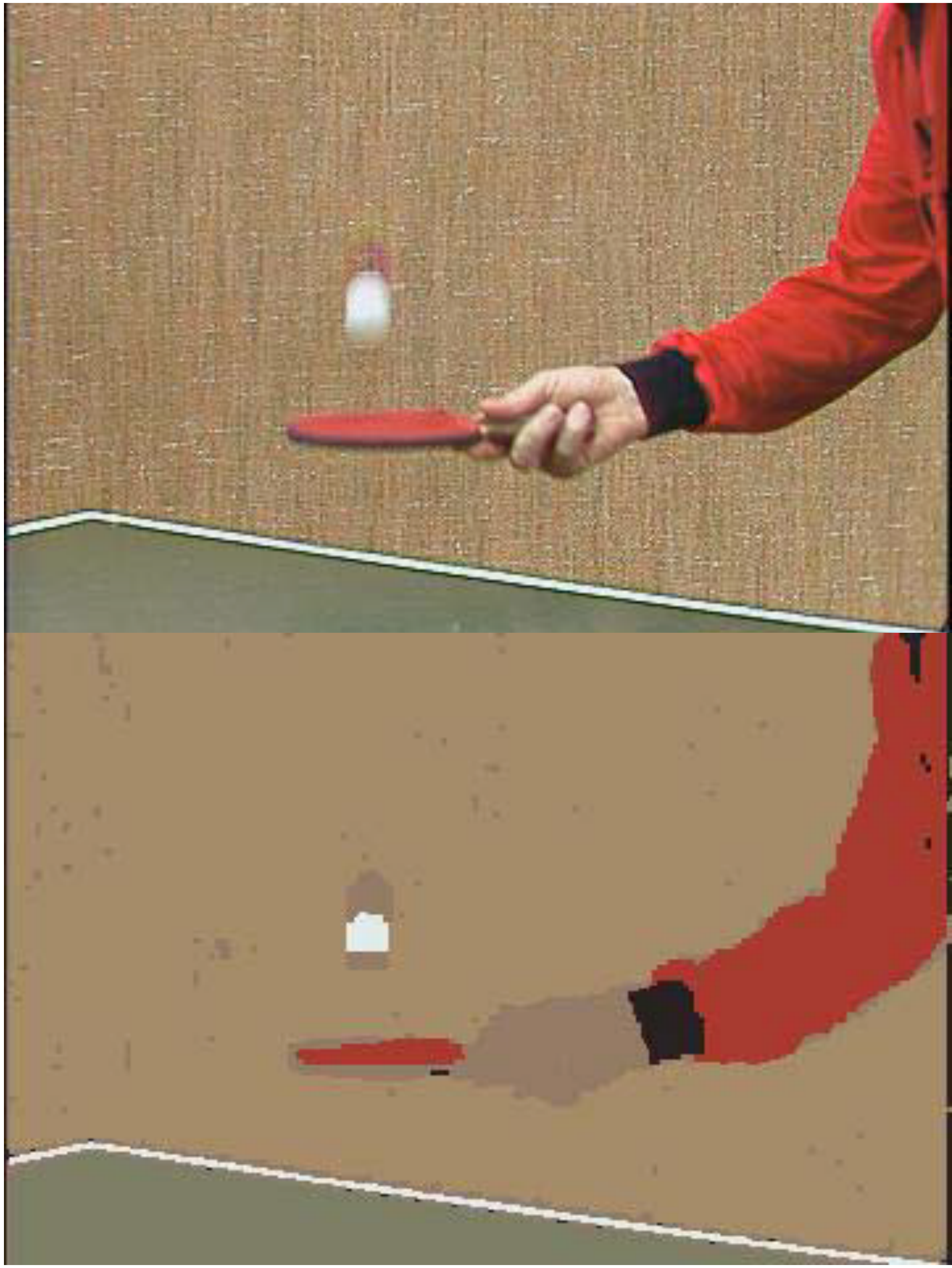
Smoothness in the approximation

Boundary smoothness

- How could you describe the type of images you'd expect to find being outputted in $\hat{f}$?

Mumford-Shah Example Results





Source: Z. Kato slides.

Piecewise Constant Mumford-Shah

- Assuming regions are constant rather than smooth
  - E.g., stereo disparities typical case

$$E_{\mathrm{PC}}(\hat{f}, \partial) \doteq \alpha \int \int_R (\hat{f} - f)^2 dx dy + \gamma |\partial|$$

Discretized Versions: Markov Random Fields

- Explicitly instantiate a graph-lattice
- Set up same energy functionals (constant or smooth)
- "Weak String/Membrane Models"

$$E_{\mathrm{WS}}(\hat{g}) = \alpha \sum_{i \in \Gamma} (g_i - \hat{g}_i)^2 + \beta \sum_{i \in \Gamma} (\hat{g}_{i+1} - \hat{g}_i)^2 (1 - l_i) + \gamma \sum_i l_i$$

Graph Lattice

Edge Presence Variable (binary)

- See Geman & Geman 1984 for example.
  - And take my course in the spring.

Example

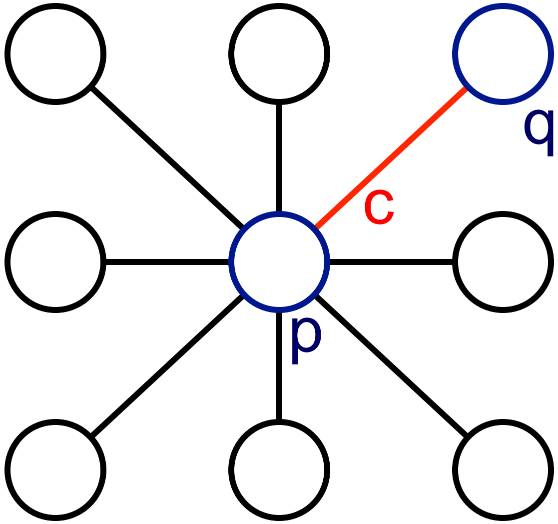
Source: Z. Kato slides.

Discrete / Graph-Based Models

- Minimum Spanning Forest Method
- Intelligent Scissors
- Min-cut
- Normalized cut
- Segmentation by Weighted Aggregation

Setting up the problem

- Treat the image as a graph



Graph

- node for every pixel **p**
- link between every adjacent pair of pixels, **p**,**q**
- cost **c** for each link

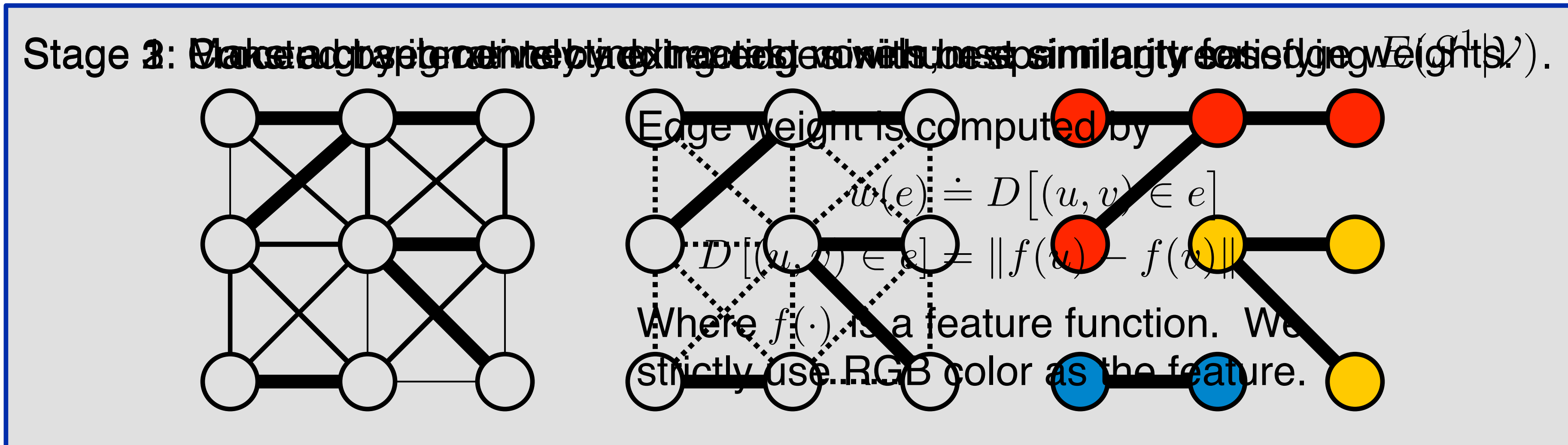
Note: each *link* has a cost

Source: S. Seitz slides.

Basic Minimum Spanning Forest Method

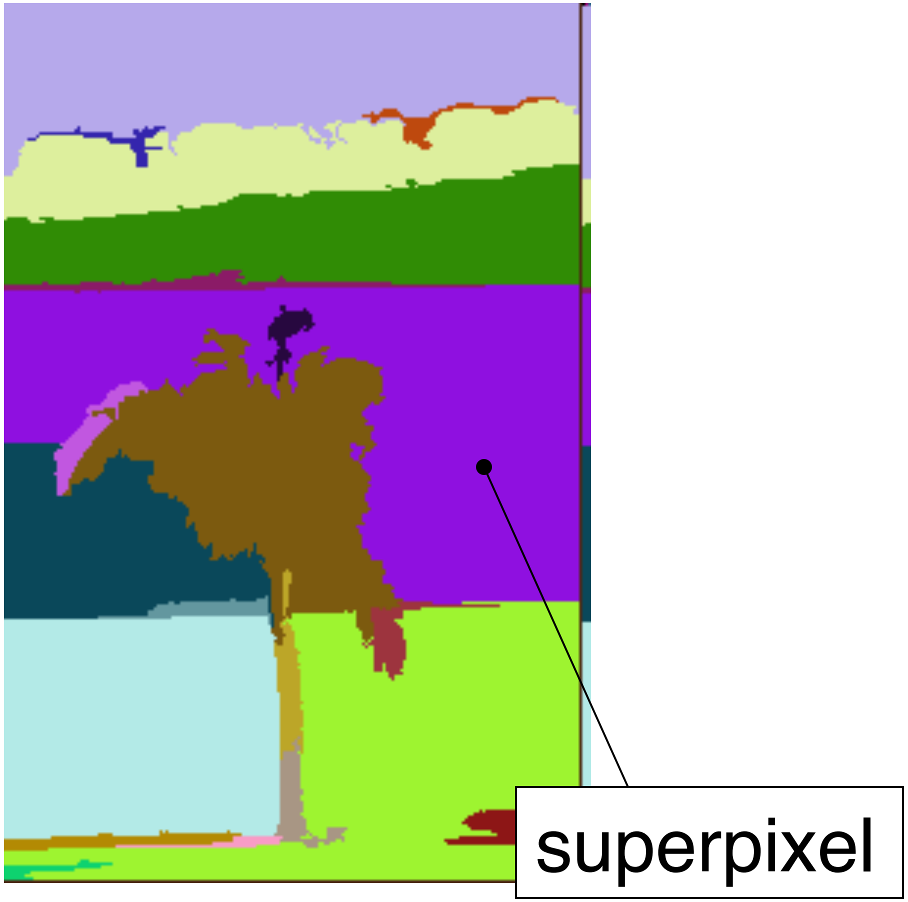
- Use the minimum spanning tree method of Felzenszwalb and Huttenlocher IJCV 2004.

$$E(S^1|\mathcal{V}) = \tau \sum_{s \in S^1} \sum_{e \in \mathrm{MST(s)}} w(e) + \sum_{s,t \in S^1} \min_{e \in \langle s,t \rangle} w(e)$$



Efficient Graph-Based Image Segmentation

Efficient Graph-Based Image Segmentation Pedro F. Felzenszwalb and Daniel P. Huttenlocher
International Journal of Computer Vision, Volume 59, Number 2, September 2004



C++ implementation
http://people.cs.uchicago.edu/~pff/segment

Source: S. Savarese slides.

Intelligent Scissors [Mortensen 95]

- Approach answers a basic question
  - Q: how to find a path from seed to mouse that follows object boundary as closely as possible?

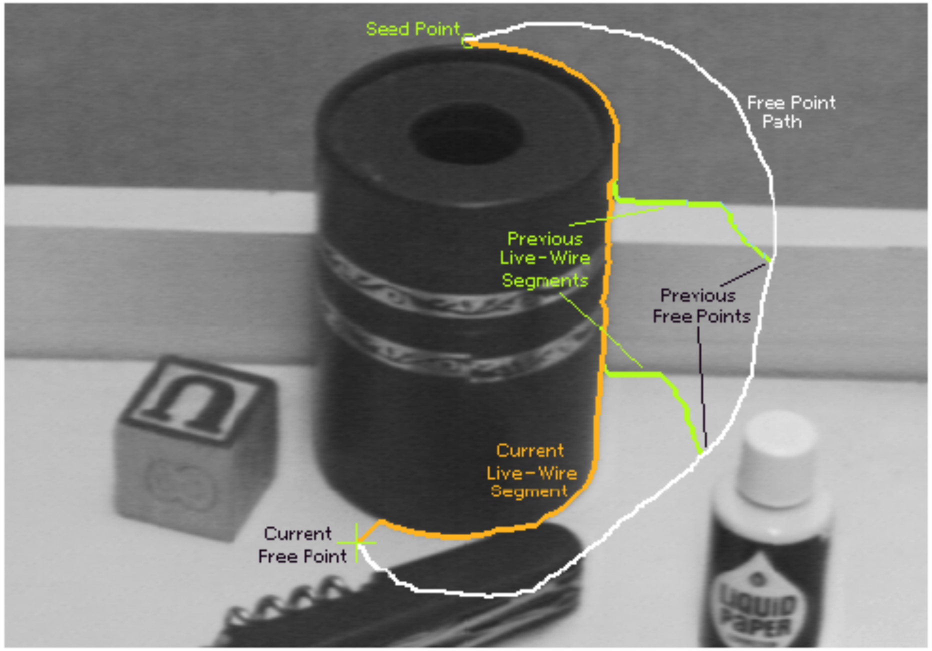
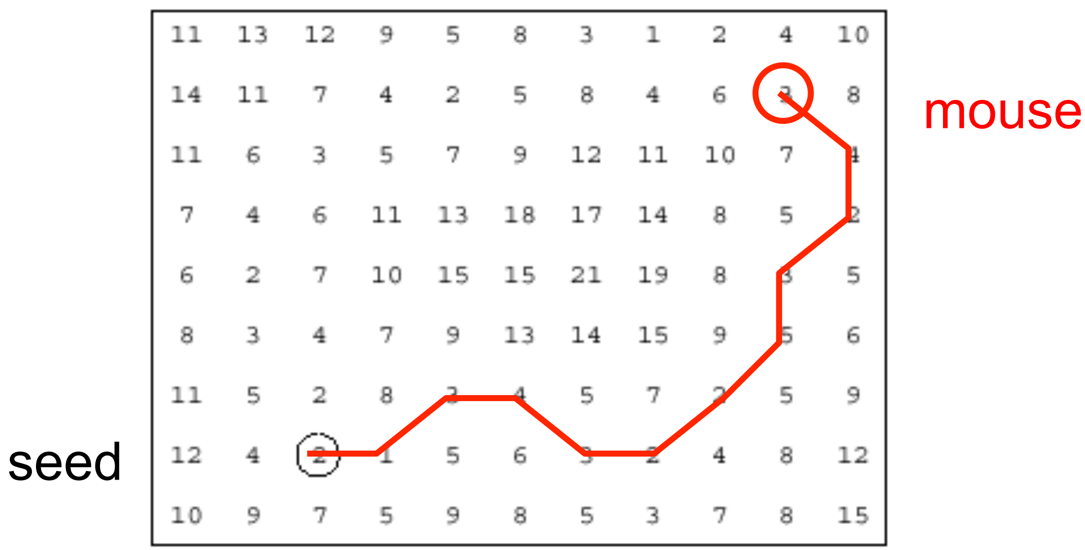


Figure 2: *Image demonstrating how the live-wire segment adapts and snaps to an object boundary as the free point moves (via cursor movement). The path of the free point is shown in white. Live-wire segments from previous free point positions (t_0, t_1, and t_2) are shown in green.*

Source: S. Seitz slides.

Intelligent Scissors

- Basic Idea
  - Define edge score for each pixel
    - edge pixels have low cost
  - Find lowest cost path from seed to mouse



Questions

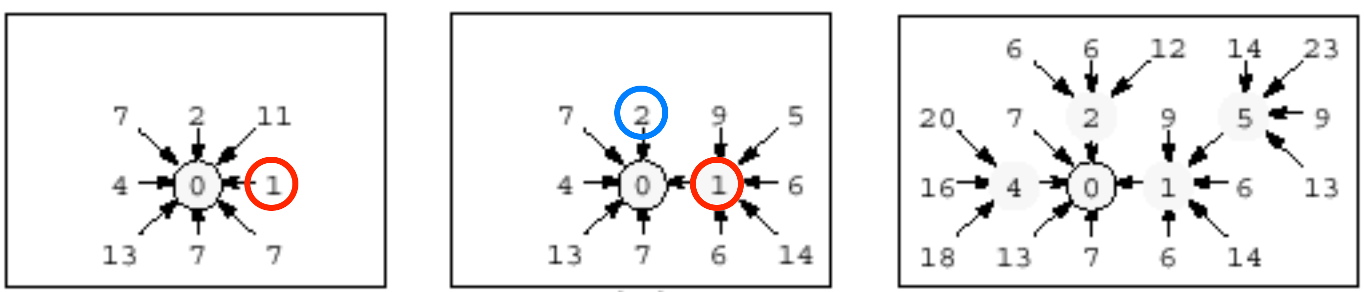
- How to define costs?
- How to find the path?

Source: S. Seitz slides.

Path Search (basic idea)

- Graph Search Algorithm
  - Computes minimum cost path from seed to *all other pixels*

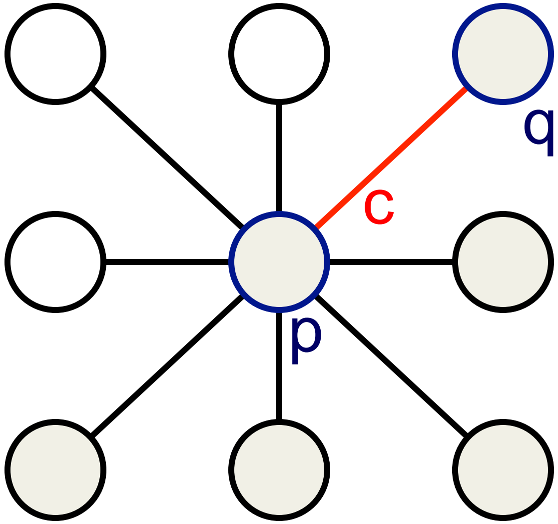
11	13	12	9	5	8	3	1	2	4	10
14	11	7	4	2	5	8	4	6	3	8
11	6	3	5	7	9	12	11	10	7	4
7	4	6	11	13	18	17	14	8	5	2
6	2	7	10	15	15	21	19	8	3	5
8	3	4	7	9	13	14	15	9	5	6
11	5	2	8	3	4	5	7	2	5	9
12	4	2	1	5	6	3	2	4	8	12
10	9	7	5	9	8	5	3	7	8	15



Source: S. Seitz slides.

How does this really work?

- Treat the image as a graph

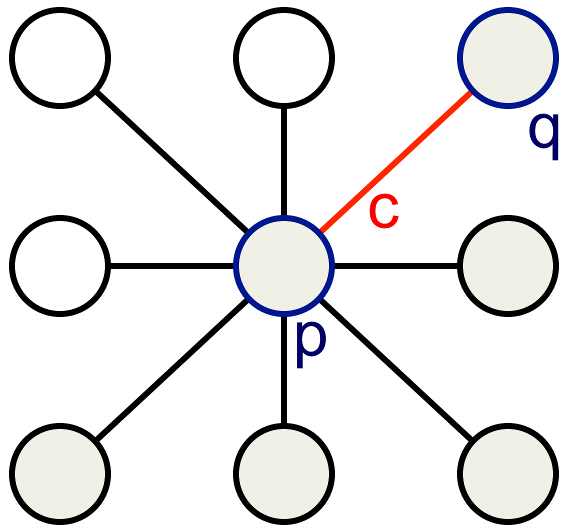


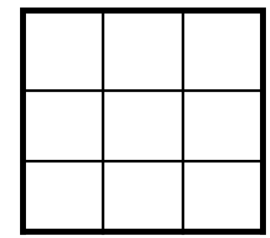
Want to hug image edges: how to define cost of a link?

- the link should follow the intensity edge
  - want intensity to change rapidly ⊥ to the link
- c ≈ - |difference of intensity ⊥ to link|

Source: S. Seitz slides.

Defining the costs



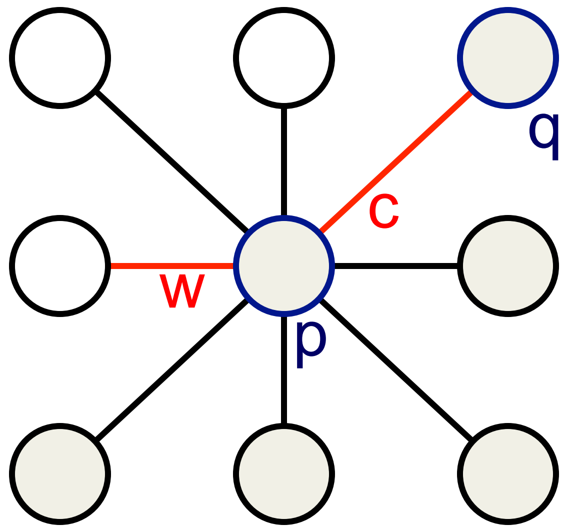


- c can be computed using a cross-correlation filter
  - assume it is centered at p
- Also typically scale c by its length
  - set c = (max-|filter response|)
    - where max = maximum |filter response| over all pixels in the image

Source: S. Seitz slides.

Defining the costs

$$\frac{1}{4}\begin{bmatrix} 1 & 1 & \\ & & \\ -1 & -1 & \end{bmatrix}$$

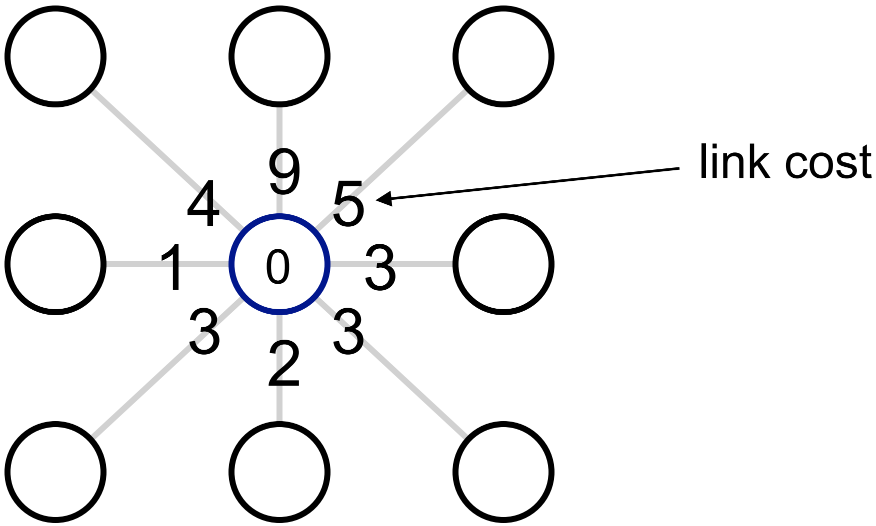


$$\frac{1}{\sqrt{2}}\begin{bmatrix} & 1 & \\ & & -1 \\ & & \end{bmatrix}$$

- c can be computed using a cross-correlation filter
  - assume it is centered at p
- Also typically scale c by its length
  - set c = (max-|filter response|)
    - where max = maximum |filter response| over all pixels in the image

Source: S. Seitz slides.

Dijkstra's shortest path algorithm



Algorithm

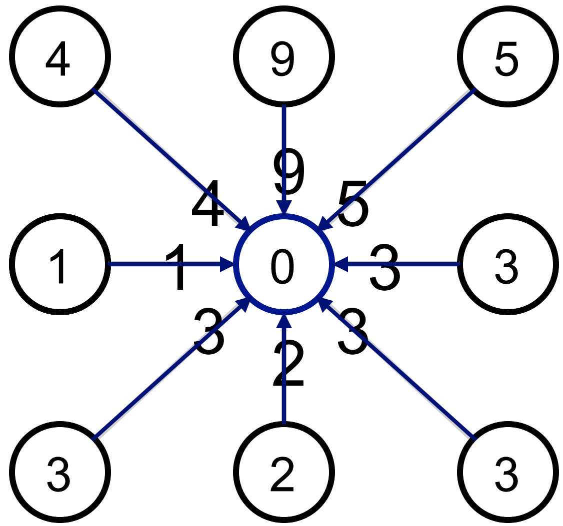
1. init node costs to ∞, set p = seed point, cost(p) = 0
2. expand p as follows:

   for each of p's neighbors q that are not expanded

   » set cost(q) = min(cost(p) + c_{pq}, cost(q))

Source: S. Seitz slides.

Dijkstra’s shortest path algorithm



Algorithm

1. init node costs to ∞, set p = seed point, cost(p) = 0
2. expand p as follows:

   for each of p’s neighbors q that are not expanded

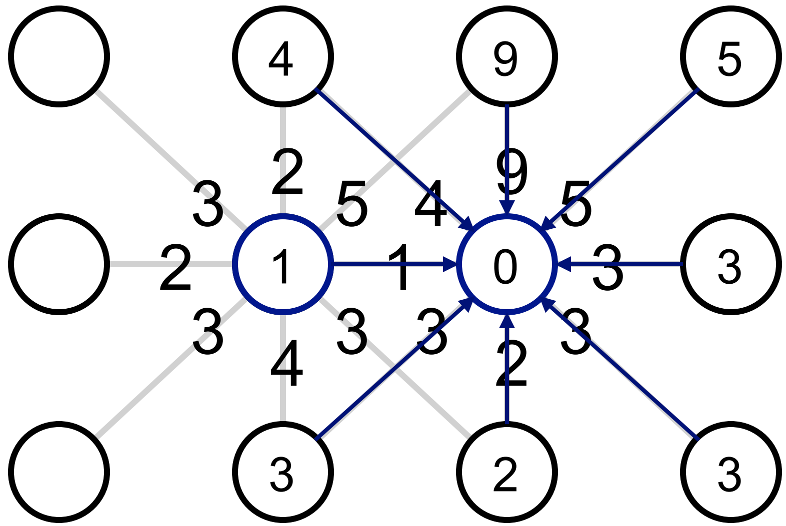
   » set cost(q) = min(cost(p) + c_{pq}, cost(q))

     » if q’s cost changed, make q point back to p

   » put q on the ACTIVE list (if not already there)

Source: S. Seitz slides.

Dijkstra's shortest path algorithm



Algorithm

1. init node costs to ∞, set p = seed point, cost(p) = 0
2. expand p as follows:

   for each of p's neighbors q that are not expanded

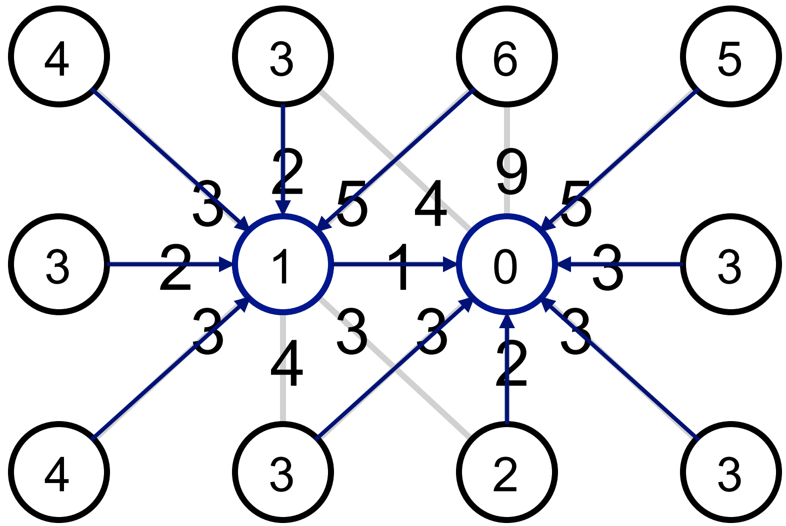
   » set cost(q) = min(cost(p) + c_{pq}, cost(q))

     » if q's cost changed, make q point back to p

   » put q on the ACTIVE list (if not already there)
3. set r = node with minimum cost on the ACTIVE list
4. repeat Step 2 for p = r

Source: S. Seitz slides.

Dijkstra's shortest path algorithm



Algorithm

1. init node costs to ∞, set p = seed point, cost(p) = 0
2. expand p as follows:

   for each of p's neighbors q that are not expanded

   » set cost(q) = min(cost(p) + c_{pq}, cost(q))

   » if q's cost changed, make q point back to p

   » put q on the ACTIVE list (if not already there)

3. set r = node with minimum cost on the ACTIVE list
4. repeat Step 2 for p = r

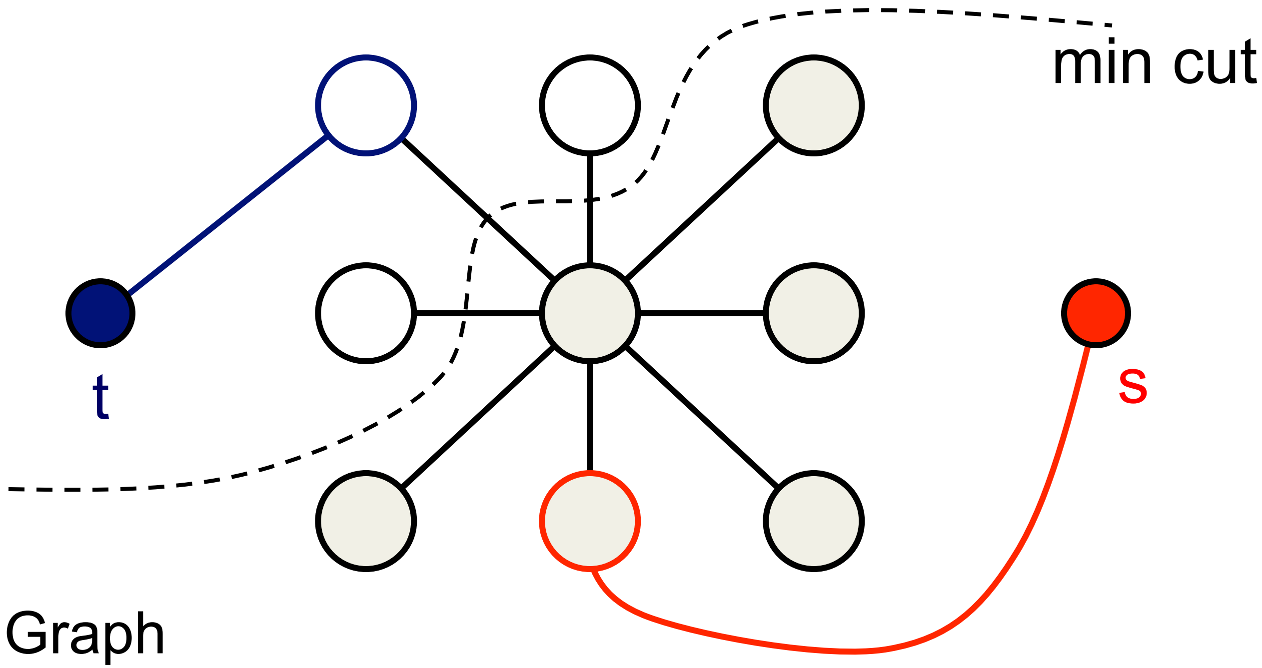
Source: S. Seitz slides.

Dijkstra's shortest path algorithm

- Properties
  - It computes the minimum cost path from the seed to every node in the graph. This set of minimum paths is represented as a *tree*
  - Running time, with N pixels:
    - $O(N^2)$ time if you use an active list
    - O(N log N) if you use an active priority queue (heap)
    - takes fraction of a second for a typical (640x480) image
  - Once this tree is computed once, we can extract the optimal path from any point to the seed in O(N) time.
    - it runs in real time as the mouse moves
  - What happens when the user specifies a new seed?

Source: S. Seitz slides.

Segmentation by min (s-t) cut [Boykov 2001]



- Graph
  - node for each pixel, link between pixels
  - specify a few pixels as foreground and background
    - create an infinite cost link from each bg pixel to the "t" node
    - create an infinite cost link from each fg pixel to the "s" node
  - compute min cut that separates s from t
  - how to define link cost between neighboring pixels?

Source: S. Seitz slides.

Grabcut [Rother et al., SIGGRAPH 2004]

Source: S. Seitz slides.

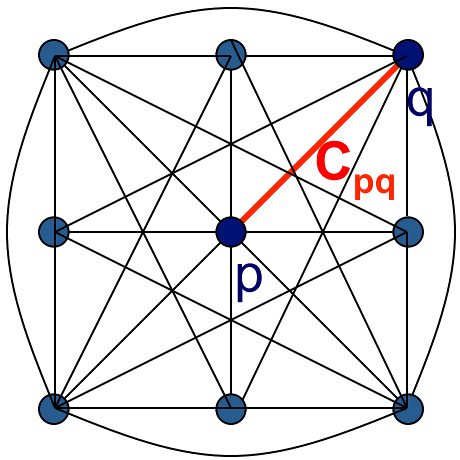
Is user-input required?

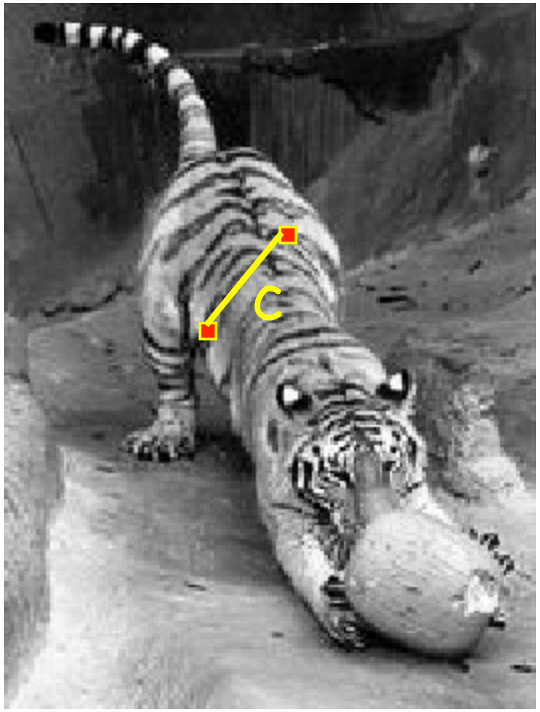
- Our visual system is proof that automatic methods are possible
  - classical image segmentation methods are automatic

- Argument for user-directed methods?
  - only user knows desired scale/object of interest

Source: S. Seitz slides.

Automatic graph cut [Shi & Malik]

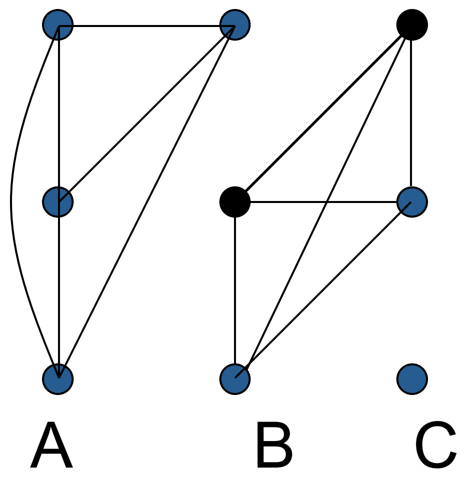




- *Fully-connected* graph
  - node for every pixel
  - link between *every* pair of pixels, **p**,**q**
  - cost c_{pq} for each link
    - c_{pq} measures *similarity*
      - similarity is *inversely proportional* to difference in color and position

Source: S. Seitz slides.

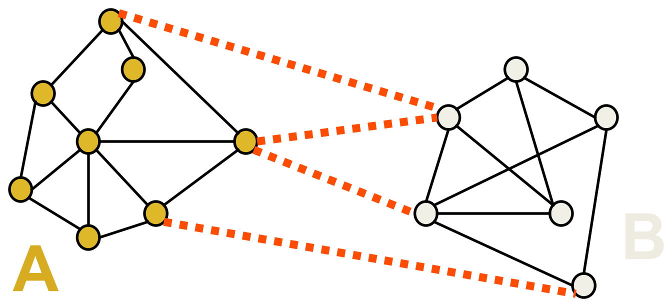
Segmentation by Graph Cuts



- Break Graph into Segments
  - Delete links that cross between segments
  - Easiest to break links that have low cost (similarity)
    - similar pixels should be in the same segments
    - dissimilar pixels should be in different segments

Source: S. Seitz slides.

Cuts in a graph



- Link Cut
  - set of links whose removal makes a graph disconnected
  - cost of a cut:

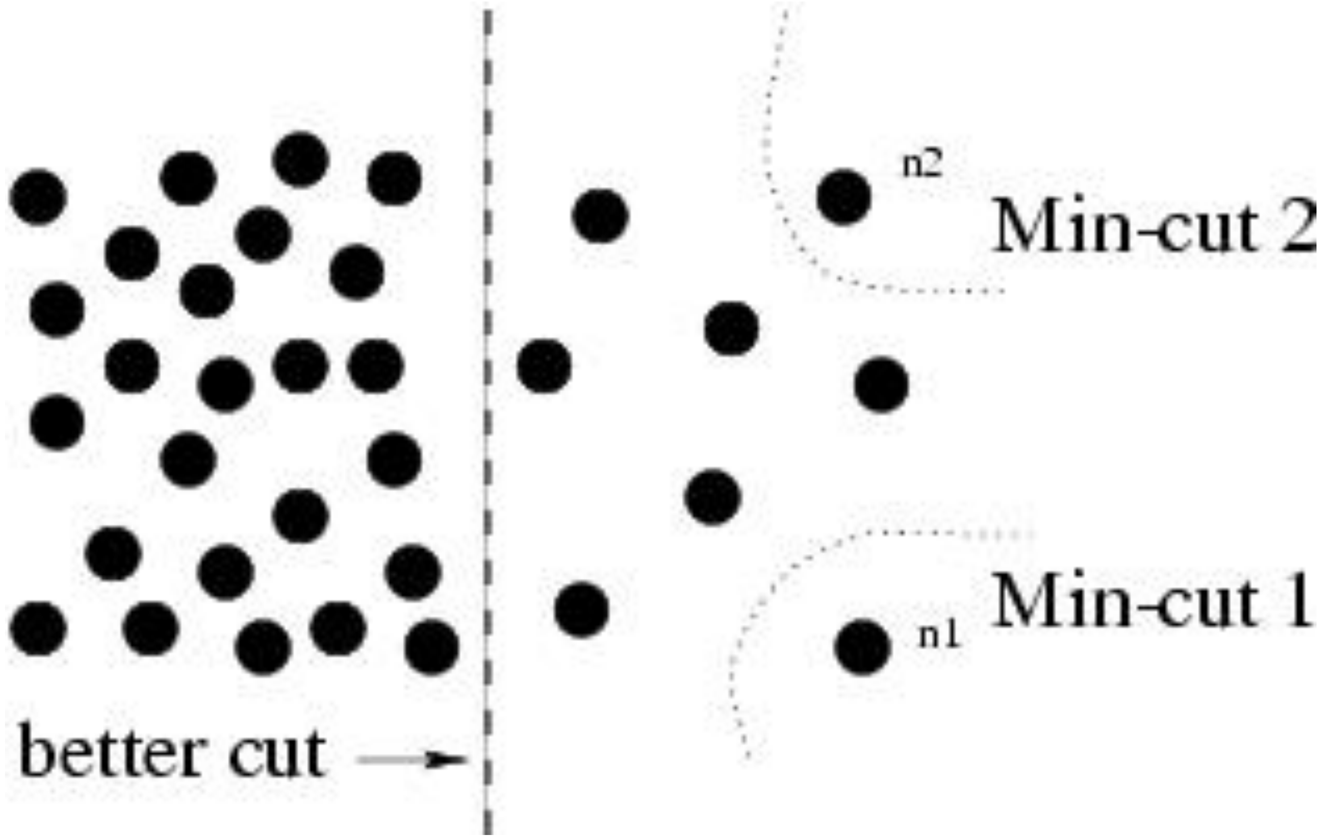
$$cut(A, B) = \sum_{p \in A, q \in B} c_{p,q}$$

Find minimum cut

- gives you a segmentation

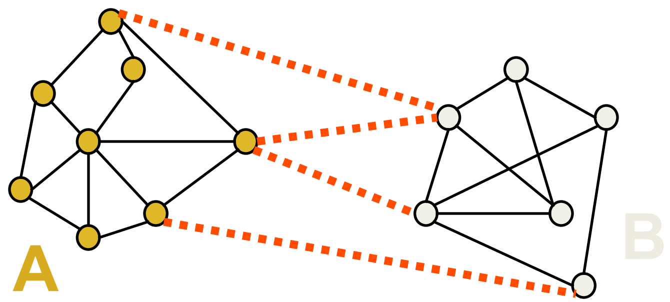
Source: S. Seitz slides.

But min cut is not always the best cut...



Source: S. Seitz slides.

Cuts in a graph



Normalized Cut

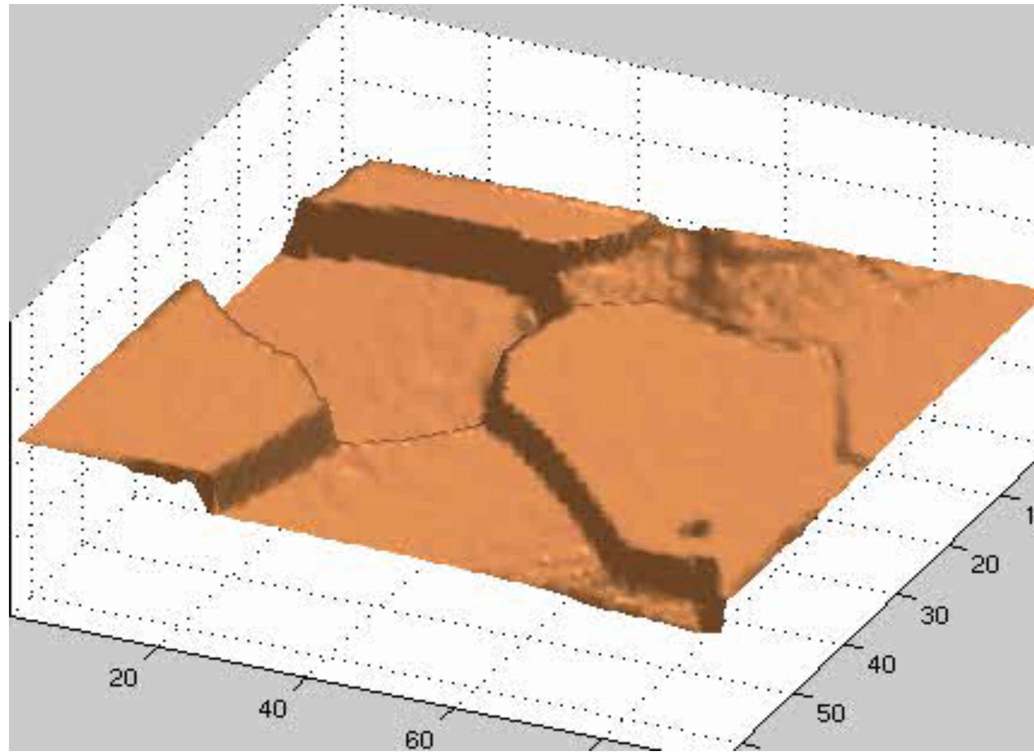
- a cut penalizes large segments
- fix by normalizing for size of segments

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

- volume(A) = sum of costs of all edges that touch A

Source: S. Seitz slides.

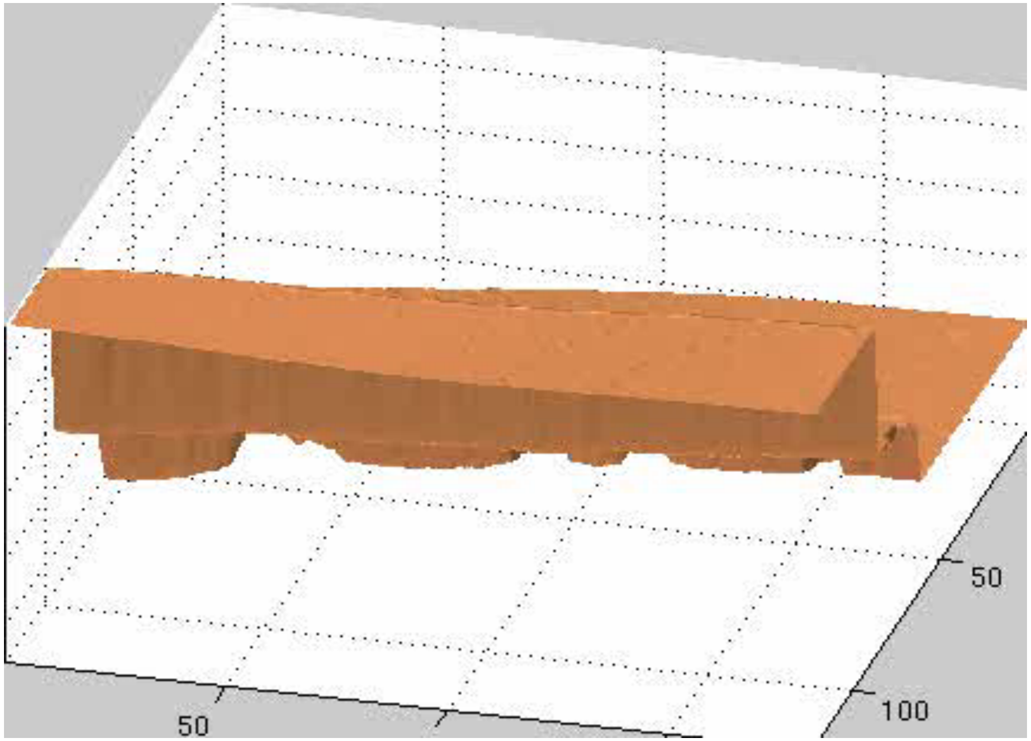
Interpretation as a Dynamical System



- Treat the links as springs and shake the system
  - elasticity proportional to cost
  - vibration “modes” correspond to segments
    - can compute these by solving an eigenvector problem
    - http://www.cis.upenn.edu/~jshi/papers/pami_ncut.pdf

Source: S. Seitz slides.

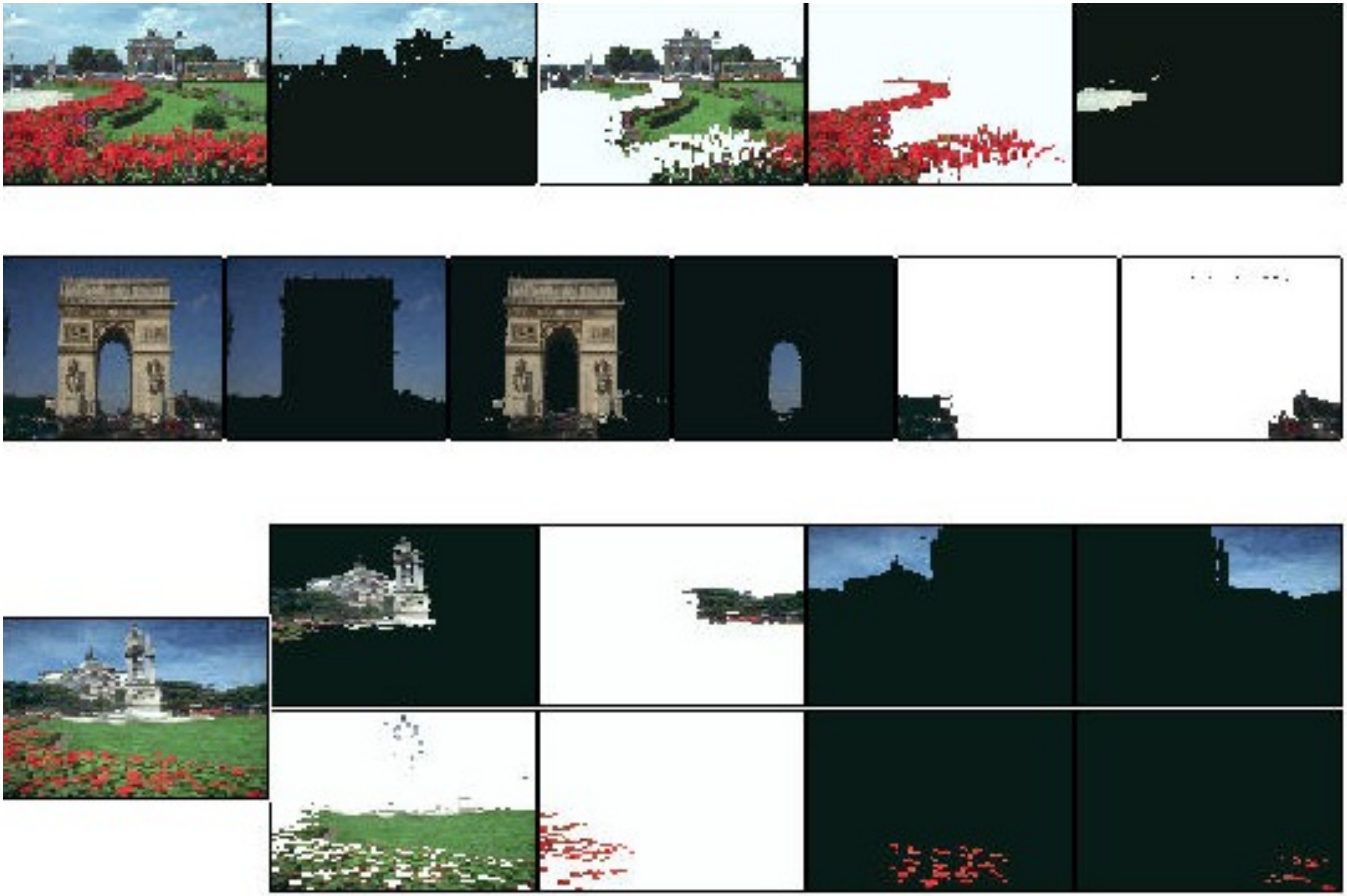
Interpretation as a Dynamical System



- Treat the links as springs and shake the system
  - elasticity proportional to cost
  - vibration “modes” correspond to segments
    - can compute these by solving an eigenvector problem
    - http://www.cis.upenn.edu/~jshi/papers/pami_ncut.pdf

Source: S. Seitz slides.

Color Image Segmentation Examples



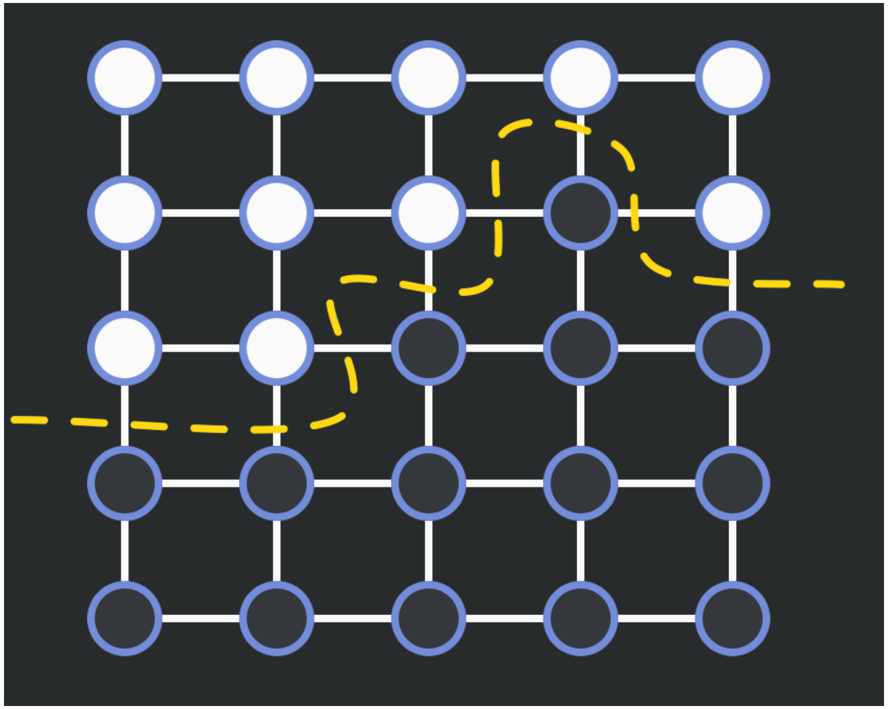
Source: S. Seitz slides.

Segmentation by Weighted Aggregation Set-Up

- Define the problem on a graph: $G = \{\mathcal{V}, \mathcal{E}\}$
  - Edges are sparse, to neighbors.
  - Each pixel / voxel is a node.
- Augment nodes, for $v \in \mathcal{V}$
  - statistics: s_v
  - class label: c_v
- Define affinity between $u, v \in \mathcal{V}$

$$w_{uv} \in \exp\left(-D(s_u, s_v; \theta)\right)$$

  - where D is some non-negative distance function and θ are some predetermined values.
- Regions are defined by cuts.

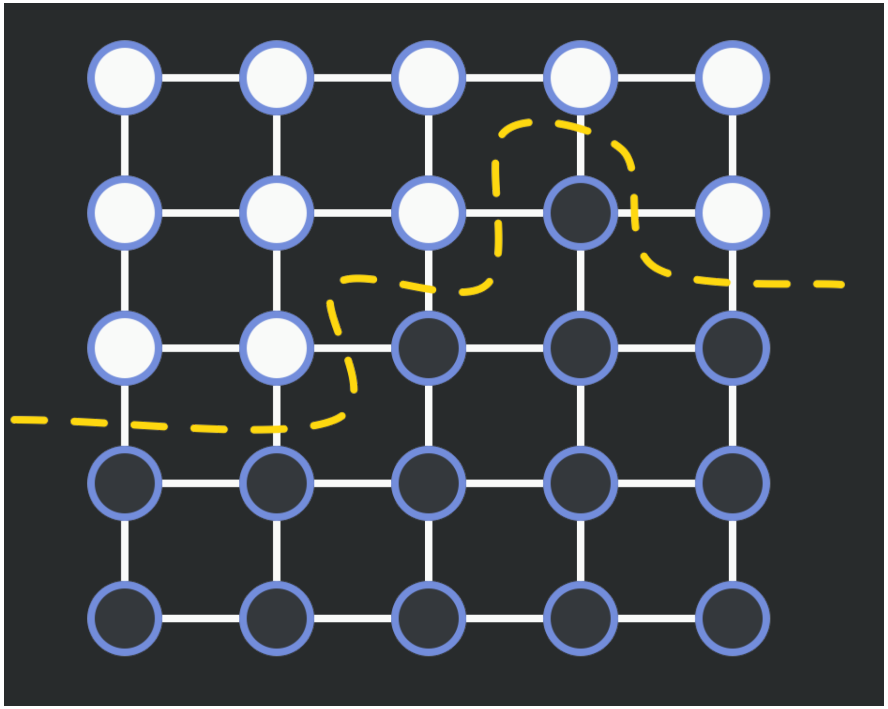


SWA Region Saliency

- Define a region saliency measure.

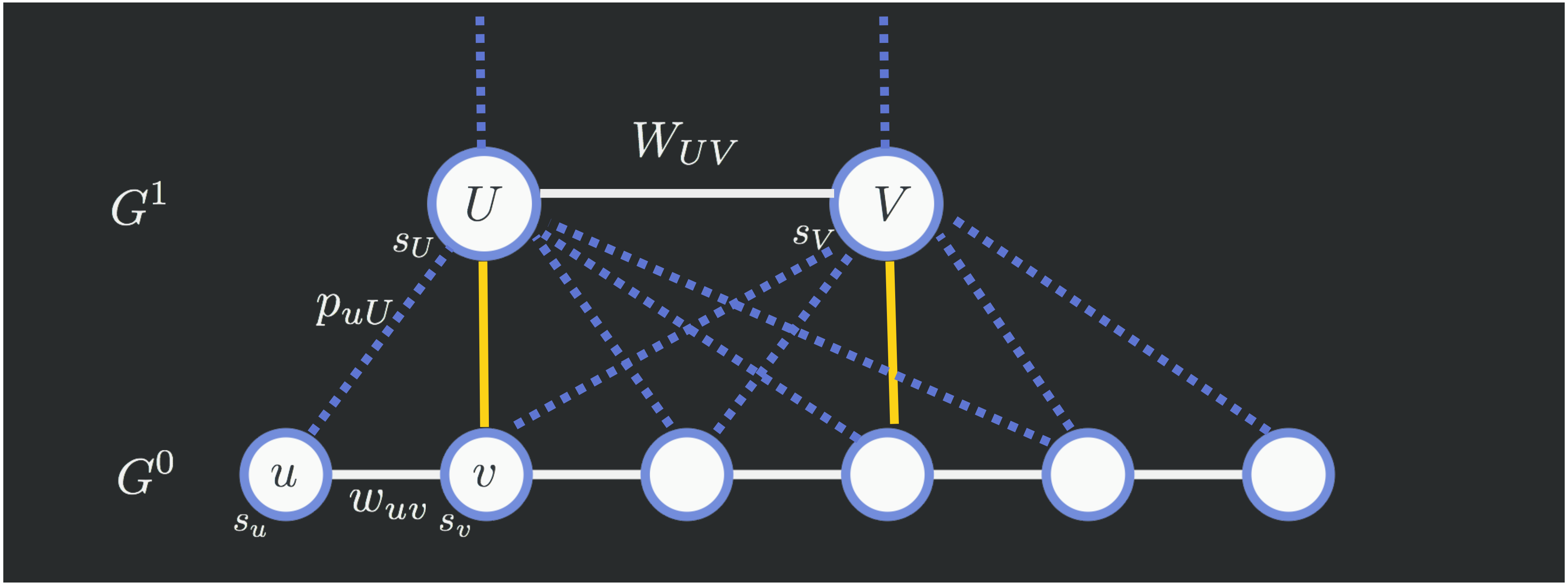
$$\Gamma(R) = \frac{\sum_{u \in R, v \notin R} w_{uv}}{\sum_{u,v \in R} w_{uv}}$$

- Low $\Gamma(R)$ means good saliency:
  - Low affinity on boundary.
  - High affinity in interior.
- Criterion is based on the normalized cut criterion (Shi & Malik)
  - Affinities at the pixel scale only.



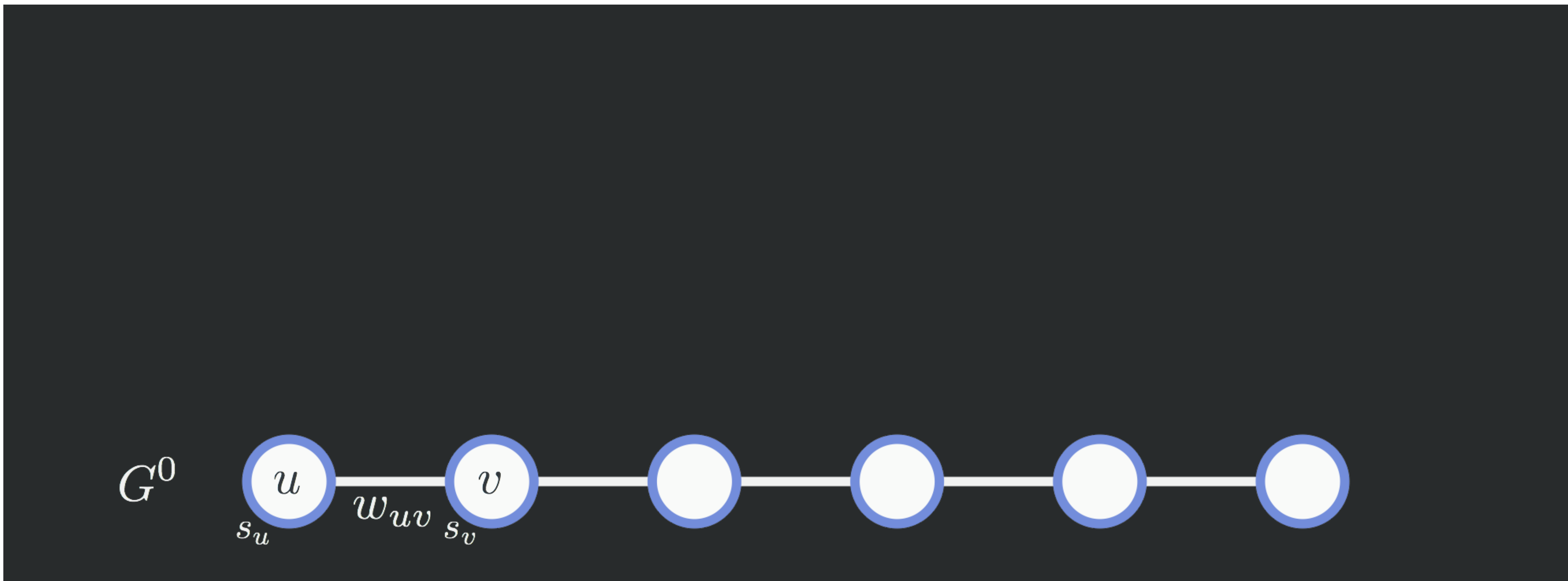
Segmentation by Weighted Aggregation

- Invented in natural image domain by Sharon et al. (CVPR 2000, 2001, Nature 2006).
- Used in medical imaging Akselrod-Ballrin (CVPR 2006), Corso et al. (MICCAI 2006, TMI 2008)
- Extended to videos Xu and Corso (CVPR 2012, ECCV 2012)
- Efficient, multiscale process inspired by Algebraic Multigrid optimization.
- Results in a pyramid of recursively coarsened graphs that capture multiscale properties of the data.
- Affinities are calculated at each level of the graph.
- Statistics in each graph node are agglomerated up the hierarchy.



Segmentation by Weighted Aggregation

- Finest layer induced by pixel/voxel lattice
  - 4/6-neighbor connectivity
  - Node properties s_u set according to multimodal image intensities.
  - Affinities initialized by L1-distance: $w_{uv} = \exp\left(-\theta|s_u - s_v|_1\right)$
- Superscripts on graph denotes level in a pyramid of graphs. $\mathcal{G} = \{G^t : t = 0, \ldots, T\}$

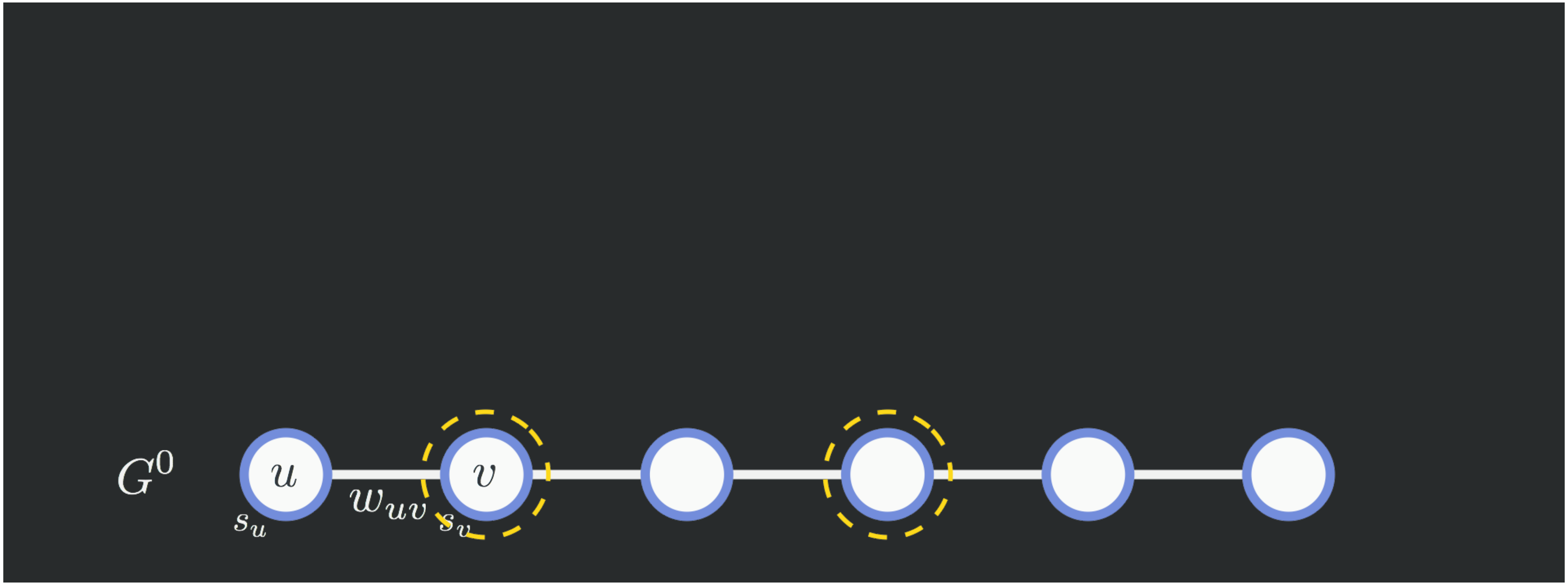


Segmentation by Weighted Aggregation

- Select a representative set of nodes satisfying

$$\sum_{v \in \mathcal{R}^t} w_{uv} \geq \beta \sum_{v \in \mathcal{V}^t} w_{uv}$$

  - i.e., all nodes in finer level have strong affinity to nodes in coarser.

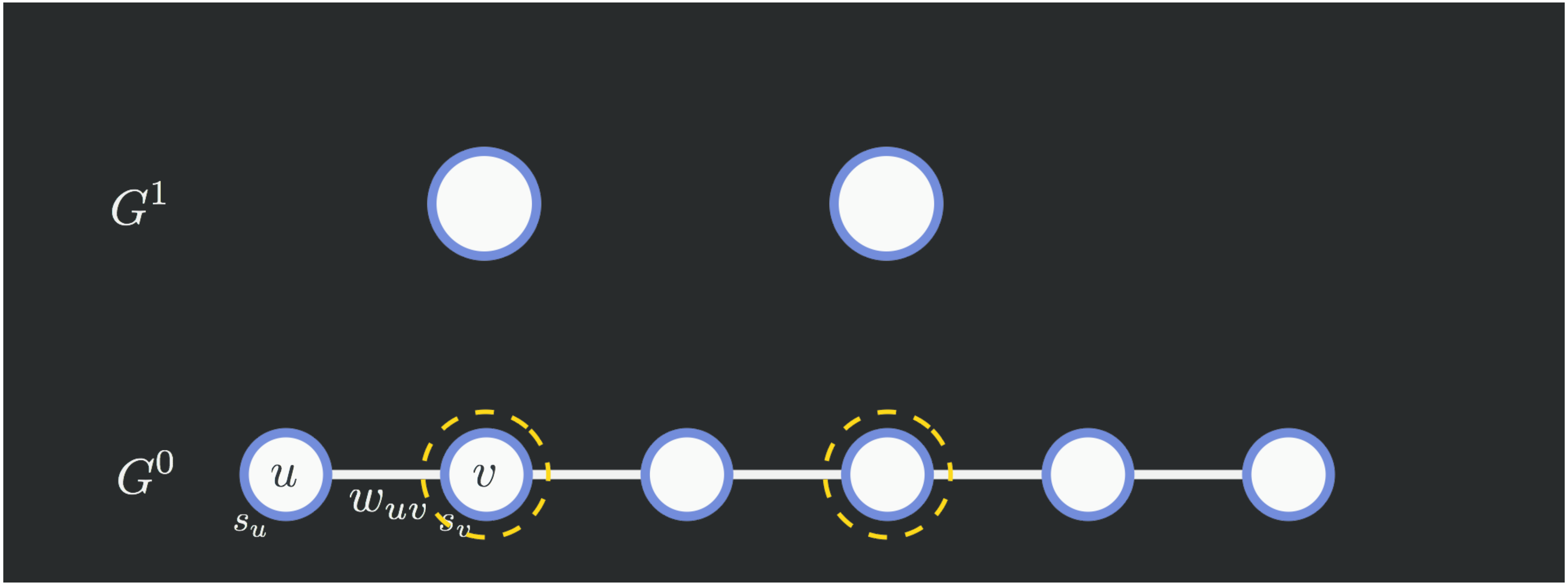


Segmentation by Weighted Aggregation

- Select a representative set of nodes satisfying

$$\sum_{v \in \mathcal{R}^t} w_{uv} \geq \beta \sum_{v \in \mathcal{V}^t} w_{uv}$$

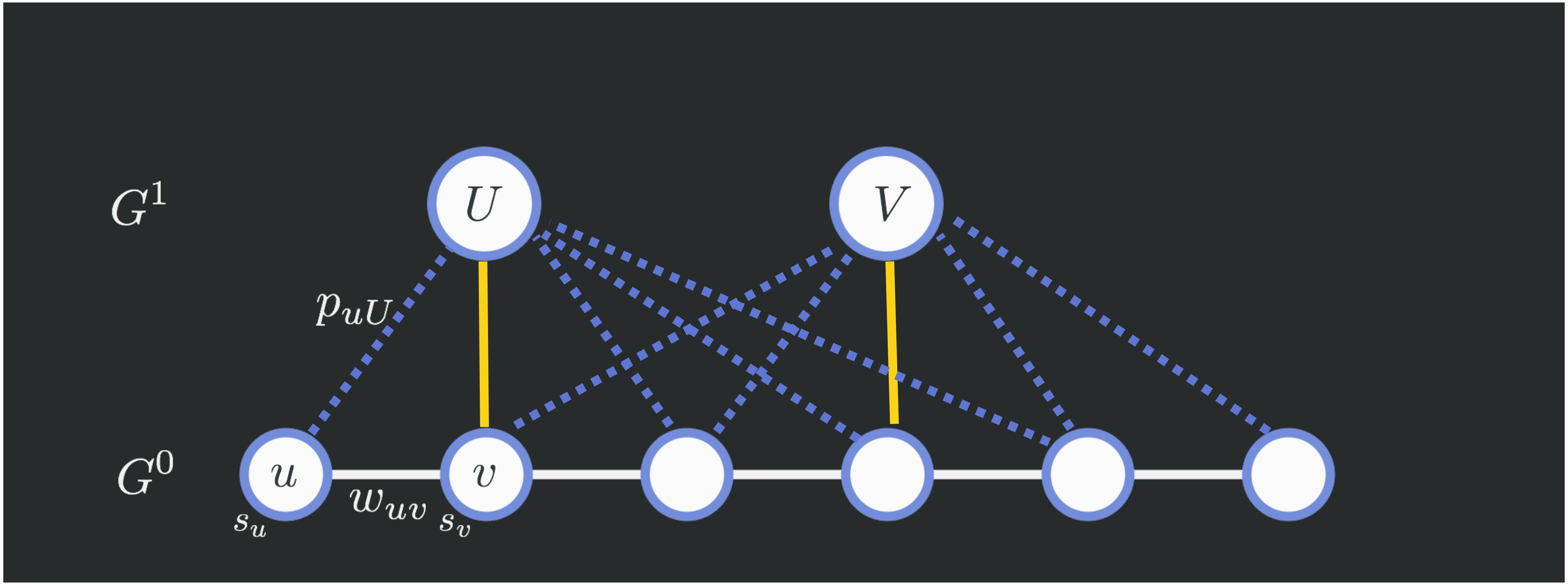
  - i.e., all nodes in finer level have strong affinity to nodes in coarser.
- Begin to define the graph $G^1 = \{\mathcal{V}^1, \mathcal{E}^1\}$



Segmentation by Weighted Aggregation

- Compute interpolation weights between coarse and fine levels

$$p_{uU} = \frac{w_{uU}}{\sum_{V \in \mathcal{V}^{t+1}} w_{uV}}$$



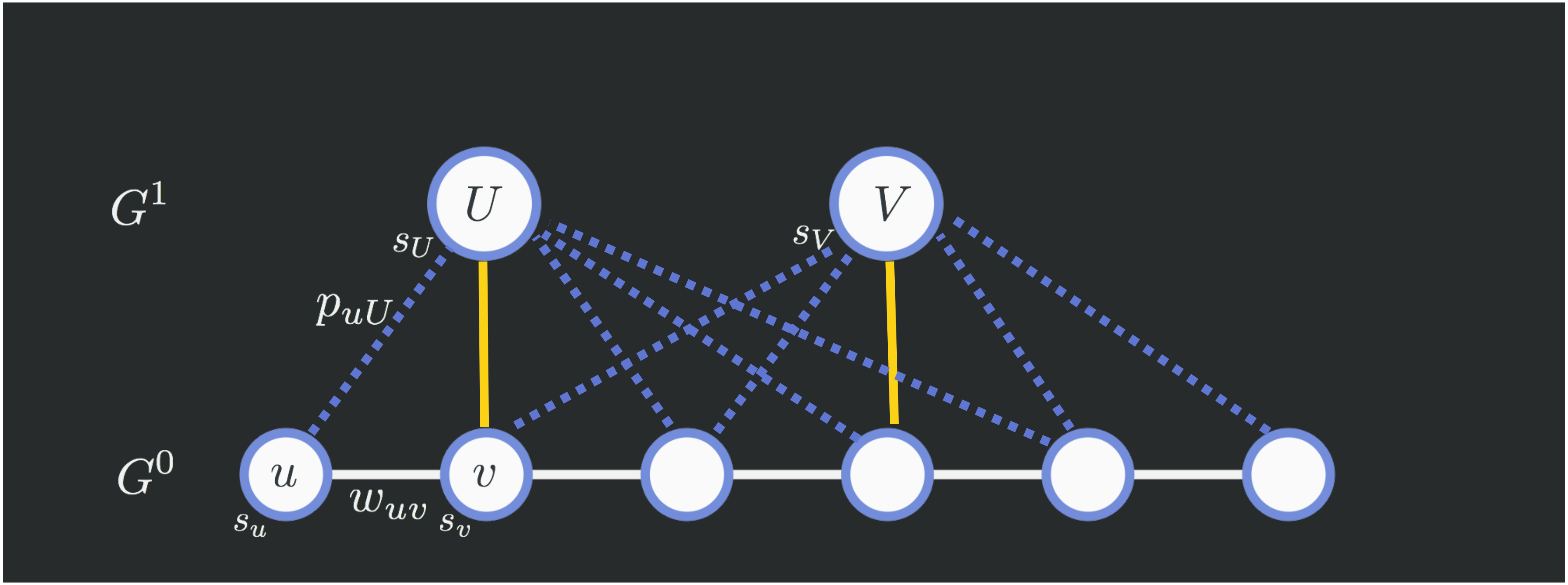
Segmentation by Weighted Aggregation

- Compute interpolation weights between coarse and fine levels

$$p_{uU} = \frac{w_{uU}}{\sum_{V \in \mathcal{V}^{t+1}} w_{uV}}$$

- Accumulate statistics at the coarse level

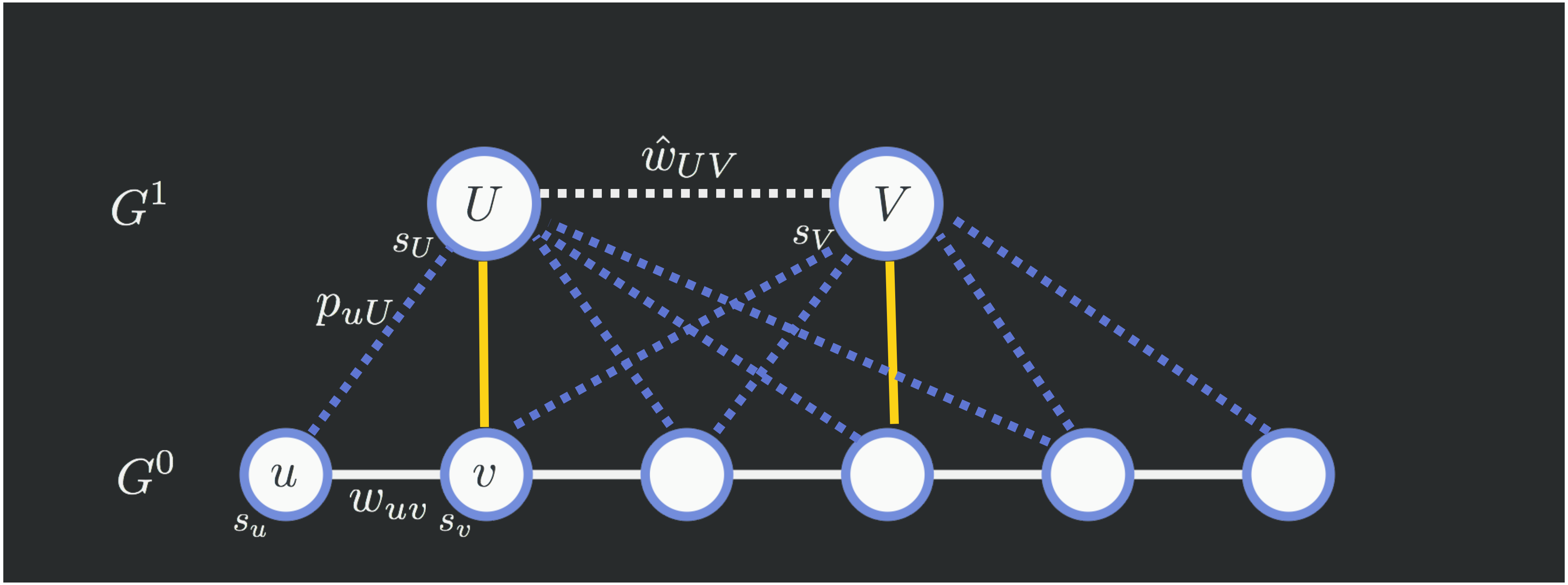
$$s_U = \sum_{u \in \mathcal{V}^t} \frac{p_{uU}\, s_u}{\sum_{v \in \mathcal{V}^t} p_{vU}}$$



Segmentation by Weighted Aggregation

- Interpolate affinity from finer levels

$$\hat{w}_{UV} = \sum_{(u \neq v) \in \mathcal{V}^t} p_{uU} w_{uv} p_{uV}$$



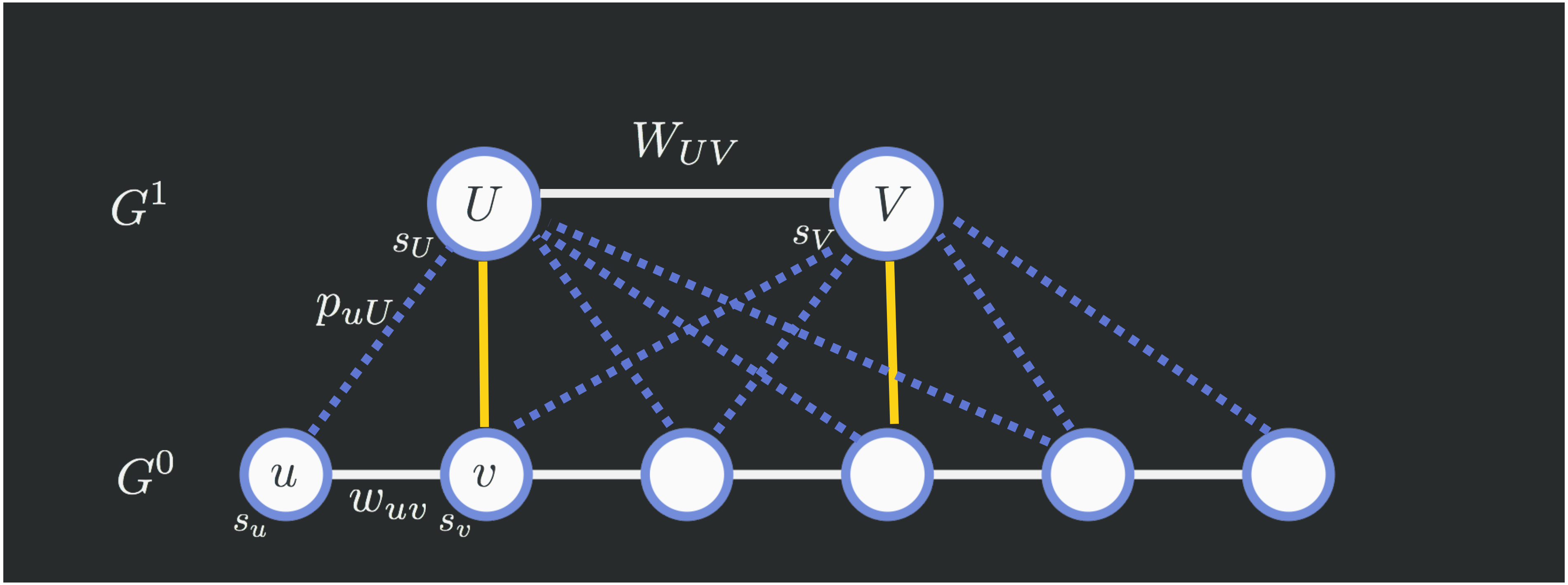
Segmentation by Weighted Aggregation

- Interpolate affinity from finer levels.

$$\hat{w}_{UV} = \sum_{(u \neq v) \in \mathcal{V}^t} p_{uU} w_{uv} p_{uV}$$

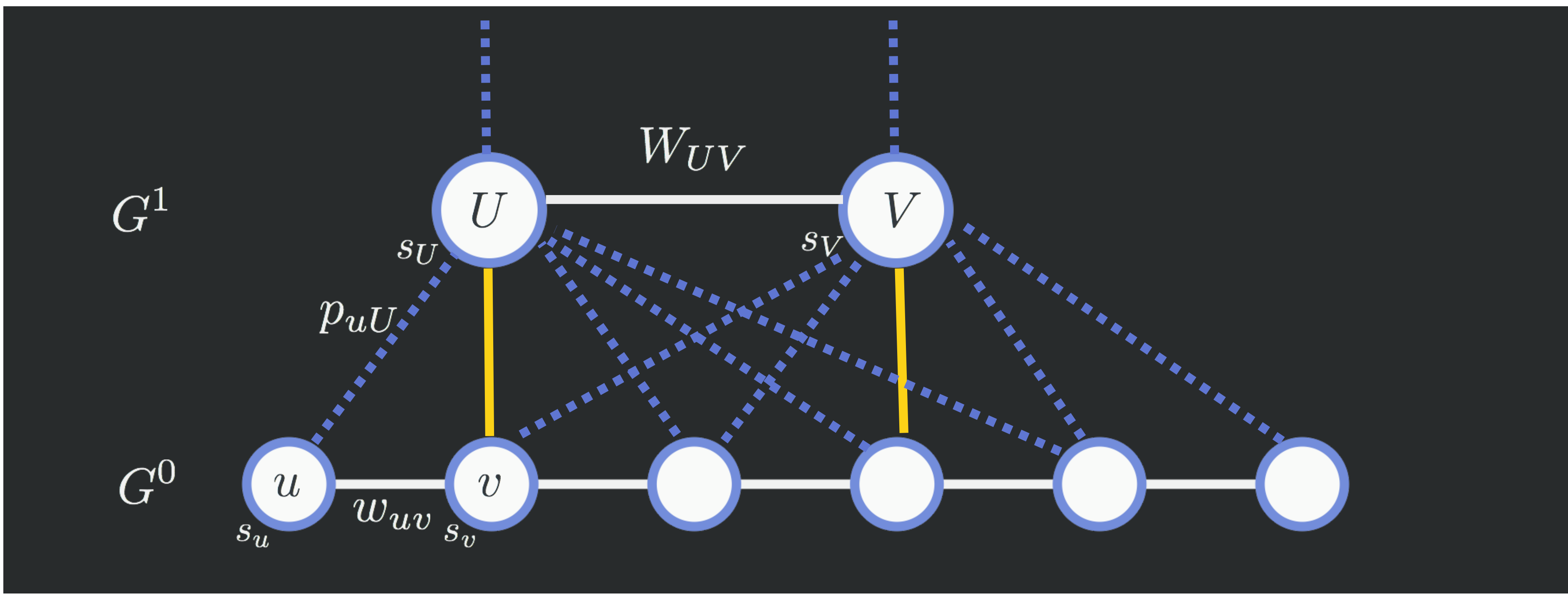
- Use coarse affinity to modulate the interpolated affinity.

$$W_{UV} = \hat{w}_{UV} \exp\left(-D(s_U, s_V; \theta)\right)$$



Segmentation by Weighted Aggregation

- Repeat ...



Bayesian Affinities

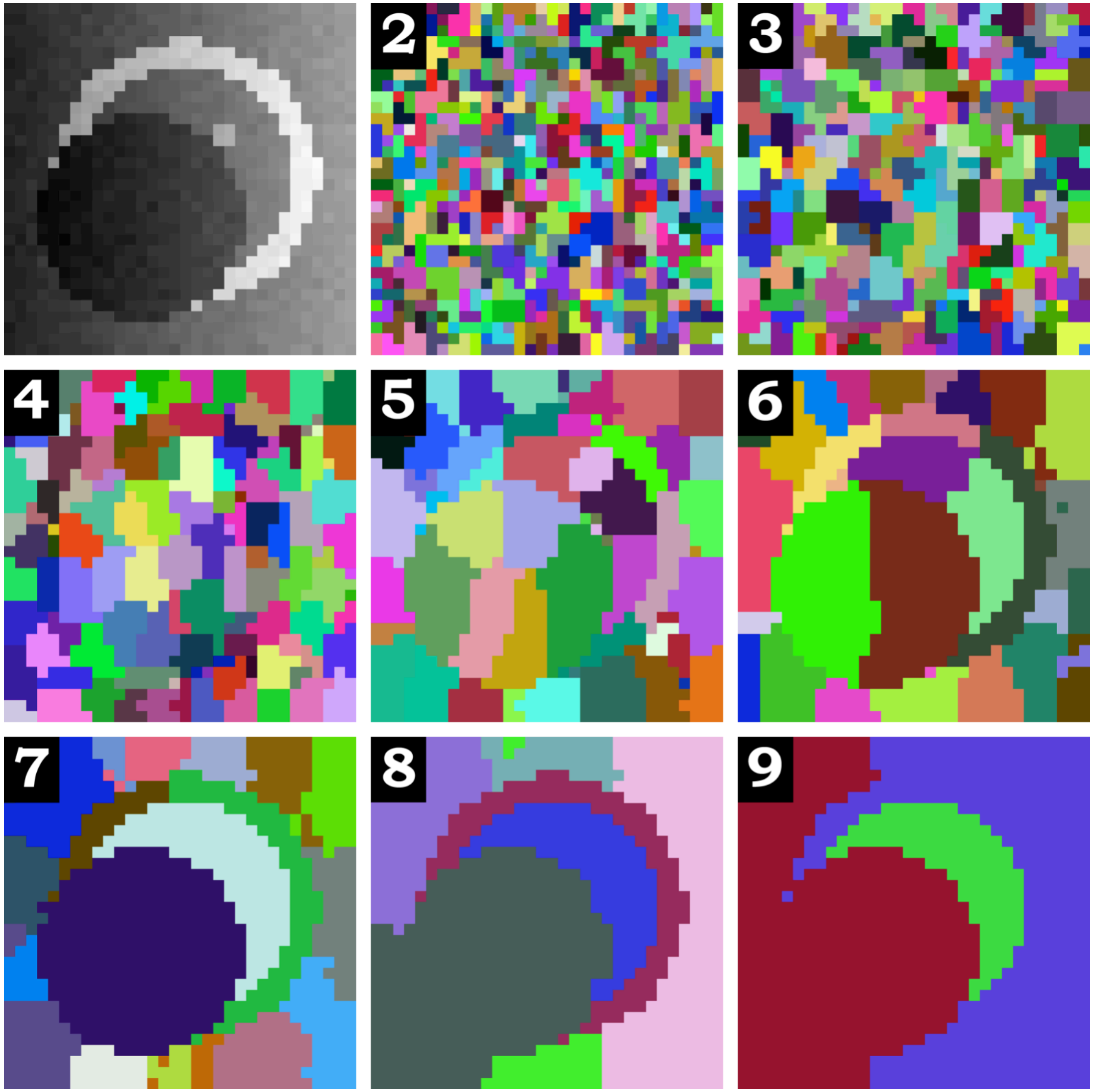
- Standard affinity calculation is based on simple features, such as the L1-distance of intensities as in the example.
- Affinity can be extended using metric learning
  - LMNN [Weinberger et al. NIPS05], ITML [Davis et al. ICML07], RFD [Xiong et al. KDD12]
- Or Bayesian view of affinity [Corso, Yuille TMI 2008]
  - Introduce a binary grouping random variable X_{uv} .

$$P(X_{uv}|s_u, s_v) = \sum_{m_u}\sum_{m_v} P(X_{uv}|s_u, s_v, m_u, m_v)P(m_u, m_v|s_u, s_v) \quad ,$$

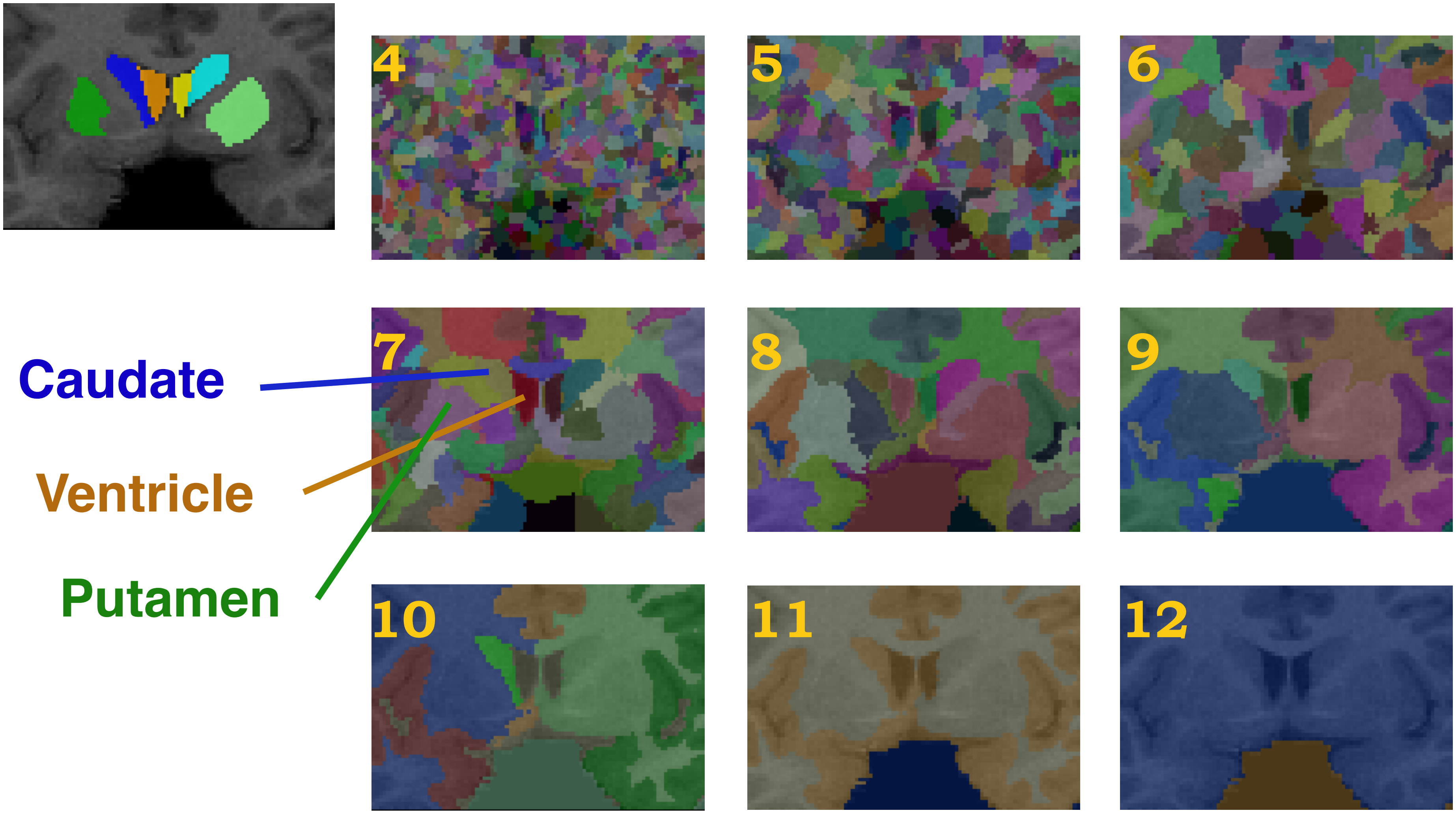
$$\propto \sum_{m_u}\sum_{m_v} P(X_{uv}|s_u, s_v, m_u, m_v)P(s_u, s_v|m_u, m_v)P(m_u, m_v) \quad ,$$

$$= \sum_{m_u}\sum_{m_v} \underbrace{P(X_{uv}|s_u, s_v, m_u, m_v)}_{\text{Model Specific Measurement}}\underbrace{P(s_u|m_u)P(s_v|m_v)}_{\text{Node Likelihoods}}\underbrace{P(m_u, m_v)}_{\text{Class Prior}}$$

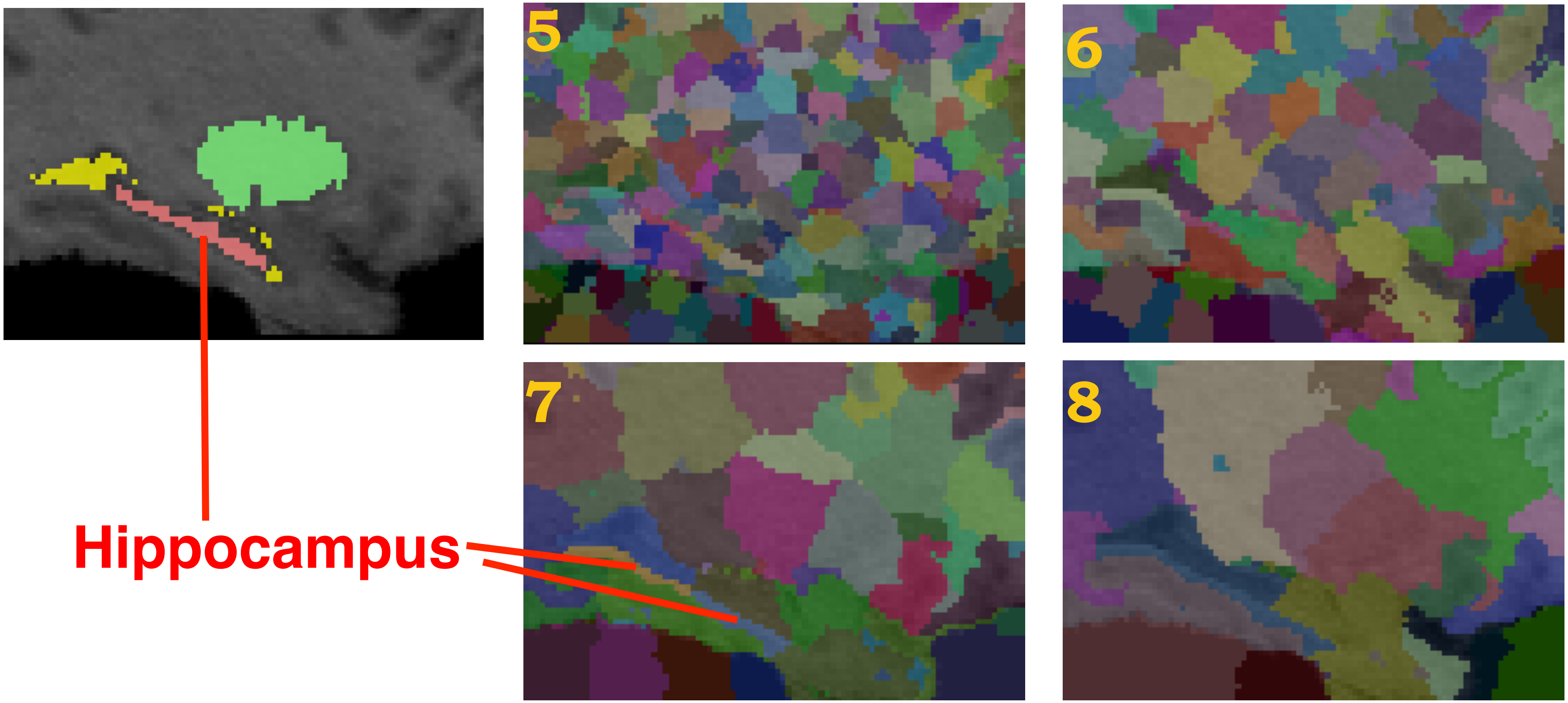
Example on Synthetic Grayscale Image



Example of the Segmentation Pyramid



Example of the Segmentation Pyramid



Next Lecture: Model-Fitting and Contours

- Readings: FP 10; SZ 4.3, 5.1