



Visual Attention: low level and high level viewpoints

Fred Stentiford

University College London

<http://www.ee.ucl.ac.uk/~fstentif/>



Outline

- Prior research
- Issues
- A top-down viewpoint
- Application to an illusion
- A bottom-up approach to attention
- Conclusions

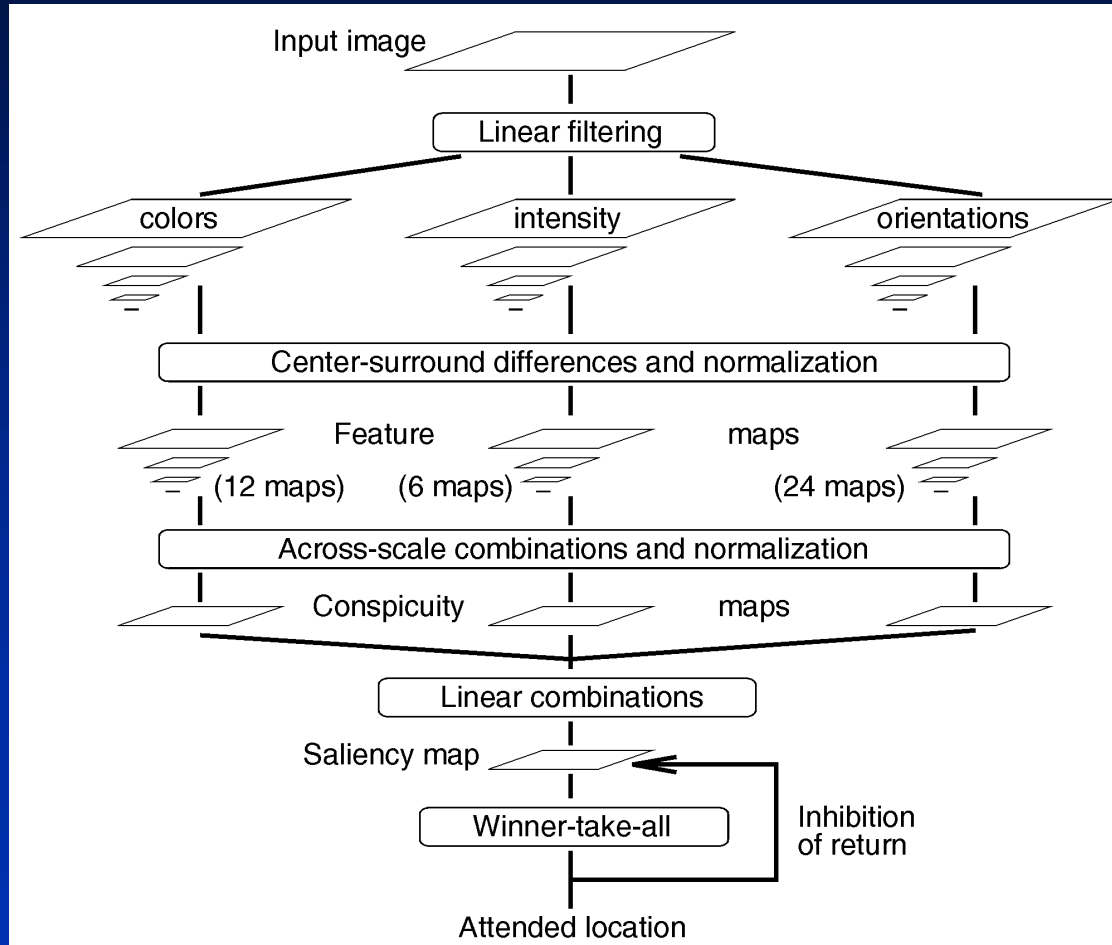


Research – bottom-up

- Saliency features
 - Osberger, Luo, Itti, Harel
- Mismatching – Stentiford, Fang
- Entropy - Kadir
- Content – Le Meur
- Mutual Information - Gao
- Training – Liu, Zhang, Bruce

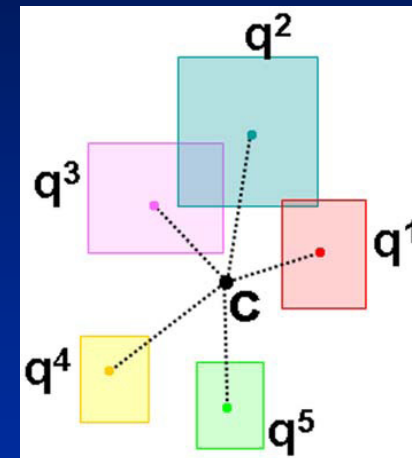


Itti Model



Research – Top-down

- Contextual features – Oliva
- Patch ensembles - Boiman
- Spectral processing - Hou





Issues

- Bottom-up vs Top-down
- Representative data
- Feature selection
- Characterising the unusual

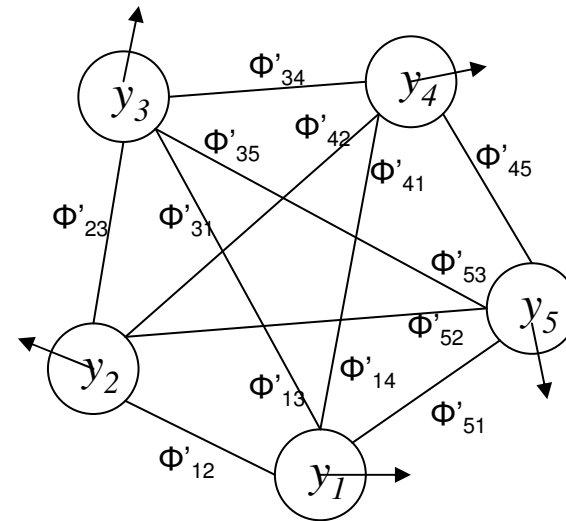
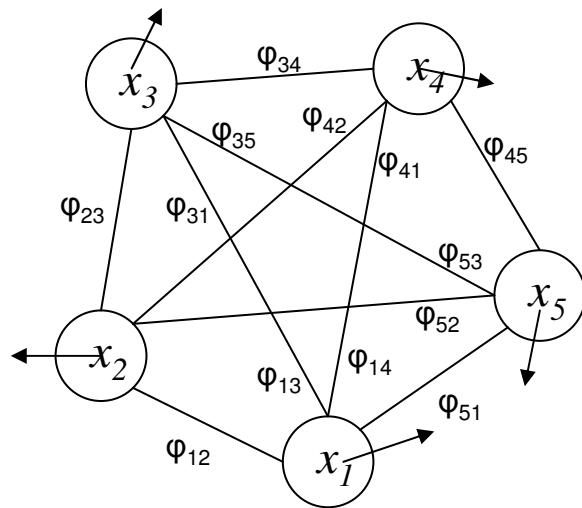


Graphical Representation

- Interest points or pixels are represented as nodes
- Nodes possess properties such as
 - Location
 - Colour
 - Brightness
 - Gradient
- A relationship exists between nodes if their properties and relative orientation match that of a pair in a second image.
- A maximal clique is the largest subset of nodes that all possess a relationship with each other



Cliques of Interest Points





Interest Points





Matching Interest Point Pairs



Image 1

Gradient $\epsilon_1 < 27^\circ$



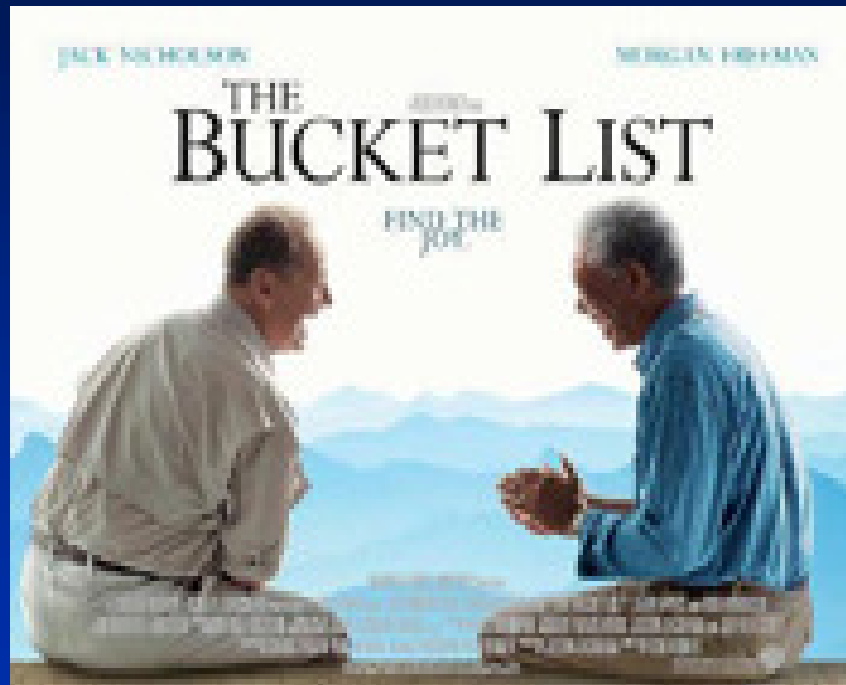
Image 2

Relative orientation $\epsilon_2 < 11^\circ$

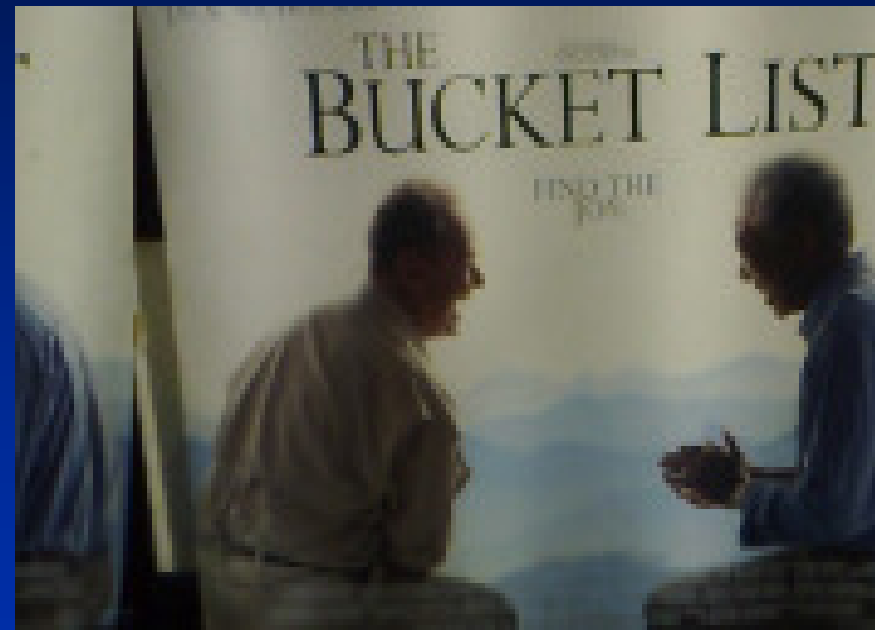
$(\epsilon_1, \epsilon_2) = (27^\circ, 15^\circ)$



Movie Posters



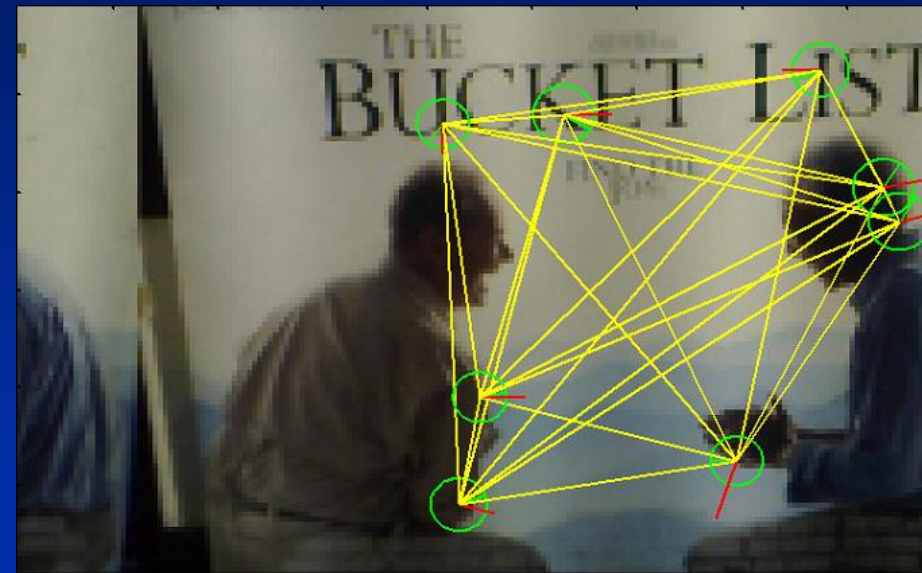
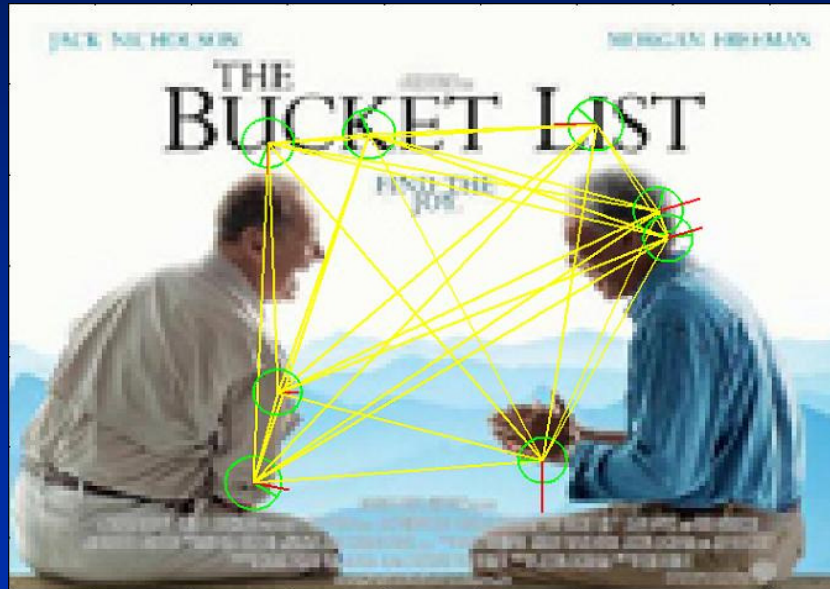
Reference



Photo

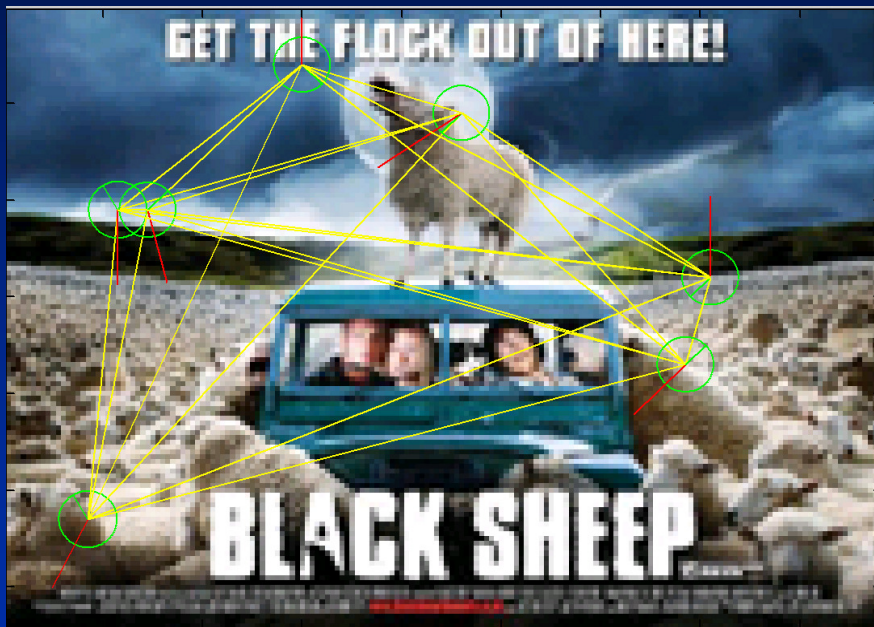


Matching Maximal 8-Clique



b channel

Matching 7-Clique

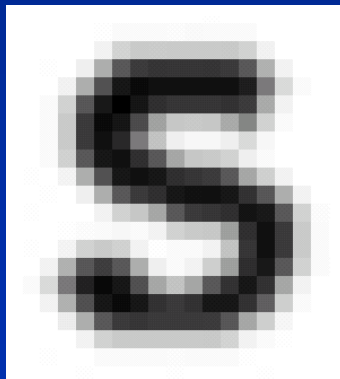
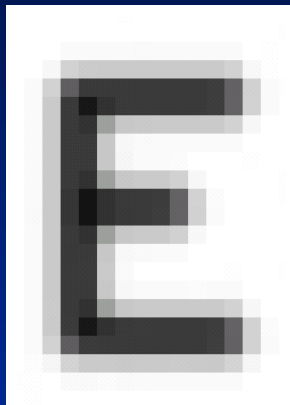


b channel

Computation for formal maximal clique extraction goes up exponentially

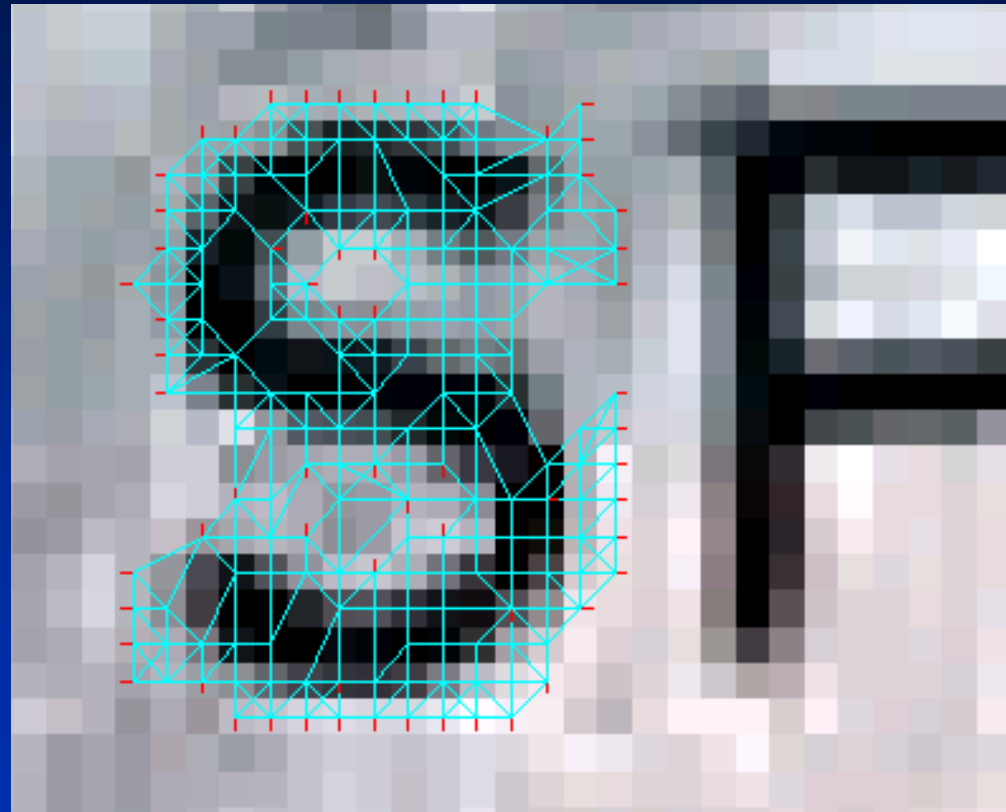
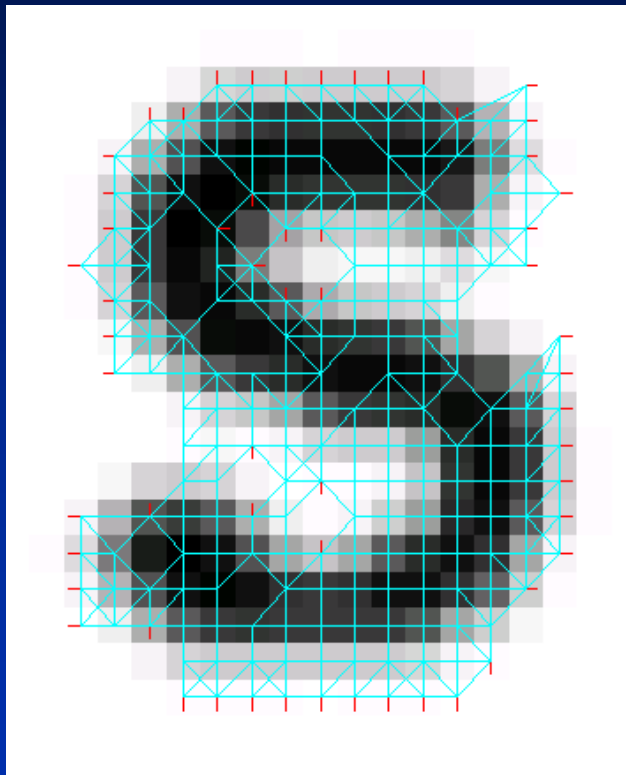


Character Location & Recognition





Maximal Clique



4 nearest points plotted

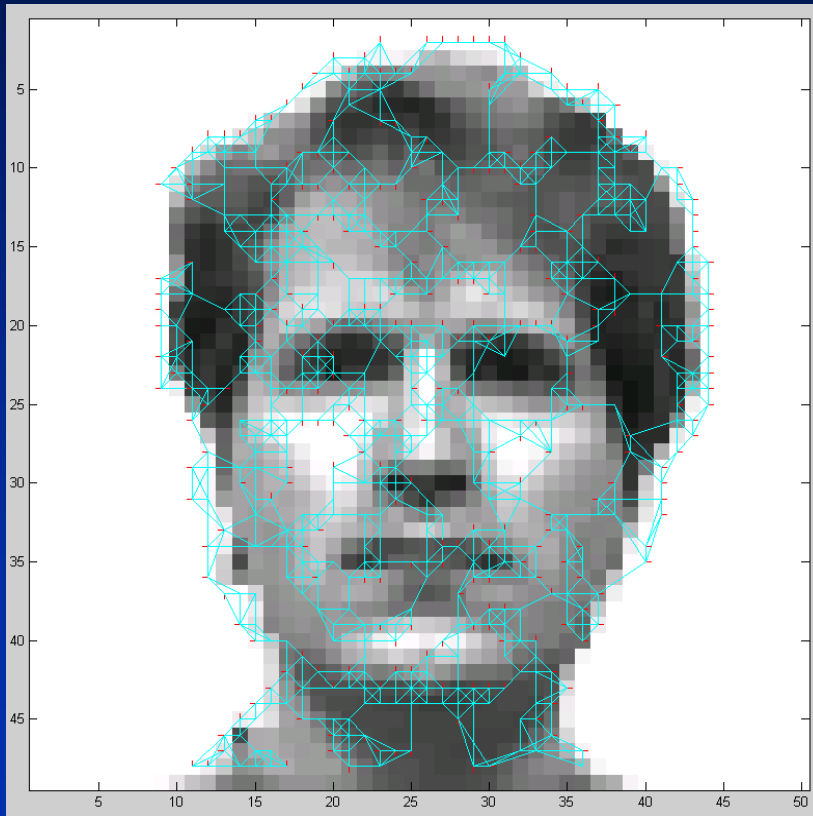


Pixels in Maximal Cliques

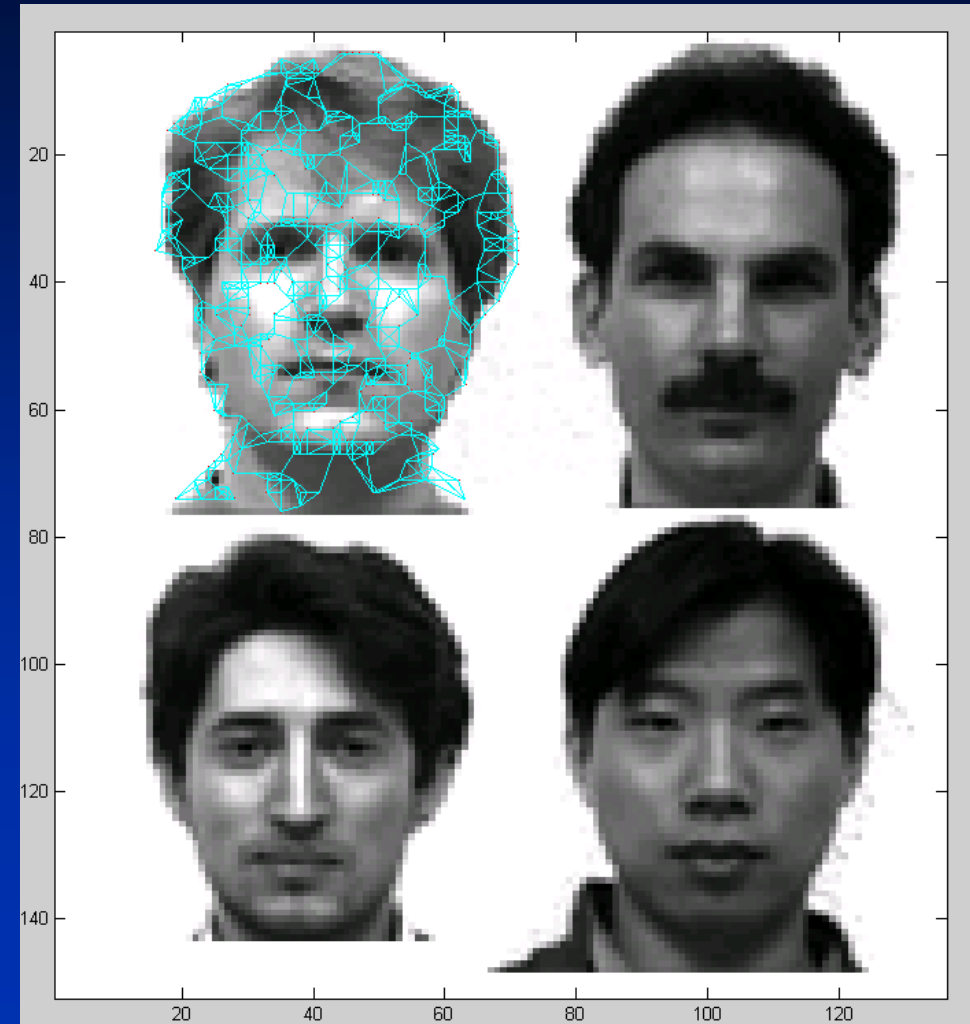




Face Recognition using Maximal Cliques

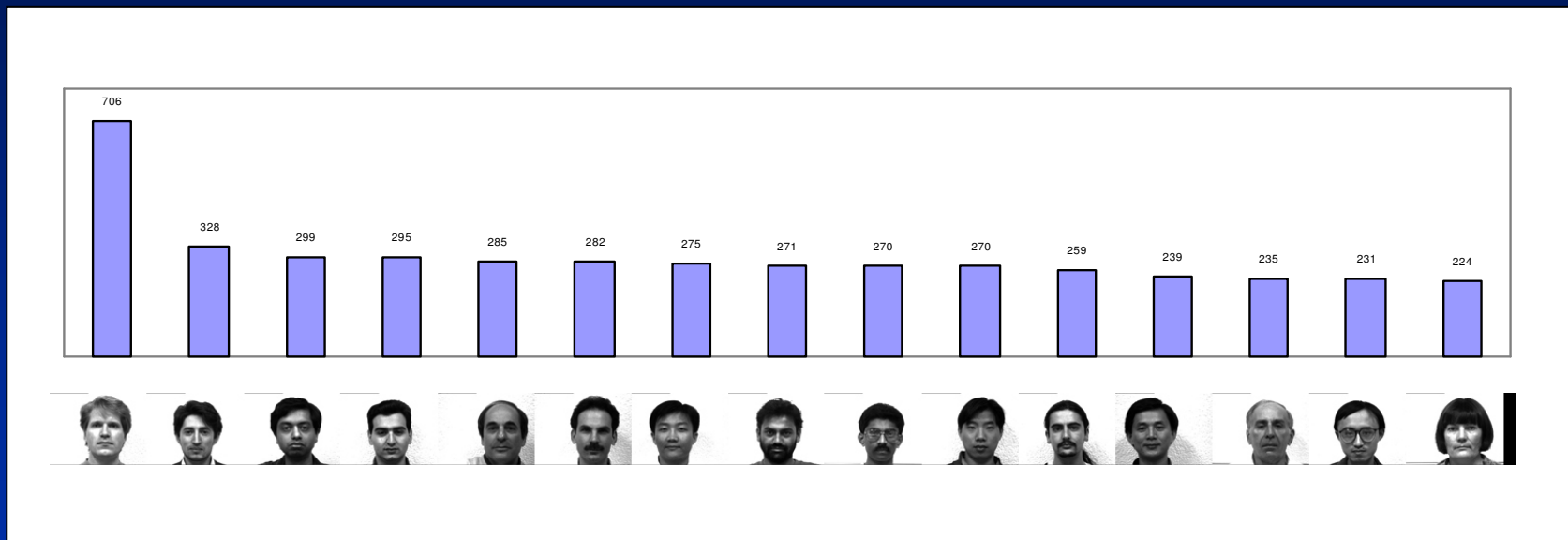


585 nodes





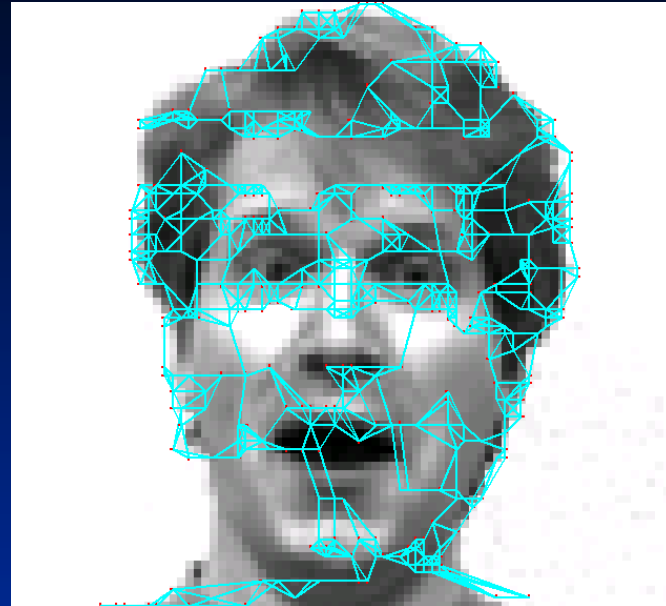
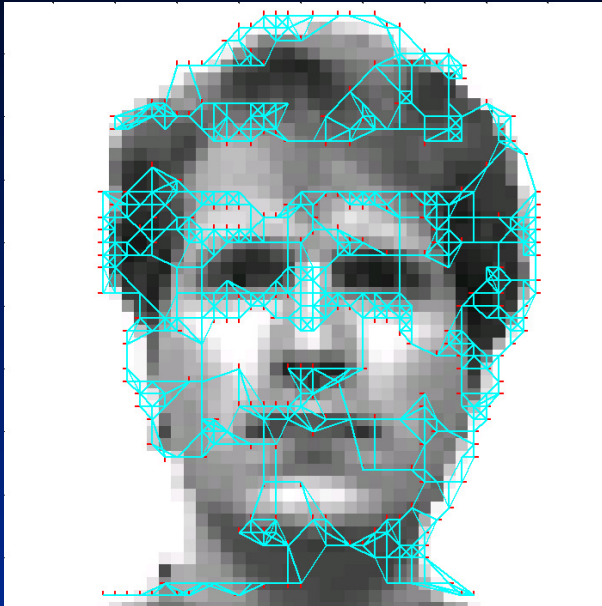
Face Recognition



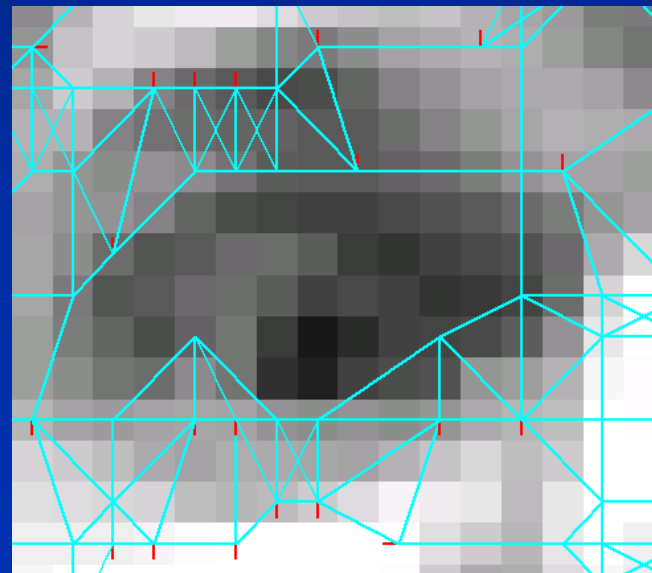
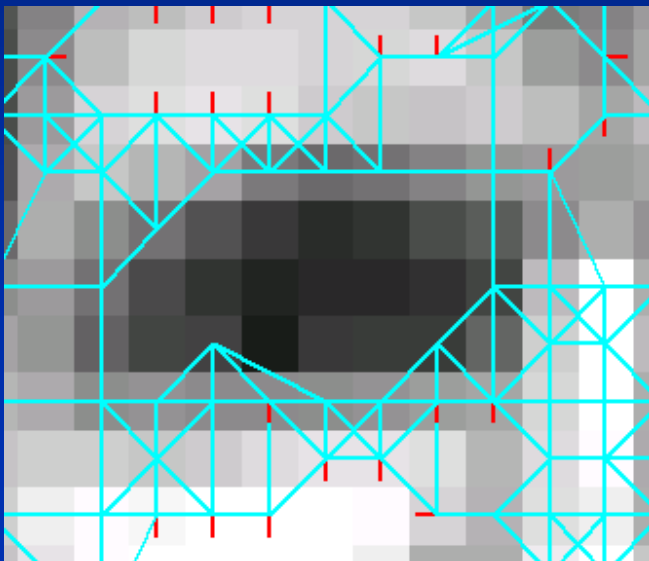
Yale faces



Expressions

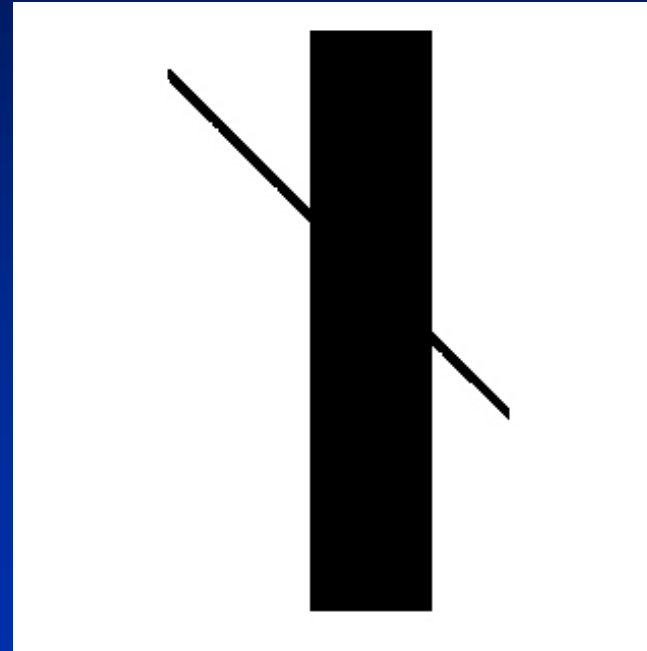
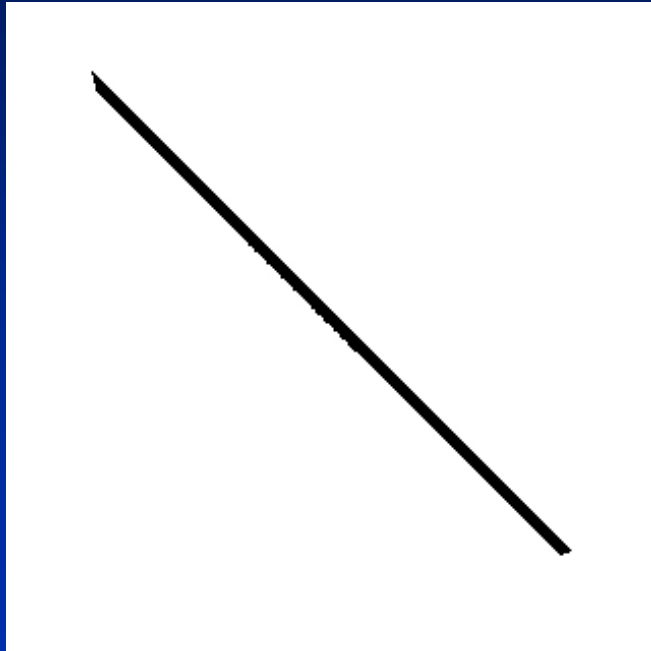


**460
points**





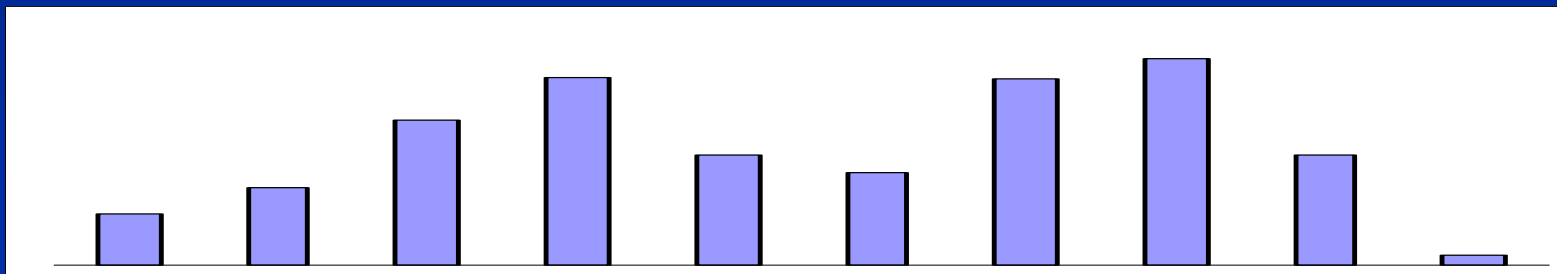
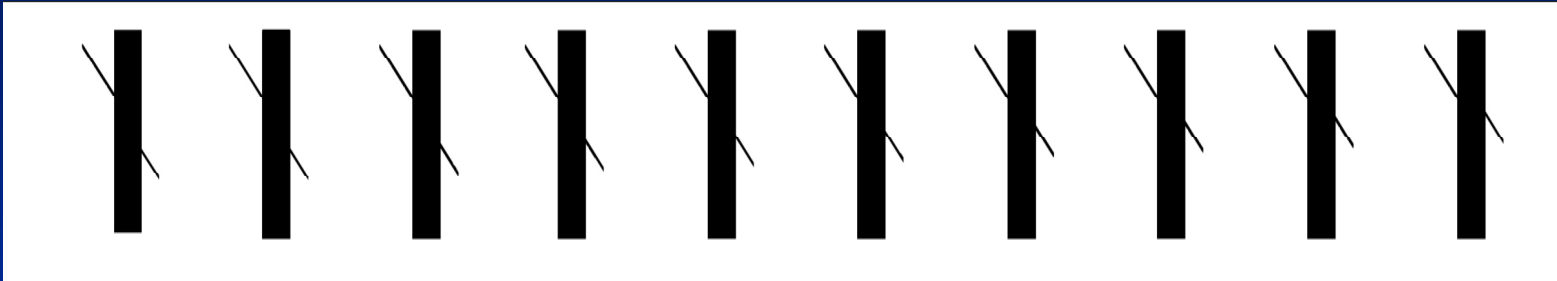
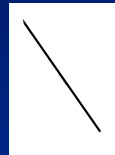
Poggendorff Illusion



Limit the maximum distance between points



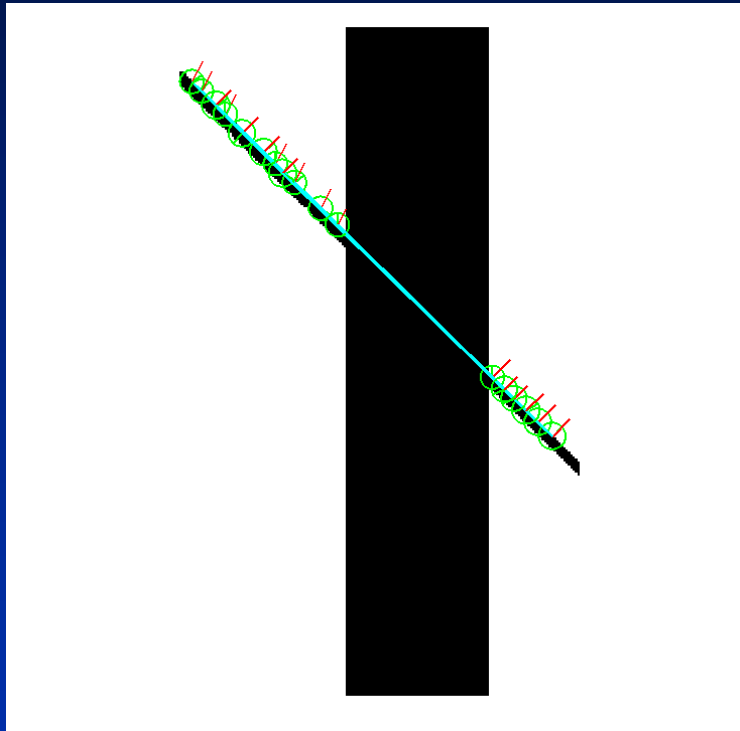
Similarity Scores





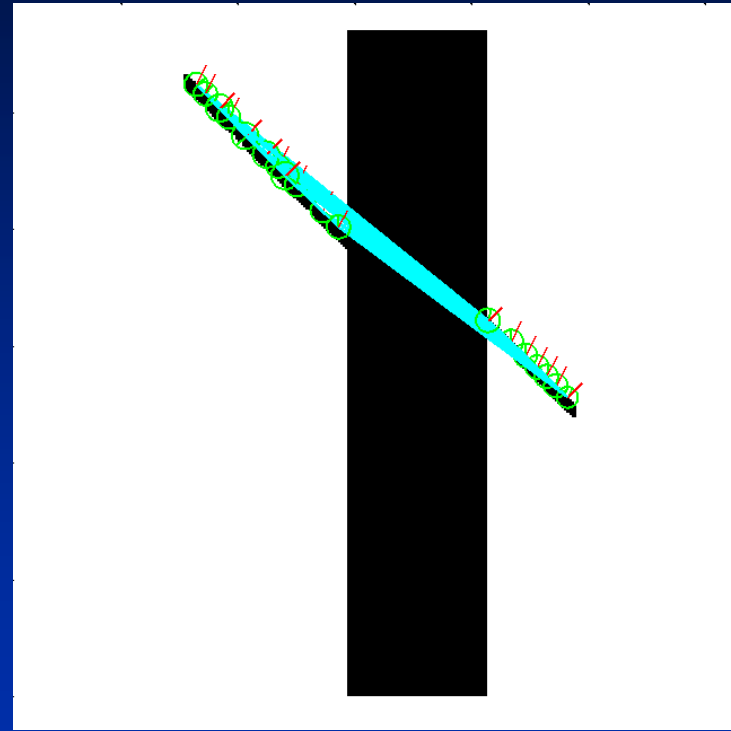
Results

co-linear



266

+4



294

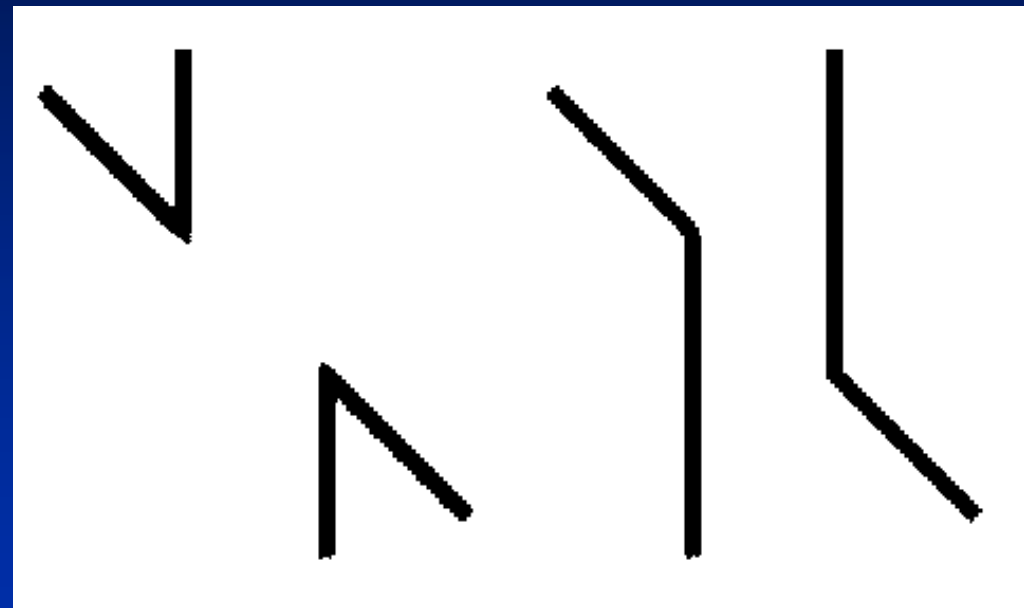
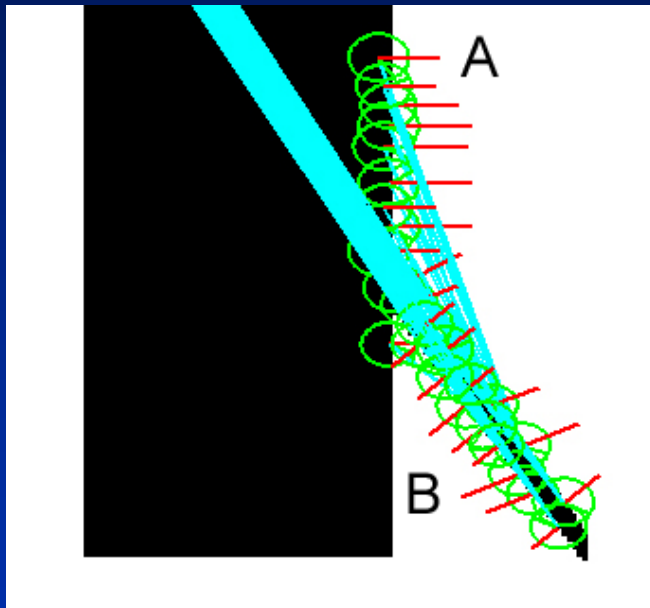


Effect at Varying Angles

	1	2	3	4	5	6	7	8	9	10
60	204	180	190	294	342	372	446	410	284	306
45	72	110	206	266	156	132	264	294	156	14
30	4	8	8	44	102	132	112	332	140	142

Peaks occur at greater shifts with smaller angles.

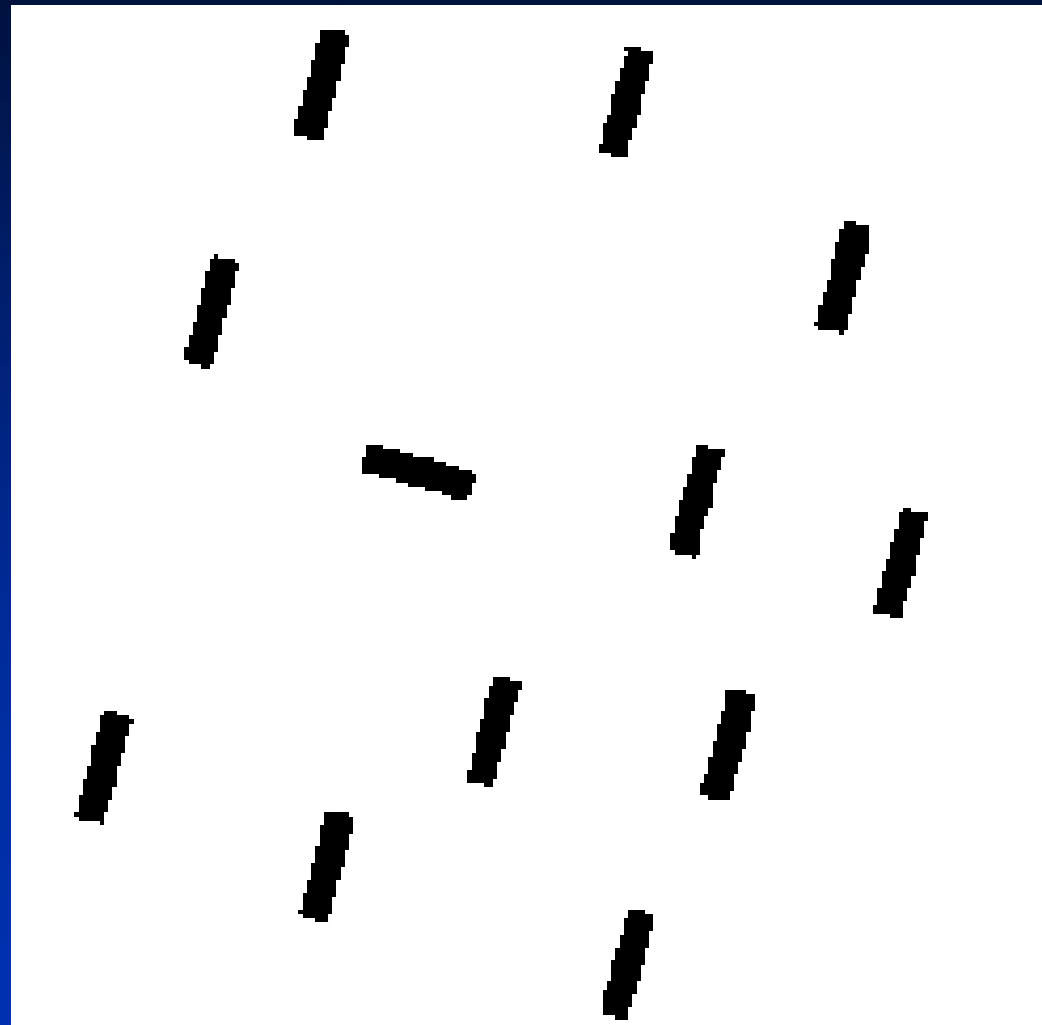
Acute and Obtuse Angled Variants



Stronger matching of interest point pairs in obtuse angles

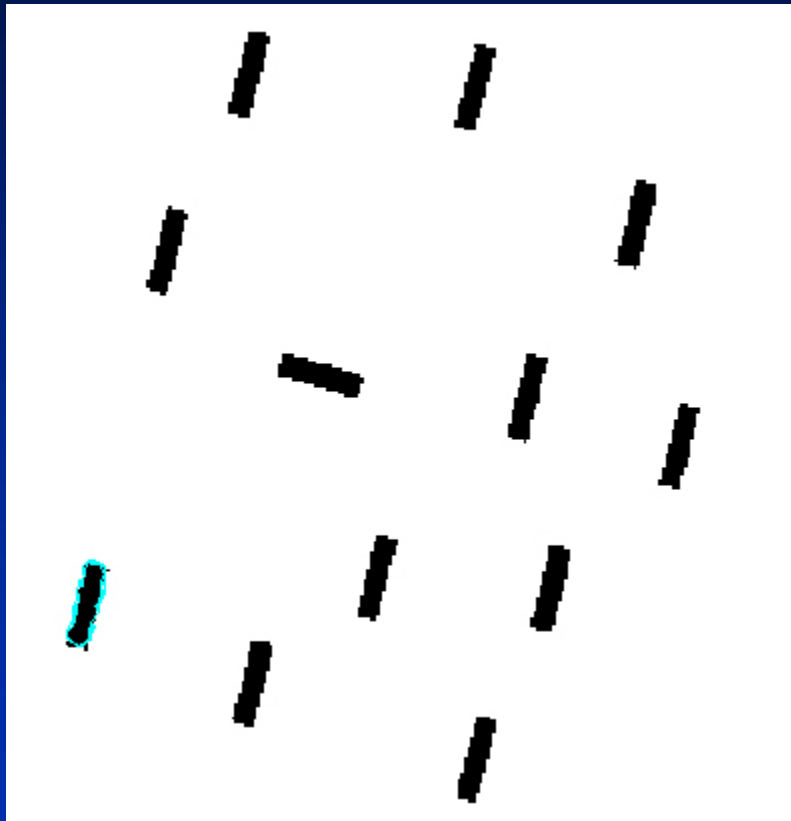


Visual Attention





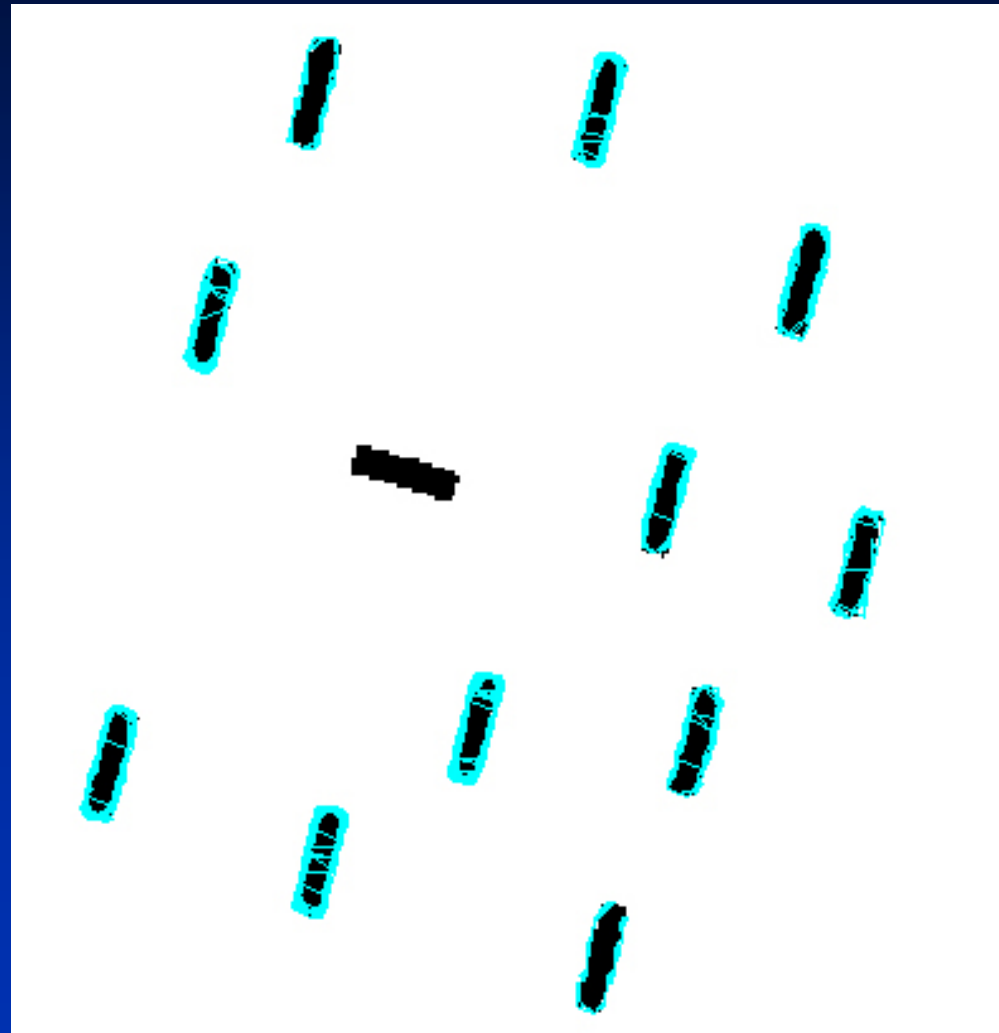
Self Matching



Constrain clique size

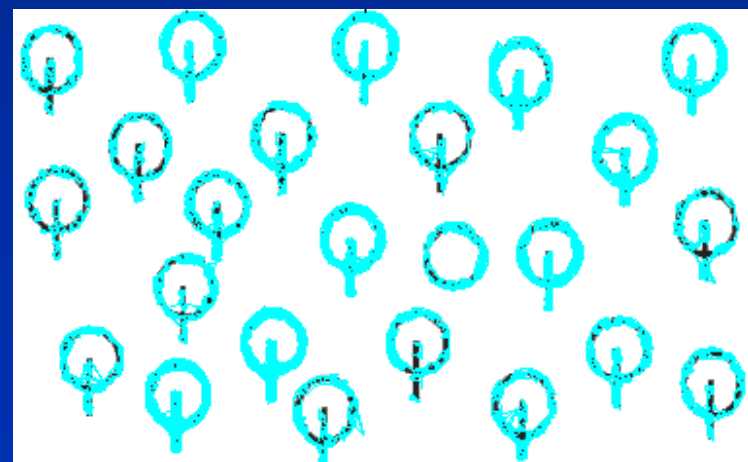
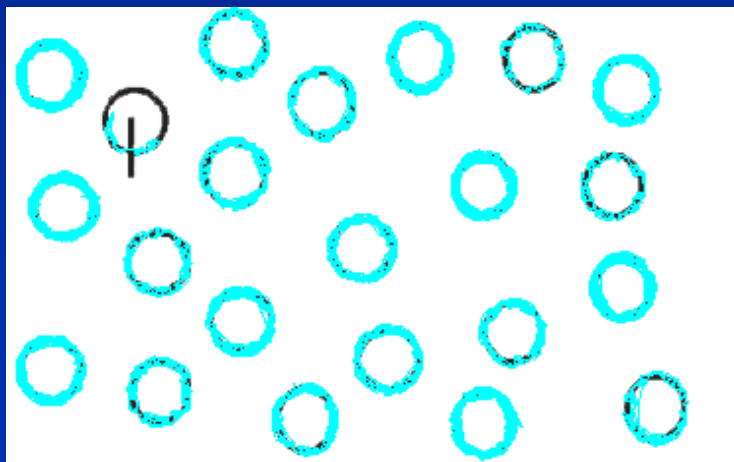
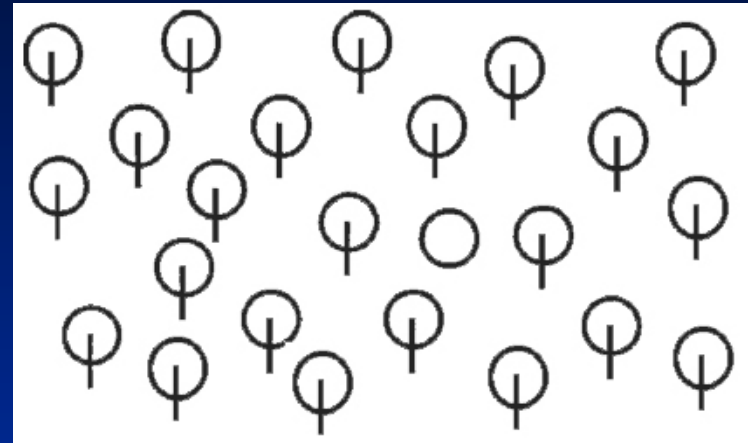
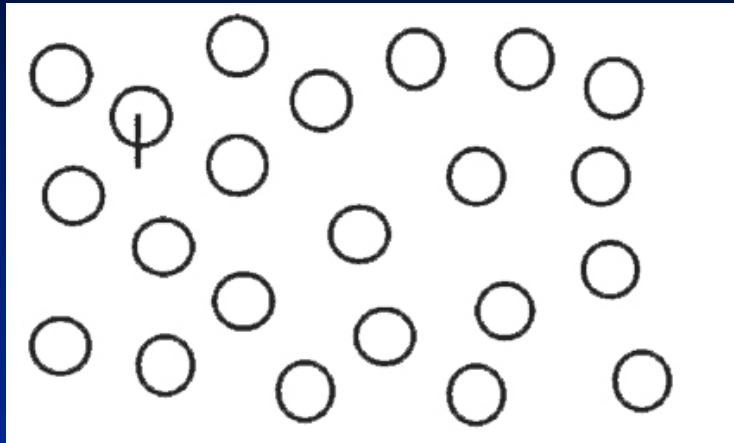


Composite Matches



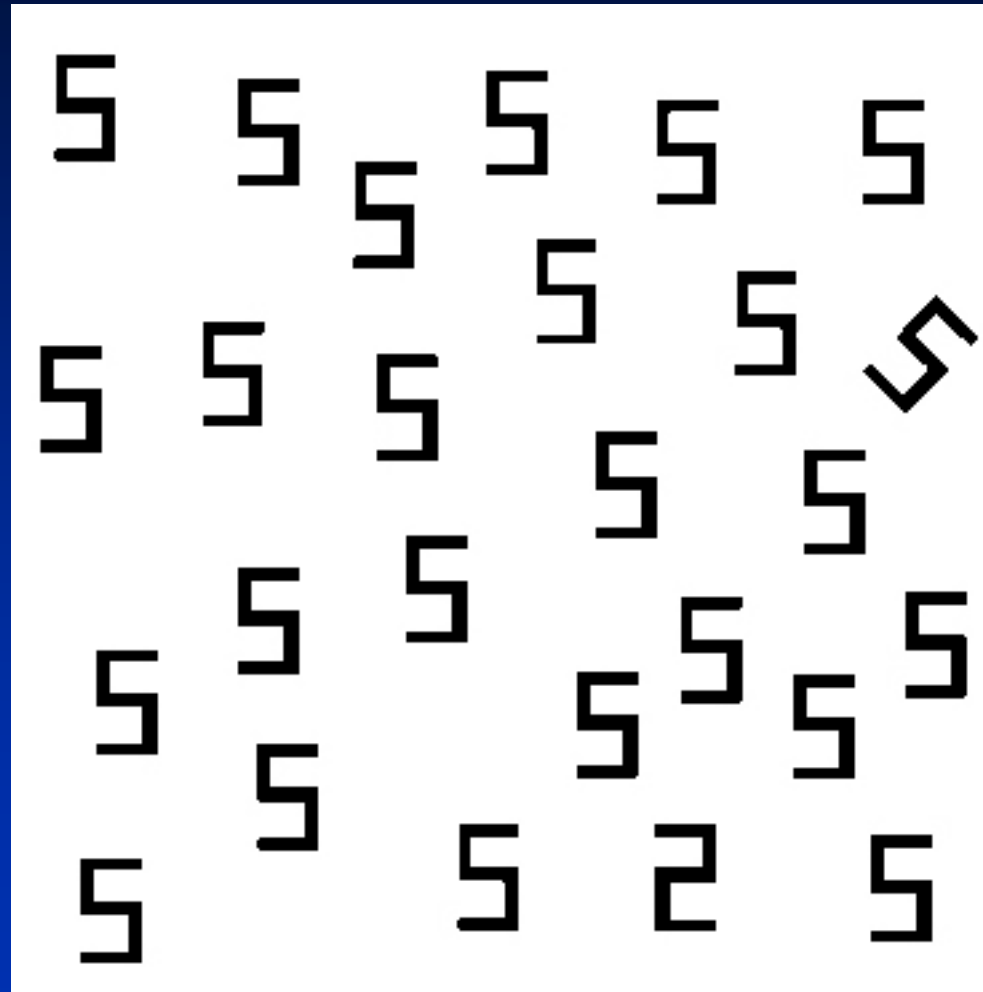


Examples



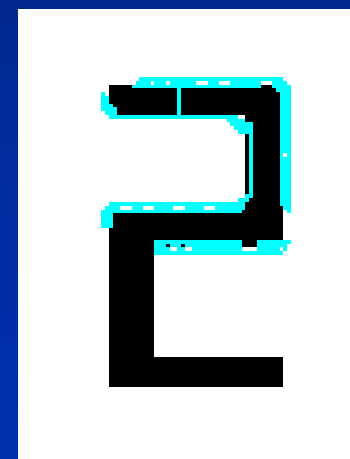
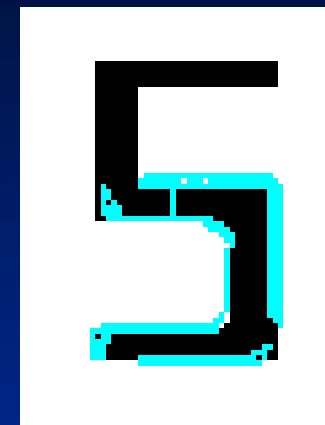
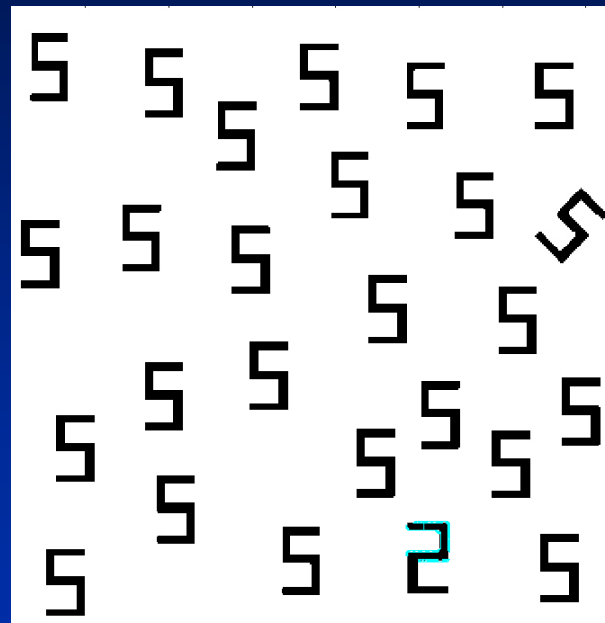
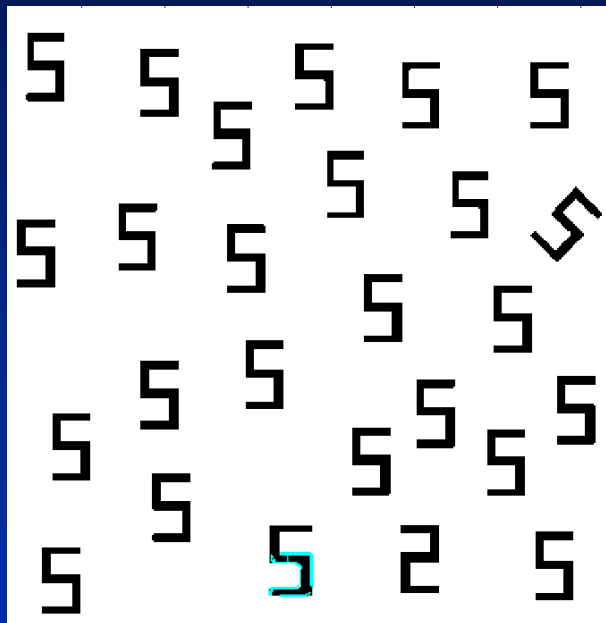


Fives & Twos



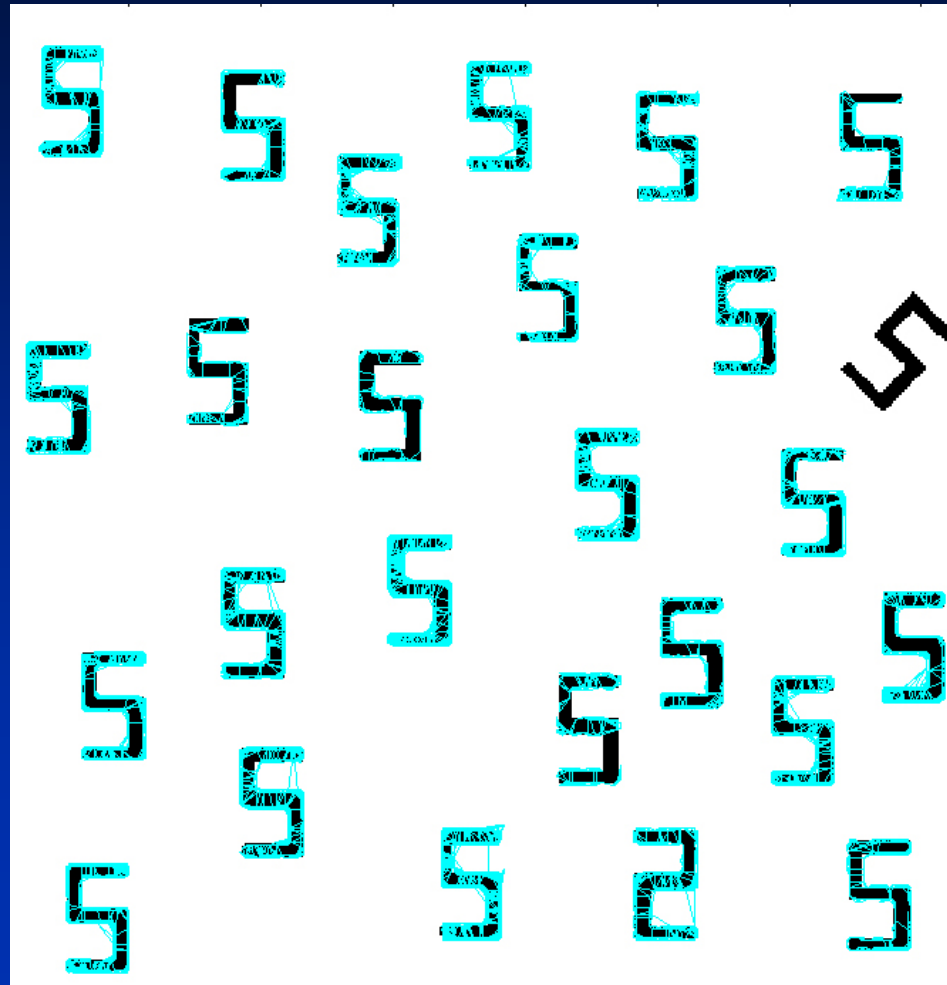


Matching Distractor



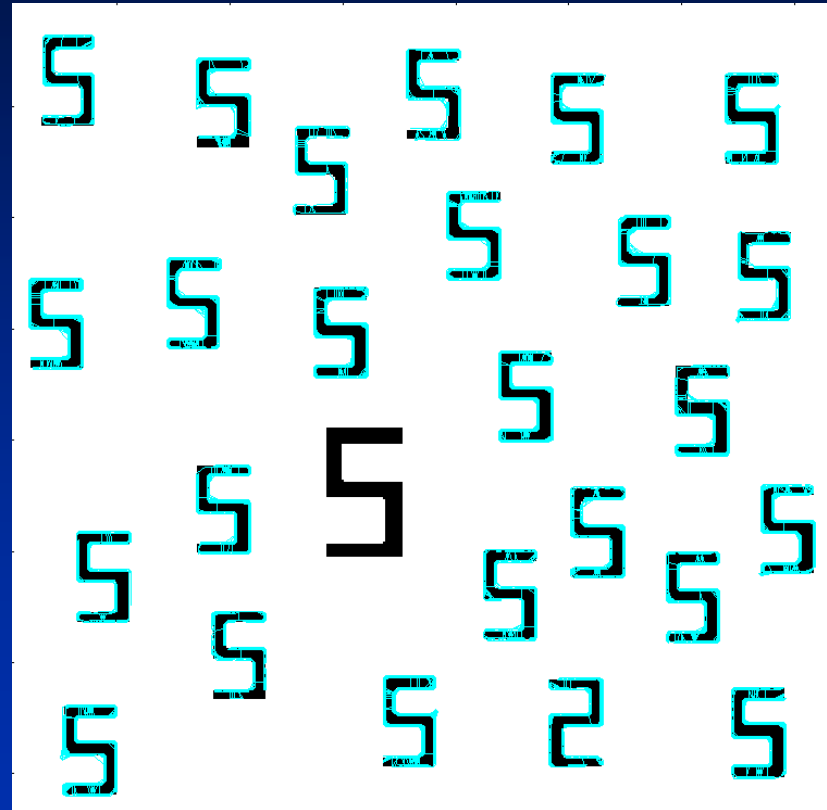
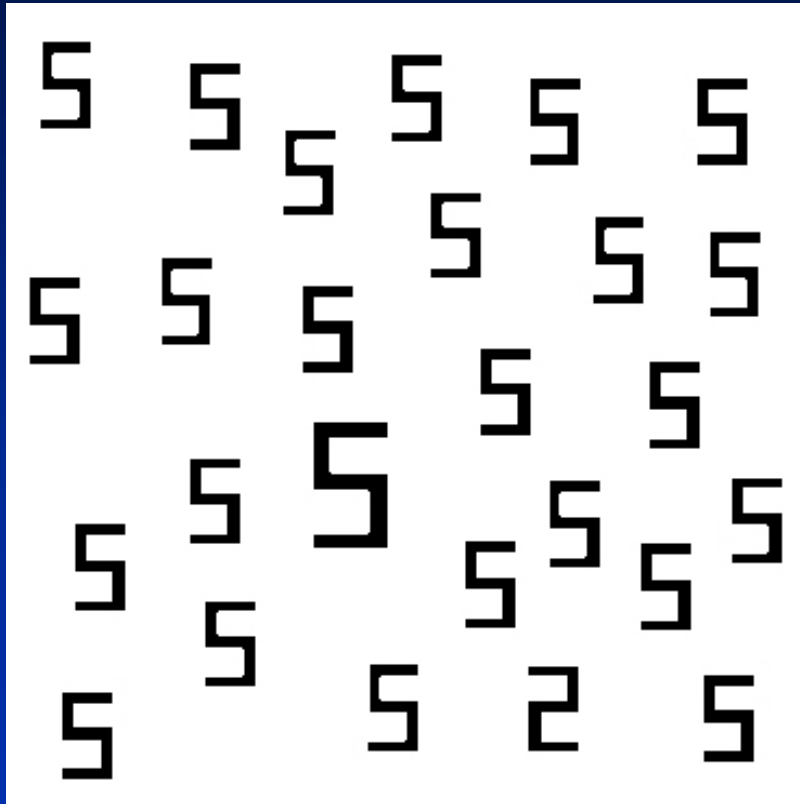


Composite Matches



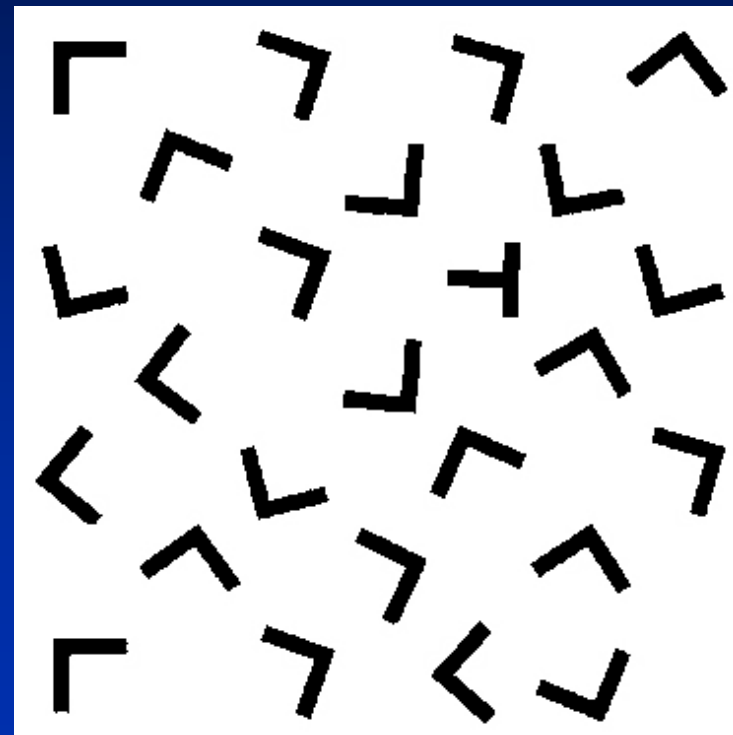
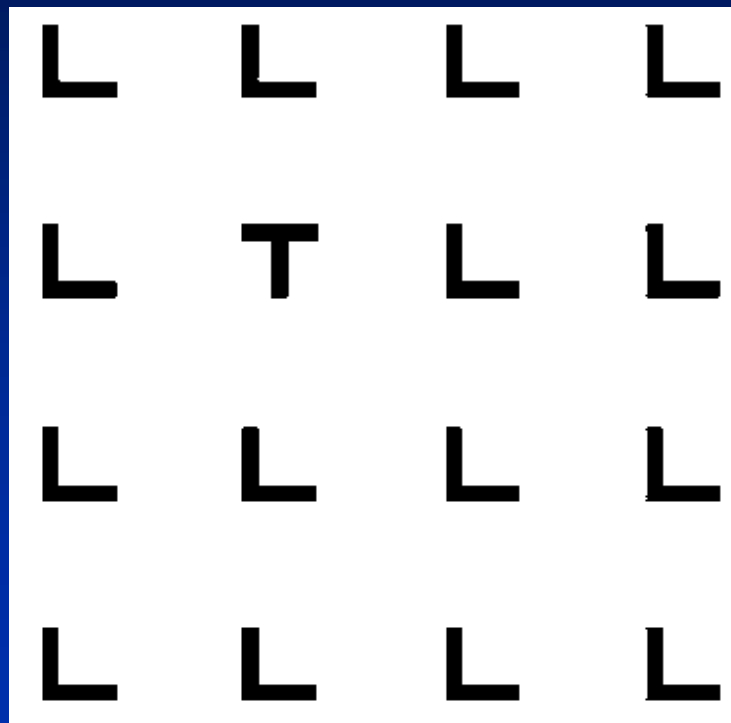


Size



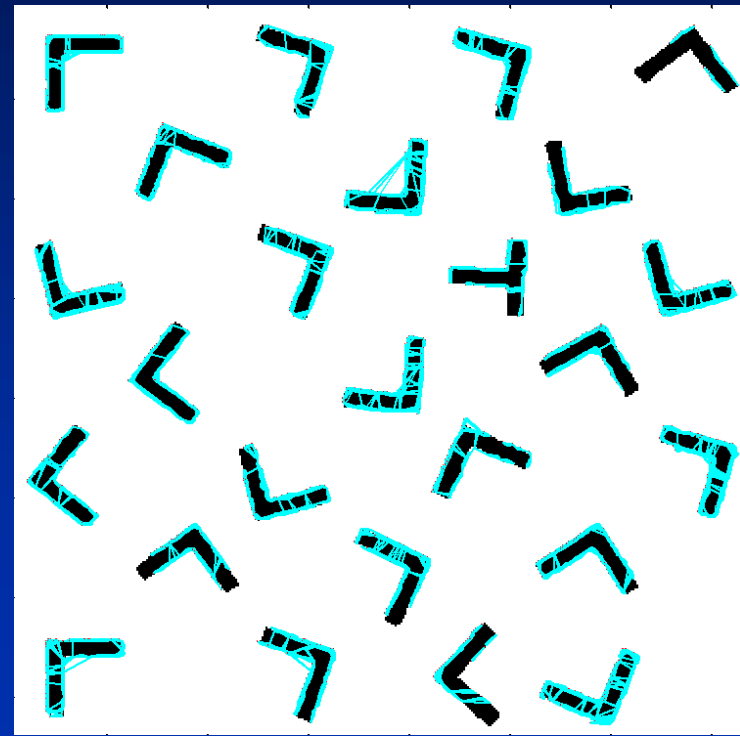
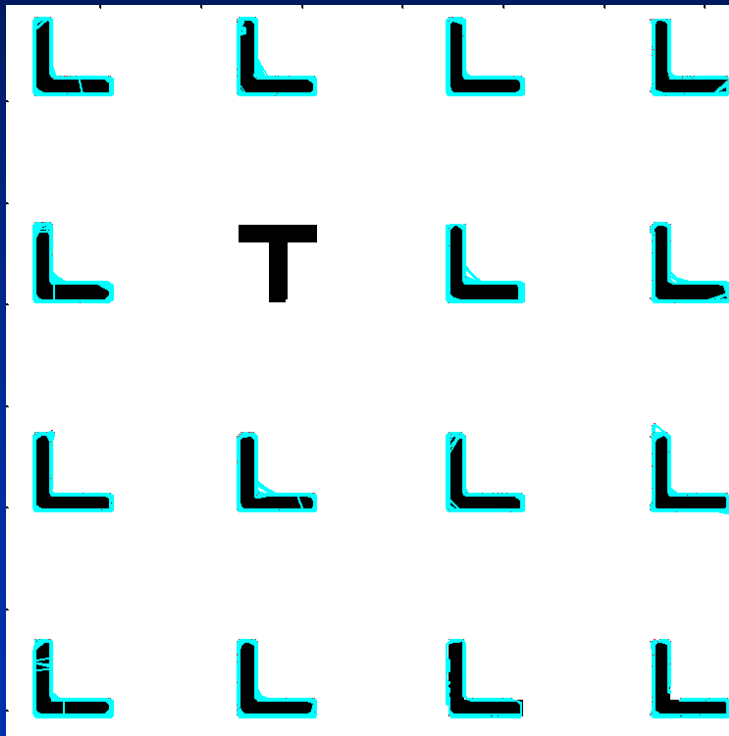


Orientation disturbs attention





Composite matches





Natural Image





Self Matching



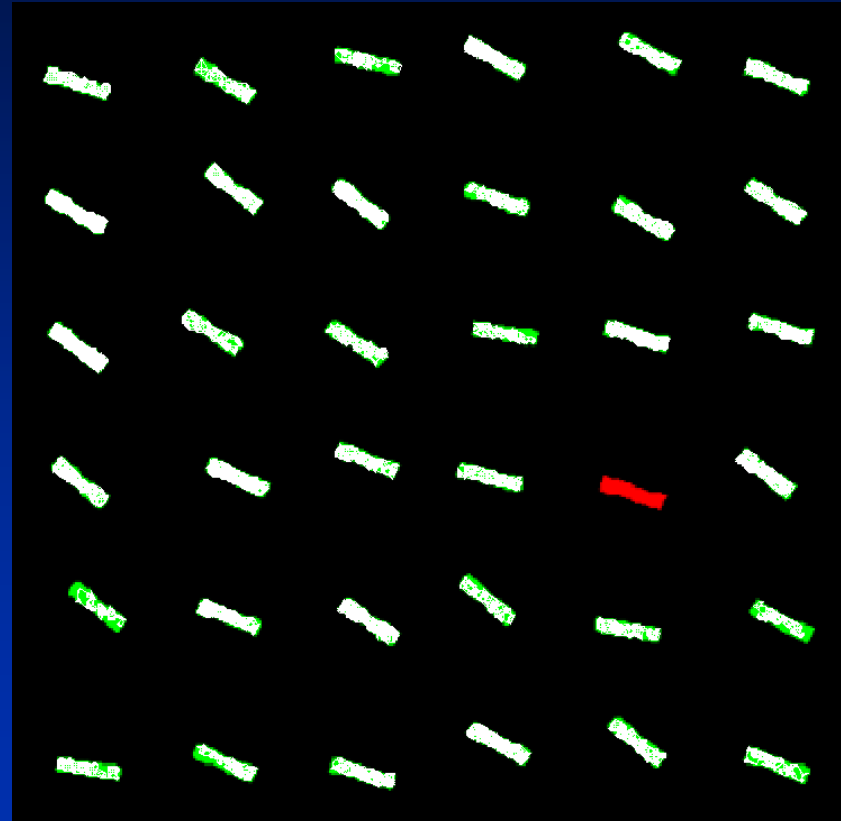
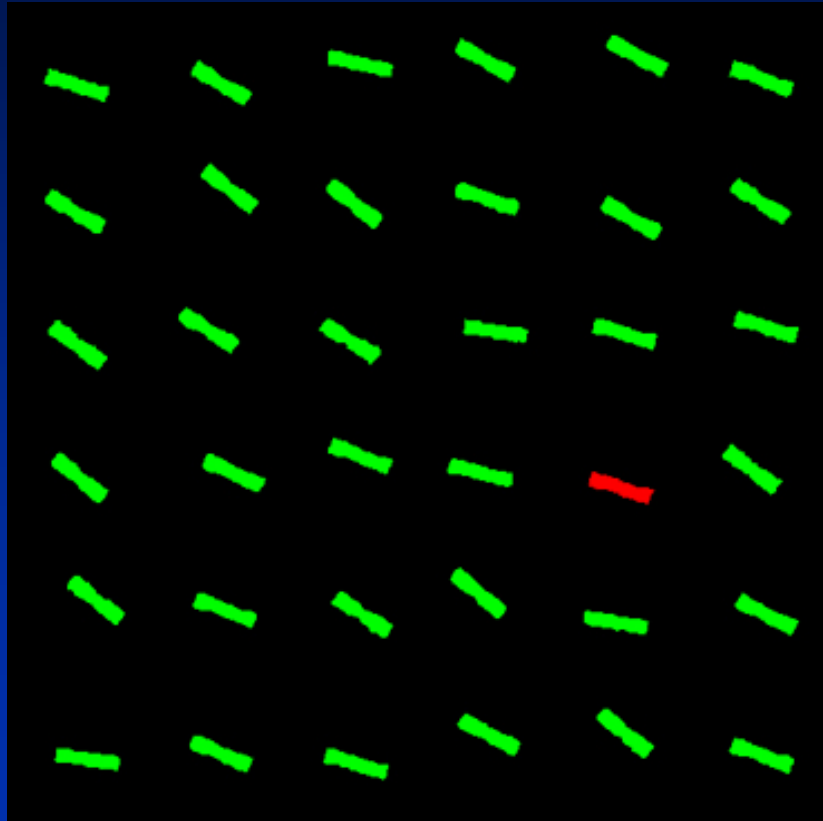


Composite



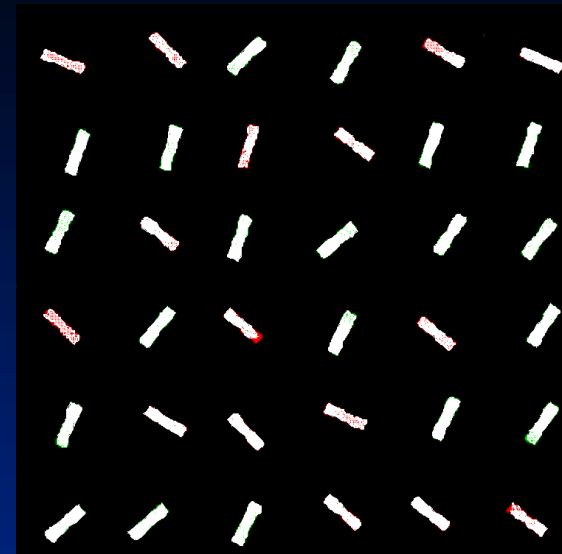
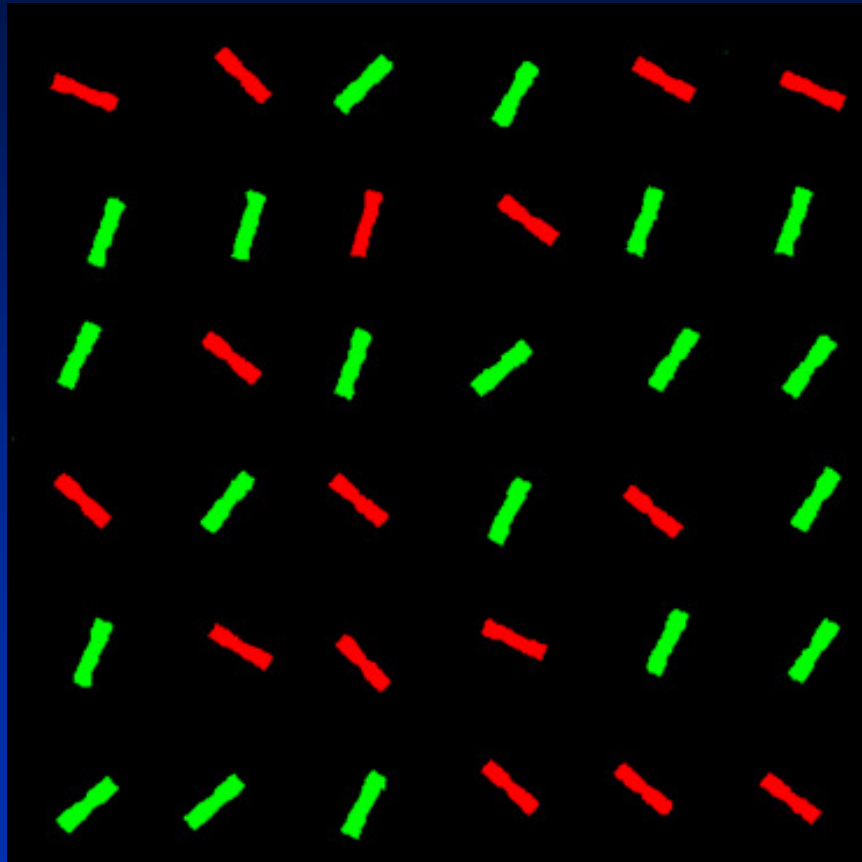


Colour

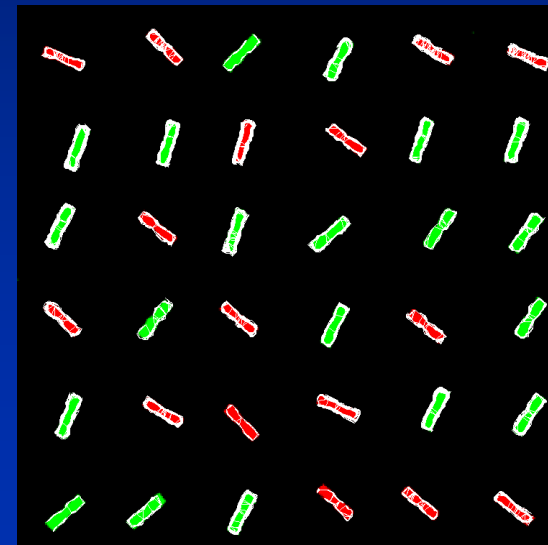




Conjunction



Colour only



Orientation only



Conclusions

- Recognition by reflection is proposed
- Background is a driver in VA model
- More work on larger datasets