Characterizing Temporal Dynamics of Visual Perceptual Grouping

Mehmet Yörüten

Supervised by: Shuchen Wu, Felix Wichmann, Eric Schulz

M.Sc Student of Neural Information Processing

University of Tübingen

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

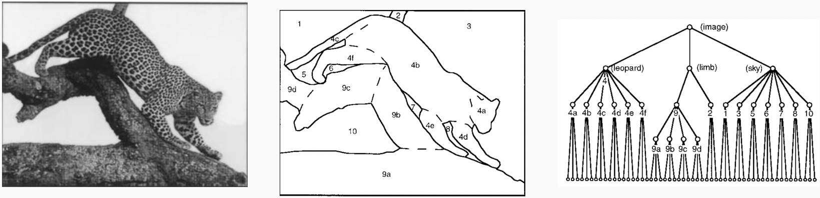
MAX-PLANCK-INSTITUT FÜR BIOLOGISCHE KYBERNETIK

CPILab

Outline

Visual Grouping

Visual Grouping



(a) (b) (c)

Figure 1: An image of a leopard (1a) can be segmented into several pieces (1b). There are endless single parts to these scene (1c) but at the first glance we perceive only the certain parts of these image as segments.

Adapted from Palmer (2002)

Gestalt Principles

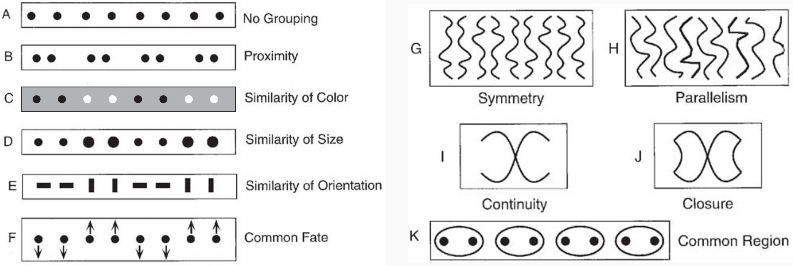


Figure 2: Some Gestalt Principles describing the way we assemble the elements into groups[1].

[1] (Stephen E Palmer [2002]. “Perceptual Organization in Vision”. In: *Stevens’ Handbook of Experimental Psychology*)

Global Precedence

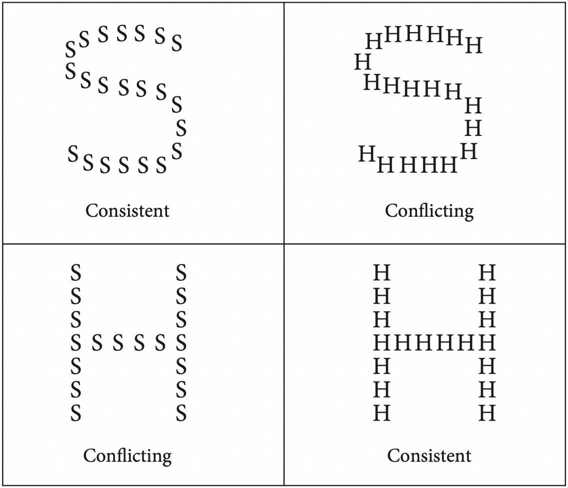


Figure 3: Faster identification of the global cues than the local cues[2].

[2] (Ruth Kimchi [2015]. "The perception of hierarchical structure". In: *Oxford handbook of perceptual organization*, pp. 129–149)

MOTIVATION

- Studying hierarchical visual grouping

Motivation

- Studying hierarchical visual grouping
- How does hierarchical perceptual grouping behavior change with increasing exposure time?

Motivation

- Studying hierarchical visual grouping
- How does hierarchical perceptual grouping behavior change with increasing exposure time?
- Relating exposure time to computational resources attributed to an image segmentation problem.

Motivation

- Studying hierarchical visual grouping
- How does hierarchical perceptual grouping behavior change with increasing exposure time?
- Relating exposure time to computational resources attributed to an image segmentation problem.

Hypothesis

As more computational resources are allocated, perceptual grouping behavior gets refined and finer segments can be detected.

Stimuli

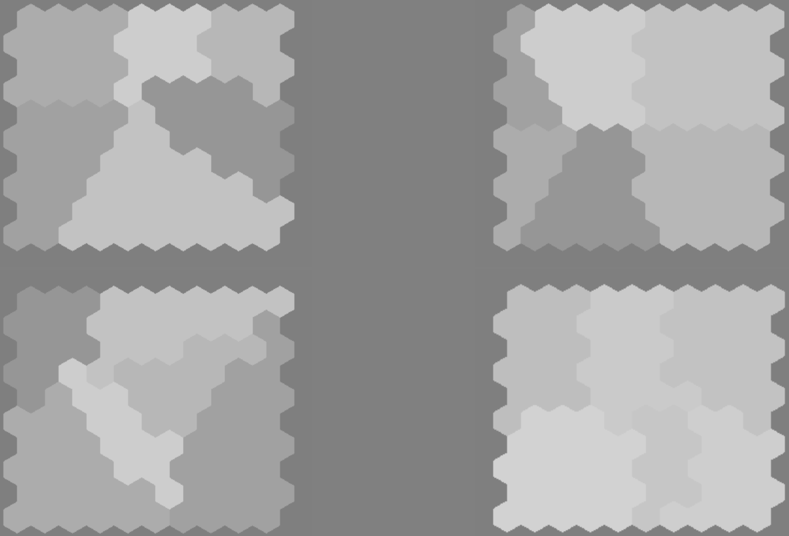
STIMULI

CRITERIA:

- We need images that induce no bias
- There should be no semantic meaning assigned to the images
- There should be no effect of object recognition

GENERATING IMAGES

GENERATING IMAGES

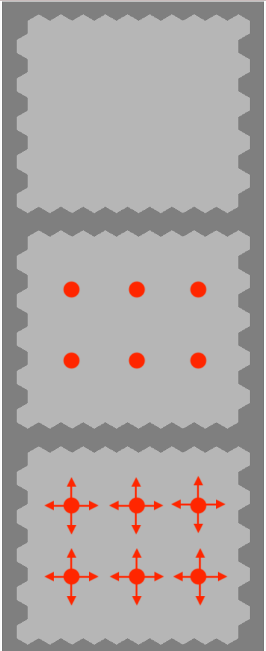


GENERATING IMAGES

1. 10 x 10 grid pattern made of hexagons.

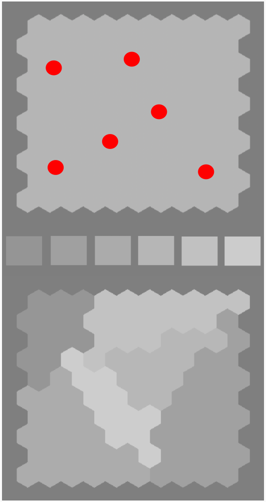
2. Assign 6 equally distant seed on the grid.

3. Sample a variance from a normal distribution and add to the seed coordinates.



4. Establish regions for each one of the seeds using Voronoi Diagram.

5. Map each of the groups to intensity values (contrast ratio 0.086).



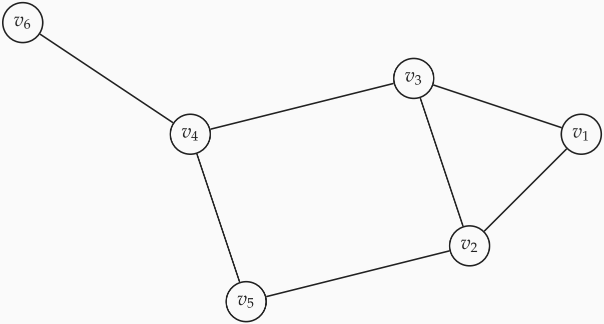
Generating Segments

We need to find an algorithm that:

- Detects global segments prior to local segments.
- Has hierarchical structure.
- Represents the computational resources spent on the segmentation process.

Normalized Min Cut Algorithm

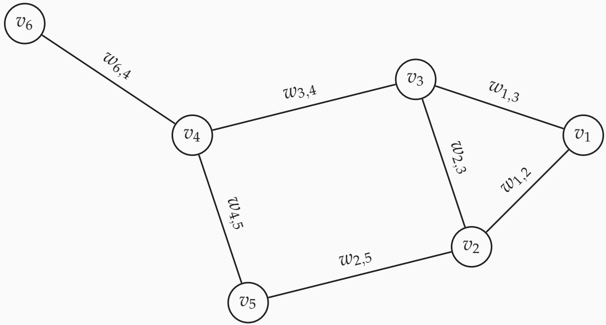
A graph $G = (V, E)$ is composed of a set *V* of vertices (nodes) and set *E* of edges.



$$V = \{v_1, v_2, ..., v_6\}$$

$$E = \{(v_1, v_2), (v_1, v_3), (v_2, v_3), (v_2, v_5), (v_3, v_4), (v_4, v_5), (v_4, v_6)\}$$

A graph can be detailed by assigning weights to the edges.



For undirected graphs:

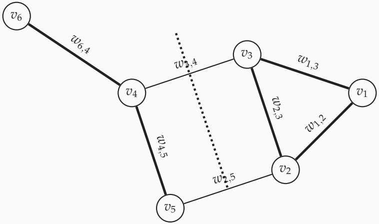
$$w_{i,j} = w_{j,i}$$

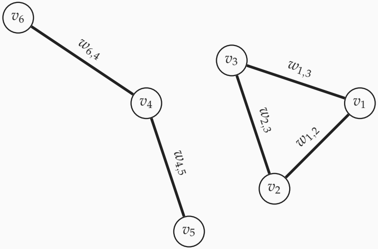
Cut

A graph $G = (V, E)$ can be partitioned into two disjoint sets A and B:

$A \cup B = V$ and $A \cap B = \emptyset$

$$cut(A, B) = \sum_{u \in A, v \in B} w(u, v) \quad (1)$$





The optimal bipartitioning is the one that minimizes *cut* values.

$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)} \tag{2}$$

$$cut(A,B) = \sum_{u \in A, v \in B} w(u,v) \tag{3}$$

$$assoc(A,V) = \sum_{u \in A, t \in V} w(u,t) \tag{4}$$

$cut(A,B)$: Total weight of the edges between two clusters
$assoc(A,V)$: Total connection from nodes in A to all nodes in the graph
$assoc(B,V)$: Total connection from nodes in B to all nodes in the graph

[3] (Jianbo Shi and Jitendra Malik [2000]. "Normalized cuts and image segmentation". In: *IEEE Transactions on pattern analysis and machine intelligence* 22.8, pp. 888–905. ISSN: 0162-8828)

Normalized Min Cut - *Optimum Bipartition*

$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)} \tag{5}$$

$$min_x Ncut(x) = min_y \frac{y^T(D-W)-y}{y^T Dy} \tag{6}$$

$$(D-W)y = \lambda Dy \tag{7}$$

- W (Adjacency Matrix) is $N \times N$ symmetrical matrix with $W(i,j) = w_{i,j}$.
- D (Degree Matrix) is $N \times N$ diagonal matrix with $d_i = \sum_j w(i,j)$ on its diagonal.

- Represent each pixel as a node.
- Connect each pair of pixels by an edge.

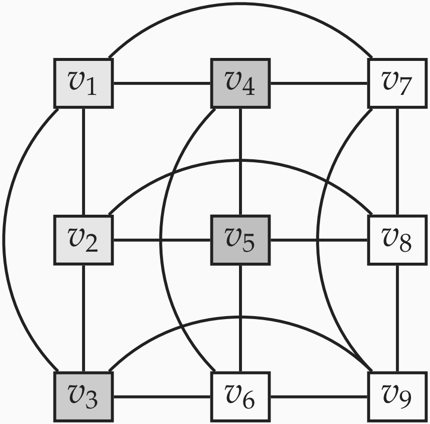


Figure 5: Example representation of 3x3 image with a graph network

- Assign weights to each edge.
- The weights represent the similarity of two pixels.

$$w_{ij} = e^{\frac{-||F_{(i)} - F_{(j)}||_2^2}{\sigma_I^2}} * \begin{cases} e^{\frac{-||X_{(i)} - X_{(j)}||_2^2}{\sigma_X^2}}, & \text{if } ||X_{(i)} - X_{(j)}||_2 < r \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

- Assign weights to each edge.
- The weights represent the similarity of two pixels.

$$w_{ij} = e^{\frac{-||F_{(i)} - F_{(j)}||_2^2}{\sigma_I^2}} * \begin{cases} e^{\frac{-||X_{(i)} - X_{(j)}||_2^2}{\sigma_X^2}}, & \text{if } ||X_{(i)} - X_{(j)}||_2 < r \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

- F is feature vector based on **brightness**, color, texture, or motion information.

- Assign weights to each edge.
- The weights represent the similarity of two pixels.

$$w_{ij} = e^{\frac{-||F_{(i)} - F_{(j)}||_2^2}{\sigma_I^2}} * \begin{cases} e^{\frac{-||X_{(i)} - X_{(j)}||_2^2}{\sigma_X^2}}, & \text{if } ||X_{(i)} - X_{(j)}||_2 < r \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

- F is feature vector based on **brightness**, color, texture, or motion information.
- r is distance between two pixel nodes i and j.
- $w_{ij} = 0$ when i and j are apart more than r
- As σ_I gets larger, the differences between the weights decrease

- Assign computed weights to the edges.

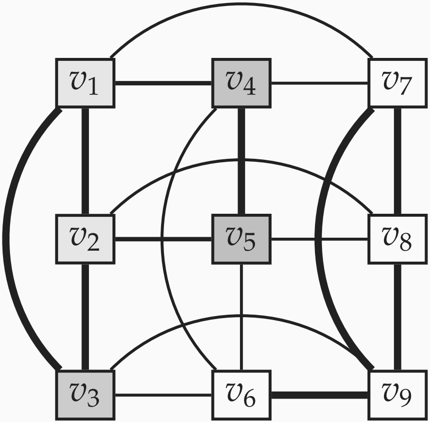
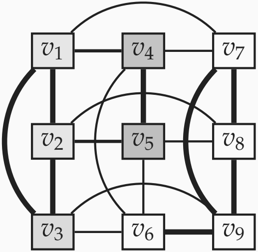
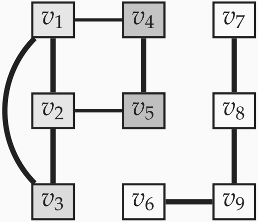


Figure 6: Representation of 3x3 image with a weighted graph network

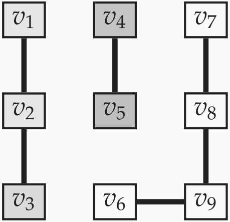
Segmentation with Norm Min Cut



(a) Graph before cut



(b) Graph after the first cut



(c) Graph after the second cut

Figure 7: An example segmentation of a graph

Getting Segments

Getting Segments

1. Apply Normalized Min Cut algorithm to find the ground truth segments.

Getting Segments - *Normalized Min Cut*

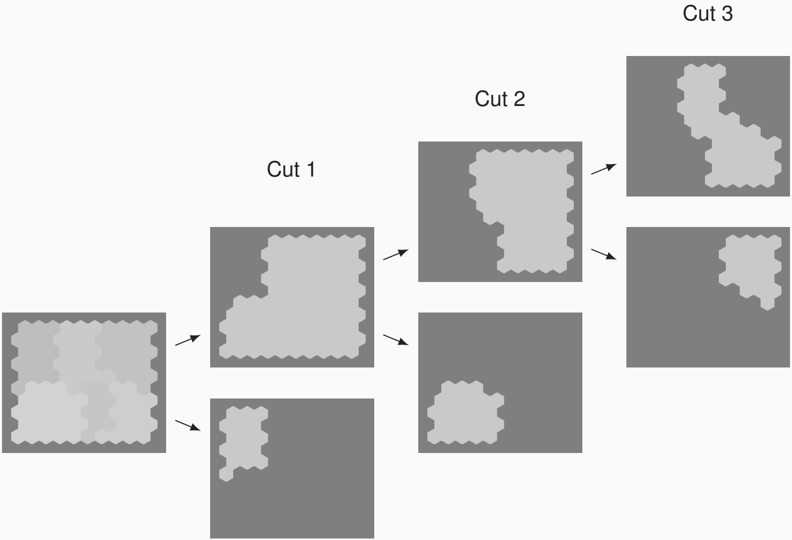


Figure 8: First three steps of segmentation

Getting Segments

1. Apply Normalized Min Cut algorithm to find the ground truth segments.
2. Assign cut orders depending on the *Ncut* value.

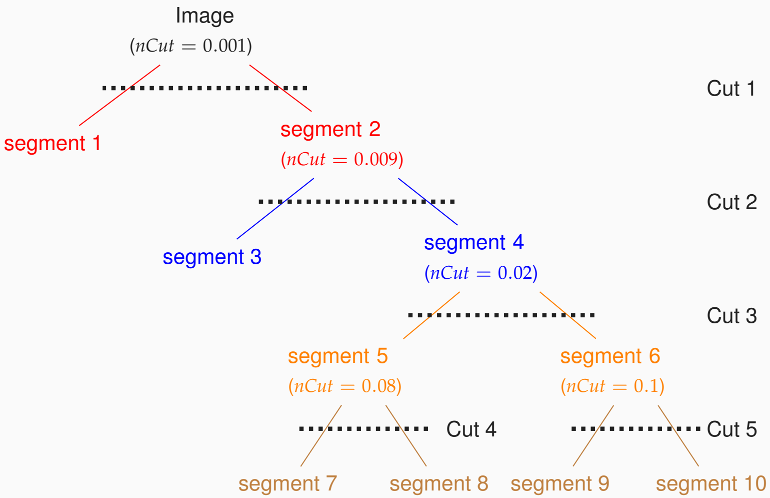


Figure 9: Found segments

Getting Segments

1. Apply Normalized Min Cut algorithm to find the ground truth segments.
2. Assign cut orders depending on the *Ncut* value.
3. Stop when all the segments are found.

Getting Segments

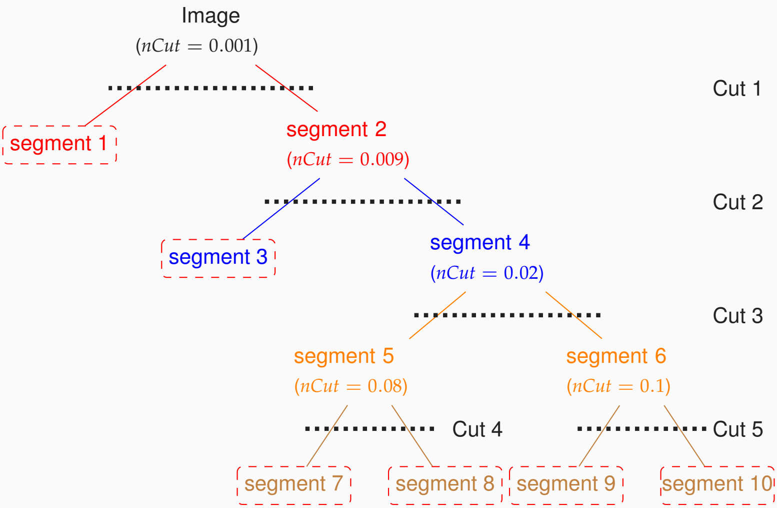


Figure 10: Found segments

Getting Segments

1. Apply Normalized Min Cut algorithm to find the ground truth segments.
2. Assign cut orders depending on the *Ncut* value.
3. Stop when all the segments are found.
4. Present segments depending on the cut order (5 cuts).
5. Display the image with the average intensity of the initial image.

DISPLAYING SEGMENTS - *Cut Order 4*

pictures/seg_5_c

Experimental Paradigm

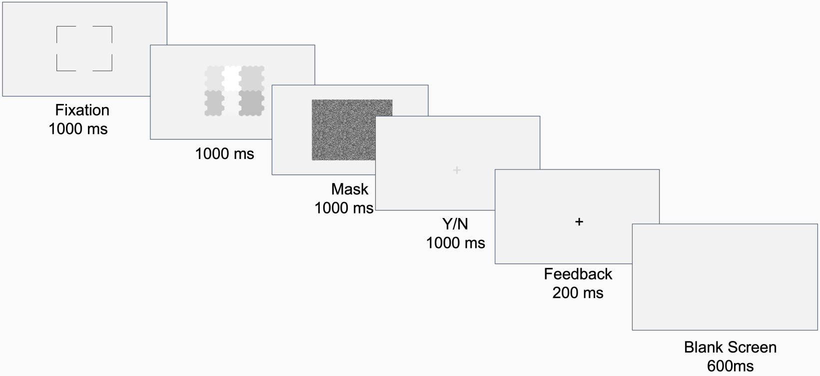


Figure 12: Contrast Sensitivity Experiment

- Present 10x10 grid pattern with six different intensities
- Segment borders remain same
- Staircase Method (QUEST) to get contrast difference between segments (δI)

Contrast Sensitivity Task - Psychometric Function

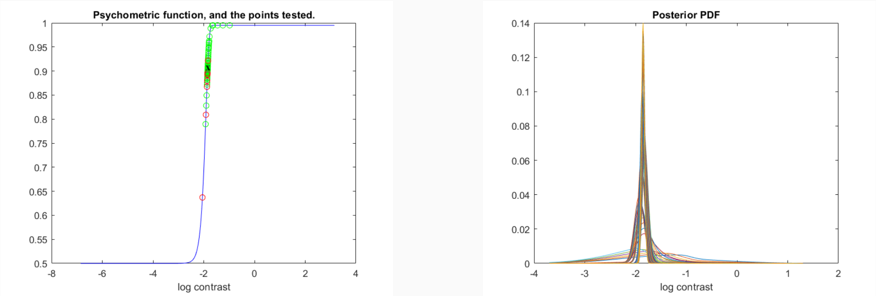
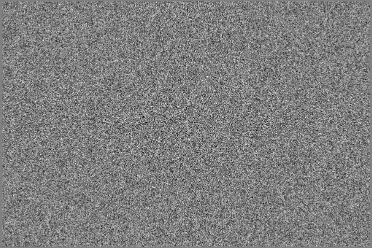


Figure 13: Psychometric and posterior distribution (PDF) functions

- Number of Trials: 100
- Threshold Criterion: 0.9
- Beta: 3.5 (Slope of the curve)
- Delta: .01 (Probability of making mistakes above threshold)
- Gamma: .5 (Yes/No Task)
- Create graded intensity levels with fixed intervals



pictures/seg_5_cut4.png

PROCEDURE

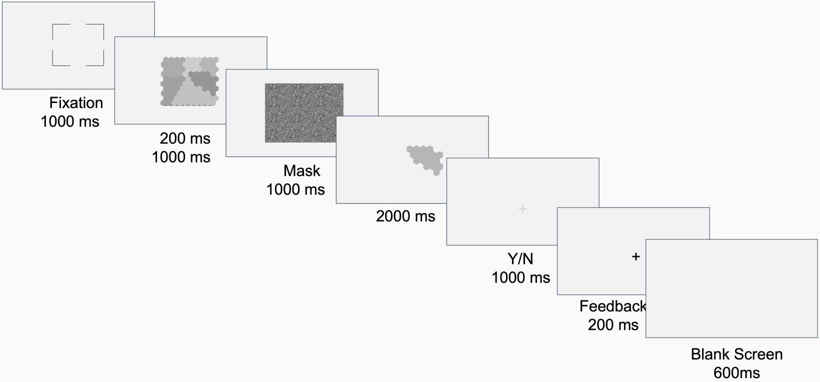


Figure 14: Experimental Procedure

- Present 10x10 grid pattern with six different intensities assigned randomly to each of groups
- Present segments in average intensity value of the image
- Ask ”Did you see the presented segment in the previously shown image?” (Y/N)
- Measure accuracy and reaction time (RT)

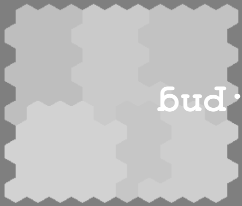
CORRECT ANSWER - *Cut Order 1*

CORRECT ANSWER - *Cut Order 3*

Control Condition

- Randomly choose a segment matching the cut order, from the pre-generated image set.
- Rotate $90°$, $180°$, or $270°$.
- Do cross correlation with the borders.
- Check if the foil borders overlap with any of the segments from the original image.
- If the overlap exceeds threshold, find another segment from another trial.

CONTROL CONDITION - *Cut Order 3*



pictures/cut3_control.png

Design

- 5 (Cut Order) x 2 (Exposure Time) x 2 (Control) Conditions
- 25 Blocks and 2000 Trials in total.
- Show average accuracy in the end of the blocks
- Experiment duration $\approx$ 4 hours
- High accuracy & shorter RT expected for the segments in the higher level

Variables

- Cut No
- Exposure Time
- Segment Size
- Segment Center’s Distance to Center of the Image
- Intensity Difference
- Intensity Sd
- Ncut Value
- Stability
- Original Segment Intensity – Presented Intensity

Parameters

Parameters

$$w_{ij} = e^{\frac{-||F_{(i)} - F_{(j)}||_2^2}{\sigma_I^2}} * \begin{cases} e^{\frac{-||X_{(i)} - X_{(j)}||_2^2}{\sigma_X^2}}, & \text{if} ||X_{(i)} - X_{(j)}||_2 < r \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

- Number of cuts can be translated into computational resources spent. More iteration is needed to detect finer segments.
- First cut refers to the most salient segments.
- σ_I: Let more distinct segments to be grouped together.
- σ_x: Effect of distance between two pixels on the weight.
- $nCut$ Value: Where to stop segmentation

PILOT - EXPOSURE TIMES

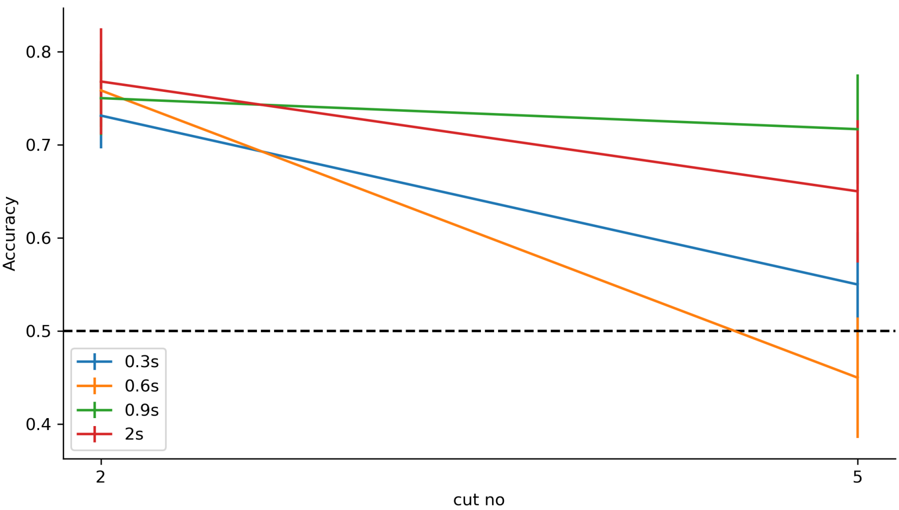


Figure 18: Compare Exposure Times

Initial Results

SW - Accuracy

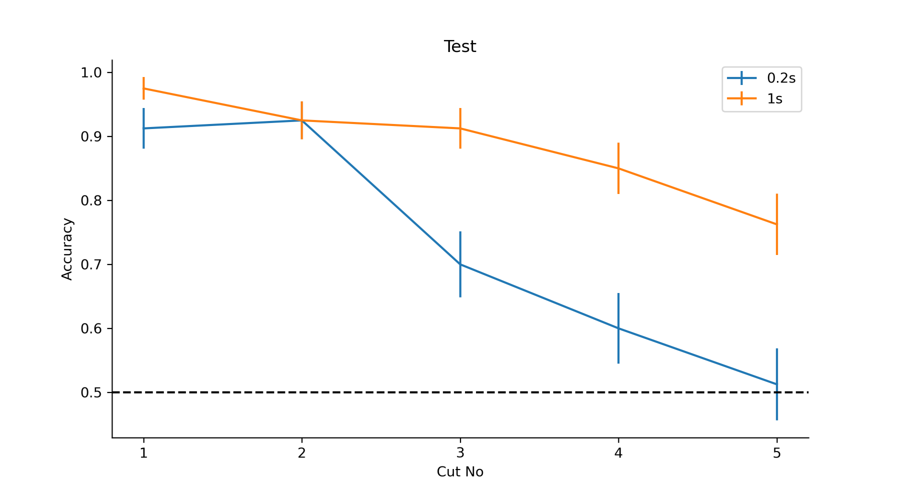


Figure 19: Accuracy over cut numbers in test trials

Effect on Accuracy

pictures/sw_coefficients_int.png

SW - Accuracy

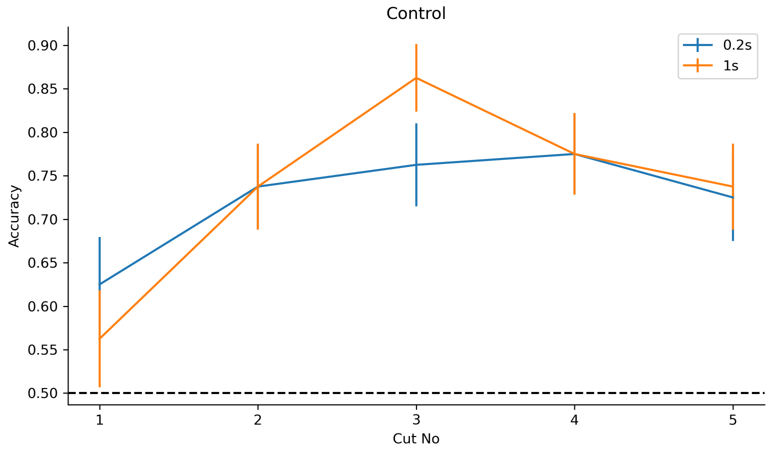


Figure 21: Accuracy over cut numbers in control trials

Effect on Accuracy

pictures/sw_cont_coefficients_othrs_sprt.png

MY - Accuracy

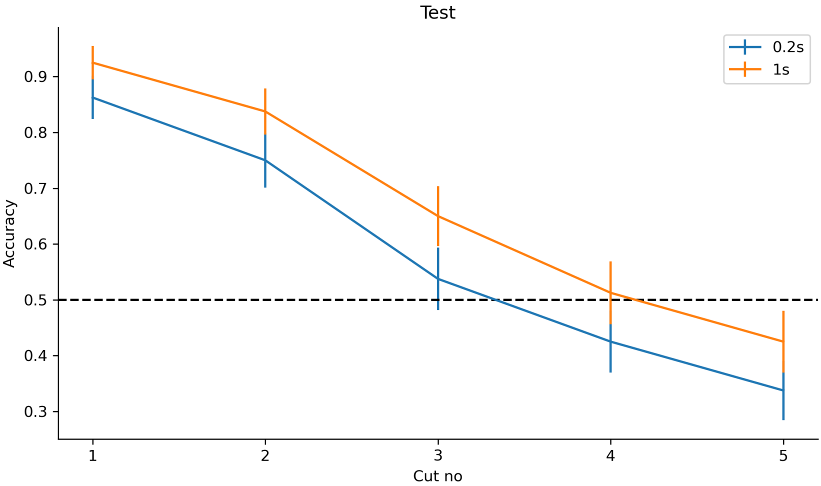


Figure 23: Accuracy over cut number in test trials

MY - COEFFICIENTS

pictures/my_coefficients_int.png

MY - Accuracy

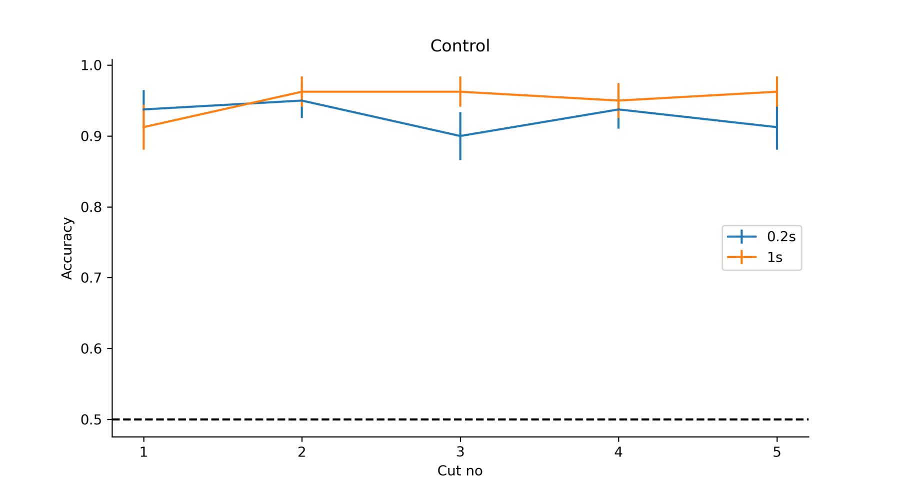


Figure 25: Accuracy over cut number in control trials

MY - COEFFICIENTS

pictures/my_cont_coefficients_othrs_sprt.png

MY - KEY PRESSES

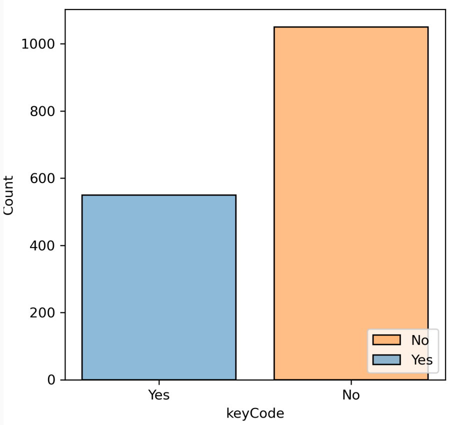


Figure 27: Key Press Distribution

References

- Kimchi, Ruth (2015). "The perception of hierarchical structure". In: *Oxford handbook of perceptual organization*, pp. 129–149.
- Palmer, Stephen E (2002). "Perceptual Organization in Vision". In: *Stevens' Handbook of Experimental Psychology*.
- Shi, Jianbo and Jitendra Malik (2000). "Normalized cuts and image segmentation". In: *IEEE Transactions on pattern analysis and machine intelligence* 22.8, pp. 888–905. ISSN: 0162-8828.

Questions

Thank You!

Questions?

Segment Analyses

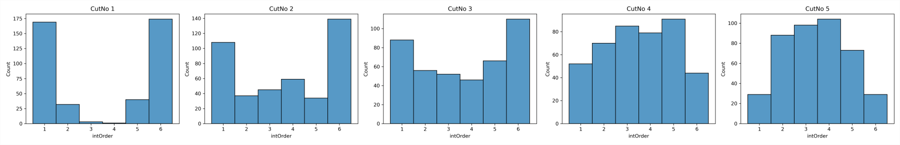
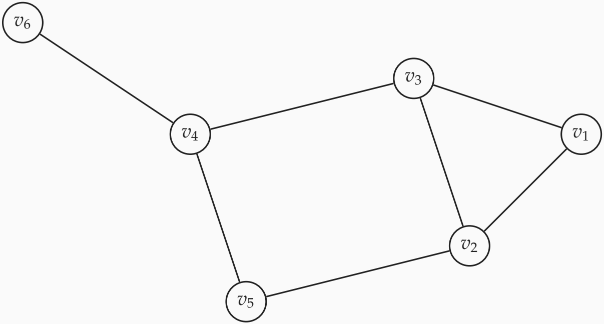


Figure 28: Distribution of intensity orders for each cut number

Graph Theory

Graph Theory

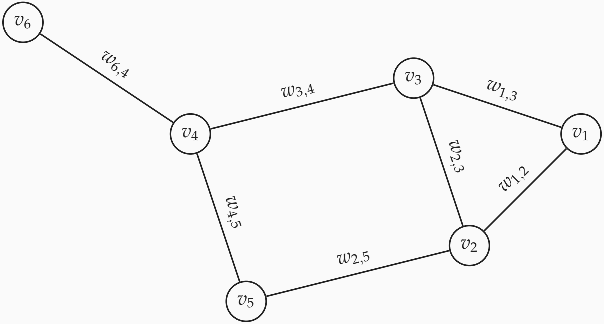
A graph $G = (V, E)$ is composed of a set V of vertices (nodes) and set E of edges.



$$V = \{v_1, v_2, \ldots, v_6\}$$

$$E = \{(v_1, v_2), (v_1, v_3), (v_2, v_3), (v_2, v_5), (v_3, v_4), (v_4, v_5), (v_4, v_6)\}$$

A graph can be detailed by assigning weights to the edges.

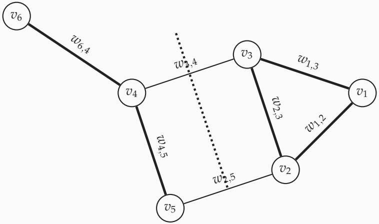


For undirected graphs:

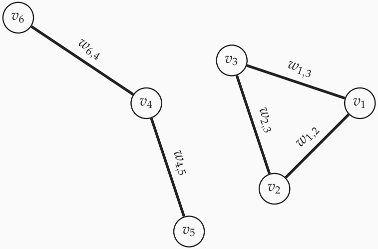
$$w_{i,j} = w_{j,i}$$

CUT

A graph $G = (V, E)$ can be partitioned into two disjoint sets A and B:



$$A \cup B = V \text{ and } A \cap B = \emptyset$$



$$cut(A, B) = \sum_{u \in A, v \in B} w(u, v) \quad (10)$$

The optimal bipartitioning is the one that minimizes cut values.

Normalized Min Cut

Normalized Min Cut [4]

$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)} \tag{11}$$

$$cut(A,B) = \sum_{u \in A, v \in B} w(u,v) \tag{12}$$

$$assoc(A,V) = \sum_{u \in A, t \in V} w(u,t) \tag{13}$$

$cut(A,B)$: Total weight of the edges between two clusters
$assoc(A,V)$: Total connection from nodes in A to all nodes in the graph
$assoc(B,V)$: Total connection from nodes in B to all nodes in the graph

[4] (Jianbo Shi and Jitendra Malik [2000]. "Normalized cuts and image segmentation". In: *IEEE Transactions on pattern analysis and machine intelligence* 22.8, pp. 888–905. ISSN: 0162-8828)

NORMALIZED MIN CUT - *Optimum Bipartition*

$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)} \tag{14}$$

$$min_x Ncut(x) = min_y \frac{y^T(D-W)-y}{y^T Dy} \tag{15}$$

$$(D-W)y = \lambda Dy \tag{16}$$

- W (Adjacency Matrix) is $N \times N$ symmetrical matrix with $W(i,j) = w_{i,j}$.
- D (Degree Matrix) is $N \times N$ diagonal matrix with $d_i = \sum_j w(i,j)$ on its diagonal.

- Assign weights to each edge.
- The weights represent the similarity of two pixels.

$$w_{ij} = e^{\frac{-||F_{(i)} - F_{(j)}||_2^2}{\sigma_I^2}} * \begin{cases} e^{\frac{-||X_{(i)} - X_{(j)}||_2^2}{\sigma_X^2}}, & \text{if } ||X_{(i)} - X_{(j)}||_2 < r \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

- Assign weights to each edge.
- The weights represent the similarity of two pixels.

$$w_{ij} = e^{\frac{-||F_{(i)} - F_{(j)}||_2^2}{\sigma_I^2}} * \begin{cases} e^{\frac{-||X_{(i)} - X_{(j)}||_2^2}{\sigma_X^2}}, & \text{if } ||X_{(i)} - X_{(j)}||_2 < r \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

- F is feature vector based on **brightness**, color, texture, or motion information.

- Assign weights to each edge.
- The weights represent the similarity of two pixels.

$$w_{ij} = e^{\frac{-||F_{(i)} - F_{(j)}||_2^2}{\sigma_I^2}} * \begin{cases} e^{\frac{-||X_{(i)} - X_{(j)}||_2^2}{\sigma_X^2}}, & \text{if } ||X_{(i)} - X_{(j)}||_2 < r \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

- F is feature vector based on **brightness**, color, texture, or motion information.
- r is distance between two pixel nodes i and j.
- $w_{ij} = 0$ when i and j are apart more than r
- As σ_I gets larger, the differences between the weights decrease

NORM MIN CUT - *Summary*

1. Represent images with graph networks.
2. Solve generalized eigensystem $(D - W)y = \lambda Dy$.
3. Use the eigenvector with the second smallest eigenvalue. Assign values into two groups by considering a splitting point.
4. Check if the current partition should be subdivided by checking stability and $Ncut$ value.
5. Recursively repartition the segmented parts if necessary.

Parameters

Parameters

$$w_{ij} = e^{\frac{-||F_{(i)} - F_{(j)}||_2^2}{\sigma_I^2}} * \begin{cases} e^{\frac{-||X_{(i)} - X_{(j)}||_2^2}{\sigma_X^2}}, & \text{if} ||X_{(i)} - X_{(j)}||_2 < r \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

- Number of cuts can be translated into computational resources spent. More iteration is needed to detect finer segments.
- First cut refers to the most salient segments.
- σ_I: Let more distinct segments to be grouped together.
- σ_x: Effect of distance between two pixels on the weight.
- $nCut$ Value: Where to stop segmentation

Generating Next Stimulus

$$I_{i+1} = \begin{cases} I_i + I_i \cdot \delta I, & \text{if } i > 1 \text{ and } i < 6 \\ I_{init} + I_{init} \cdot \delta I, & \text{if } i = 1 \end{cases} \tag{19}$$

- Each trial get a new δI
- Generate 6 intensity values using δI and I_{init}
- $$I_{init} = 0.745 \tag{20}$$