Science of Visual Photo Design

Metro Continuing Education
Livio Fent

Course Objectives

- λ Gain an overall appreciation of some of the scientific aspects in composing photographs
- λ Gain an appreciation of why and how the eye is predisposed to viewing certain objects (using the 'bottom-up' approach to perception)
- λ Gain an understanding of the concept of order and information and how it is associated with visual patterns
- λ Apply the principles to images and visual design
- λ Become aware of how these scientific components can be effective in our own photography

Course Outline

Section 1. Introducing photocomposition

- Scientific principles that affect how we see: emphasis on form not content
- A photography experiment: what type of images do we prefer?

Section 2. The 'why' of eye movements: applications to the visual field

- Basic physiology of the retina: receptors, bipolar, horizontal and ganglion cells
- The “on-center, off surround” and “off-center, on surround” perceptual function
- A biological system for detecting information

Course Outline – Cont'd

λ **Section 3. Information in the visual field**

- The concept of 'information', order, entropy in the visual field
- Information and the retinal mechanism
- Information content: High and low information visuals – what works, why.

λ **Section 4. From biology and physics to visual design**

- Eye movement in images, fixation points, saliency
- Elements of design with applications

Background

Some guiding principles:

- Try to suppress your *interpretation* of the images
- Try to think graphically rather than thematically
- Think point(s) or areas of interest in the photograph
- Think where do your eyes enter the photograph
- Think what are they doing within the frame
- Think where do they exit (if at all)
- How long is your interest maintained;
- How simple or complex is the photograph *visually*
- How is the image organized, why does it work or not

Introduction

The basic question that we will be exploring is: 'What makes a photo 'better' than another?

We apply this question to photocomposition and investigate it from a number of angles:

- Our intuitive preferences for images
- Our eyes' preference for form
- Our desire to look for information and order
- How our predisposition translates to eye movements and visual design principles.
- How we can use these elements to improve our art, visual statement and communication.

Introduction

How we see – two basic theories

Constructivist (top-down)

- Indirect, cognition driven
- Much information that reaches the eye is lost -90%
- The brain has to guess what a person sees based on past experiences.
- Our perceptions of the world are hypotheses based on past experiences and stored information.
- Heavy reliance on memory functions, the focus of this approach is in post-sensory processes

Ecological (bottom-up)

- Direct, eye initiated
- Data driven, information reliant
- Sensory information is analyzed in one direction: from simple analysis of raw sensory data to ever increasing complexity of analysis through the visual system.
- This is the approach we will investigate in some detail.

Source: McLeod, S. A. (2007). Visual Perception. http://www.simplypsychology.org/perception-theories.html

Feature Integration Theory.. a reconciliation

- Developed by Anne Treisman (1980), the theory is best characterized as a *sequenced* application of both the bottom-up and top-down approaches to perception
- Objects are analyzed in two steps:
  - Preattentive stage where preception is analytical and unconscious, related to primal processing of edges, contrast, lines, color, orientation in the retina and the visual cortex.
  - Focused attention stage where perception is context related and the individual features of an object combine in order to perceive the whole object.
- The result of the process is a 'Feature Map' and has been used extensively in computer vision algorithms. We'll look at applications of this theory later in the course.

Constructivist or Ecological or Both?

Source:The perception of art and the science of perception by Robert Pepperell

Human Vision and Electronic Imaging XVII

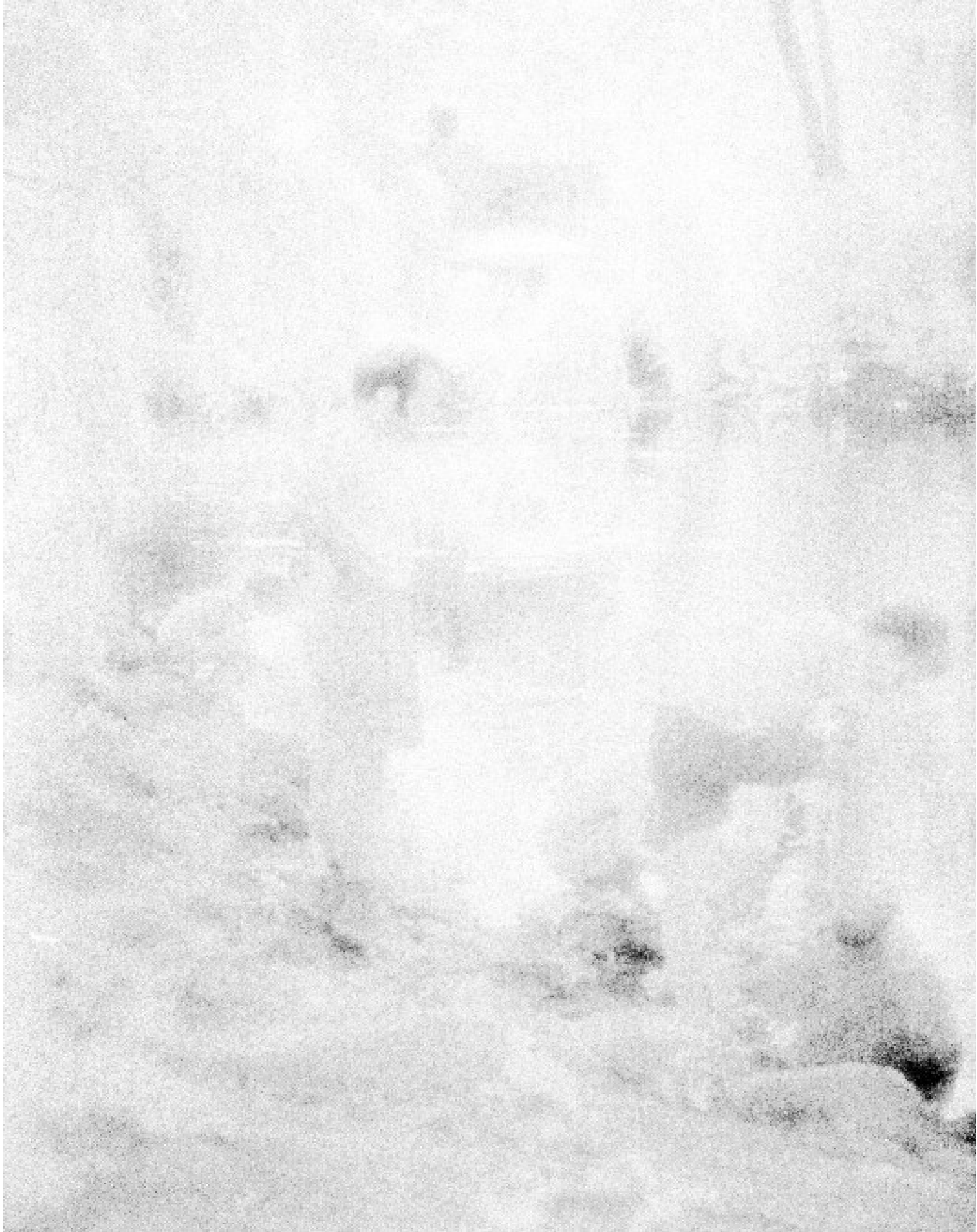
SPIE-IS&T/ Vol. 8291 829113-1

We begin with an experiment on preference...

- View and rate the photographs from 1 (not at all pleasing) to 10 (extremely pleasing)
- Review results: which image was most pleasing, which was least.
- Discussion: what is working, what is not working
- Can we formulate a pattern, a rule, a process to replicate?

Photograph 1

Photograph 2

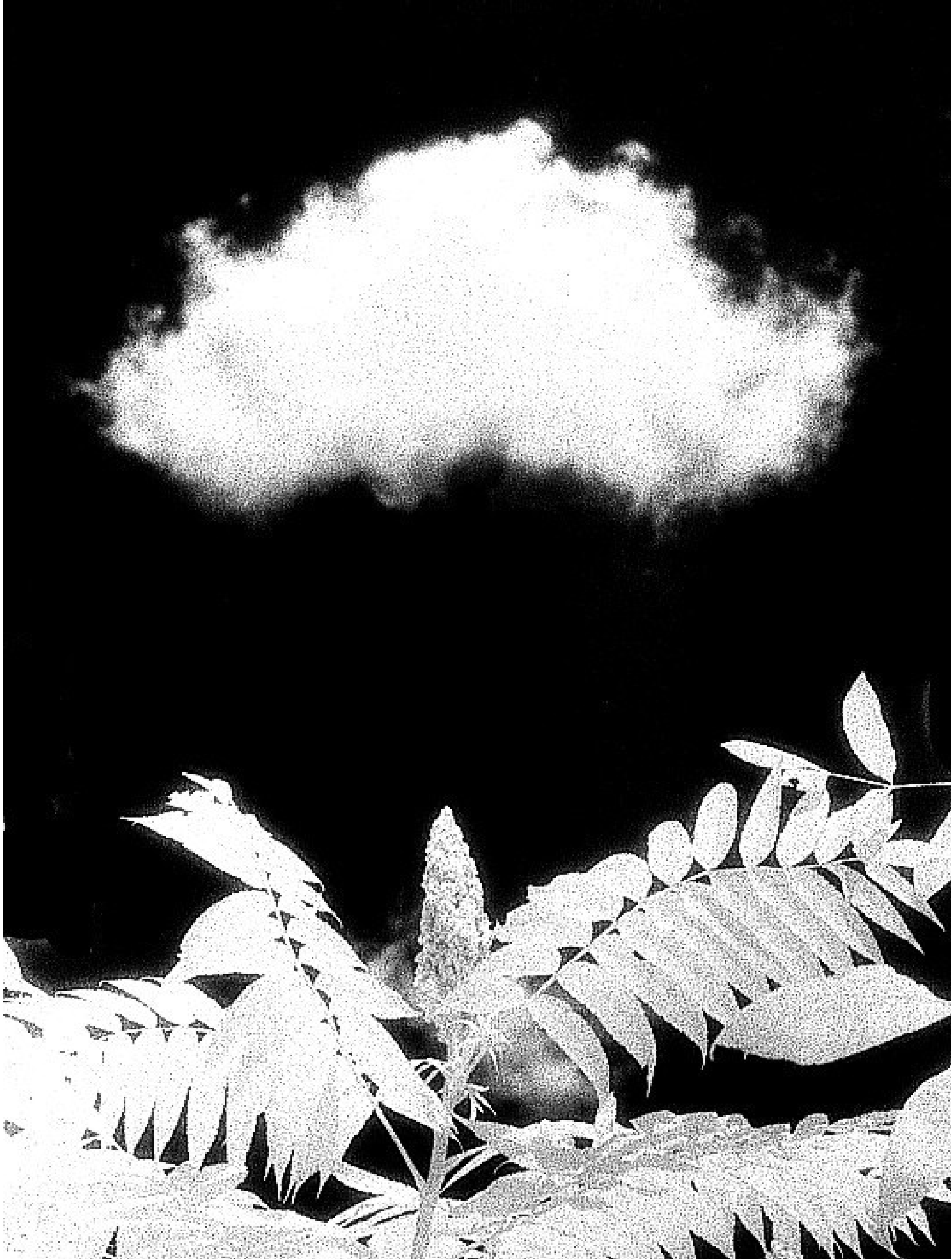


Photograph 3

Photograph 4

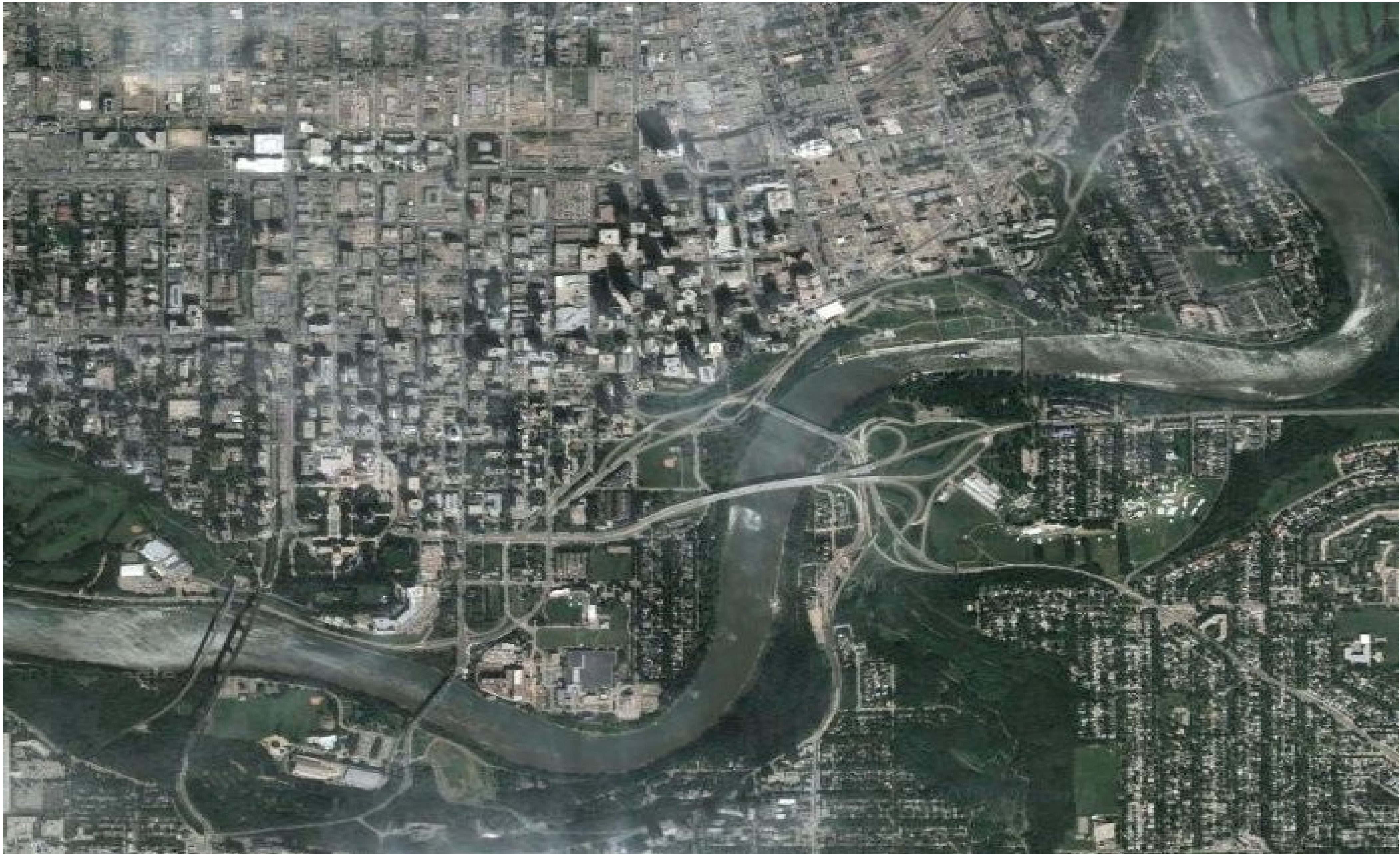
Photograph 5

Photograph 6



Photograph 7

Photograph 8



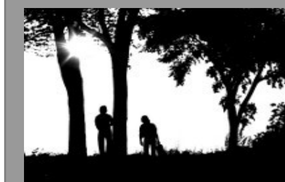
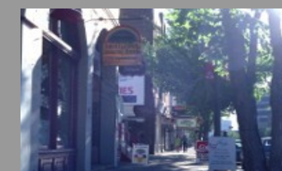
Photograph 9

Photograph 10

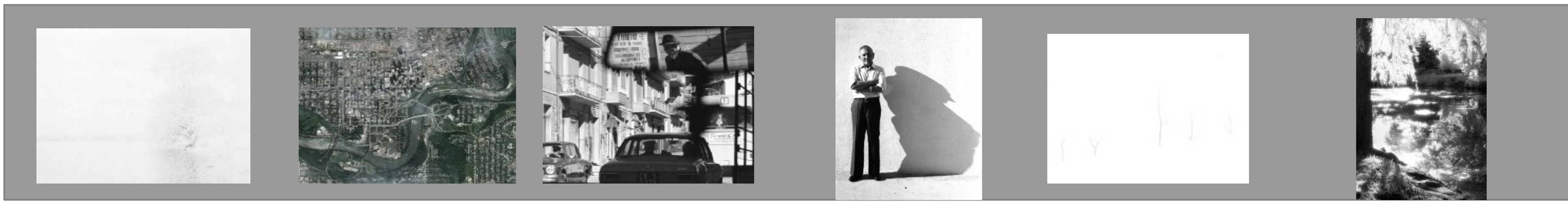
Photograph 11

Photograph 12

Preference Summary

						
	7					
	7					
	8					
	5					
	6					
AVG	6.6	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
VAR	1.30	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!

Preference Summary



	7					
	8					
	5					
	6					
	3					
AVG	5.8	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
VAR	3.70	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!

Discussion

- Assess the ratings, which image rated high, which rated low
- Take the highest rated image, consider it from the the 'form' viewpoint: what are the eyes doing in this image, how long is interest maintained; how complex is the image
- Take the lowest rated image, do the same
- What's common to both the highly and lowly rated images
- Which images had the largest spread or disagreement (variance), which had the lowest. What can we say about these inconsistencies and consistencies for preference.
- Is there a pattern, process, guide: what conclusions can we come to from this exercise?

How We See: Eye Circuits

In this section

- Cells involved in seeing
- Cell connections
- Cell excitation and inhibition
- On-off center, on-off surround circuits
- Feature detection and application to photography

We start this section with another small experiment...

Discussion

1. The uniformity of the single tone images seem less interesting than the dual tone image

2. The eye is attracted to the edge of the dual tone, less interested in the tones themselves

3. There is also more 'information' in the dual tone image which the eye is picking up; we'll investigate this later on.

What causes the eye to dwell on this edge?

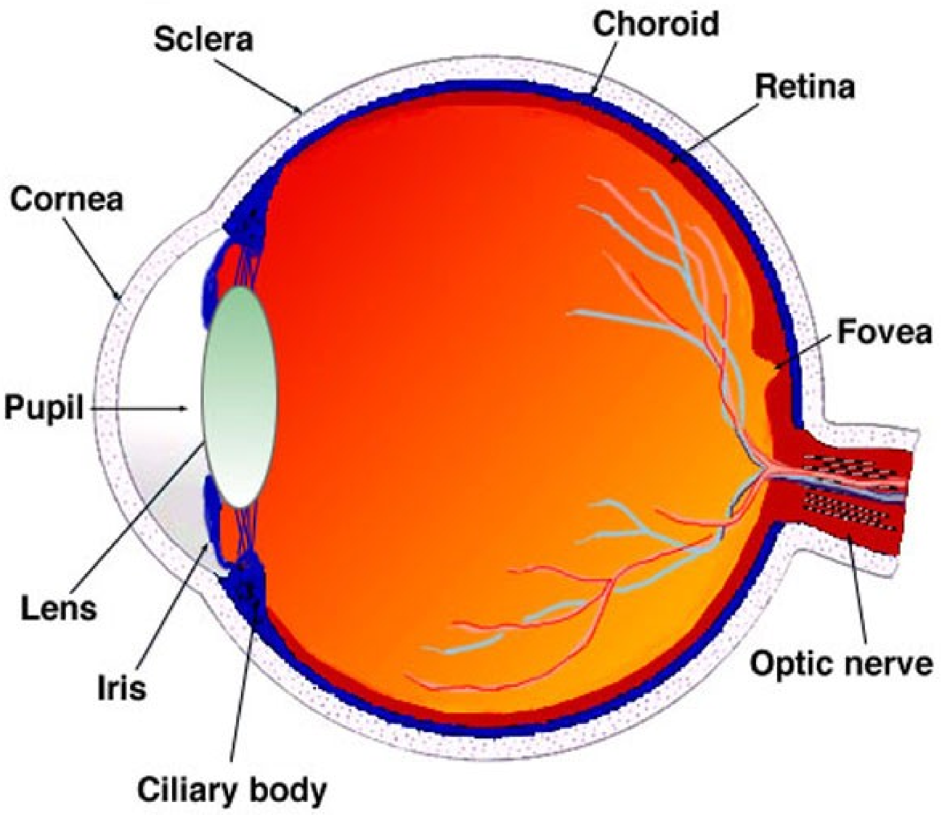
Vision

Vision is one of the five senses, it enables us to convert electro-magnetic radiation (light) into information that we can use for our needs.

The ability to see has two general components:

- The eye and its mechanisms to sense light, and
- The brain and its mechanisms to interpret light signals.

We will be concentrating on the eye component, and specifically on the retina.

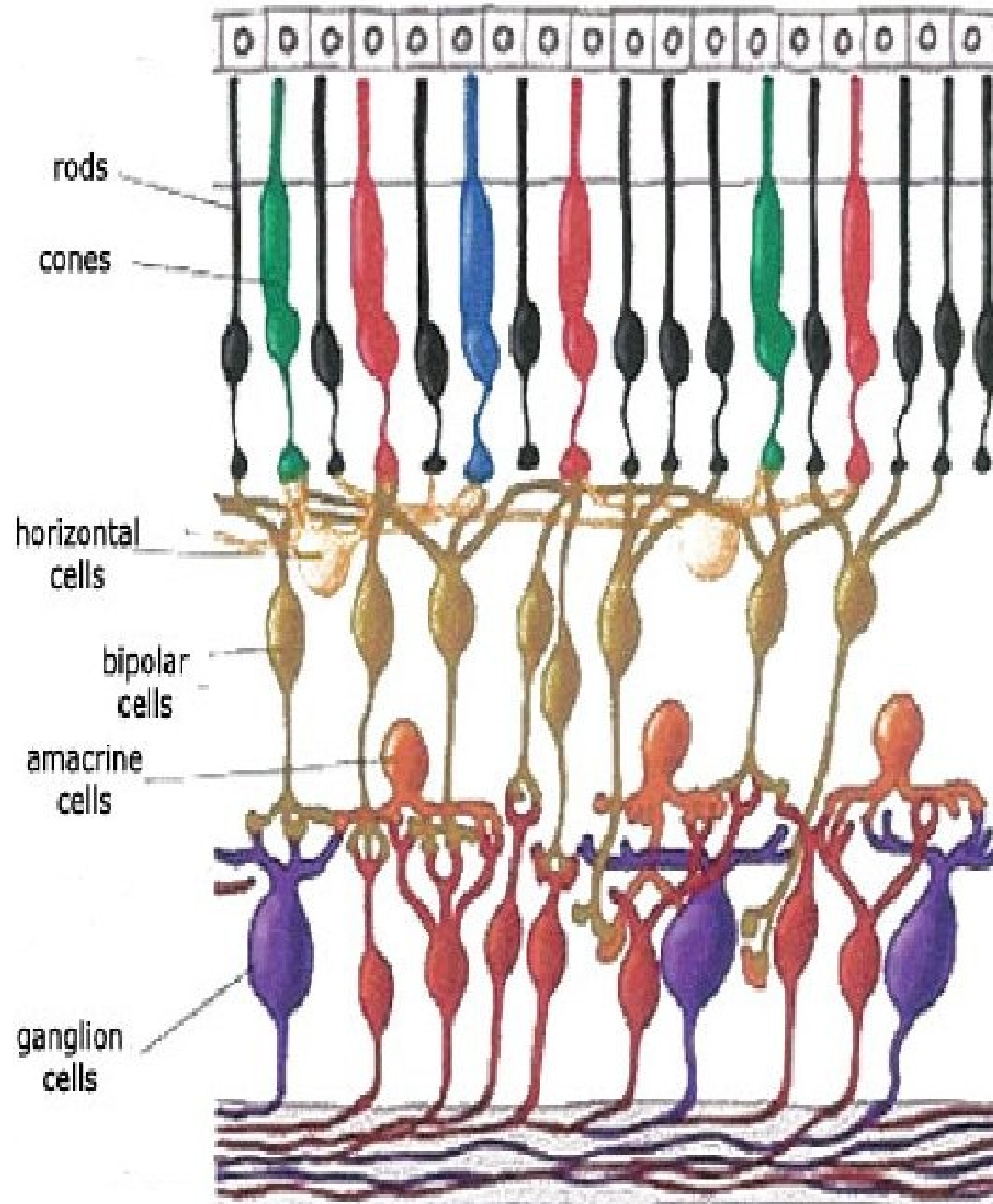


http://webvision.med.utah.edu

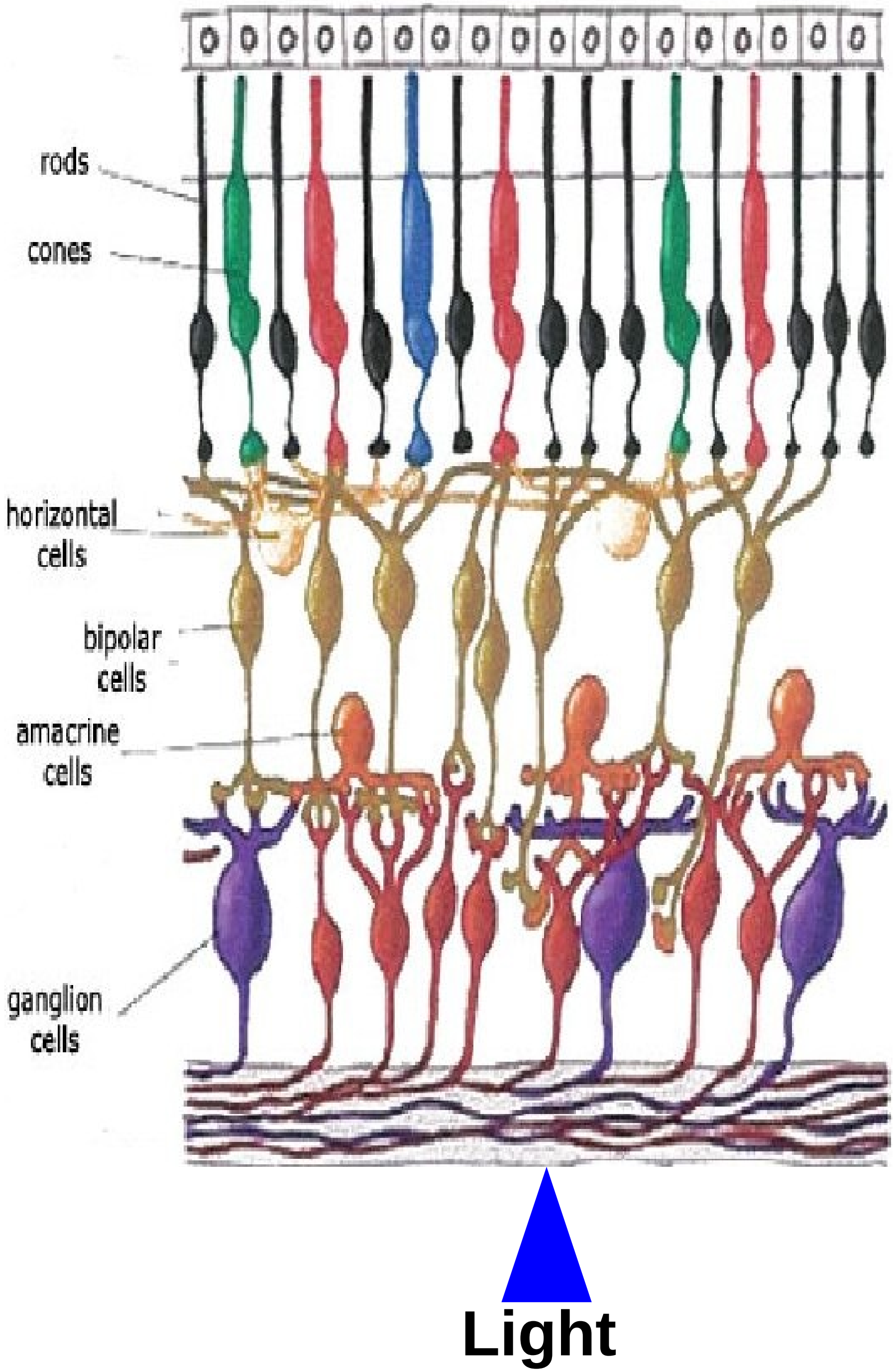
Eye cells involved in seeing

- The area of the eye we are investigating is the retina; it is at the back of the eye and it is where light entering the eye is sensed
- The cells in the retina are (in order of processing):

1. Cone and rod cells
2. Bipolar cells
    (a) Horizontal cells
3. Ganglion cells
    (a) Amacrine cells
4. To the optic nerve



Light and retinal structure

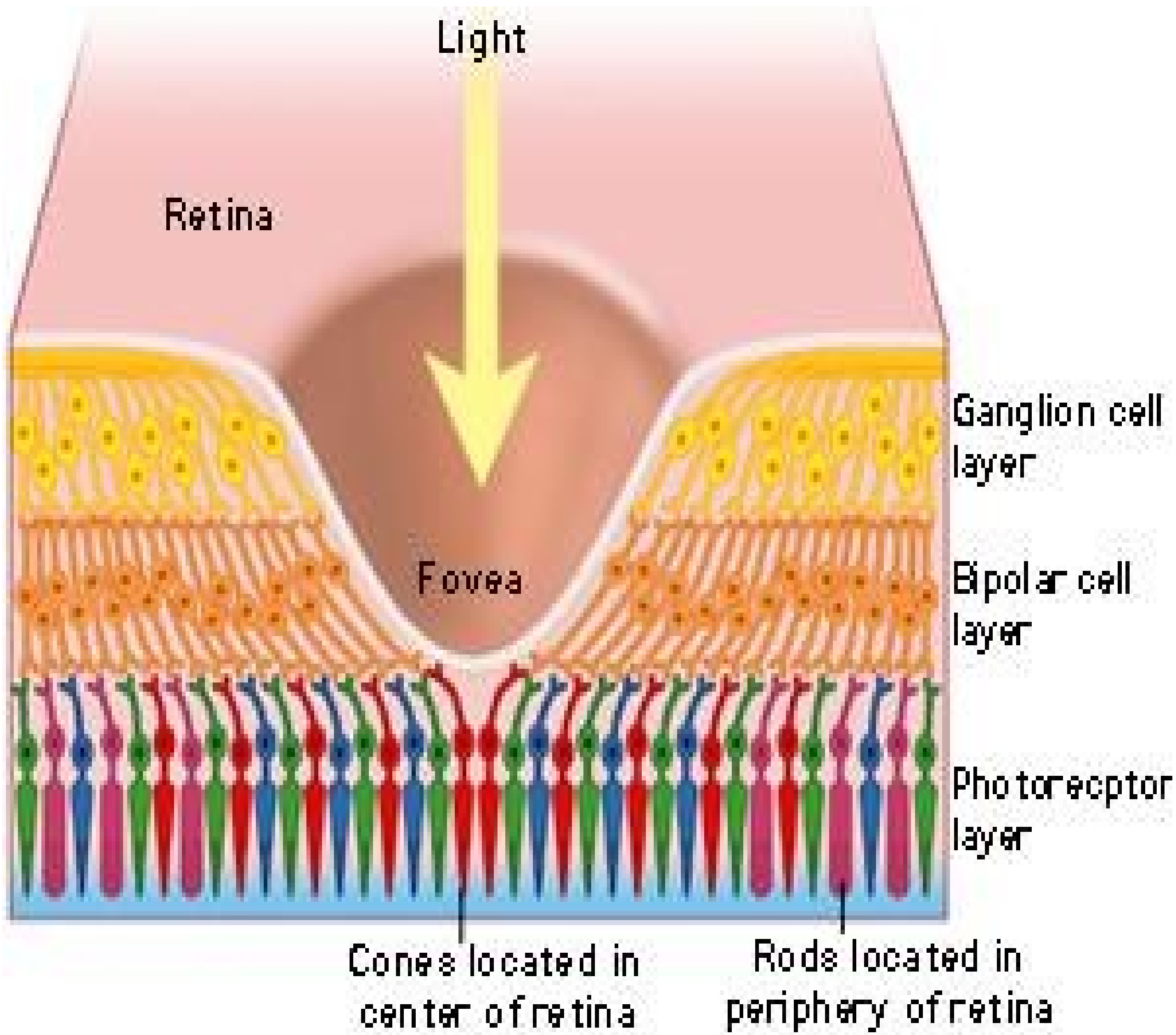


Note the direction of the light and the counter intuitive location of the photoreceptors at the **back** the retina (and eye). Why, and does this not affect vision?

1. Nutrient source: the epitial layerprovides a supply of retinene (retinaldehyde), a light-sensitive derivative of Vitamin A.

2. Reduces glare: The dark epithelial layer absorbs photons not absorbed by the photoreceptors thus reducing stray light effects

Light and retinal structure 2



Source: https://sites.google.com

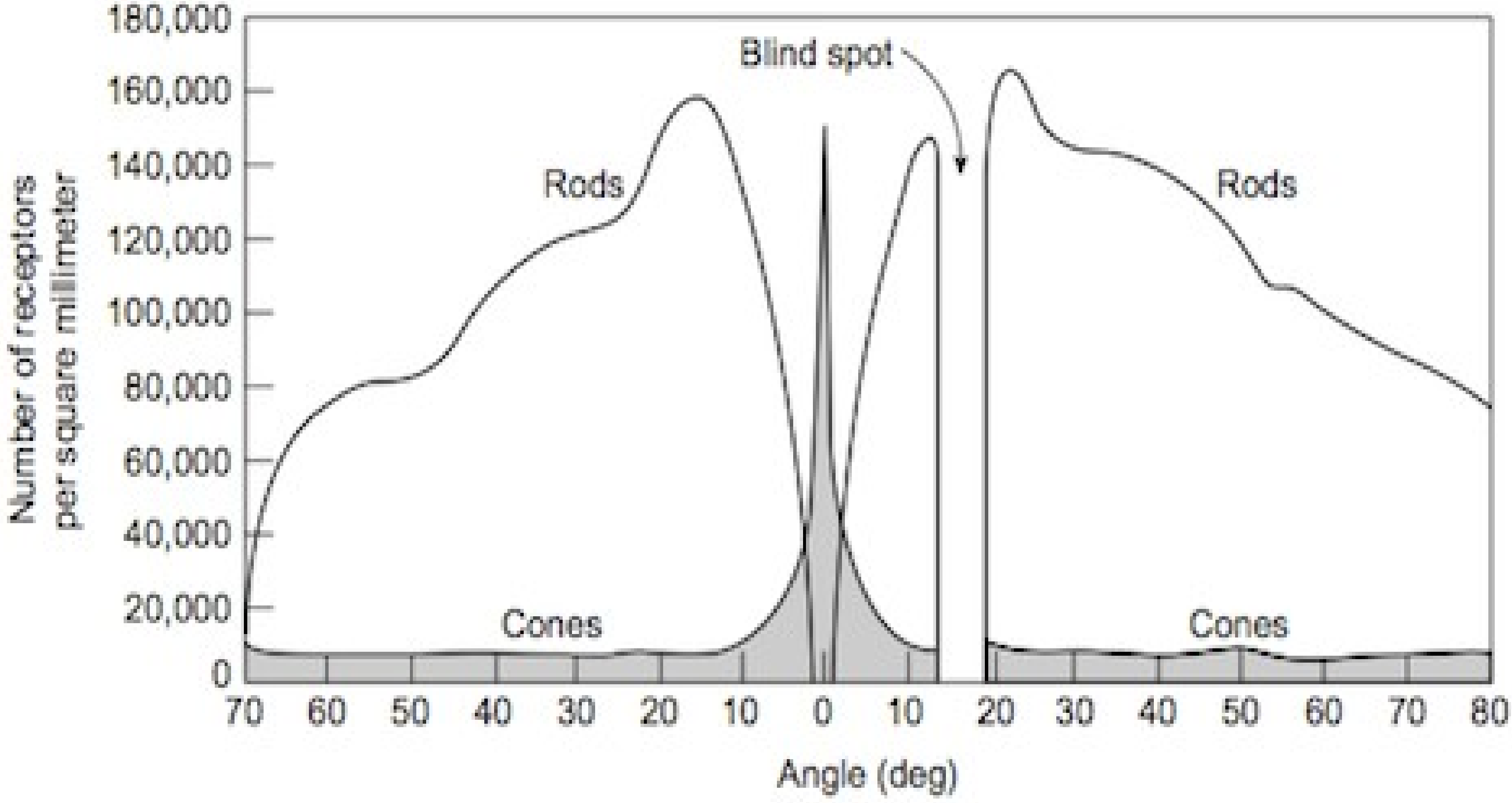
If you noted in the diagram of the eye earlier there was a region in the retina called the the fovea, looks like a depression in the retinal layer.

This region has:

- The highest density of photoreceptors (cones)
- Has the overlying layer of blood vessels and other retinal cells shoved aside so that light hits the photoreceptors directly.

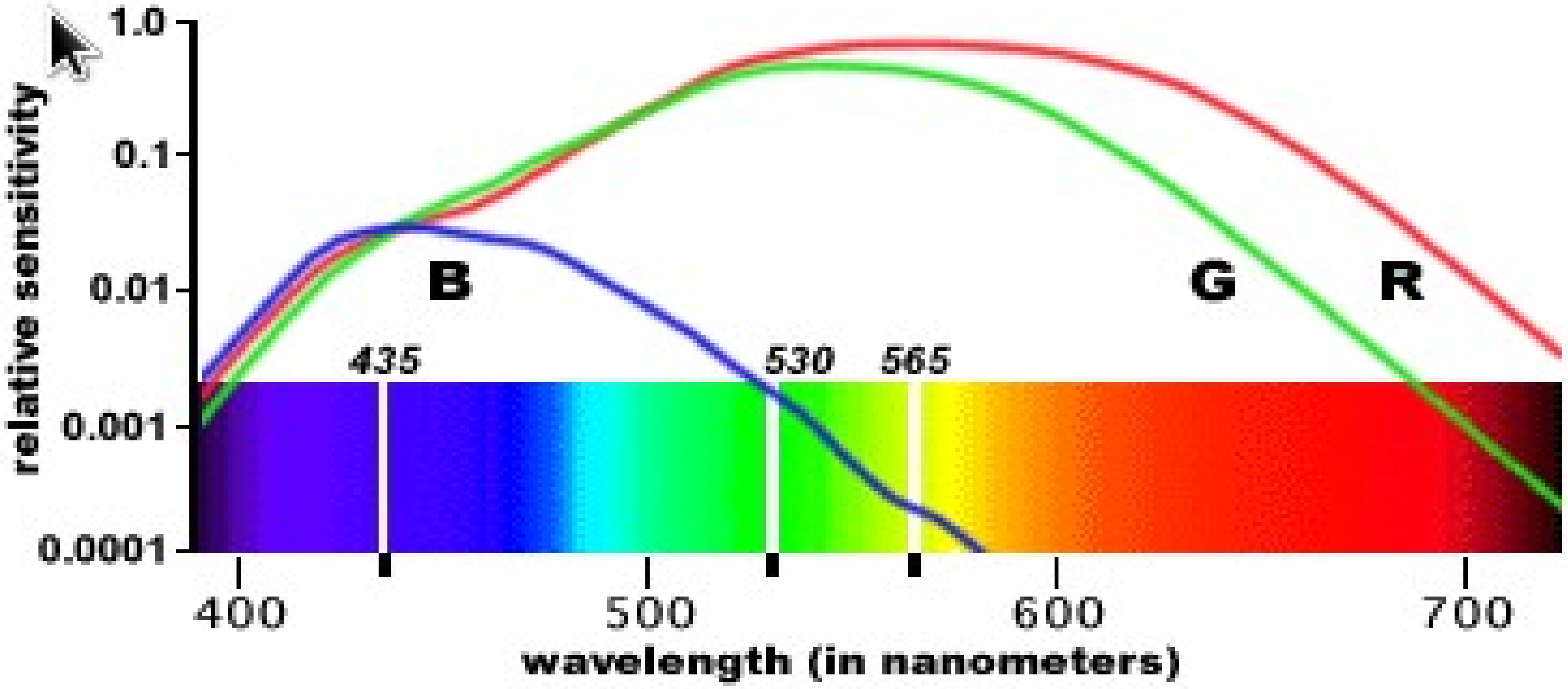
Cones and Rods: the basics

- Rods sense monochrome light and are more sensitive than cones; they are mainly active in low light conditions
- Cones require more light for activity and are sensitive to color (3 types – Red, Green, Blue)
- Photoreceptors transform photons to chemical signals through a series of processes known as photobleaching and **hyperpolarization**



Sensitivity of Cones and Rods

Cones are the daytime sensors, there are three types of cones: short, medium and long wavelenghts (also blue, green, and red sensitive), but not really. Blue cones peak at 435nm, green at 530nm and red at 565nm. Rods (monochrome) peak at 505nm.



Source: http://www.handprint.com/

Photobleaching and Hyperpolarization of Photoreceptors:

General sequence

1. Photon hits a rod And is absorbed by Pigment rhodopsin
2. Rhodopsin photo-bleaches, splits into Retinal and odopsin
3. One opsin molecule Activates many transducin enzymes
4. Transducin enzymes reduce cyclic Guanosine mono-Phosphate (cGMP)
5. Lower cGMP reduces Photoreceptor cell Pore openings
6. Lower cell pore Openings reduce Na ion flow into cell
7. Less Na ion = less Current potential, From -40 to -70 mV
8. Cell membrane Hyperpolarizes, Transmitter declines
9. Neurotransmitter change sends signal To bipolar cell

For more detail see Wikipedia reference: Visual phototransduction

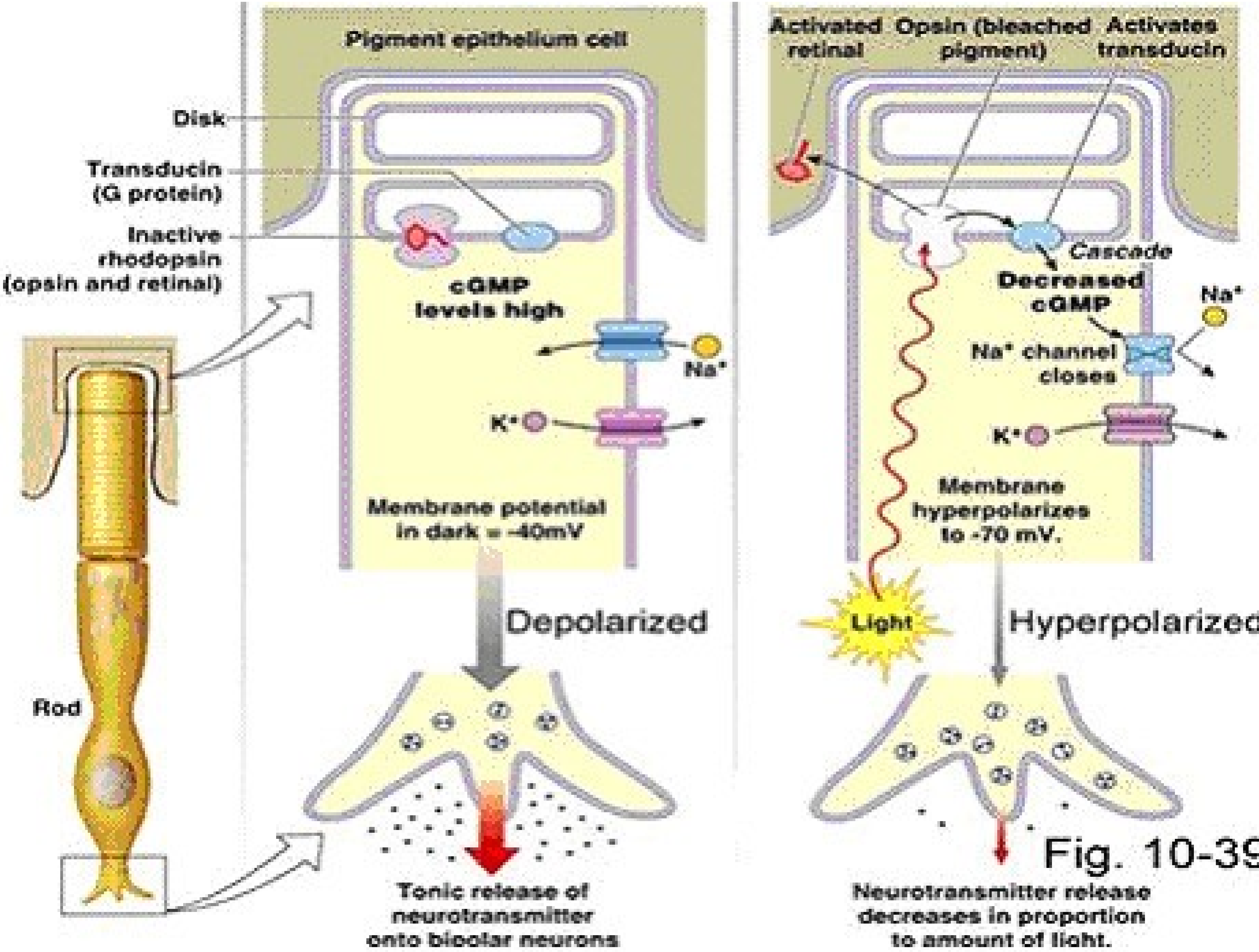
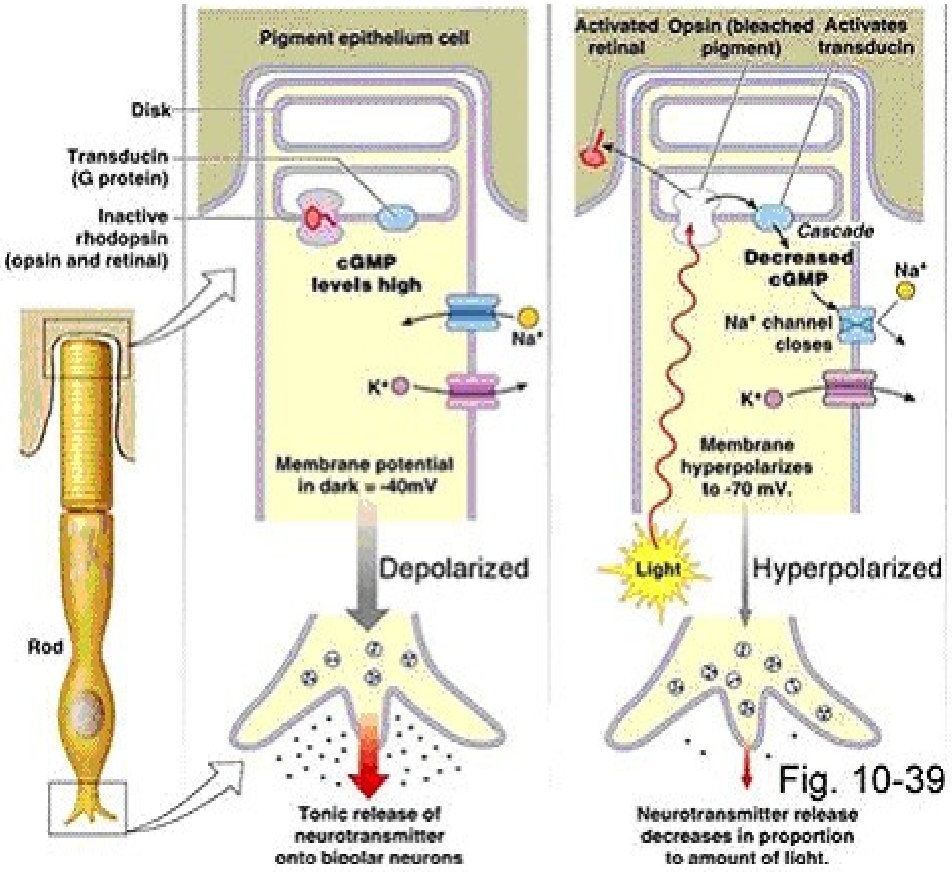

Pigment epithelium cell
Disk
Transducin
(G protein)
Inactive
rhodopsin
(opsin and retinal)
cGMP
levels high
Na⁺
K⁺
Membrane potential
in dark = -40mV
Depolarized
Rod
Tonic release of
neurotransmitter
onto bipolar neurons
Activated
retinal
Opsin (bleached
pigment)
Activates
transducin
Cascade
Decreased
cGMP
Na⁺
Na⁺ channel
closes
K⁺
Membrane
hyperpolarizes
to -70 mV.
Light
Hyperpolarized
Neurotransmitter release
decreases in proportion
to amount of light.

Fig. 10-39

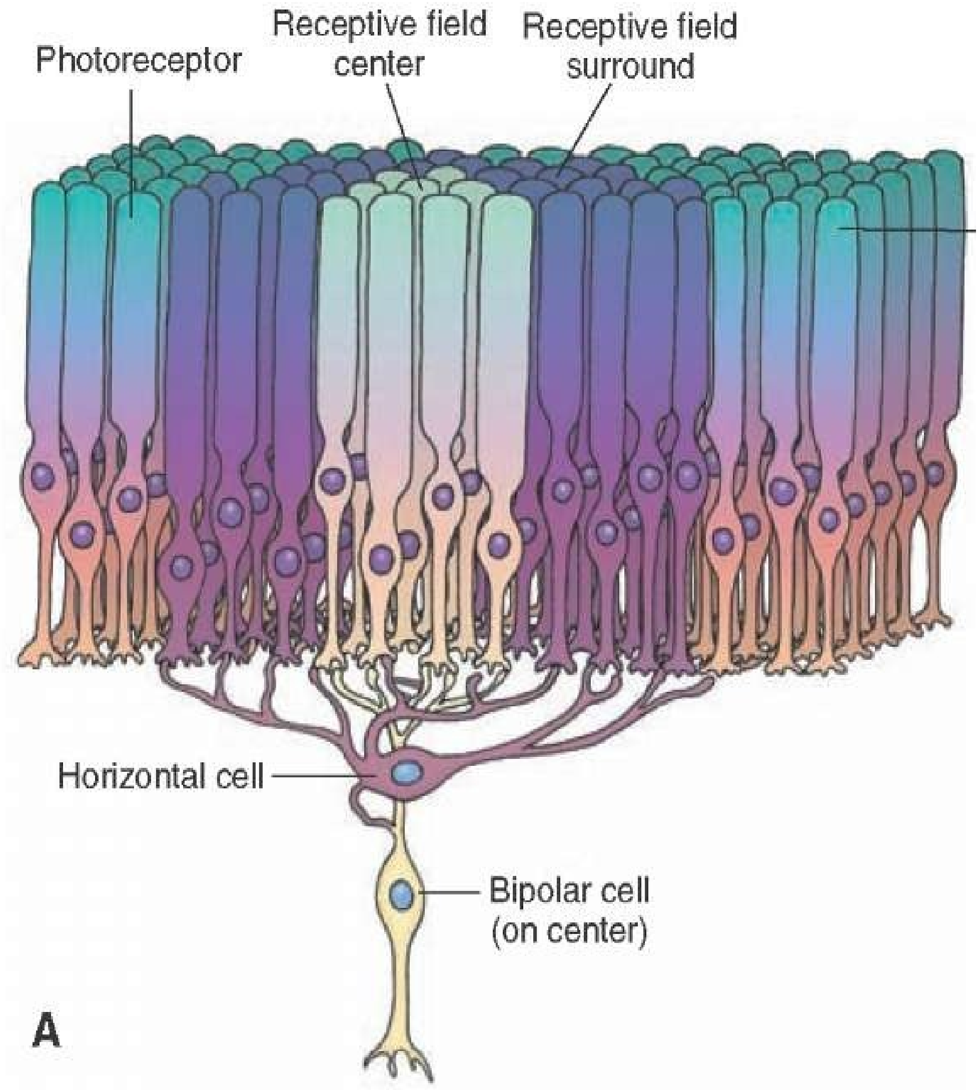
Photobleaching and Hyperpolarization of Photoreceptors: Key points



- Light <u>reduces</u> the release of neurotransmitter from a rod or cone cell
- A light stimulated rod or cone changes **from** a <u>depolarized</u> state **to** a <u>hyperpolarized</u> state.
- More light stimulus less signal transmission from rods and cones.

The Retinal Bipolar Cell

- Bipolar cells exists between preceptors (rod cells and cone cells) and ganglion cells.
- Signals from the photoreceptors <u>must</u> pass through bipolar cells in their eventual destination to the brain
- Bipolar cells receive their signals either directly from from photoreceptors or indirectly via horizontal cells (we'll look at the function of horizontal cells shortly)
- Cones are connected one-to-one with bipolar cells, rods are connected many-to-one



Bipolar Cell Organization

- Bipolar cells receive their signals from <u>groupings</u> of photoreceptors. The groupings are known as the <u>*receptive field*</u> of the bipolar cells.
- These bipolar cell receptive fields are generally circular (diameter of 2.5-10μm) and are structured to work with the <u>*centre*</u> area and the <u>*surrounding*</u> area having two opposite effects.
- These effects correspond to two types of bipolar cells known as 'on' or 'off' cells.
- Consider: we have two areas in the receptive field and two types of bipolar cells, what are our possible combinations?
  - Centre 'on' cells <u>and</u> surround 'off'
  - Centre 'off' cells <u>and</u> surround 'on”
- Remember, the centre and surround areas of the receptive field is always opposite, we cannot have both centre and surround 'on' or centre and surround 'off'.

Bipolar Cell Function

The Excitatory Circuit

- Light 'on'
- Photoreceptors decrease neurotransmission
- Bipolar cell connected to this photoreceptor is activated (depolarized).
- **This is an 'on' bipolar cell** (Activated by light 'on').
- It is situated in the centre or the surround (but not both) of the receptive field

The Inhibitory Circuit

Light 'on'

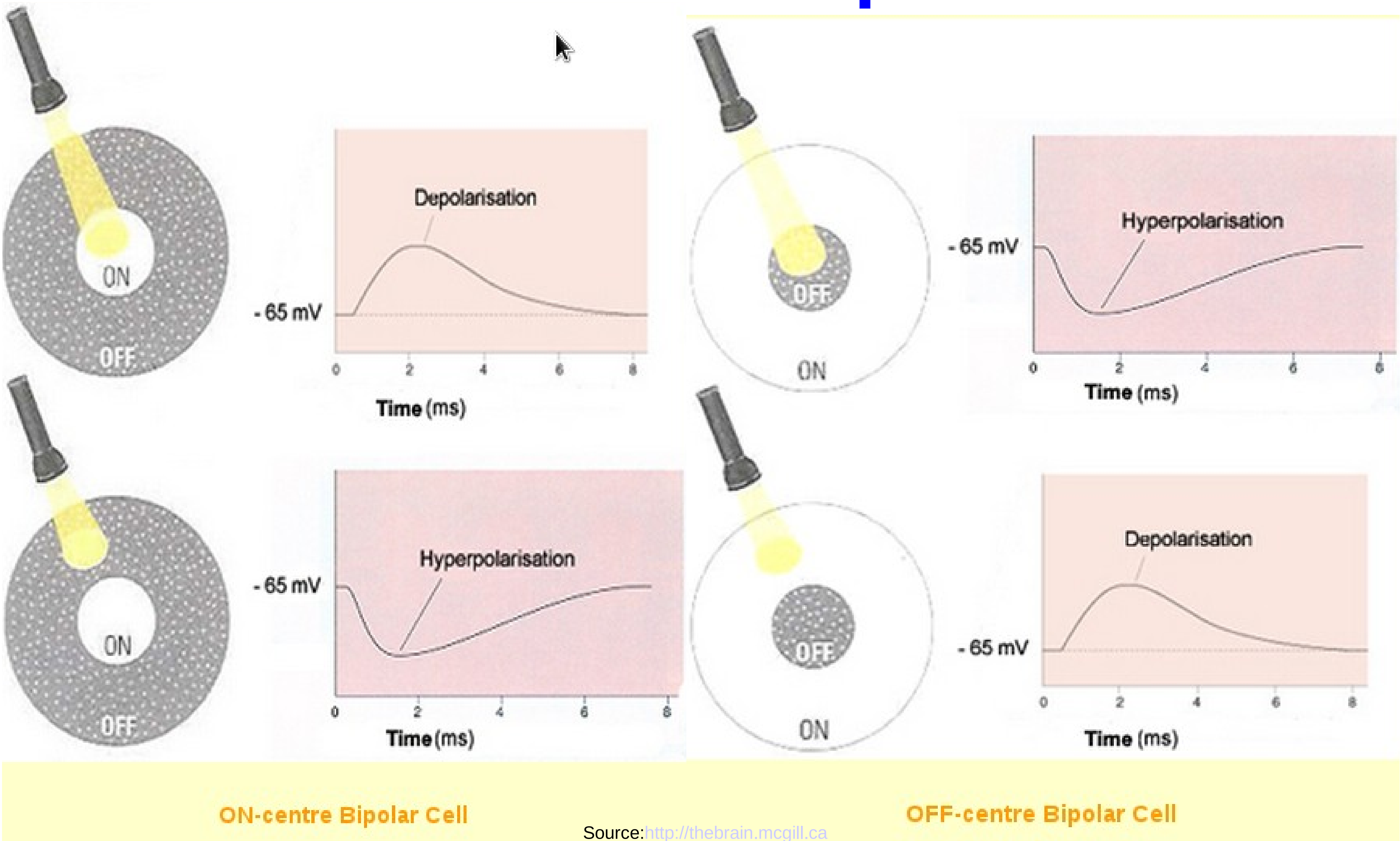
Photoreceptors decrease neurotransmission

Bipolar cell connected to this photoreceptor *decreases* activity (hyperpolarized).

This is an 'off' bipolar cell (deactivated by light 'on')

It is situated in the centre or the surround (but not both) of the receptive field

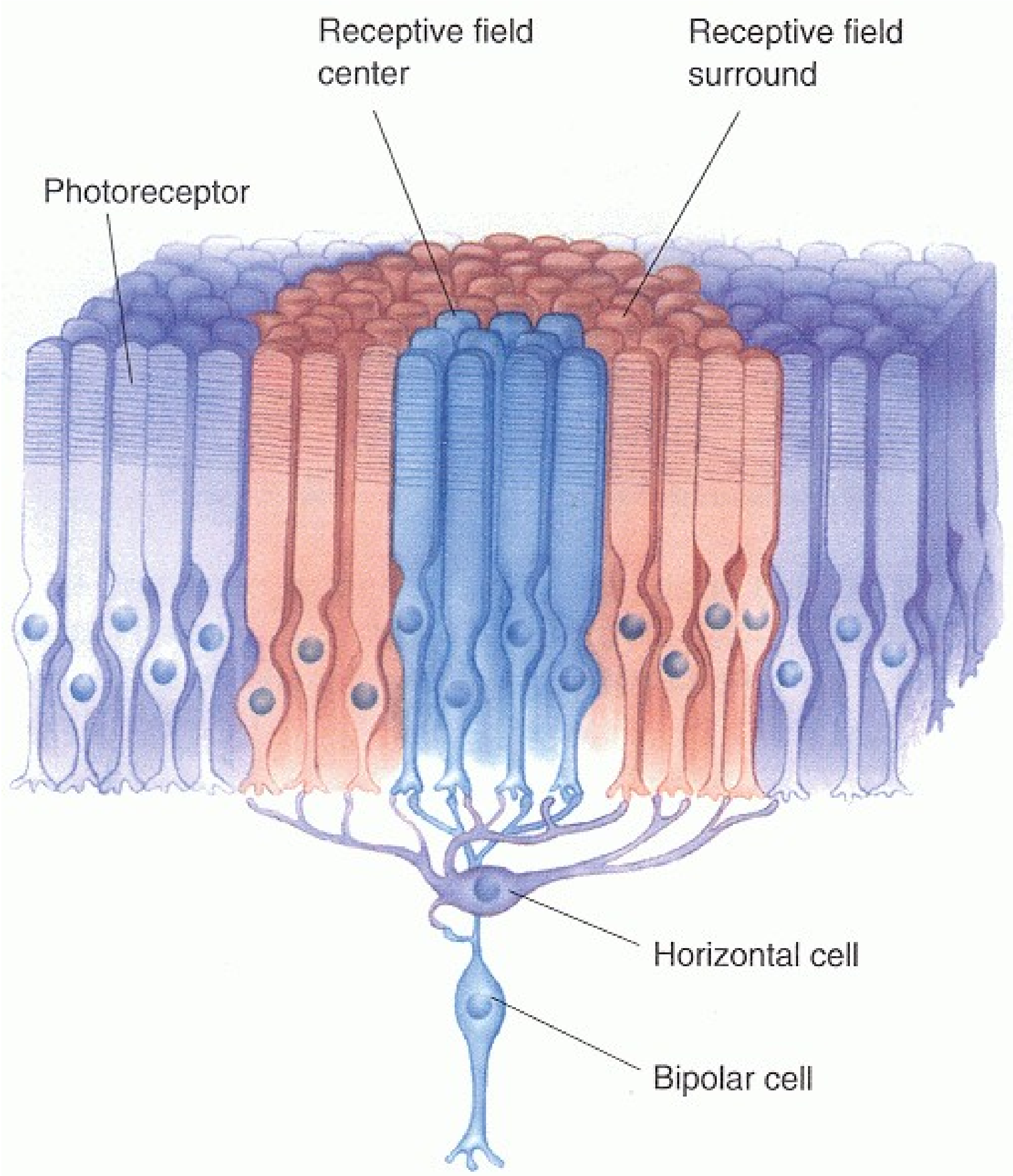
A diagrammatical view of 'on-off', 'centre-surround' bipolar cells



Source:http://thebrain.mcgill.ca

Horizontal Cells

- Horizontal cells process process signals **from/to** photoreceptors and send signals **to** bipolar cells
- The represent an **indirect** pathway for visual signal transmission.
- They are connected to many photoreceptors and enhance the centre-surround effect of bipolar cells. They release GABA (gamma Aminobutyric acid) when stimulated.
- They allow dark objects to be detected against bright backgrounds or vice-versa



http://webvision.med.utah.edu

Horizontal Cell Function

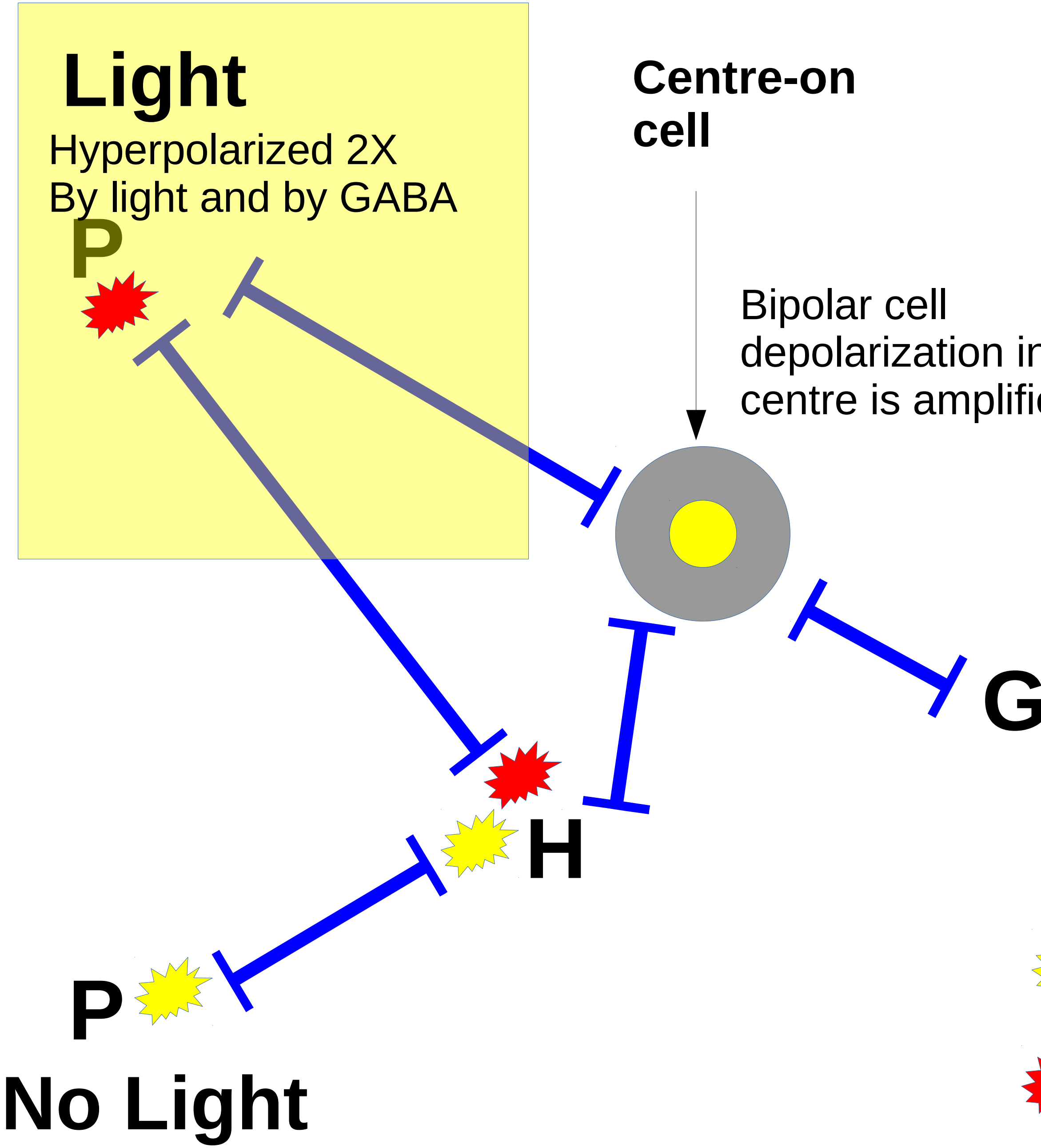
The major functional roles of horizontal cells in the retina are

1) Initiate the organization of spatially opponent receptive fields ie. centre-surround organization (integration function), via lateral connections with photoreceptors.

2) Modulate of the photoreceptor signal with different lighting conditions ie. reduce or increase the activity of neighbouring photoreceptors (**lateral inhibition**)

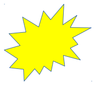
3) Adjustment of synaptic gain. (items 2 and 3 are regulation functions)

Horizontal Cells: Integrators and modulators of the receptive field

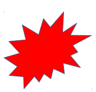


Visual Cell Actions

1. No light = Depolarized cell = release of neuro-transmitter.
2. Light = Hyperpolarized cell = no neuro-transmitter.
3. Photoreceptors depolarize horizontal cells
4. Horizontal cells hyperpolarize photoreceptors

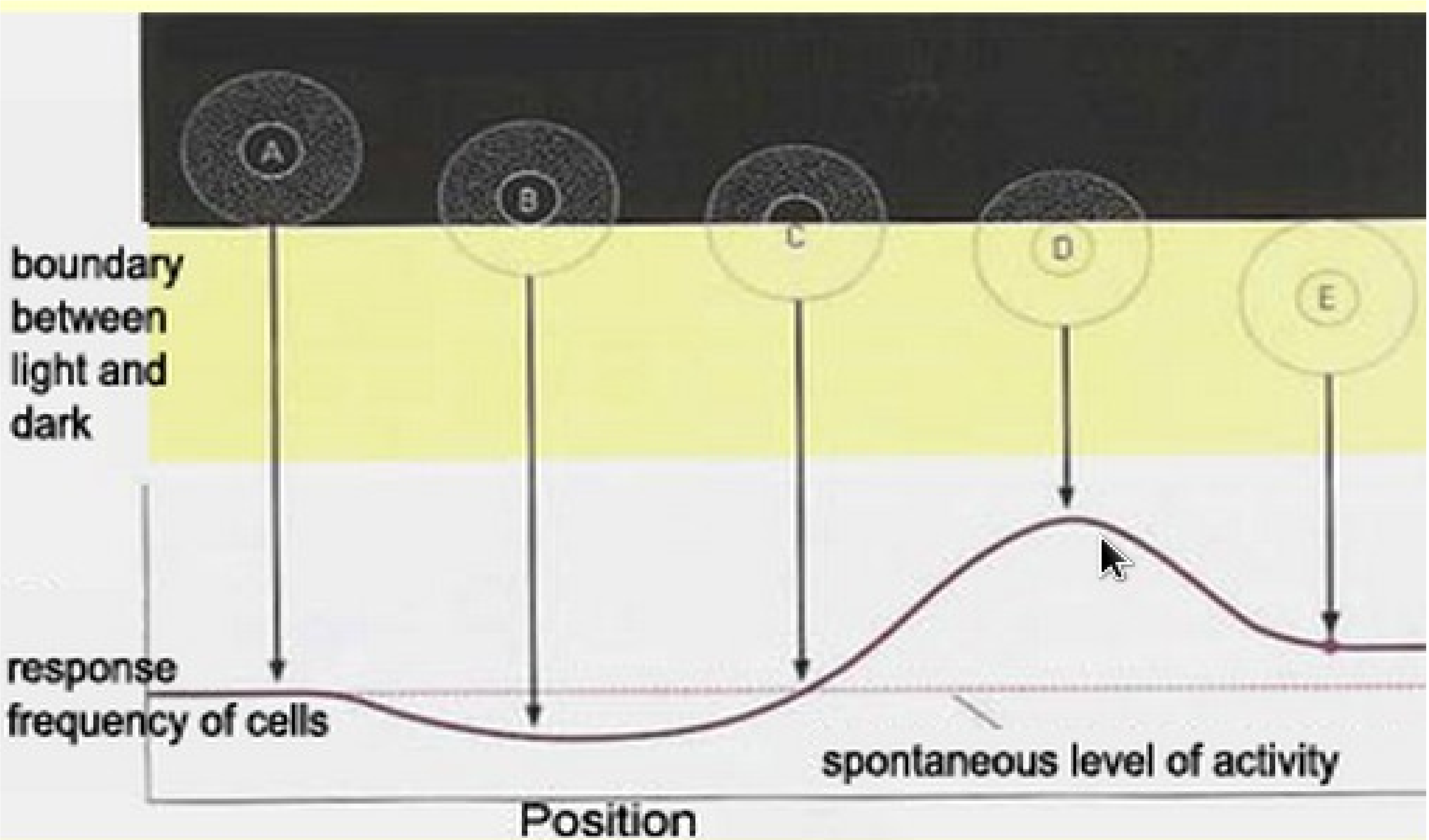


Glutamate (excitatory neurotransmitter)



GABA (inhibitory neurotransmitter)

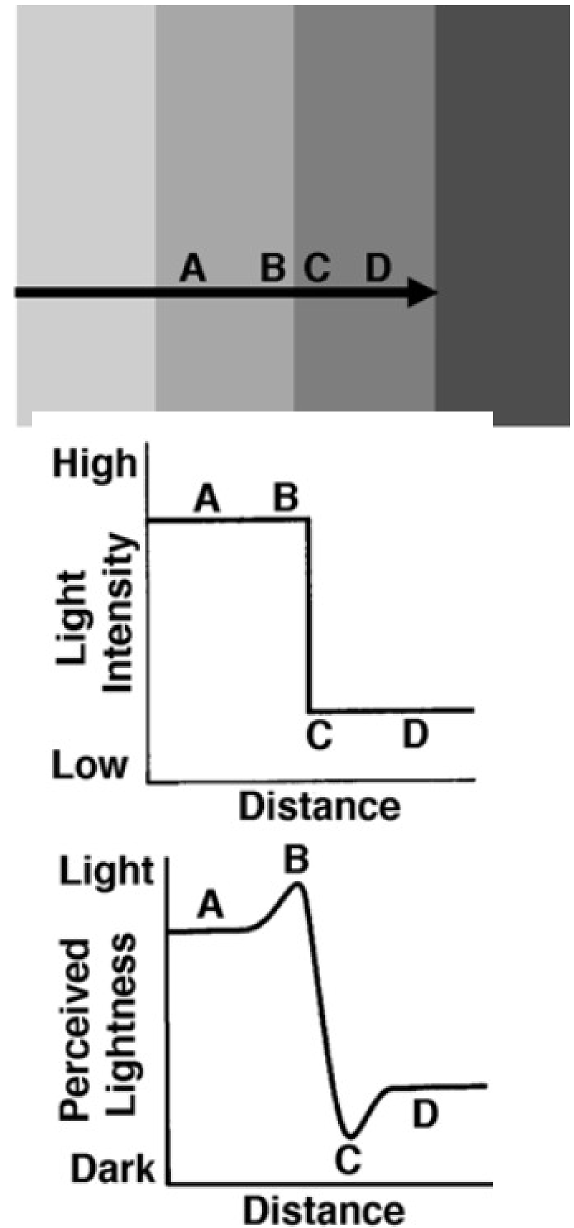
Centre-Surround Edge Detection



Source: http://thebrain.mcgill.ca

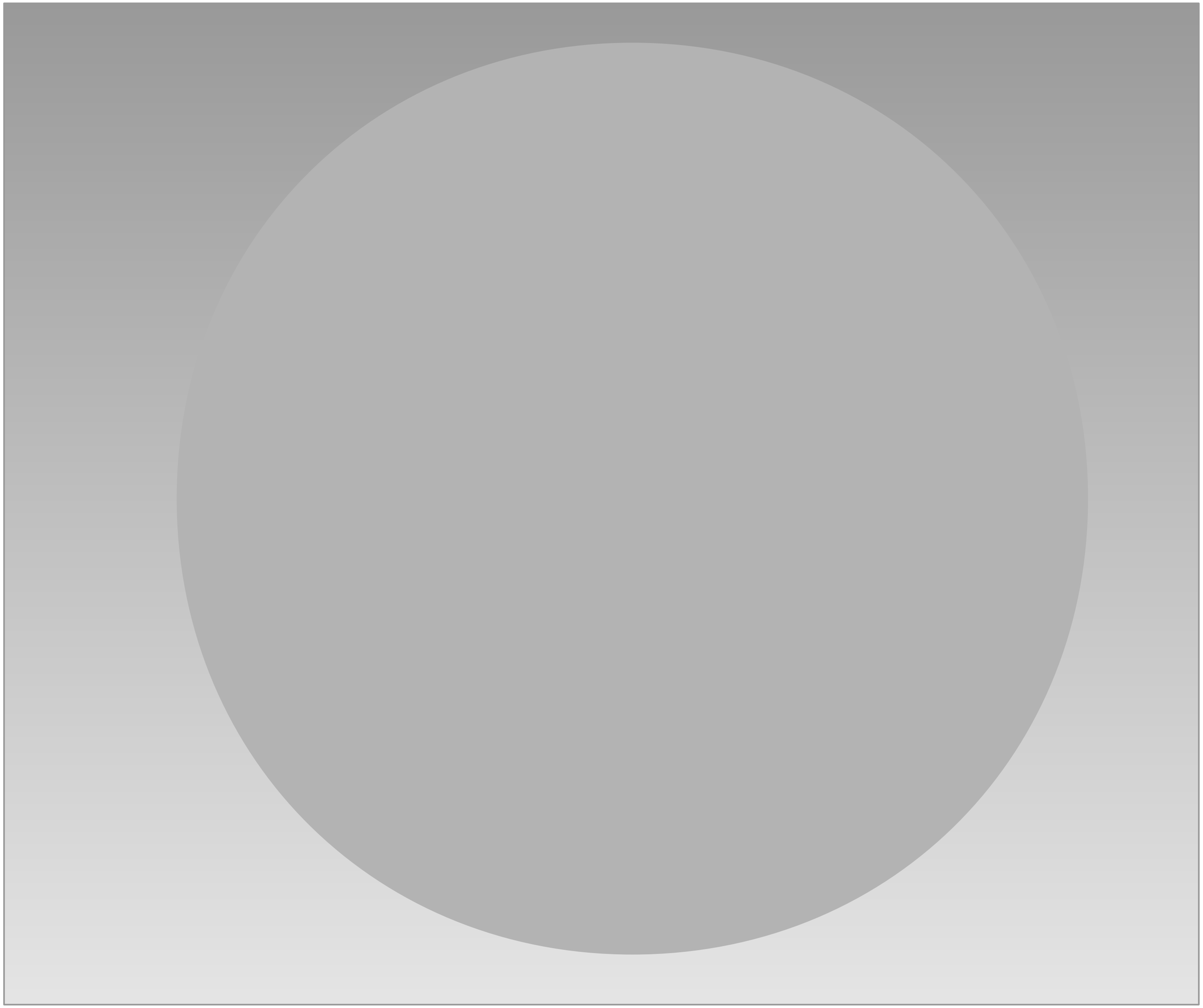
Horizontal Cell modulators of the receptive field

As lateral connectors to many rods, cones, and bipolar cells, horizontal cells' role is to **inhibit** the activity of neighbouring cells. This selective suppression of certain nerve signals is called **lateral inhibition**, and its overall purpose is to increase the acuity of sensory signals. When light reaches the retina, it may illuminate some photoreceptors brightly and others much less so. By suppressing the signal from these less illuminated photoreceptors, the horizontal cells ensure that only the signal from the well lit photoreceptors reaches the ganglion cells, thus improving the contrast and definition of the visual stimulus. (example is Mach effect)



Center-Surround and Luminance

Center-Surround: example of enhanced edges



Centre-Surround Bipolar/Horizontal cell stimulation fully engaged.

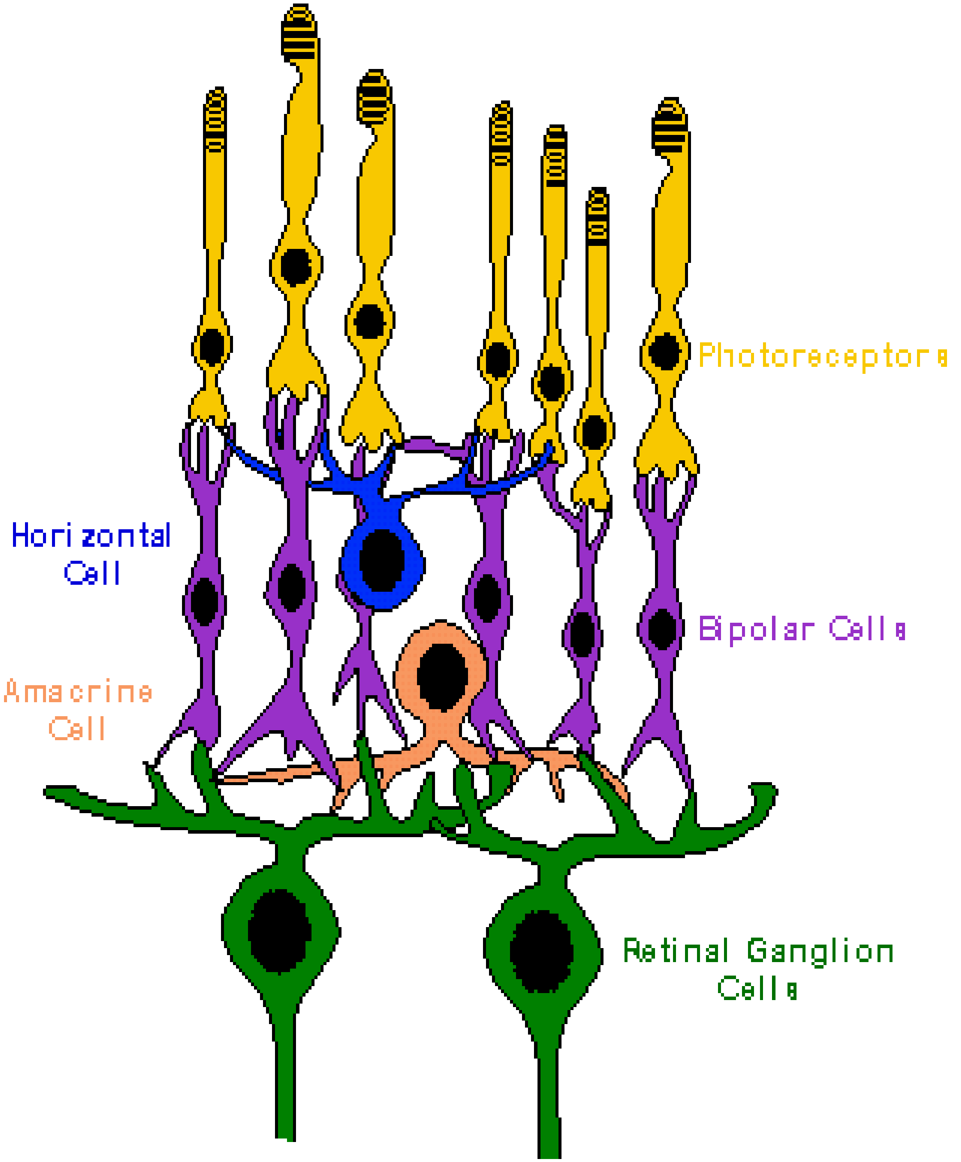
Note the almost 3D effect in the Shading.

How would you use this feature of the visual system in your photography?

What effect is now being enhanced by your eyes?

Ganglion Cell Organization

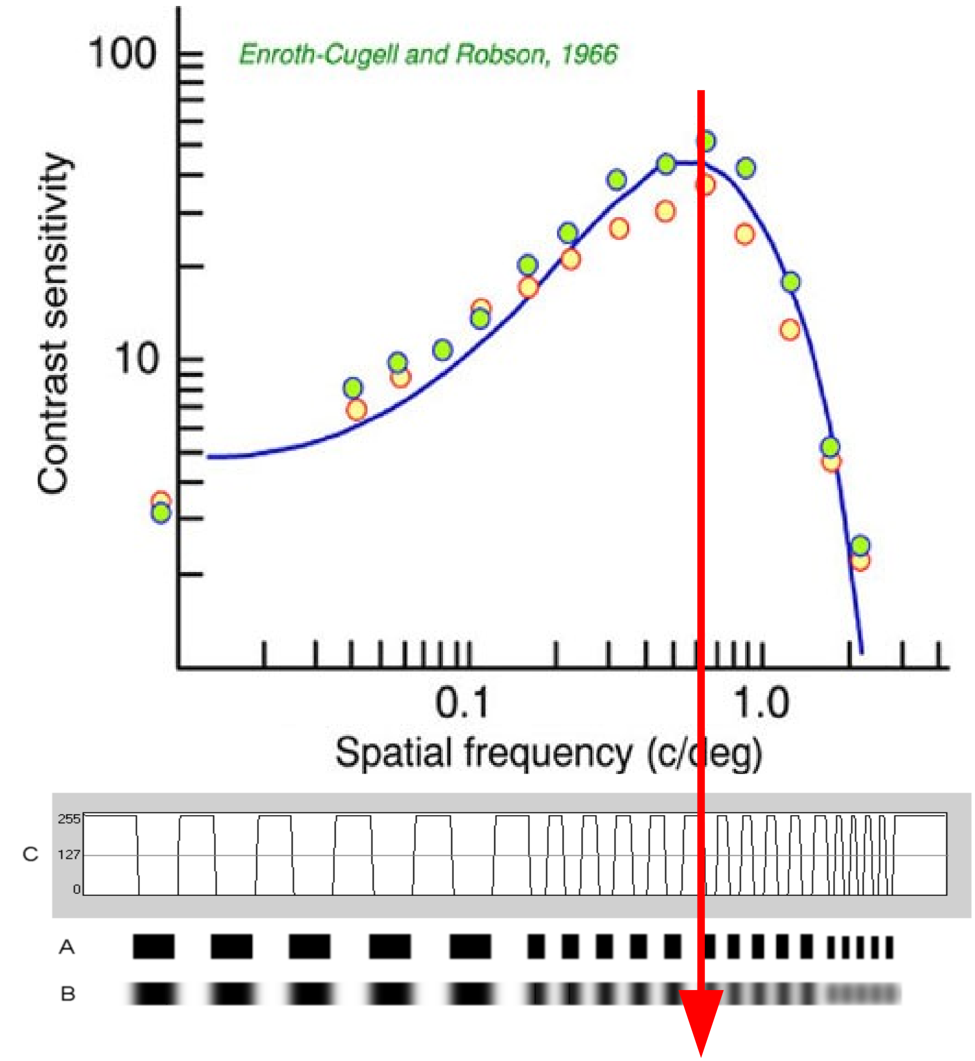
- Last point of retinal visual processing before signal is sent to the brain via the optic nerve.
- Also configured as centre-surround reflecting the bipolar cell on-off circuits.
- Along with Amacrine cells integrate and refine contrast differences of bipolar cell responses.
- Provide information important for detecting the shape (P-type) and movement (M-type) of objects (not reviewed) also bi-stratified (combination)



Ganglion Cell Functions

- 'Spatial tuning' or size selectivity of ganglion cell receptive fields is reflected in peaked contrast sensitivity functions.
- Separate rod and cone pathways converge
- **Tonic or Midget** ganglion cells respond to weak contrast and fine grain, **phasic or parasol** types respond to high contrast and wide areas.
- As already noted: centre-surround enhancement and 'on-off' refinement.
- Ganglion cells also refine color responses.

Spatial Tuning and object size



~ 1cm at 5m

Type of Ganglion Cell Characteristics

	Midget	Parasol	bistratified
Nature of receptive field	Centre-srnd	Centre-srnd	Centre-srnd
Percent of ganglion cells	70%	20%	10%
Speed of conduction	Slow	Fast	Very slow
Sensitivity to details (acuity)	High	Low	Low
Sensitivity to light	Low	High	Low
Sensitivity to wavelength	Yes	No	Yes

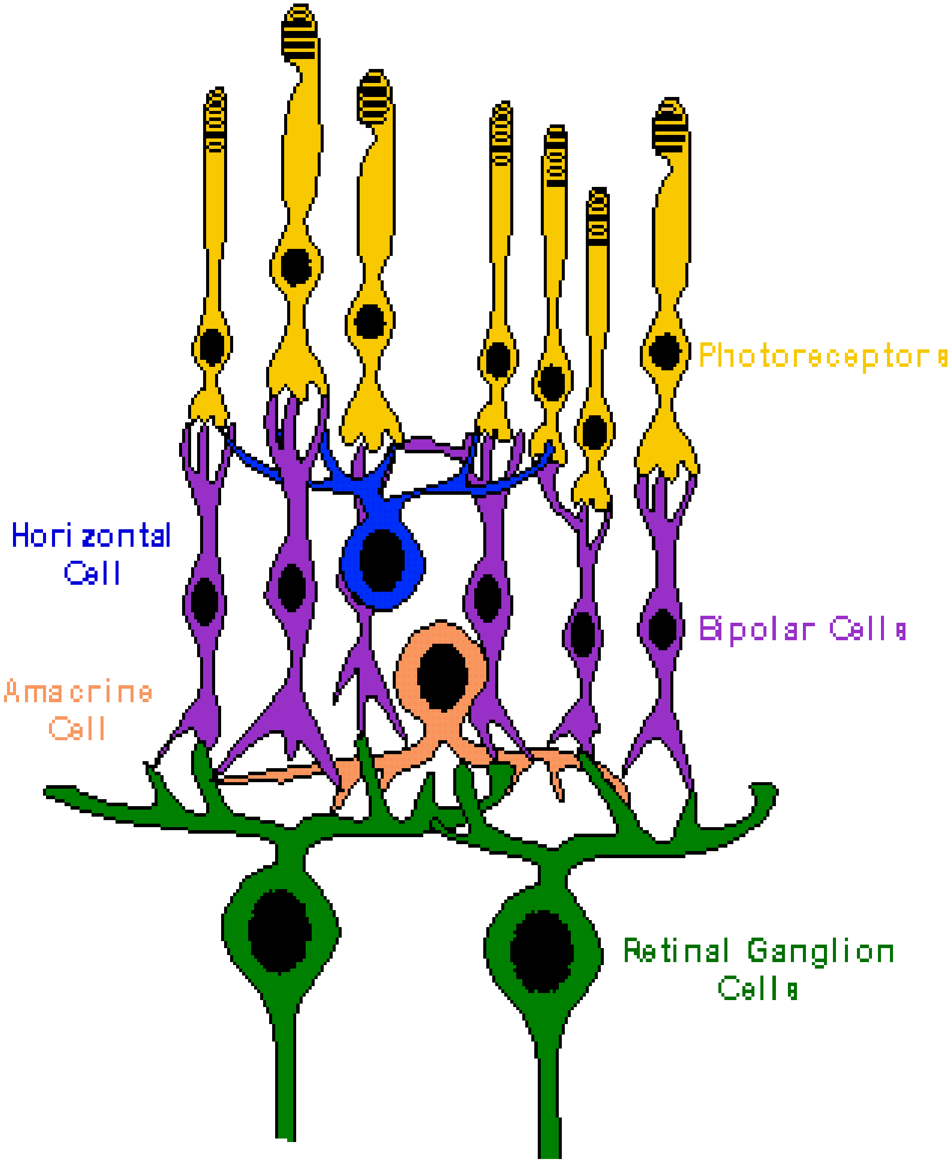
Amacrine Cells

Like horizontal cells, amacrine cells *enhance* basic visual stimulus. And, like horizontal cells, they tend to be connected laterally at the ganglion cell level.

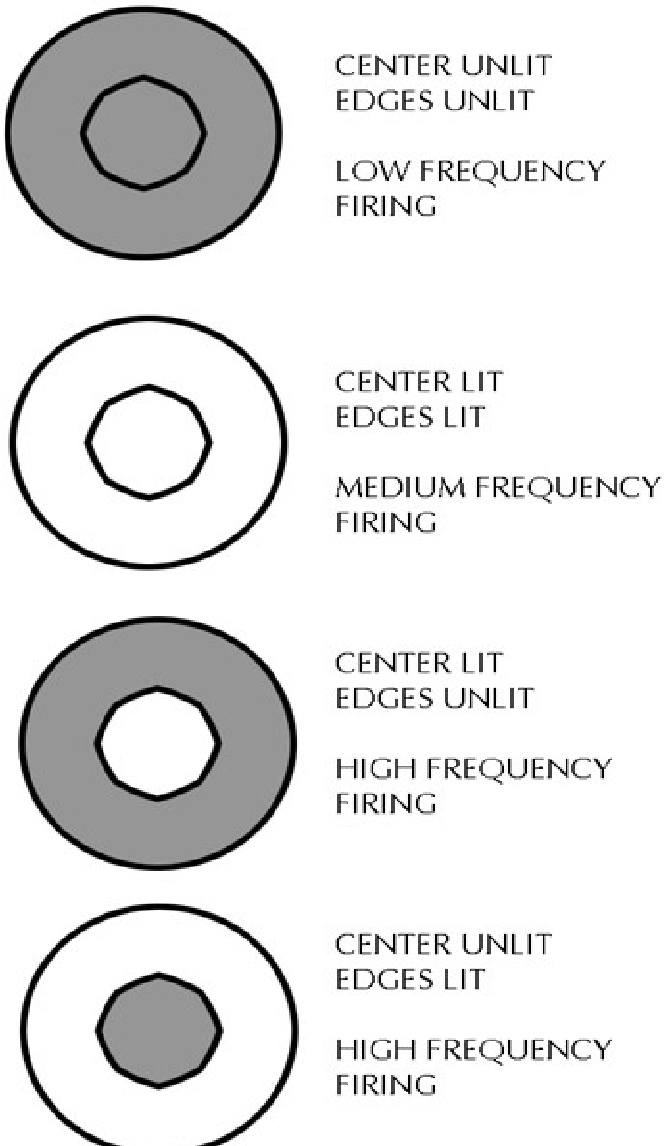
Play an important role in transmitting information from rod photoreceptors to ganglion cells.

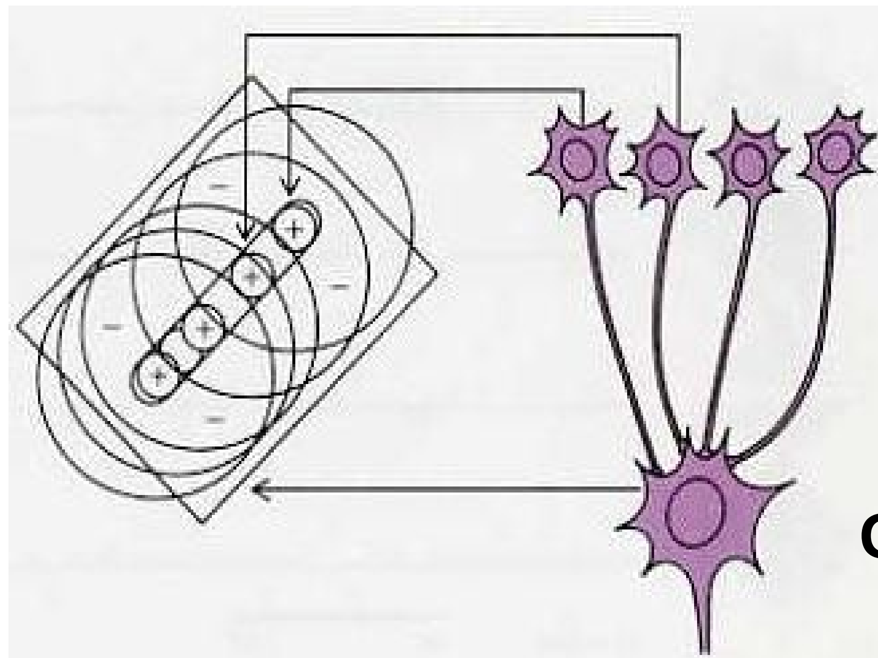
Collect messages from many rod-connected bipolar cells, allowing the perception of very dim light.

Amacrine cells feed information directly to OFF ganglion cells.



More challenges..



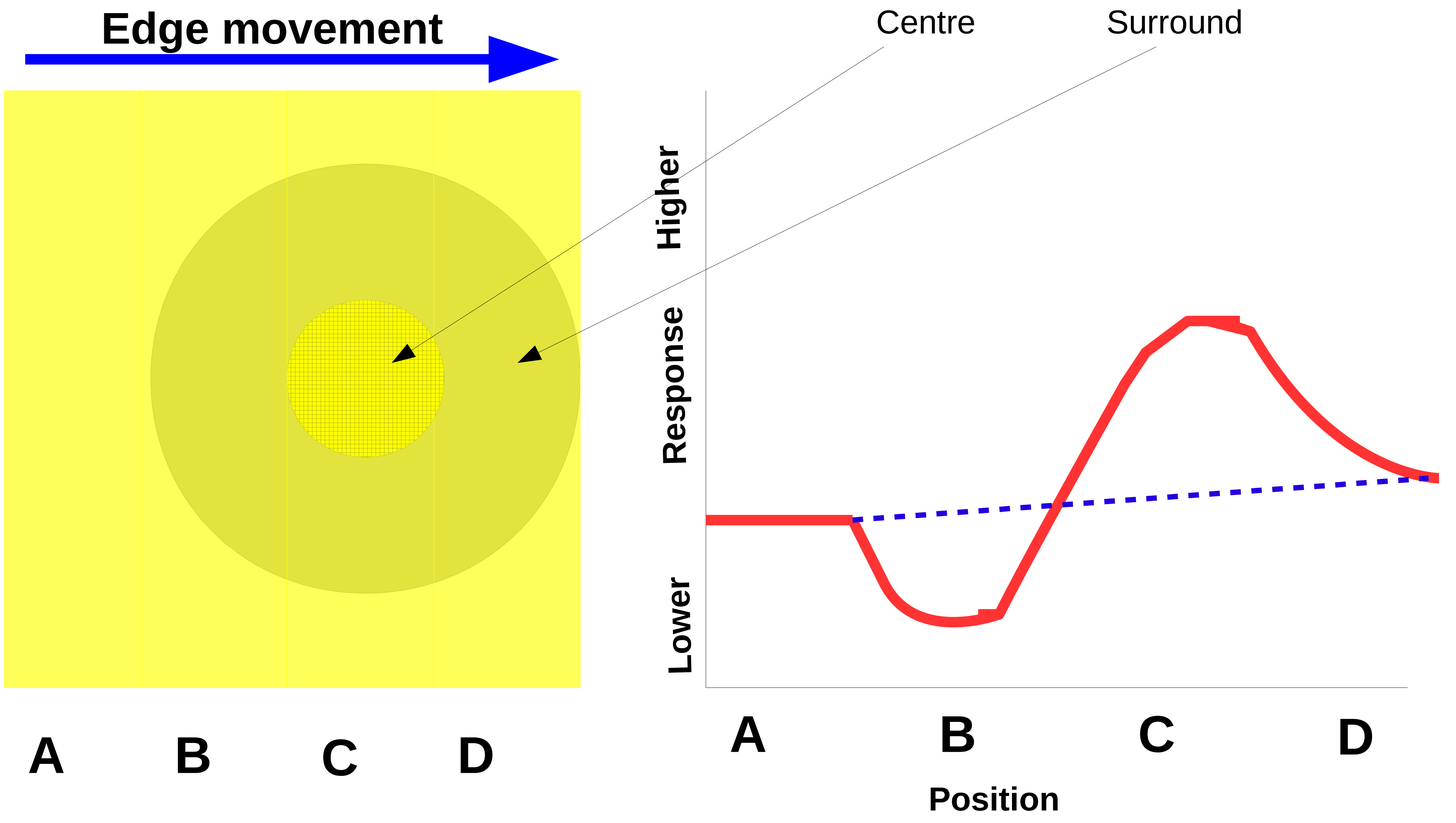


Bipolar cells

Ganglion cell

1. What type centre-surround are the bipolar/ganglion cells to the left?
2. What type of illumination or visual field would cause the conditions on the left?
3. Note the overlap of adjacent receptive fields on the right.

Centre-Surround Edge Detection



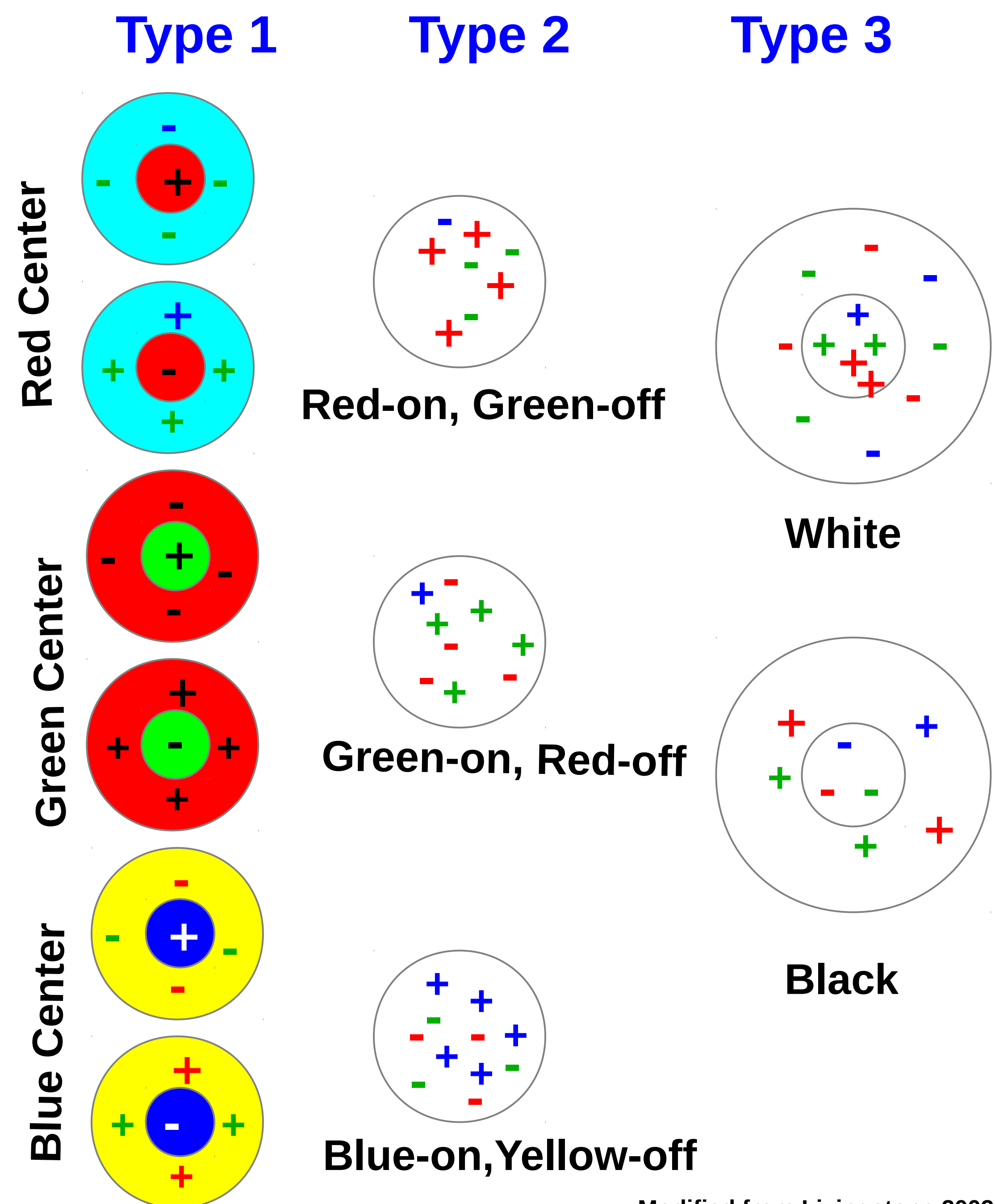
Color.. briefly

Color is detected by cones and aggregated in P-type ganglion cells.

Type 1 cells are structured with center-surround, on-off effects for color edge detection.

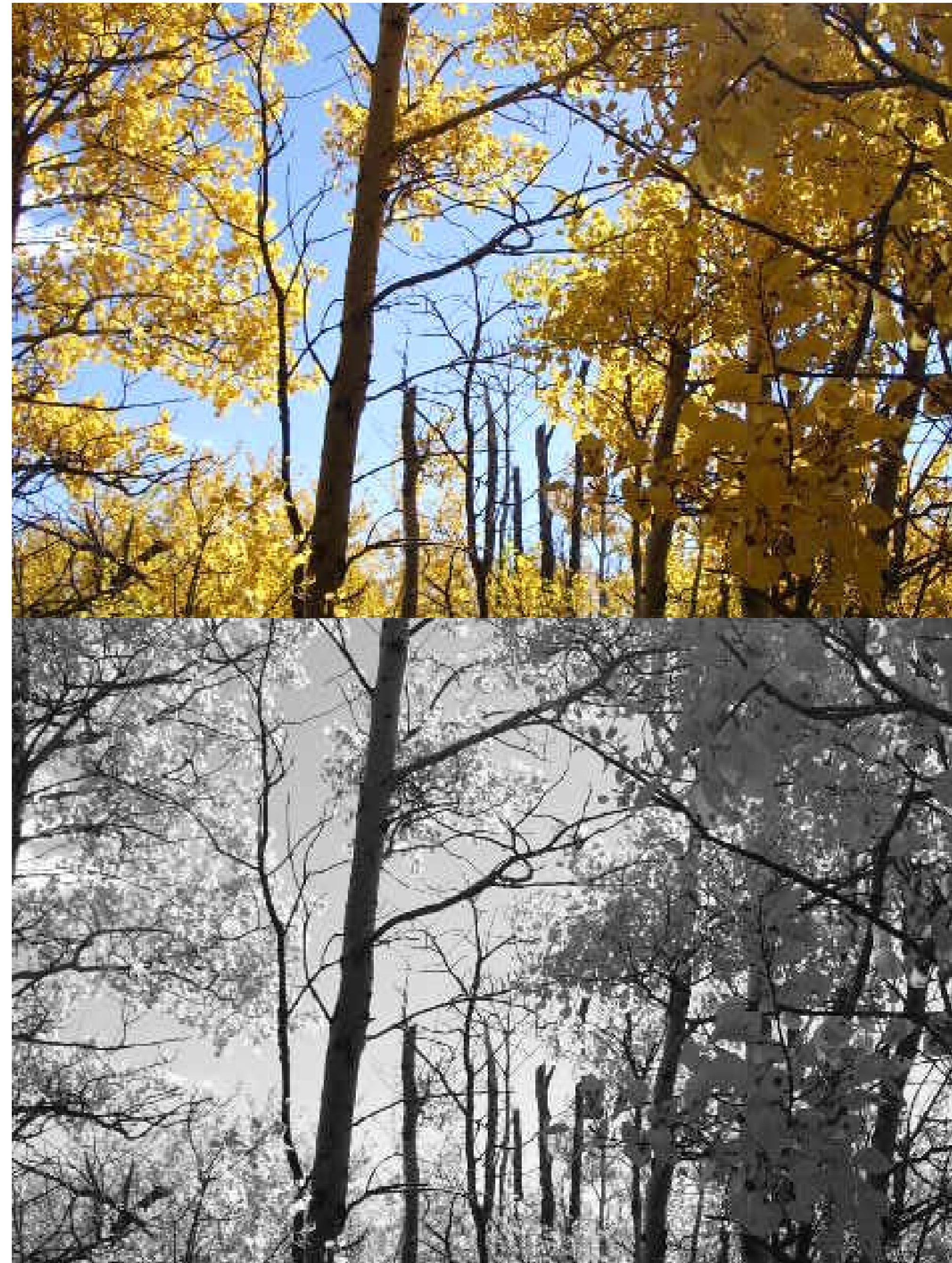
Type 2 cells (color opponent) are excited by one color and inhibited by another; they have color opponent centers and **<u>no</u>** surrounds (color enhancers)

Type 3 cells are activated by strict luminosity they represent the cells we've been used to seeing in most of our discussion (respond to white-grey-black)



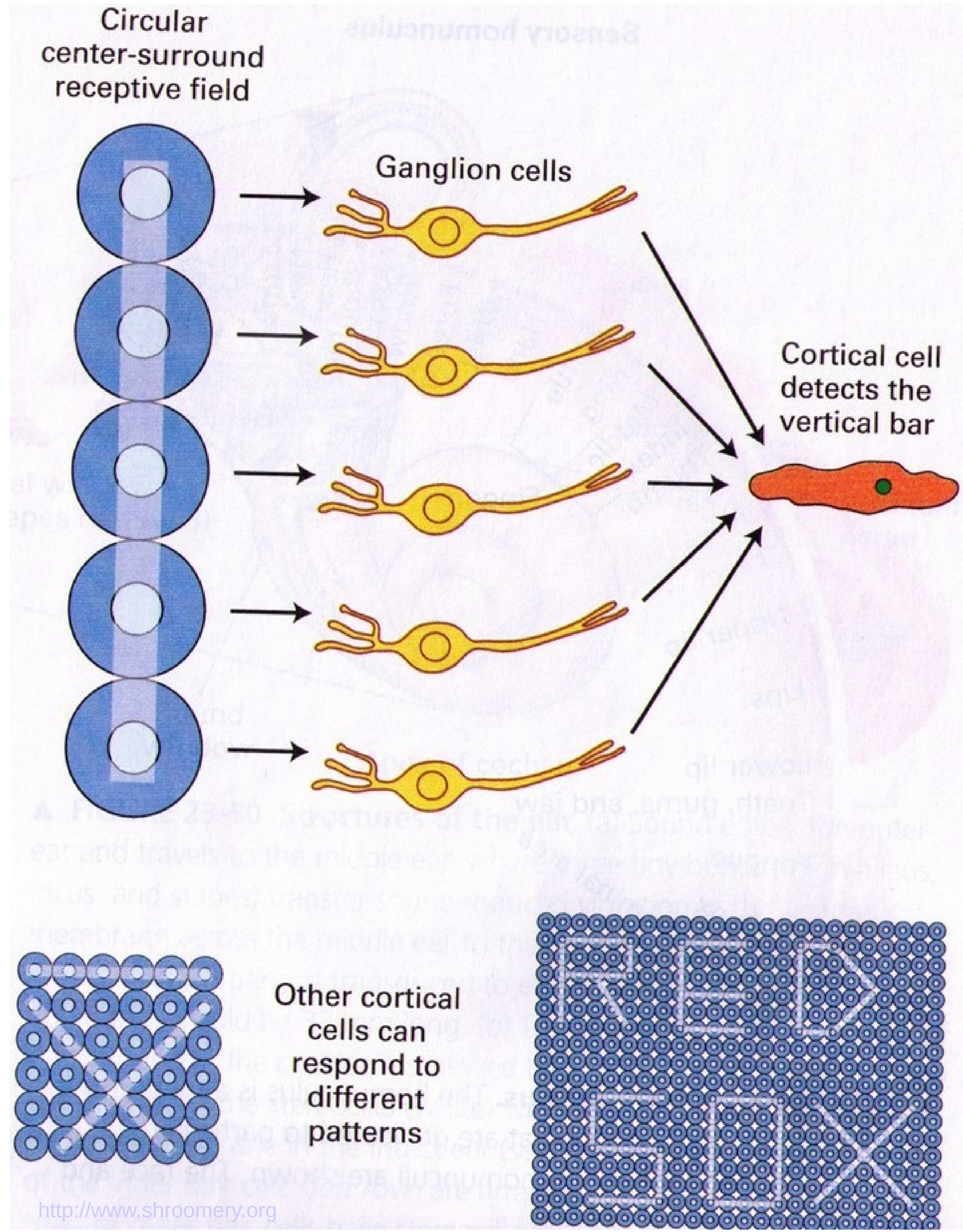
Modified from Livingstone 2002

Edge Enhancement with Color

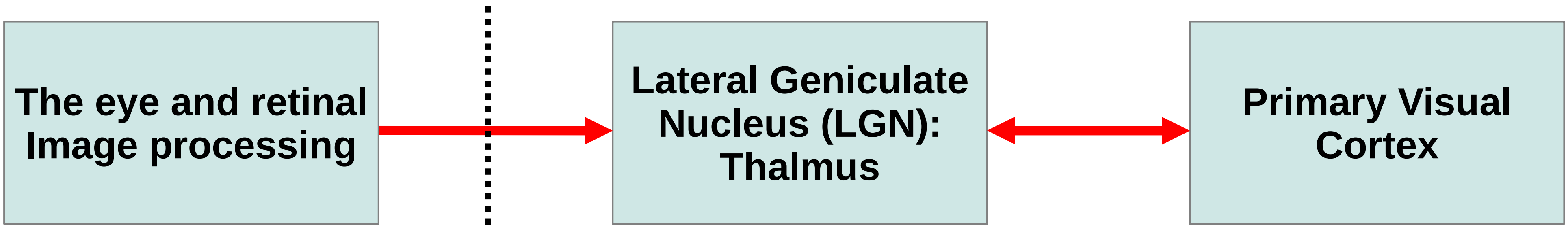


And one step beyond the ganglion cells..

Complex pattern recognition. Visual cortex cells responds to the sum of multiple ganglion cell "center", thus detecting the vertical bar shown. Other cells respond to different combinations of ganglion fields or combinations of cortical neuron fields to recognize more complex patterns.



The visual system beyond the eye



Image(lens)
Receptors
Bipolar
Ganglion

Intensity
Contrast
Color
Edge,lines

Thalmic/
Geniculate cells

No visual information processing
1. Input from two eyes are registered
2. Feedback zone to also receive signals back from the cortex

Simple, complex, and Hypercomplex cells. Levels V1, V2, V3.. + columns

Edge,corner refined
Movement
Color 2Xopponent
Orientation
Contours
Shapes

Examples of visual processing....

Which 'on' or 'off' circuit in which cell(s) are being activated as you read these words

Examples of visual processing....

Which on-off circuit in which cell(s)
are being activated with these words?

Examples of visual processing....

What about this grey area below?

Examples of visual processing....

And this example that we have seen before?

Examples of visual processing....

What do you think is occurring here?

Examples of visual processing....

And what do you think is happening with this photograph we've seen before?

Examples of visual processing....

And one final one....

Outcomes

Which 'on' or 'off' circuit in which cell is being activated as you read these words?

we read black letters against a white background using the OFF-centre bipolar-ganglion cells

What about with these words?

we read white letters against a black background using the ON-centre bipolar-ganglion cells

Off centre-surround and on centre-surround both activated, stimulus is muted

On-centre, off-surround stimulus high at edge, in fact we see the edge more prominent than what it really is, this is called the Mach banding effect

Outcomes

This image is engaging on-centre, off surround cells, edge detection fully engaged. Because of the dark forms there may be Mach banding around the figures to further enhance their visibility

This image is approximating the uniform grey tests we noted earlier, we have predominance of off-on centre-surround stimulation but this attenuates full signal transmission. The eye goes into scanning mode and will tend to rest on the little there is to get full engagement of the on-off bipolar-ganglion configuration.

This image is rich with many visual elements but the eye's physiological preference for maximum signal transmittance will engage the viewer along the leaf edges, along the two slopes at the left and right of the waterfall, and along the waterfall's horizontal and vertical edges. Processing is mostly by bipolar and type 1-2 ganglion/cortical cells

Order and Information in Images

In this Section

- Concept of order/disorder: Entropy
- Information and its occurrence in the visual field
- (Un)Certainty, (un)predictability, complexity
- The Simplicity Rule in perception: Structured Information

Making 'sense' of imagery: Order in photography

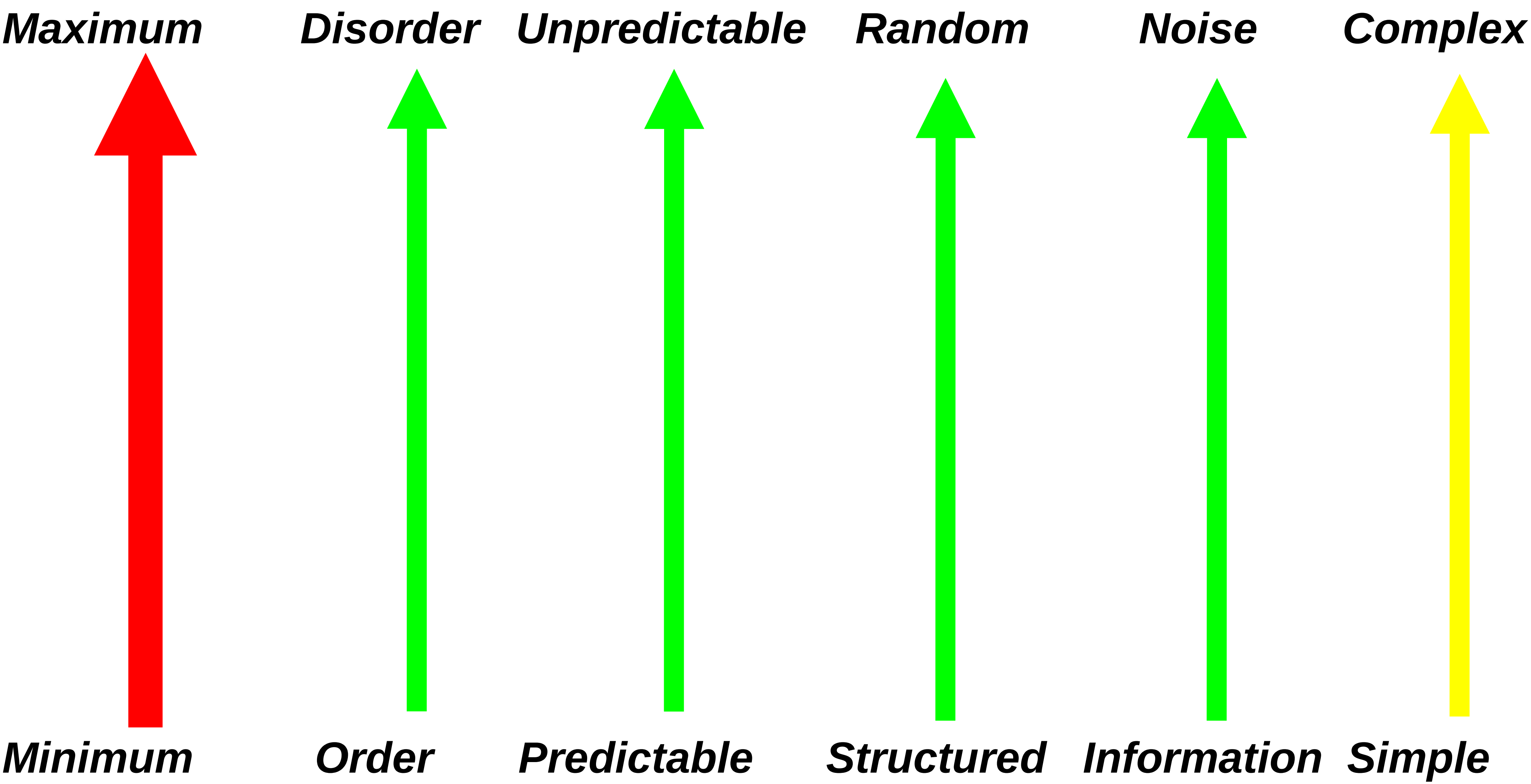
- The examples in the preceding slides point to a fundamental quality to both human vision and making photographs; *structure* in the visual field.
- Structure implies an *orderly* arrangement of sorts, a sense of predictability, of knowing, as opposed to chaos, randomness, uncertainty.
- If the visual system has evolved to detect some 'order' in the visual field as we have seen, there must be some fundamental quality that defines it in nature. And there is..

..Entropy

- Entropy in nature is defined by the second law of thermodynamics and classically expresses the state of randomness or 'disorderliness' of a molecular arrangement. The law states that entropy in the universe increases *overall.*
- This fundamental law is in action whenever we are using our eyes to see; we are scanning the visual field for areas where there is more order or less entropy - the edges, lines, contrasts are these areas of lower entropy – our eyes have evolved to maximize our search for order. Biological systems, in fact, are mechanisms for entropy reduction – at least in the short term (open systems) - all life eventually 'randomizes' ;)
- Entropy as an 'order' measure, has been ingeniously related to predictability, certainty, complexity and information.

Entropy relationships

Entropy



Entropy, more formally defined..

- As noted entropy defined a state of disorderliness, the more disorder, the greater the entropy and the less certain we would be about that state. Total randomness would be likened to complete disorder whereas order would be equated to total predictability.
- The entropy law defines these states of certitude by probabilities (or predictability).
- For example, if x is the digital state of <u>one</u> pixel for a greyscale image (256 grey levels) then the probability, p(x), of that pixel having a given digital number (DN) is 1/256.
- ***Entropy is defined as the logarithm of the probability, $S = \log_2 p(x)$***,
- $S = \log_2 p(1/256)$, $S = -8$ bits, also the level of uncertainty about the pixel
- If we were more certain that the pixel was darkish, say DN <32, then $S = \log_2 p(1/32)$, $S = -5$ bits, our level of certainty lowered the entropy by 3 bits

The consequence of entropy: Information

- In the previous example, the knowledge of a pixel being in the 0-32 range as opposed to the 0-255 range made the predictability of the DN greater; we had more **information** on hand.
- The example serves to make an important link between the level of entropy (or disorder) and information.
- We define information by stating that *Information is the reduction of uncertainty*, or, the more information we have on hand about an event, the less uncertain we are about the outcome of that event.
- So, we now have information related *inversely* with uncertainty and disorder, or, as information increases entropy decreases. ($S = 1/H$ or $H = 1/S$)

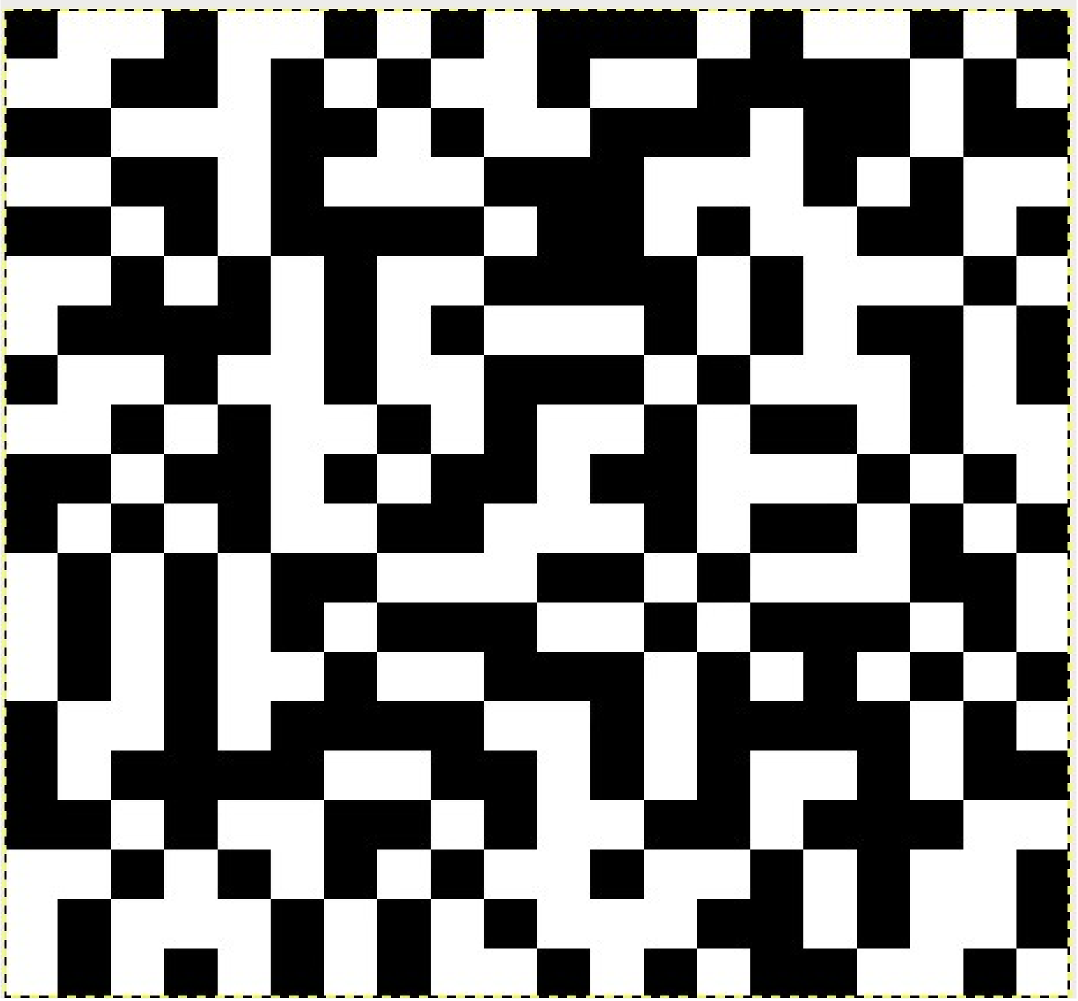
Entropy, Information, and Images

So, what does 'disorder' in imagery look like? And what characterizes more information in an image? We'll answer these questions with an image example but before we can quantify these basic qualities we need an example, let's look at all the pixels of an image.

We can use the following version of entropy/information relationship so that we can sum up all the probabilities for all the pixels in an image ($\sum$ means sum)

$$S = -\sum_{i=1}^{n} p(x_i) log_2 p(x_i)$$

A simple image..



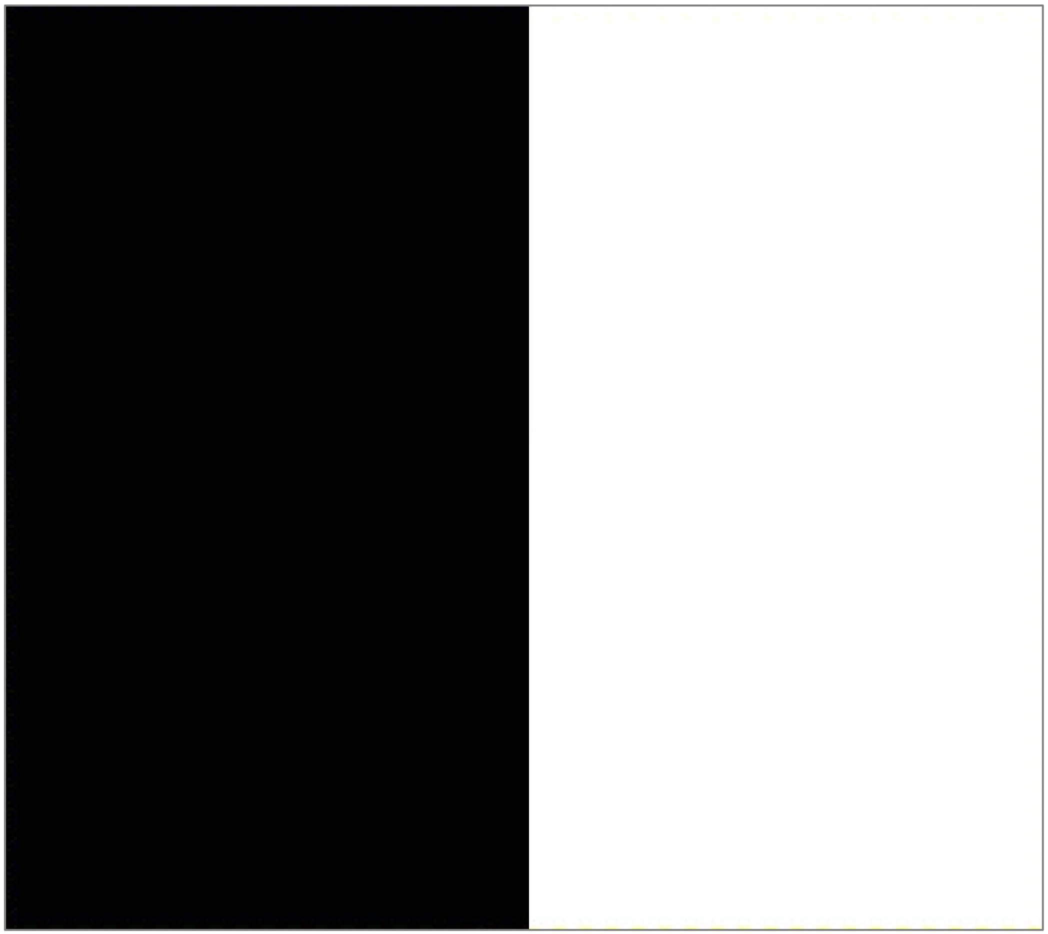
- We start with a basic picture with only 400 pixels (20X20).
- The pixels are either black or white.
- The probability of a pixel being black or white is 1/2 or 0.5. Since we cannot predict whether the pixel will be black or white each pixel will have a half chance guessing it of being black or white.
- Applying the preceding formula, Entropy, S = 200 bits

The value for pixel #1 (*i*=1) is (0.5) X $\log_2$ (0.5), or (0.5)X(-1), or (-0.5).

The value for pixel #2 (*i*=2) is (0.5) X $\log_2$ (0.5), or (0.5)X(-1), or (-0.5).

Continuing on and summing for all 400 pixels we get a total value of -200.
Applying the negative sign the final value is 200 bits

Another variation of the simple image..



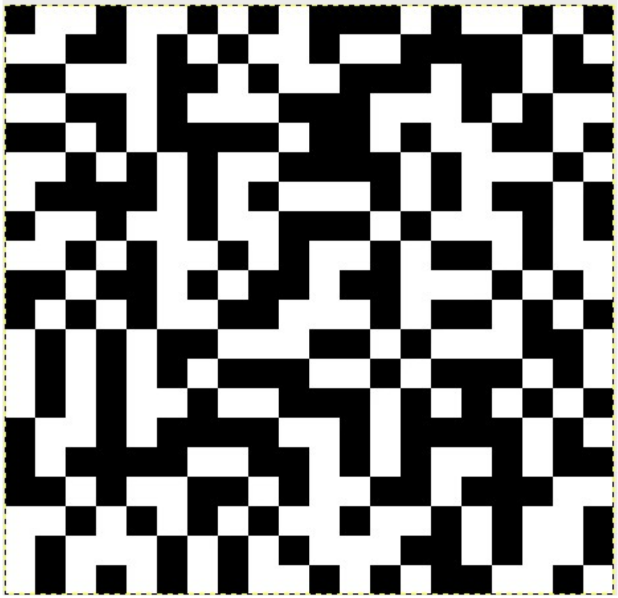
- We still have a basic picture with only 400 pixels (20X20).
- The pixels are either black or white but the arrangement indicates a pattern or order, infact once the pattern is known the probability of the pixel being either black or white increases to certainty (probability is 1).
- Using the entropy formula below we can see that the ordered pixels have decreased the entropy of the image to nil.

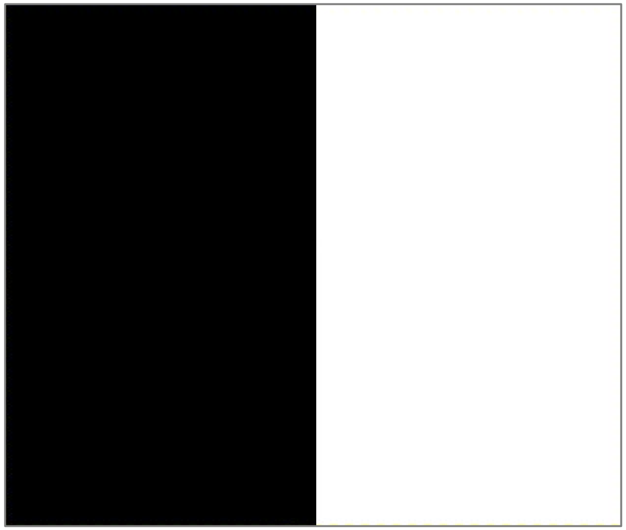
The value for pixel #1 (*i*=1) is (1) X $\log_2$ (1), or (0)X(0), or (0).

Continuing on and summing for all 400 pixels we get a total value of 0.
The final value is 0 bits

Some observations on the two image examples

- A completely random pixel arrangement (total disorder) had an entropy of 200 bits.
- When the pixels were ordered such that the 'blackness' or whiteness' was completely predictable, entropy dropped to zero; we gained 200 bits of information (remember information is the inverse of entropy)
- The ordered image, which could well be the edge of some form, was extremely more informative and, not surprisingly, more visually stimulating (our eyes 'know' what to do with this pixel arrangement)





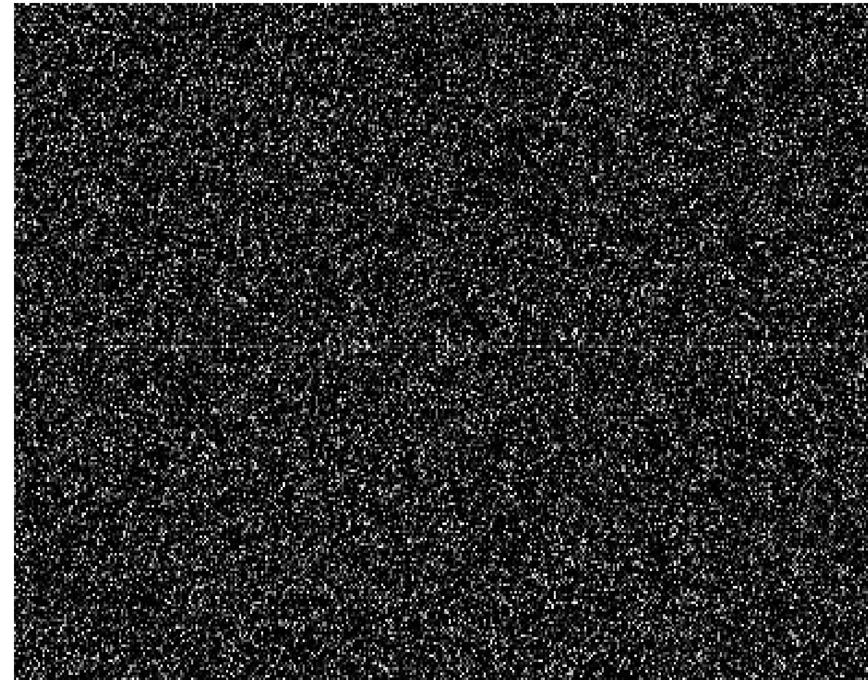
The Information Paradox

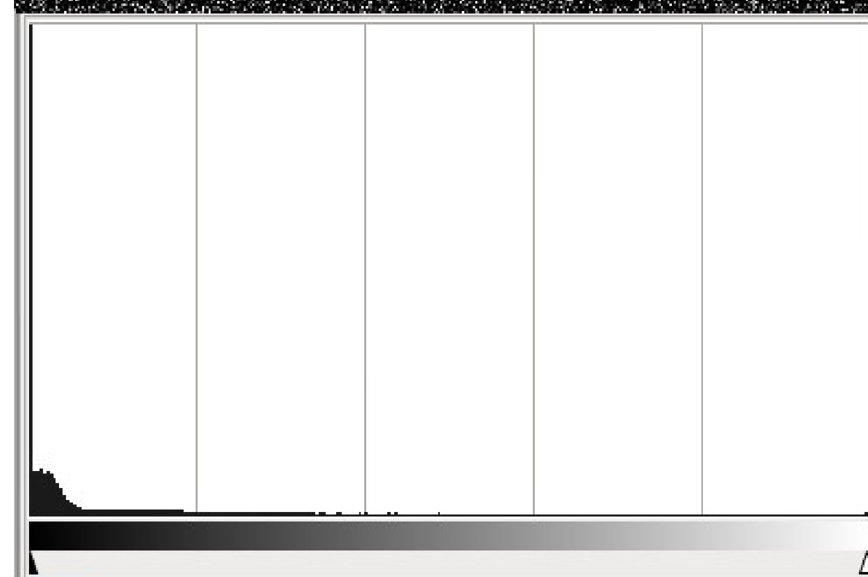
- So far we have noted that entropy and information are inversely related: a more random visual arrangement has less information than one that is more ordered. In the Information jargon it is more 'noisy'.
- In general this is true, but consider our random 20X20 B&W pixel image again. What if there was a *code* such that the apparent random arrangement of BW pixels were actually following a pattern or a formula, *and you knew the code* would it still be noisy? Consider the problem from the probability perspective: You now are 100% certain that the pixel value is either black or white because you can now decipher the code.

The Information Paradox 2

- The pixels in the right image were completely randomized – total disorder. The eye, as an information processor, cannot 'anchor' to anything that it is designed to process. What is its 'Information' content? Its entropy?

The top image is considered 'noise': undecipherable. But, what if the code to decipher the pixel arrangement was provided? The 'information' provided would allow one to re-arrange or *re-order* the pixels into a less complex, more predictable pattern.

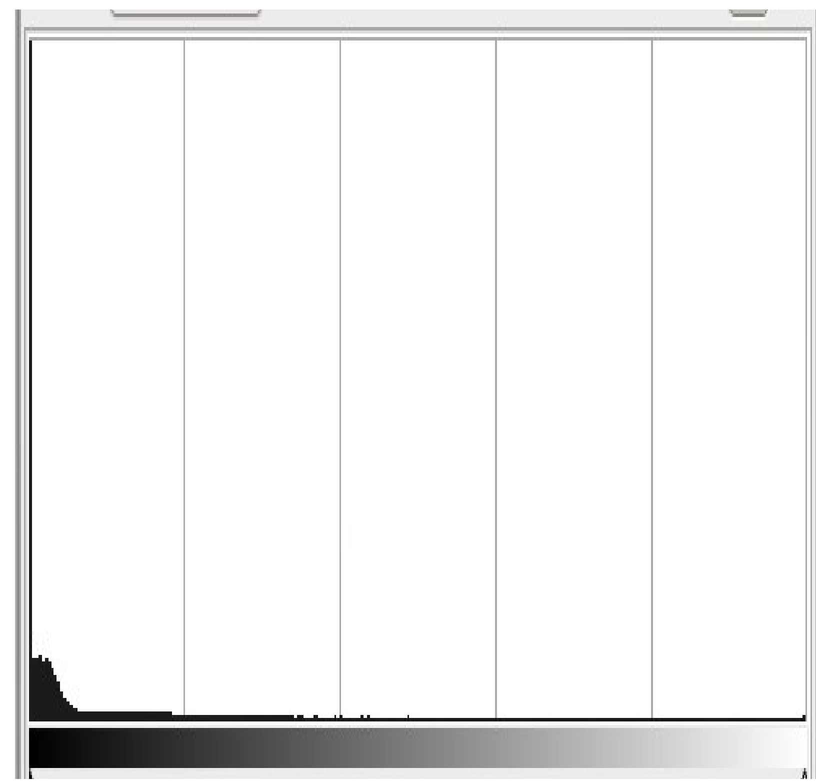




Mean = 37, Std. = 61, Median = 6

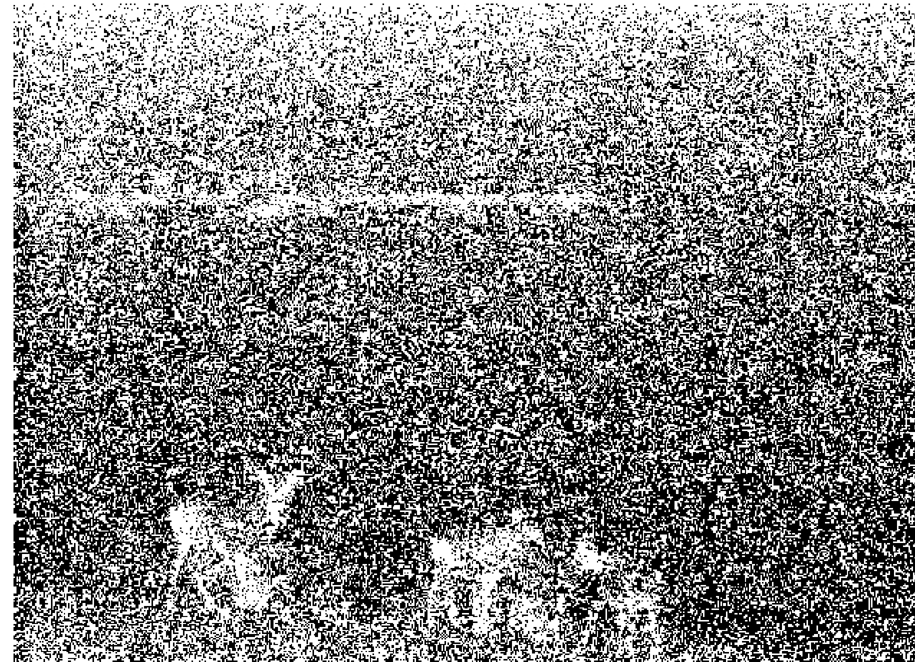
Deciphering 'noise' with coded information

Here is the image fully re-ordered reversing the fractal random code generator (x1=35); a consequence of the information gained. Our eyes, as information scanners, can now the detect the informative structures that they are designed to see: the edges, lines, and contrast. Remember, there is no 'meaning' involved here!



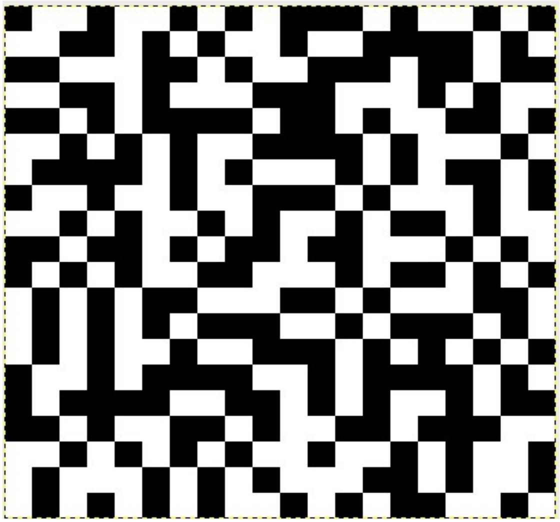
Mean = 37, Std. = 61, Median = 6

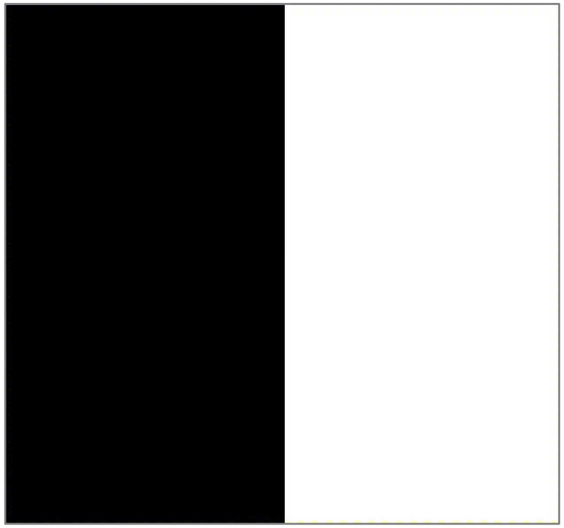
Using technology to decode information

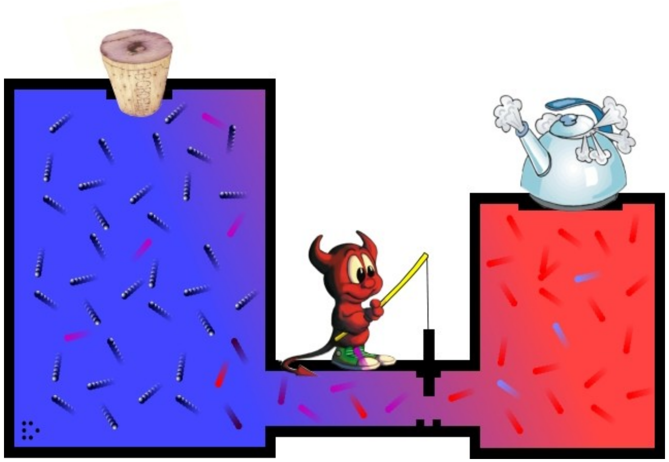



Constraints to be aware of..

- Information as discussed is related mostly to FORM
- Information in FORM is portrayed as 'repeating' or 'predictable'
- Content, context, and certainly meaning are ***not*** considered.
- An image state with maximum entropy (disorder) would be analogus randomizing all its pixels.
- Image entropy decreases as form takes shape.
- The image example is the visual analogy to Maxell's Demon, an imaginary creature Maxell created to decrease thermal entropy







Pixel value randomness and order in micro-image details

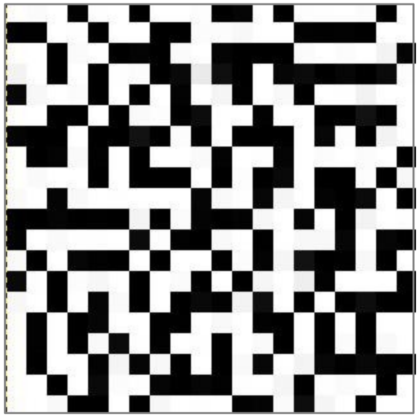
Lowering micro-image entropy

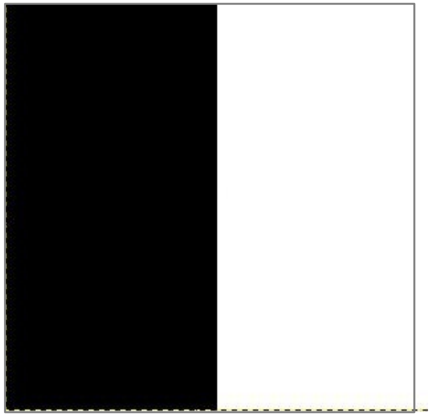
Area of high micro-information content –--> eyes are more stimulated ----> flower *shape* becomes the focal point of image.

- Photographic effect: subject (edelweiss) was placed against dark background to highlight the flower. Rocks in foreground provide secondary 'balance'.
- Retinal effect: Ganglion cell neural firing maximized around flower edges, less pronounced around rock edges, and minimal in dark background.
- Entropy effect: At maximum in dark background and at minimum around flower form.

Image entropy applied: compression algorithms

- The concepts that we've investigated are applied every time an image is compressed.
- The algorithm in creating a 'JPEG' or .jpg, file typically looks at the probability of each pixel being a certain DN, the higher the probability, the better the compression ratio (Huffman coding)
- Consider our earlier basic two images which is more efficiently compressed? By how much?





17.2 kB **4.0 kB**

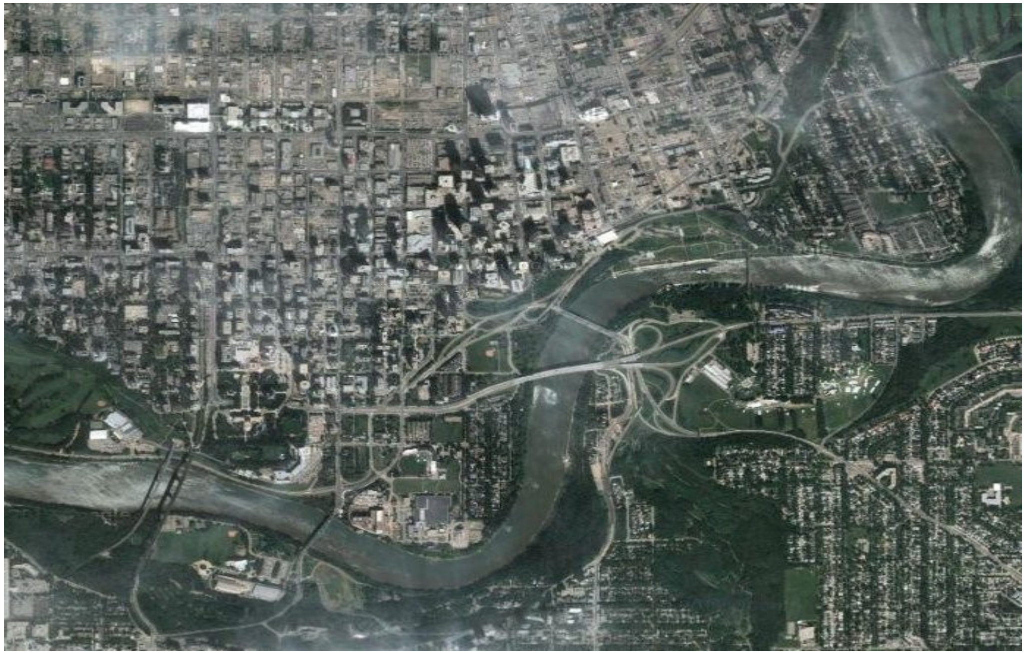
From the micro to the macro scale

Ask yourself the questions regarding, certainty/uncertainty, Information, and order/disorder, complexity/simplicity

Which images intuitively have more/less of these attributes?

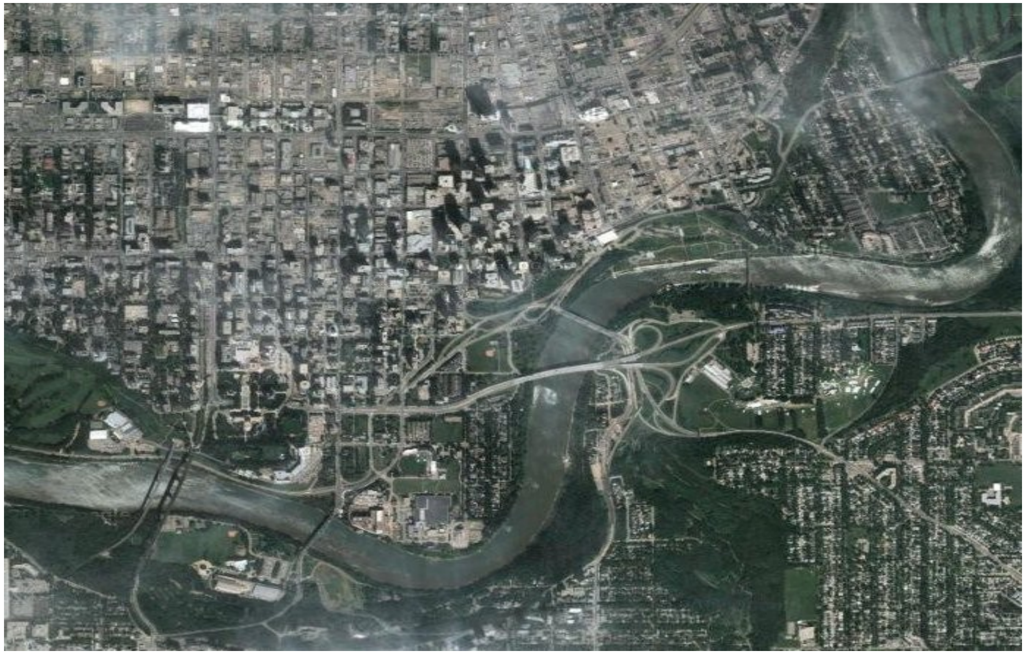
Applying the ideas to photography

A more challenging example..



We have an aerial photograph and an out-of-focus image. Which image has more information (less predictable)? which image has more entropy (more disordered)? Which is 'simpler', which is more complex in structure? To what extent is the aerial photo tending towards 'noise'? If you were told that this is a photo of downtown Edmonton how does this 'code' change things?

Preference for the Simple



The photo of downtown Edmonton and the the one below it are examples of imagery tending towards 'noise' on the macro level, mostly because the information load is high and it becomes more difficult to order and make sense of the fundamental visual stimuli (our edges and lines). Images tending toward simpler, more ordered arrangements tax our visual system less and are perceptually preferred. This fusion of simple structural preference and Information content has given rise to **Structural Information Theory....**

Structural Information Theory
SIT: The basics

From Wikipedia:

"... a formal coding model starting from the assumption that the perceptually preferred interpretation of a stimulus is the one with the simplest code. **A simplest code is a code with minimum information load**, that is, a code that enables a reconstruction of the stimulus using a minimum number of descriptive parameters. Such a **code is obtained by capturing a maximum amount of visual regularity** and yields a hierarchical organization of the stimulus in terms of wholes and parts"

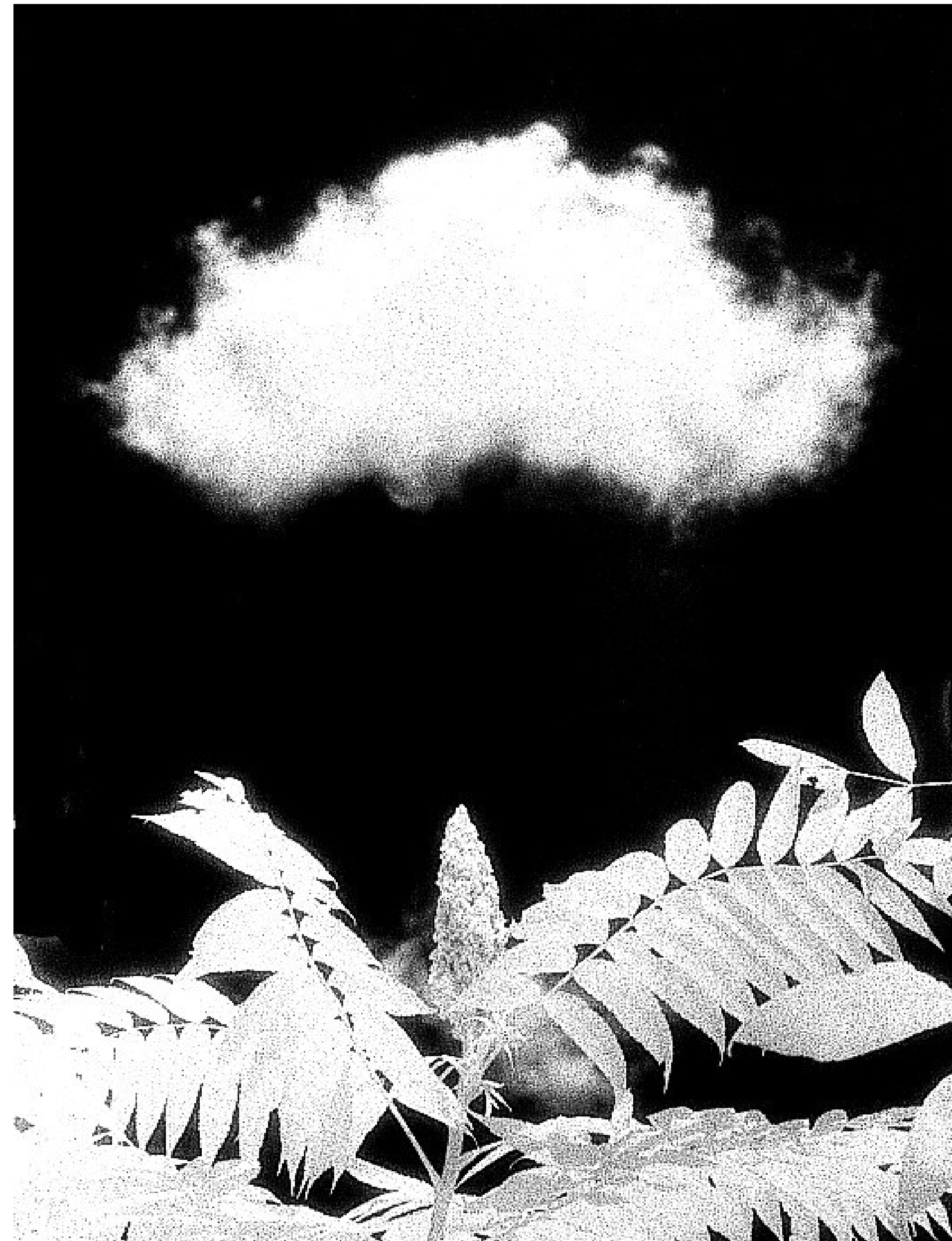
From this definition we can clearly see the entropy principle presented earlier influencing preference in basic perception.

Structural Information Theory
The basics

- SIT is based on coding, in imagery the transmitted code is the variation and structure of pixel levels, or in analogue imagery, silver grains/dyes.
- Central to SIT is the simplicity principle, which implies that the visual system is assumed to prefer the simplest interpretation among all possible interpretations of a stimulus.
- A simplest code is a symbol representation that enables the reconstruction of the stimulus using a minimum number of descriptive parameters; it is obtained by capturing a maximum amount of regularity.
- We can consider that codes and parameters in photography are assemblages of points, lines, edges, space, texture, etc.

SIT image examples..

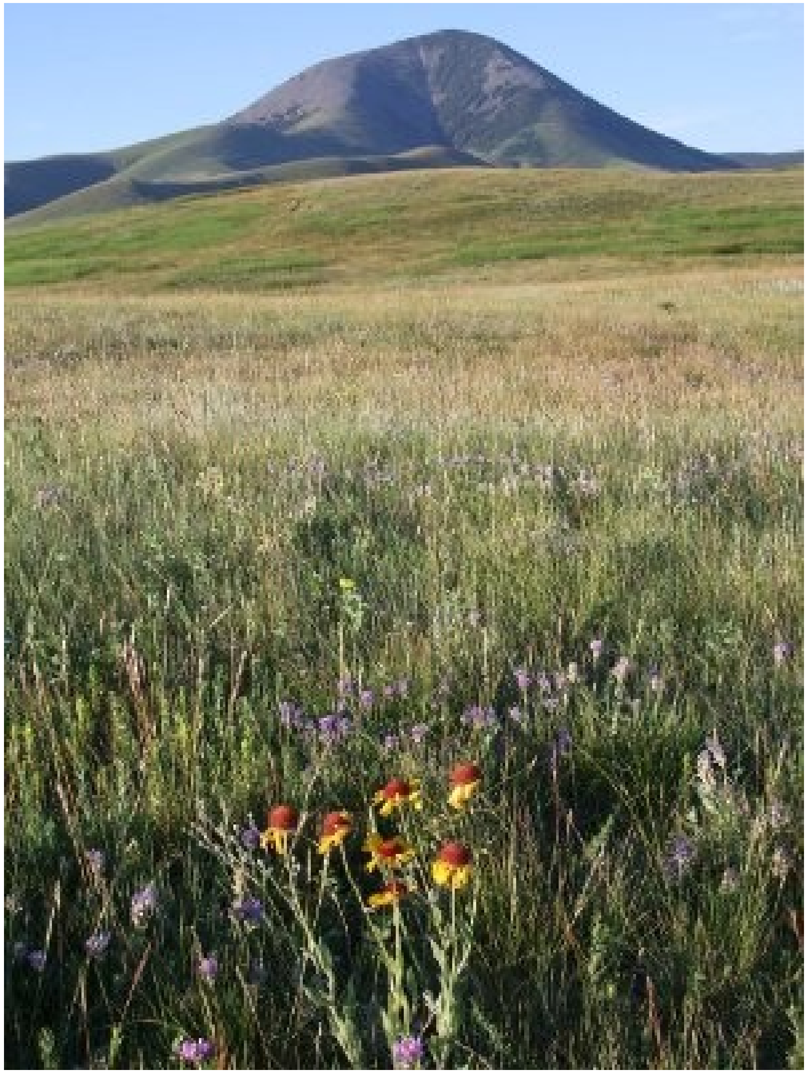
applying simplicity and regularity



Which one is more predictable and ordered?
Which is more simple, complex?
Which contains more 'information'?

SIT image examples..

applying simplicity and regularity, with a challenge..



Consider the questions of order-disorder, simplicity-complexity, predictability-randomness, certainty-uncertainty, information-noise: Can these qualities be comparatively distinguished in these two images? Why or why not?

Coming full circle..

- We noted preference for types of graphics and contrast, particularly images with definite lines and edges.
- We investigated eye biology, established that, indeed, the eye *naturally* is attracted to such changes in contrast.
- We moved on to information and entropy and noted that higher information and order are characteristics of such contrast changes.
- So, if a physical law such as entropy is a determinant in assessing visual signals, then is the eye an instrument for detecting visual fields where entropy is low?
- And bringing this all back to composition, are we naturally adapted to prefer images with lower entropy and higher simplicity?
- It would seem that the answers to the questions is 'yes',
- Let's see how the research supports what we've seen so far...

Themes in image perception research

Quantifying image 'clutter'

- Subband Entropy*
- Edge Density
- Feature Congestion

Eye movement tracking

- Saliency feature identification and mapping, including heat maps*
- Uncertainty reduction

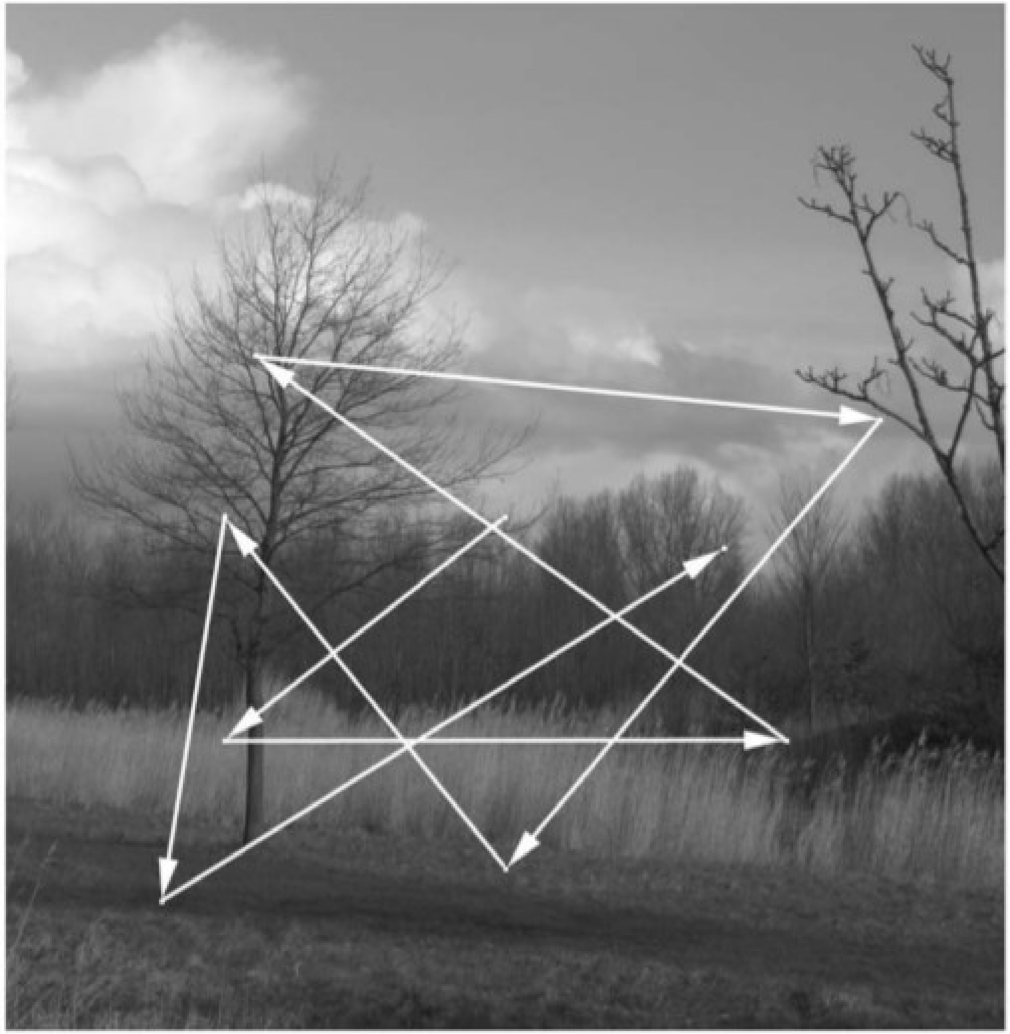
What does the research say?

- Researchers in perceptual psychology and eye physiology have developed apparati to characterize eye preferences in scenes.
- Eye fixation and movement in scenes are used to collect data on where the eye preferences tend to land.
- This type of research attempts to correlate eye fixations for an image with areas of the image where contrast is enhanced; the hypothesis being that the bipolar and ganglion retinal cells are 'programmed' for this visual stimulus.

Eye fixation research

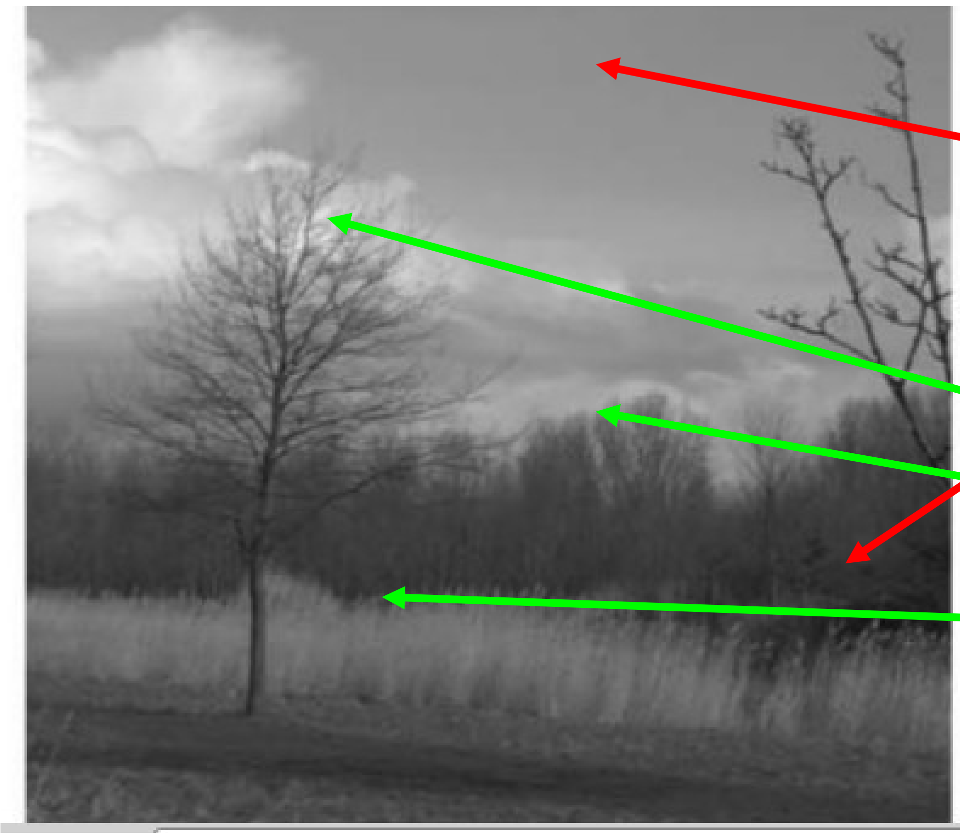
Quantitative assessment of eye Fixation points and eye movement has been of major research interest in perception psychology. Do the neurobiological hypotheses hold up to actual quantitative scrutiny? It seems that they do.

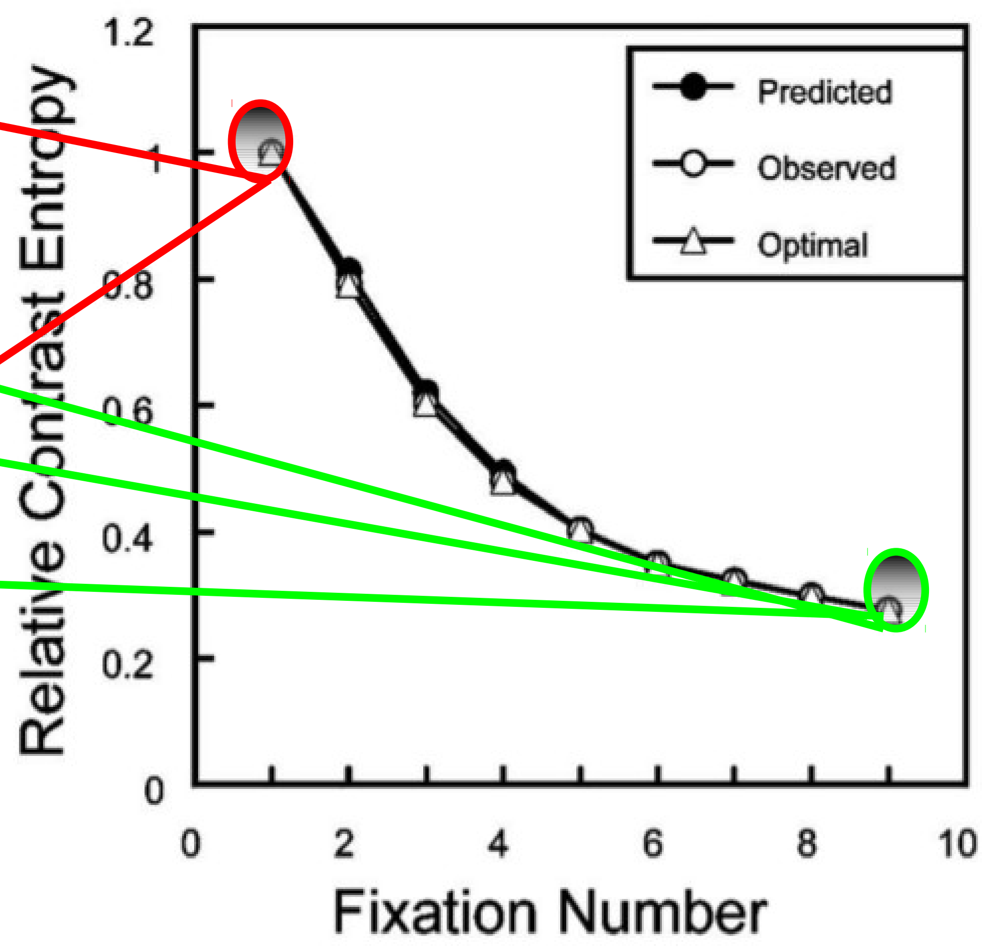
These type of investigations have led to research in computer vision, face recognition, marketing strategy just to name a few. It has also anchored many of the principles in visual design theory in science.



Raj et al. Vol. 22, No. 10 / October 2005 / J. Opt. Soc. Am. A

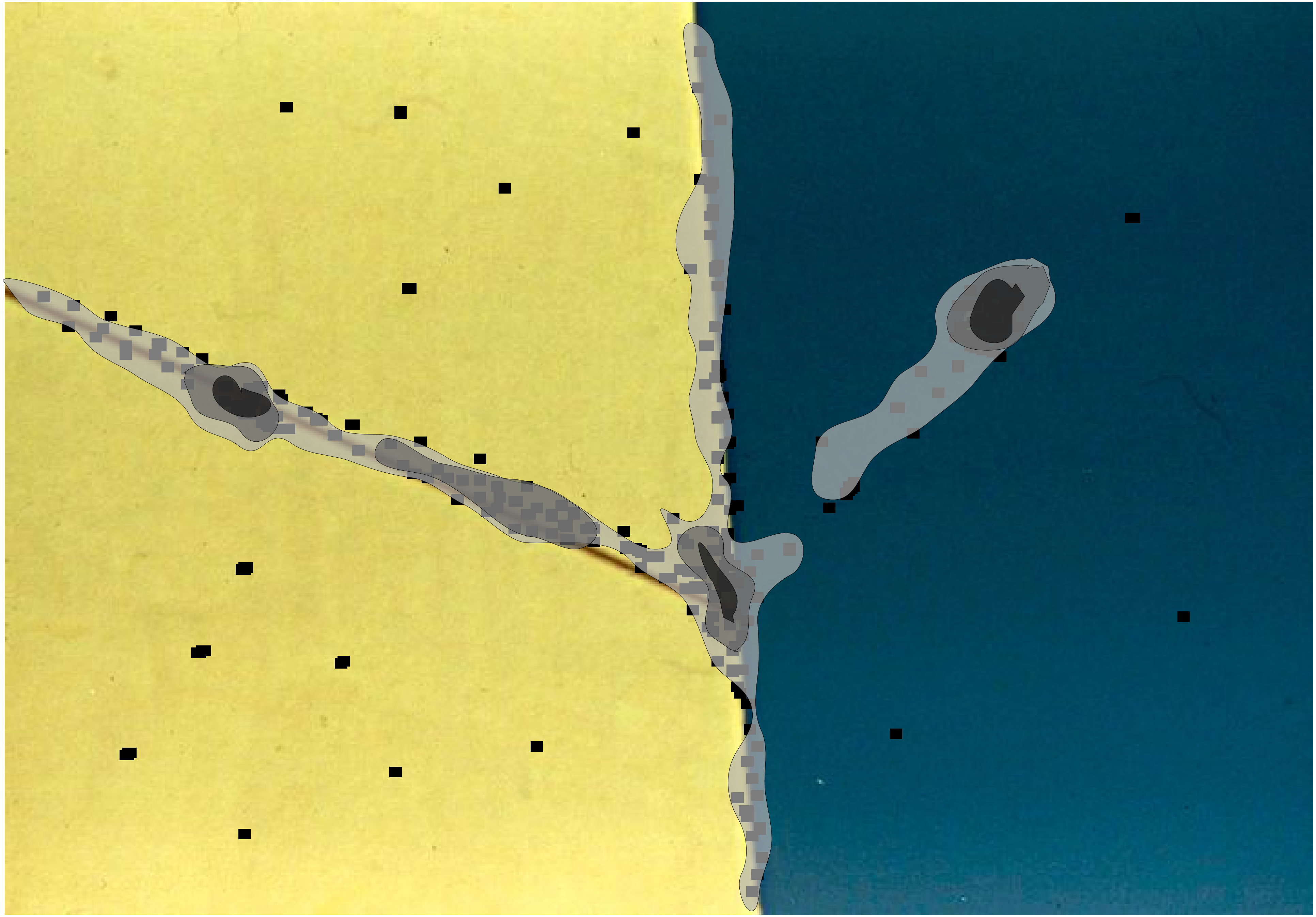
Relationship between fixations and image entropy





Modified from Raj et al. Vol. 22, No. 10 / October 2005 / J. Opt. Soc. Am. A

Topographic eye fixation distribution



Outcomes of eye movement and fixation: The Saliency Map

The *saliency map* is a topographic spatial distribution of the density of fixation points for a given scene over a given time interval.

Saliency maps are derived from both experimental data (actual eye fixations) and modeled data based on the *saliency* of visual features.

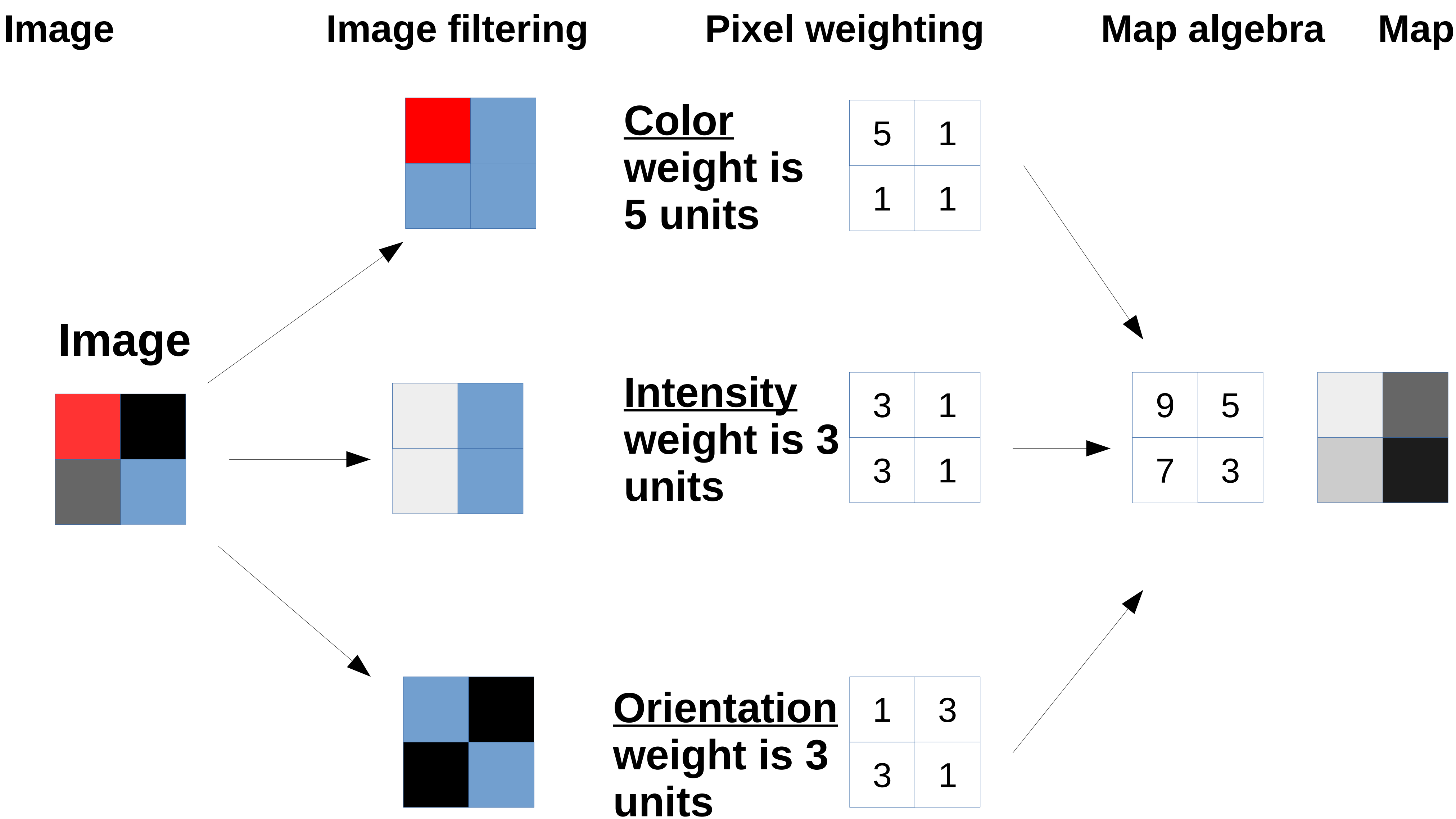
Salient features are those features that:

1. Stimulate eye circuitry
2. Have inherent information
3. Have cognitive relevance

The three levels of saliency generally relate to both the bottom-up and top-down perceptual functions. Saliency maps try to incorporate both in their function.

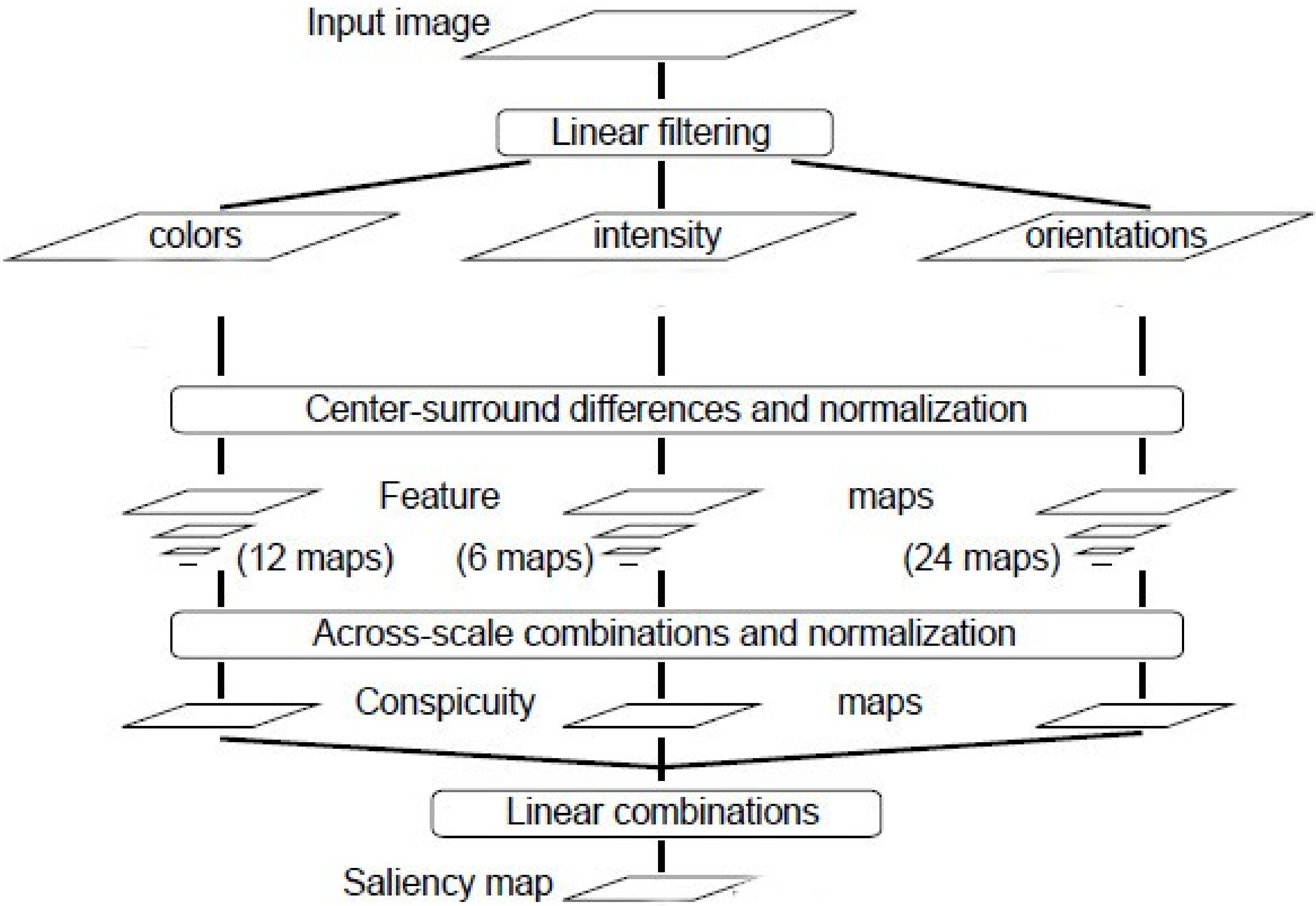
See: http://www.scholarpedia.org/article/Visual_salience by L.Itti

A simple derived saliency map

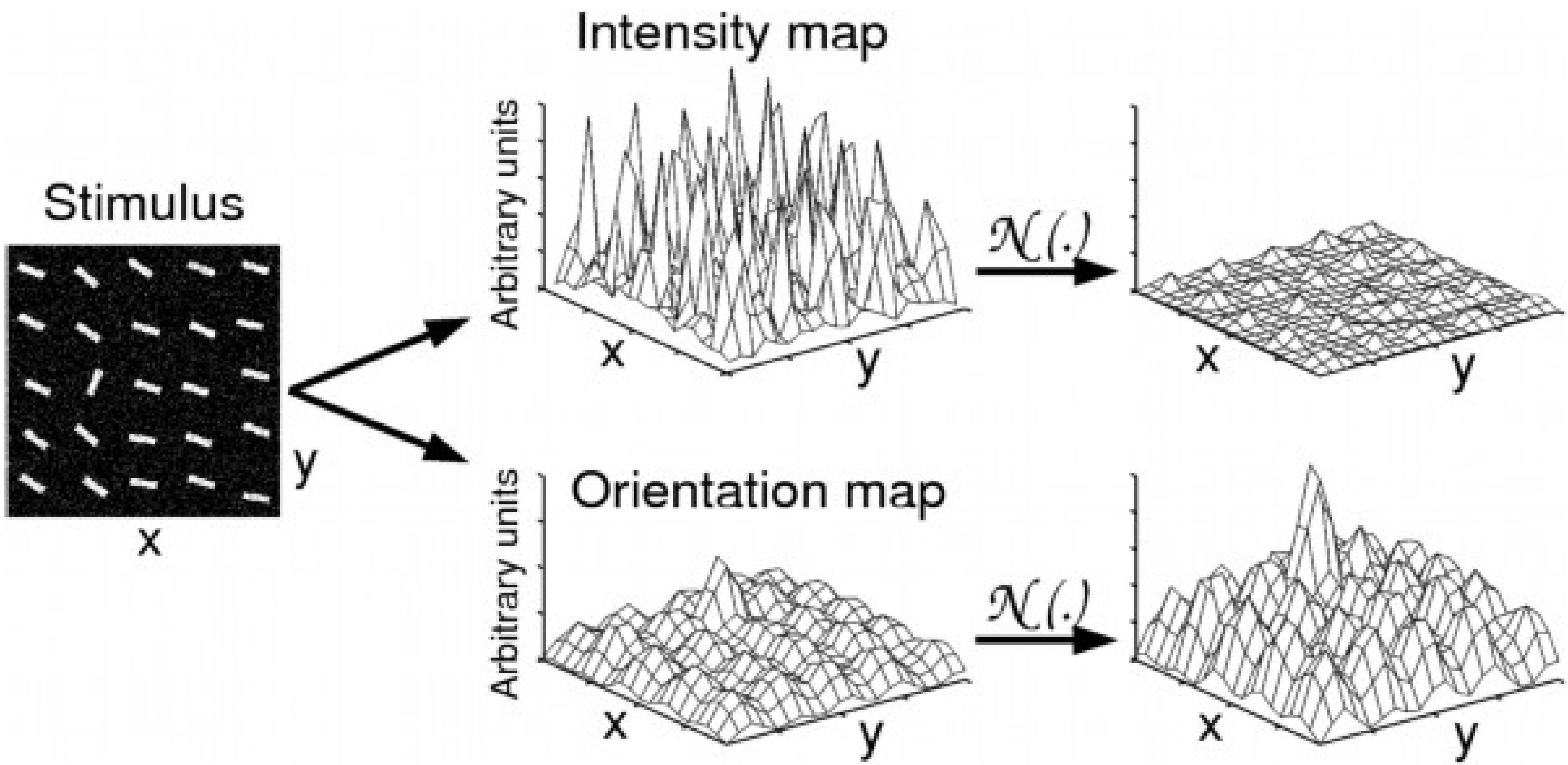


A 'bottom-up' Saliency Model

A Model of Saliency-Based Visual Attention for Rapid Scene Analysis



Typical saliency processes



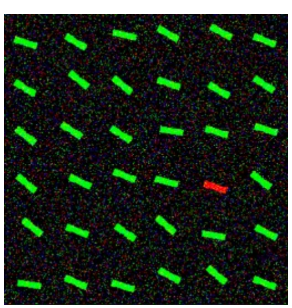
Map algebra integration

I = Pixel x1,y1..

+

O = Pixel x1,y1..

+



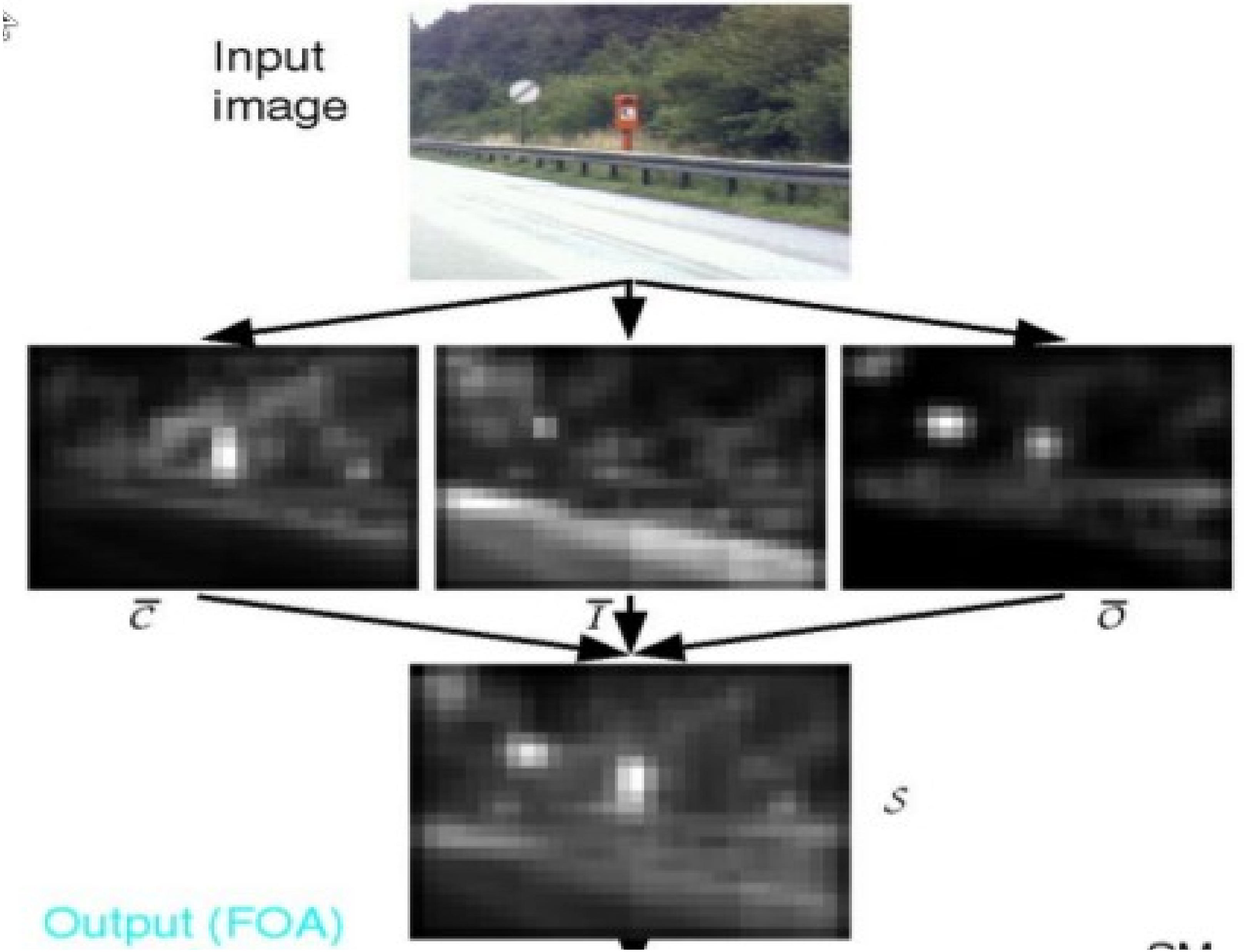
Color map (RGB combined)

C = Pixel x1,y1..

= I + O + C

Source: A Model of Saliency-Based Visual Attention for Rapid Scene Analysis, Laurent Itti, Christof Koch, and Ernst Niebur

Assembling the saliency map

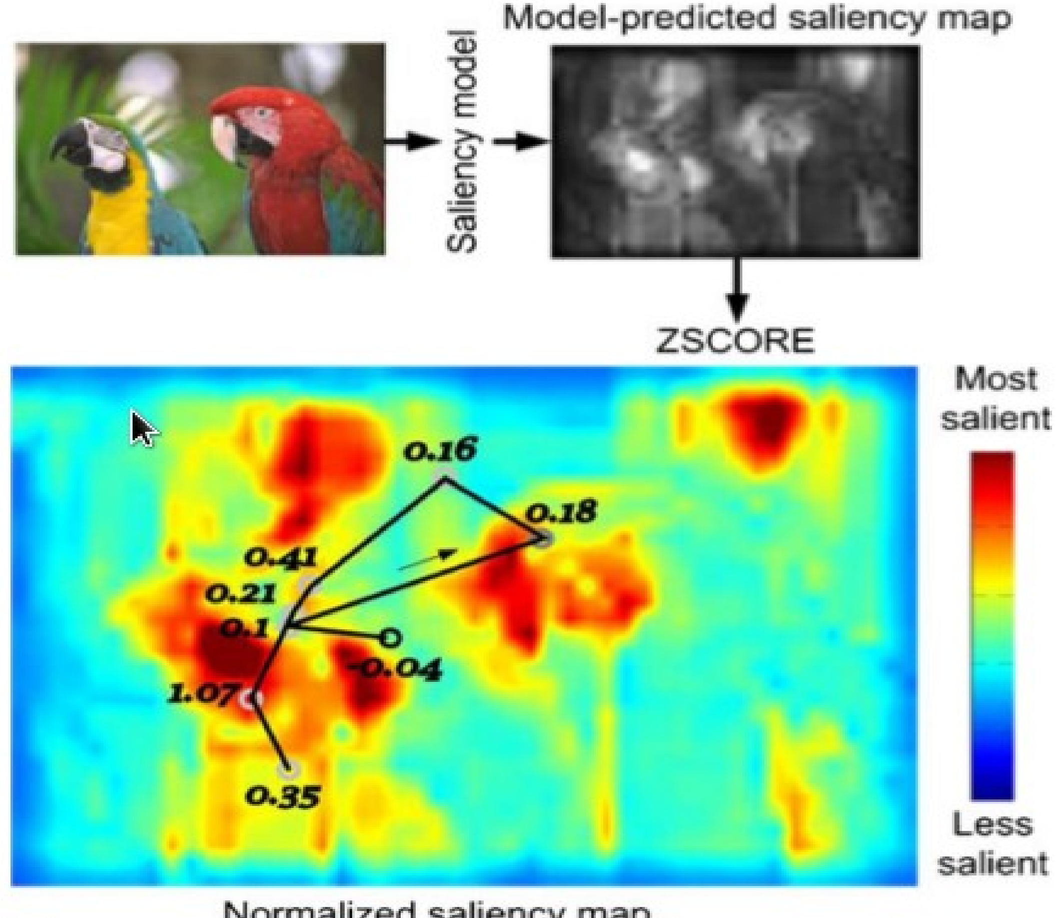


Source: A Model of Saliency-Based Visual Attention for Rapid Scene Analysis, Laurent Itti, Christof Koch, and Ernst Niebu

Saliency Maps

Fixation data versus model data

Top row demonstrates an experimentally derived saliency map based on eye fixations. On the right is a modeled derivative of the salient map.



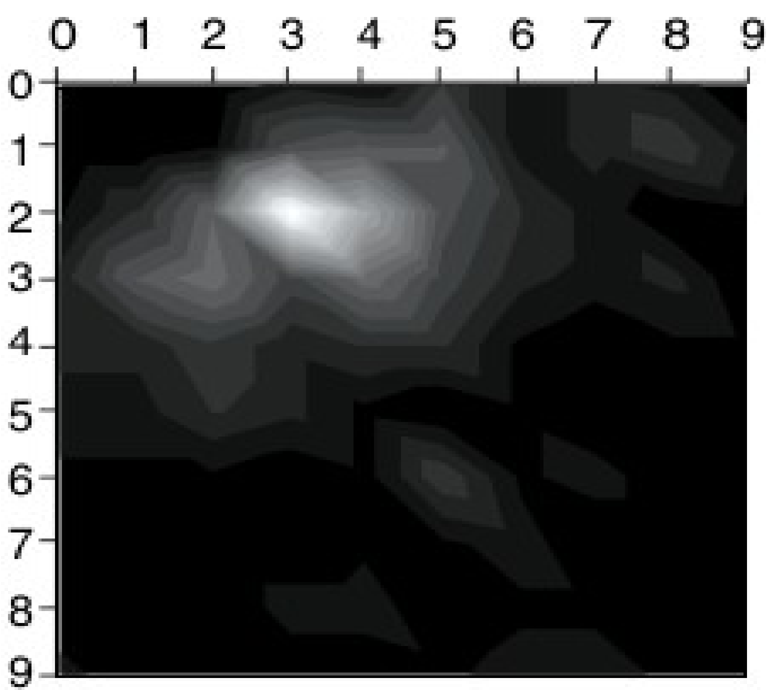
The modeled approach is very popular today, they are expressed as heat 'maps'. More on this shortly.

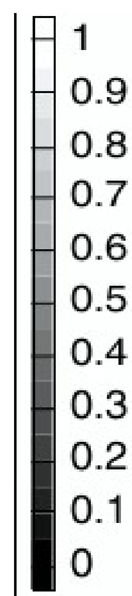
Source: Methods for comparing scanpaths and saliency maps: strengths and weaknesses, O.Le Meur,T.Baccino

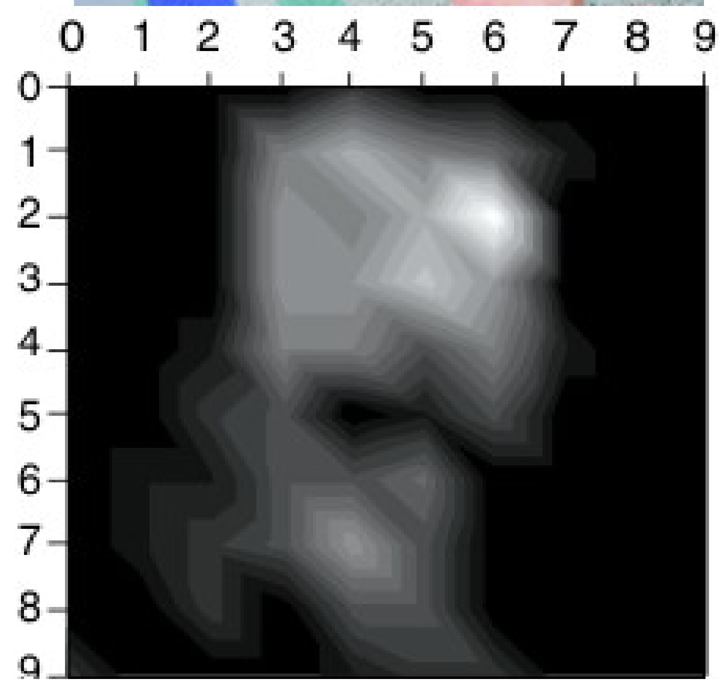
Saliency Maps

Analysis of the Eye Movements and its Applications to Image Evaluation

Chizuko Endo, Takuya Asada, Hideaki Haneishi and Yoichi Miyake







Some questions pertaining to saliency maps..

1. Does each eye fixation signify the same *meaning?* ie. Is it a predominantly more contructivist or an ecological fixation, is it bottom-up or top down?
2. As a corollary, are equal density areas of the saliency map equated with the same *information?*
3. Can fixation duration equate to some metric related to meaning or information content?
4. Can salient features be under-represented, or worse, missed?

Heat Maps:

Algorithmic applications of eye fixations

Computer generated visual 'intensity' maps based on the physiological and information principles of saliency

Intended to mimic actual eye movement equipment used in research

Typically used in business (advertising) to maximize the viewers' interest in images.

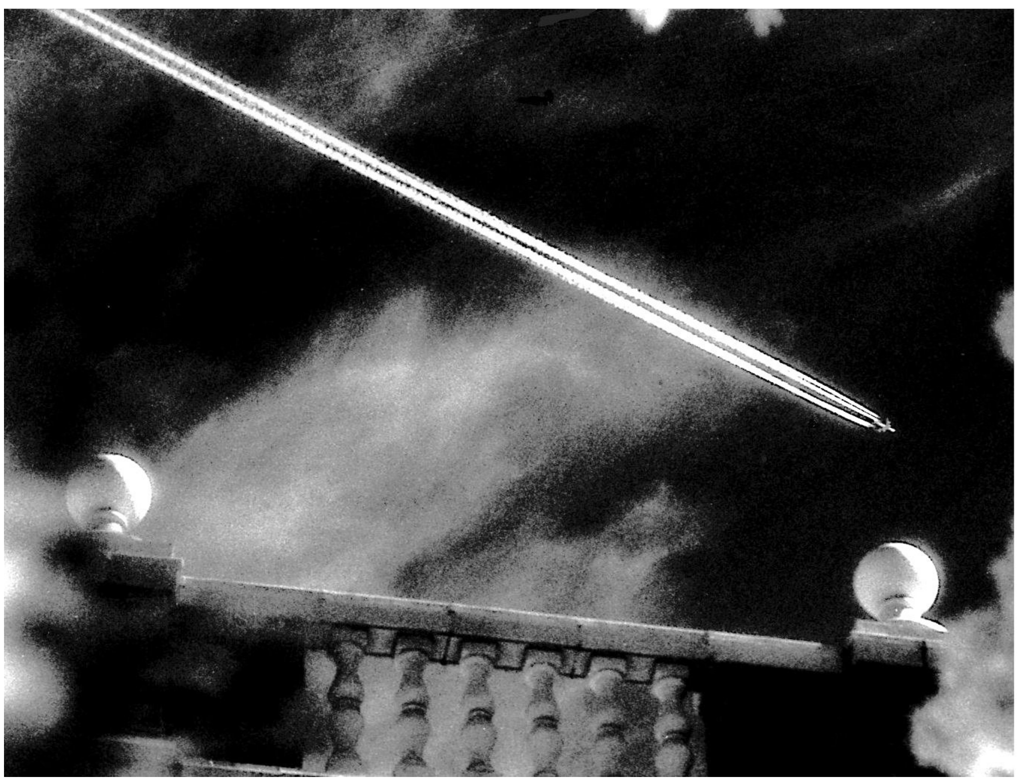
Also used in web design to ensure that viewers' interest correspond to objects of interest to the designer (may be associated with mouse movements and clicks).

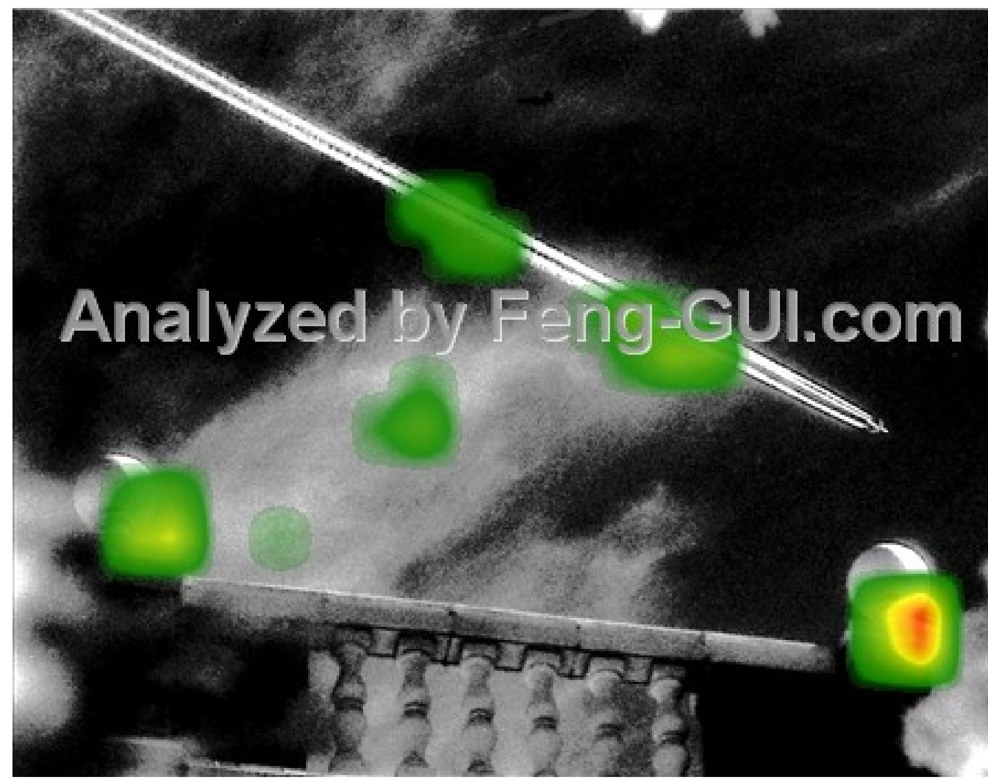
Heat Map example1

Attention & Atraction Analysis

Generated from Feng-GUI

Feng-GUI simulates human vision during the first 5 seconds of exposure to visuals, and creates heatmaps based on an algorithm that predicts what a real human would be most likely to look at. The free version is noted to be 60% accurate (85%)





How well do you think the software is characterizing visual attention in the image?

Heat Map example2

Generated from Feng-GUI

Attention & Atraction
Analysis

What do you think of this result?

How has it failed or, possibly succeeded?

What do you think is the most salient feature in this photo?

Do you think that the perceptual emphasis is on bottom-up or top-down processing?

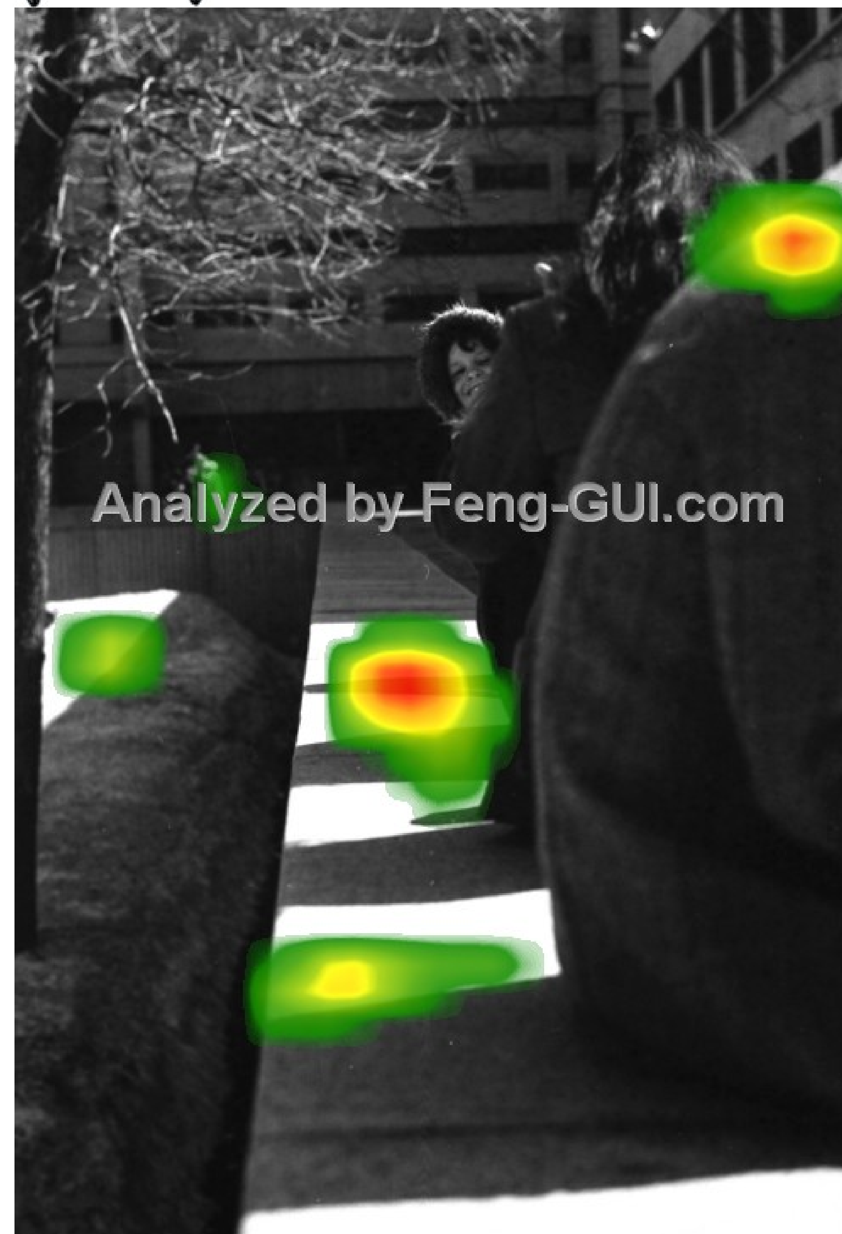


Photo-composition Decomposed

In this Section

- Containing the eye
- Creating and directing eye movement
- Visual design strategies/eye motion dynamics
- Cropping and retargetting
- Wrap-up

Objective: Low entropy-high SIT images - Visual design principles

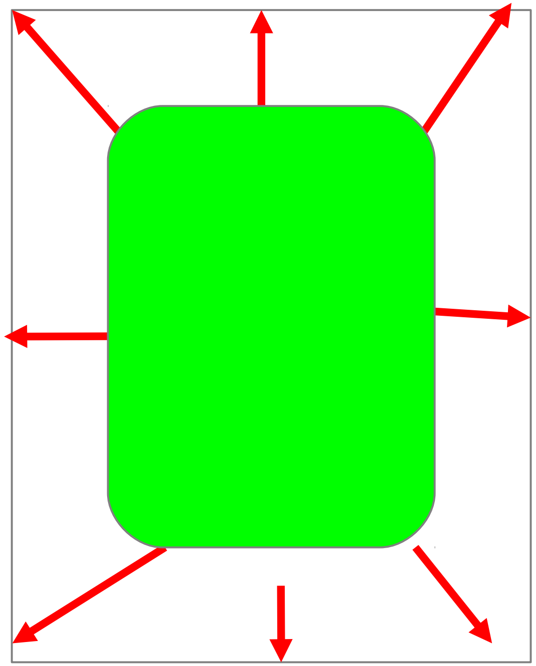
- The focus to our exploration of photo-composition has been on form and strategies in the natural world to maximize its detection.
- Arrangement, pattern, balance, tension, and other design basics of form have been for the most part ignored, other than the simplicity factor we've encountered with SIT.
- With our understanding of what the eye is designed to do and with our knowledge about perceptual information and order, we will now apply design dynamics to our photo-composition framework.

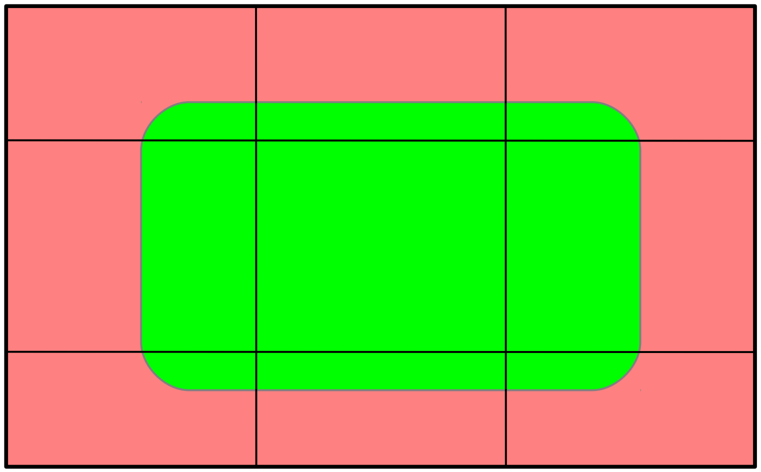
Containing the eye

- Images have a basic assumption; they contain borders that encompass the elements of the image.
- Generally we want to contain the eyes within the borders of the image (although in a creative process the opposite may be the objective).
- *The object, then, is to use elements that the eye is programmed to detect (our edges and lines) and use them to lead eye fixations within the image frame.*
- Because the image frame is itself an edge(s) our eyes are naturally attracted to it, the general rule is to compose elements away from the frame edges.

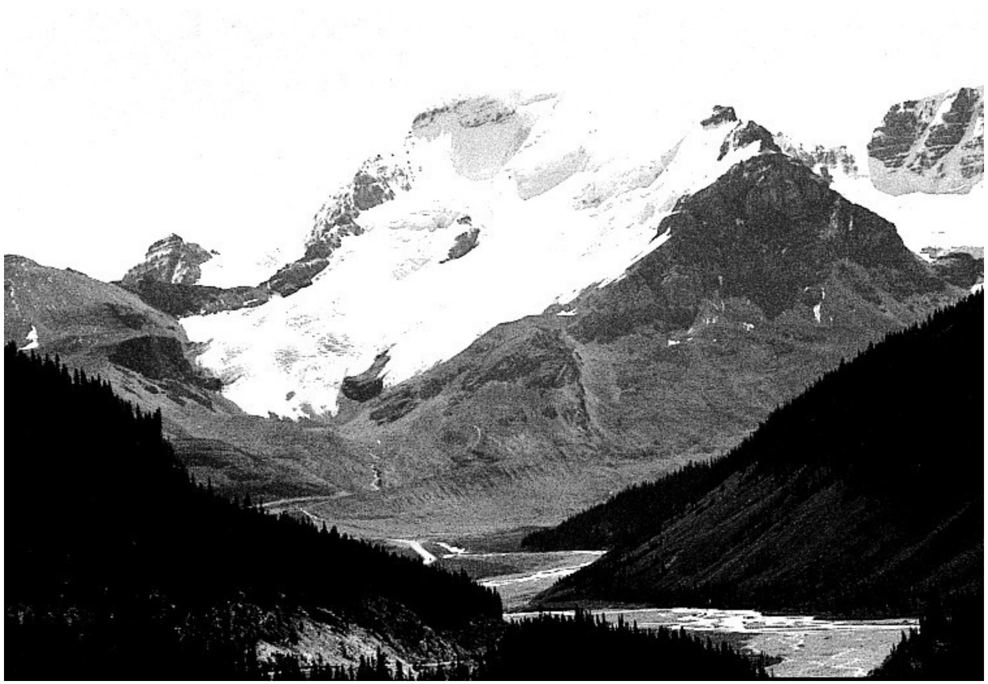
Strategies for containing the eye

- Frame containment dynamics include the so-called 'rule of thirds' or the more challenging triangular arrangement.
- Main objective of these fundamental strategies is to keep the eye away from the frame edge; you want *your* edges and lines to tell the visual story.



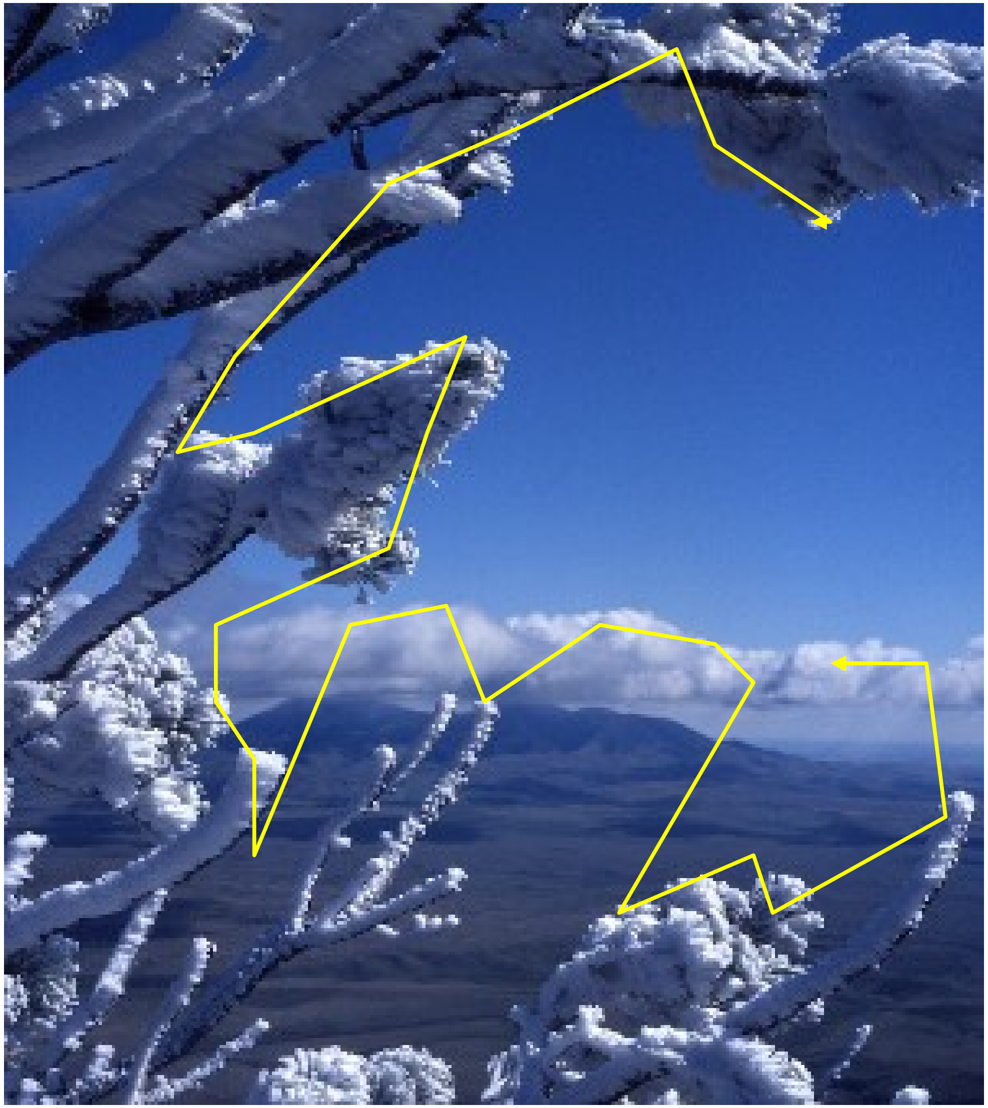


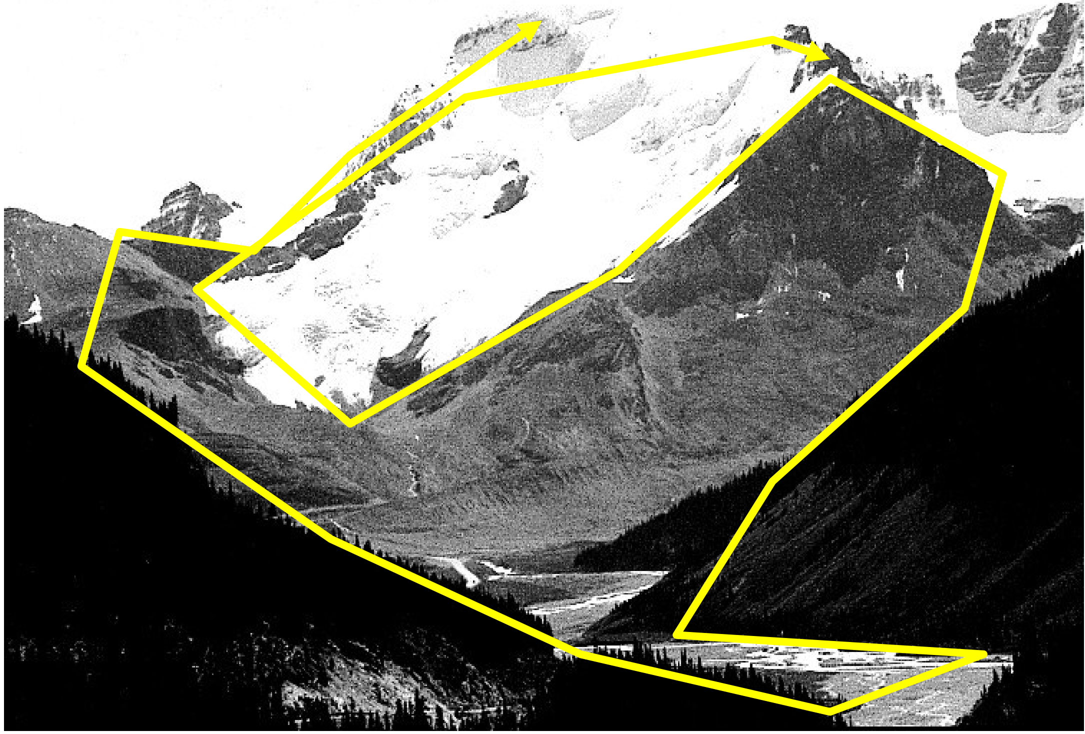
Examples of eye containment



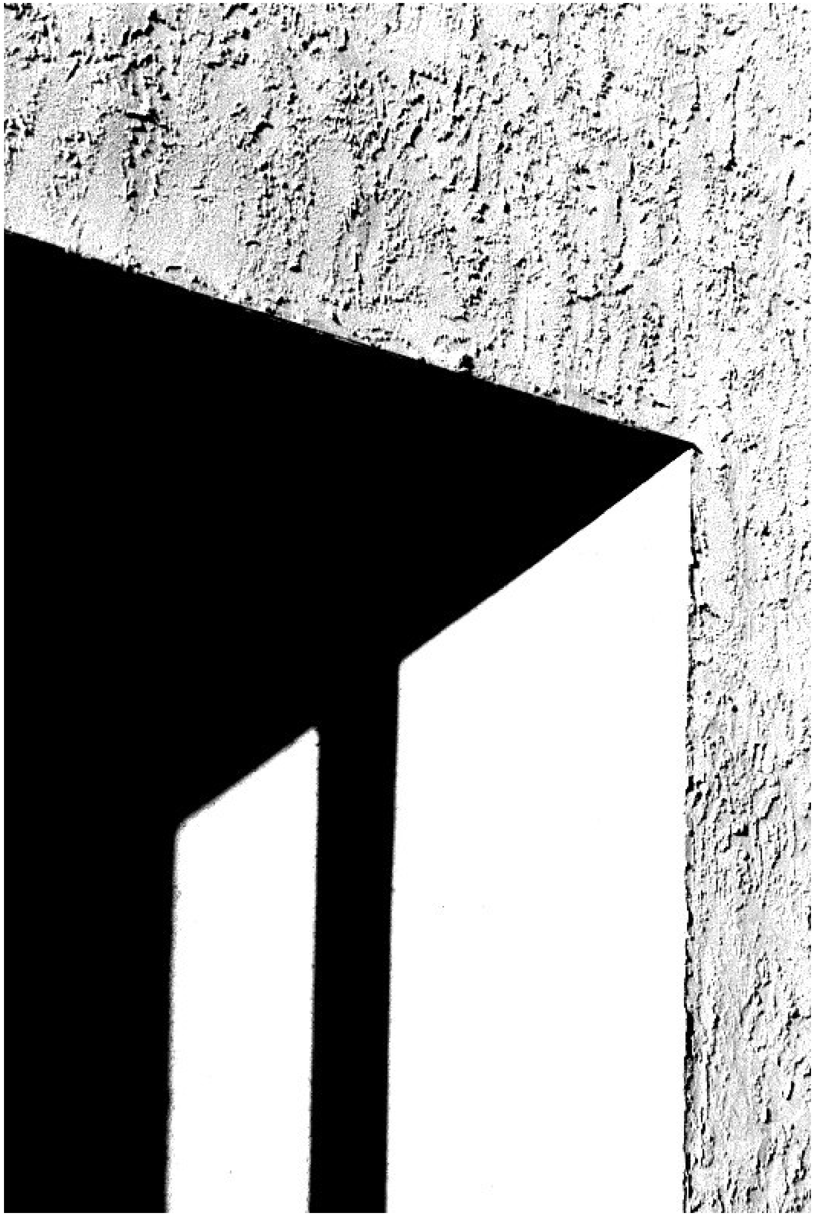
The two images generally work to contain our eyes away from the frame edge but could be improved. Which visual elements are at play in guiding your eyes? What are the faults in the image from this perspective?

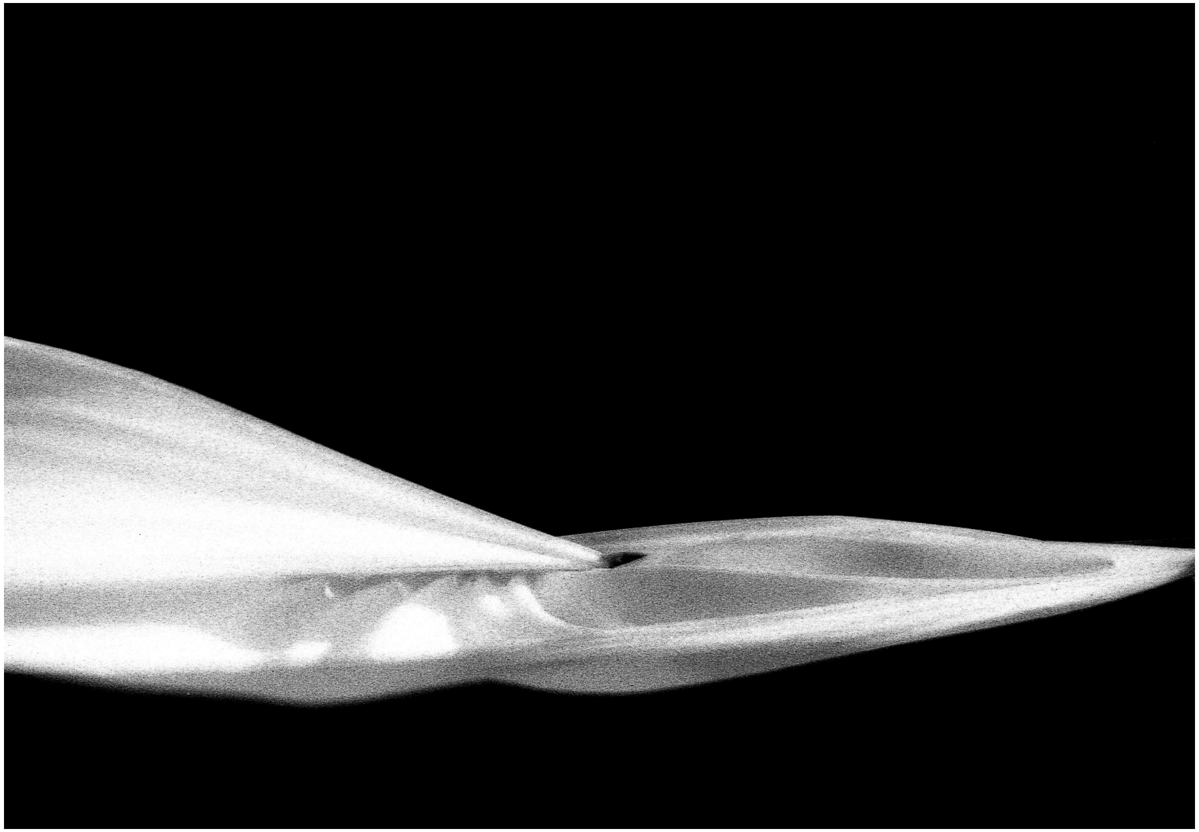
Examples of eye containment





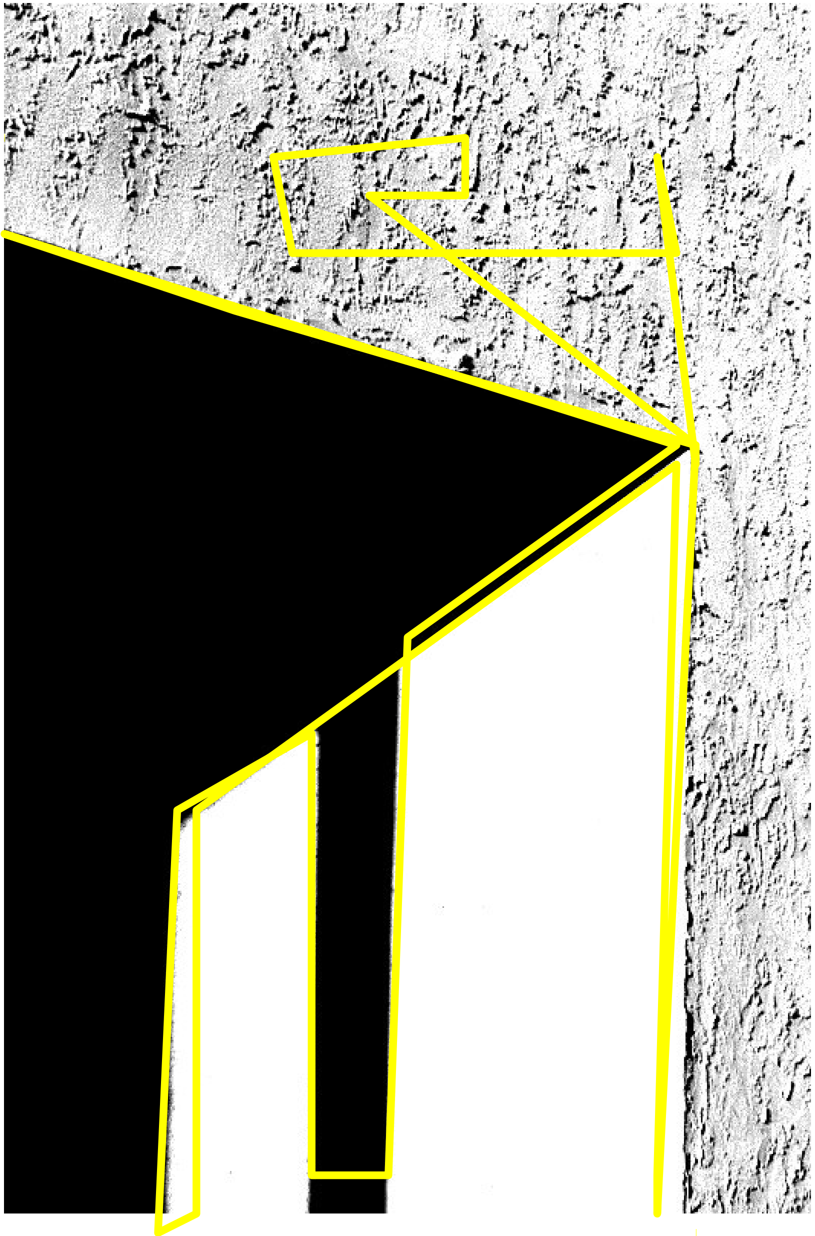
Examples of eye containment 2

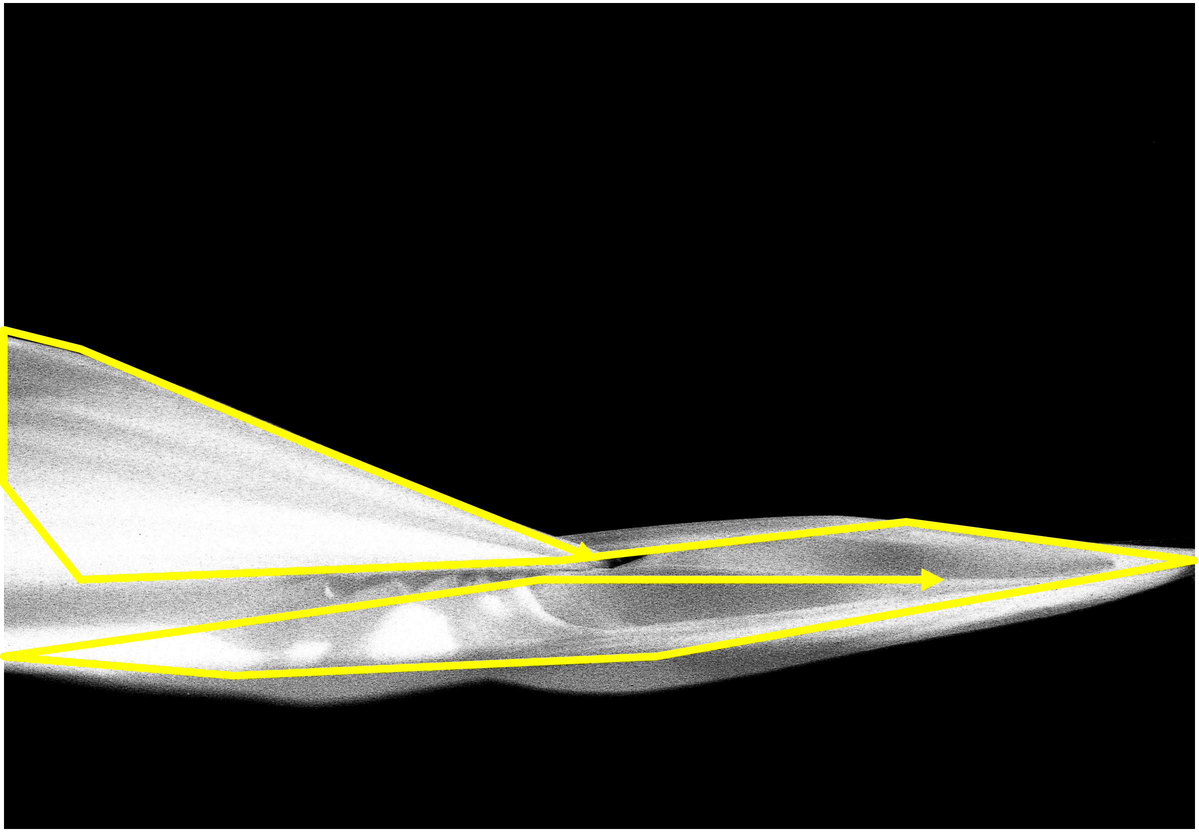




Both images have strong edges leading the eyes ***to*** the frame border.
Are the eyes being contained within the frame at all?
What elements at play here making these images somewhat interesting?

Examples of eye containment 2





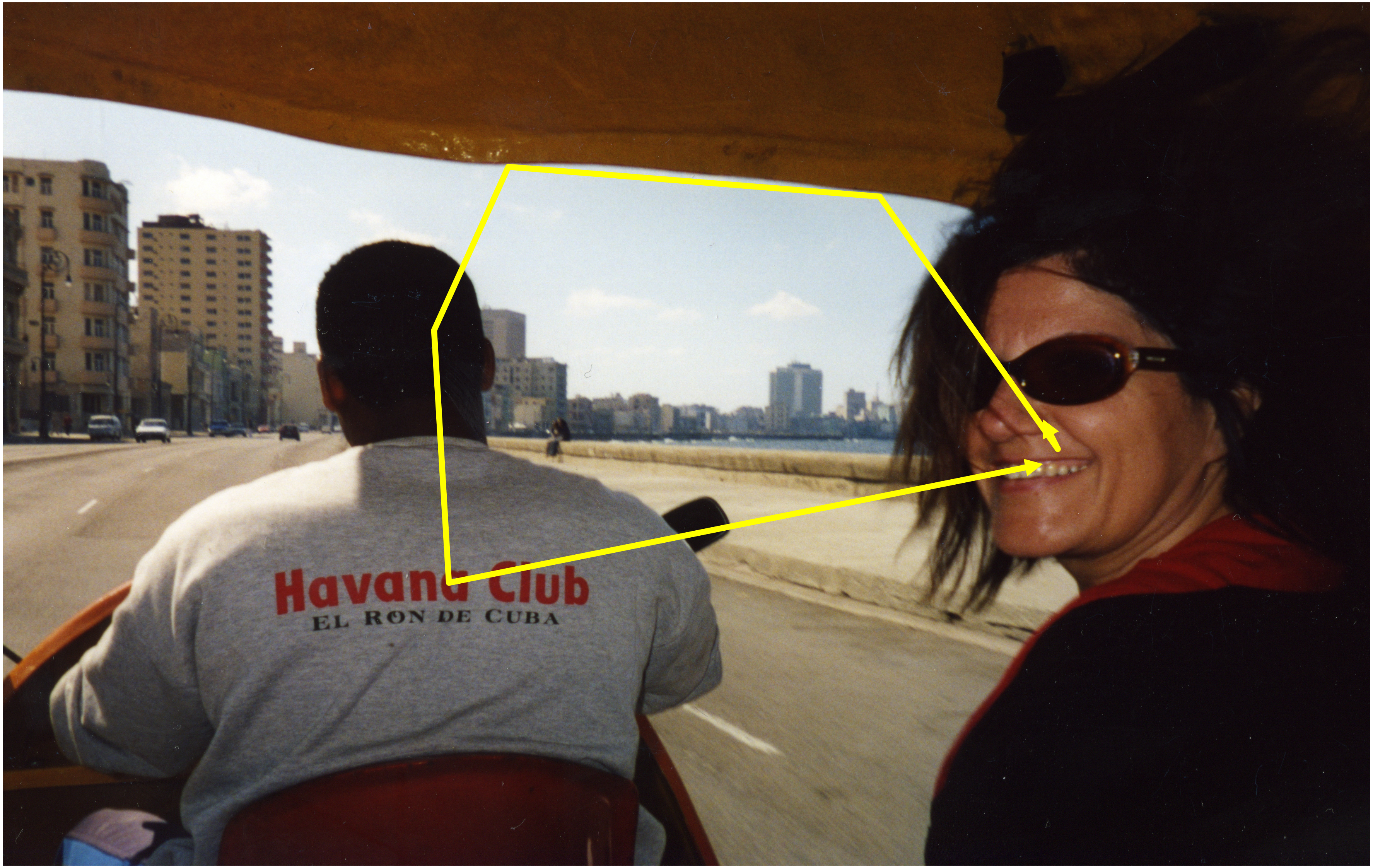
Creating eye movement

- Eye movements and fixations in the image are dependent on the visual elements we place. At least two visual elements need to be in the image to create this movement.
- The objects also create both balance and tension in the image
- What are two basic elements in this image? What's the 'motion'? (Remember follow your eyes!)

Creating eye movement
Dynamic Tension

- Tension in the frame is created when eye movement oscillates between two or three key elements in the frame.
- The more complex the image is the less the tension. SIT's simplicity cornerstone enhances dynamic eye tension.
- The strongest edges in the image are in the facial features, the words on the driver's back (also more information) and along the roof of the cab. These edges set the dynamics for the eyes. Can you see the triangular composition in action?

Creating eye movement
Dynamic Tension



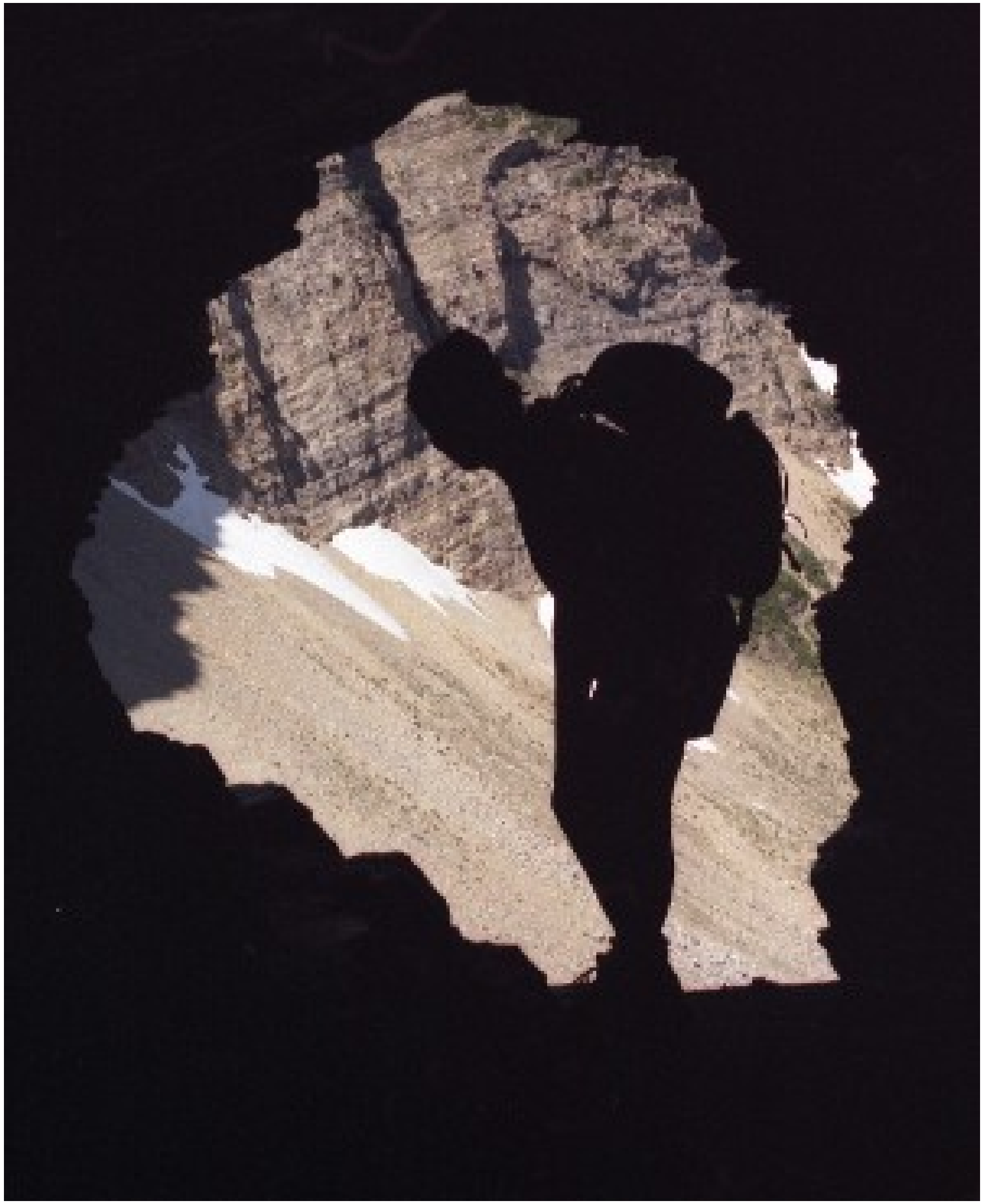
Creating eye movement
Contrast

- We've already seen that contrast (as in an edge) is fundamental to stimulating the eye and we can use this basic response to create the eye movement that we desire. But contrast in images is much more:

- **Large/Small, high/low, point/line, broad/ narrow, horizontal/vertical, smooth/rough, circular/straight, continuous/intermittent etc.**

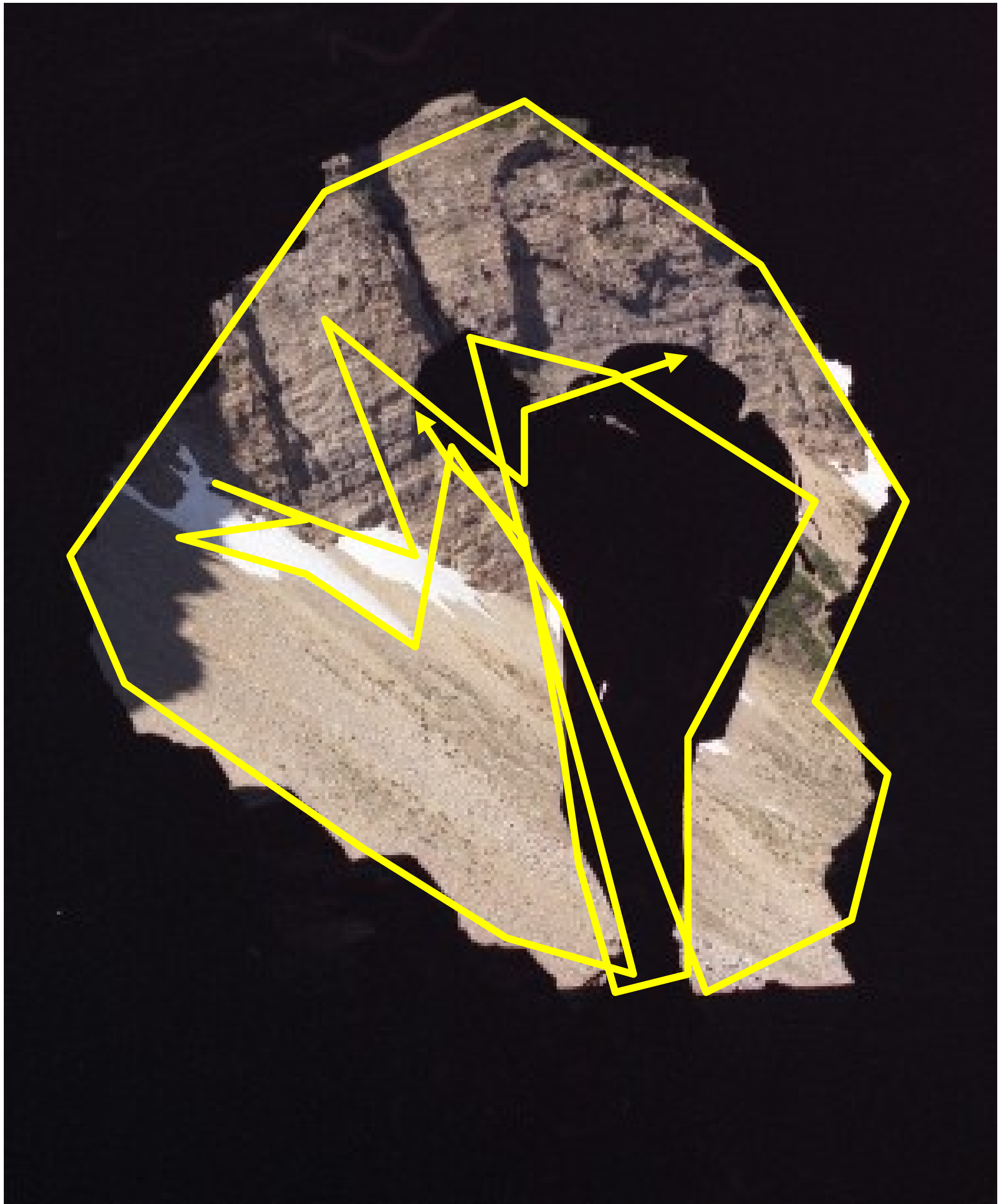
- Consider how you can create eye movement by creating tension between these qualitative contrasts.

- What contrast element(s) are portrayed in the image? What is the eye doing?



Creating eye movement

Contrast



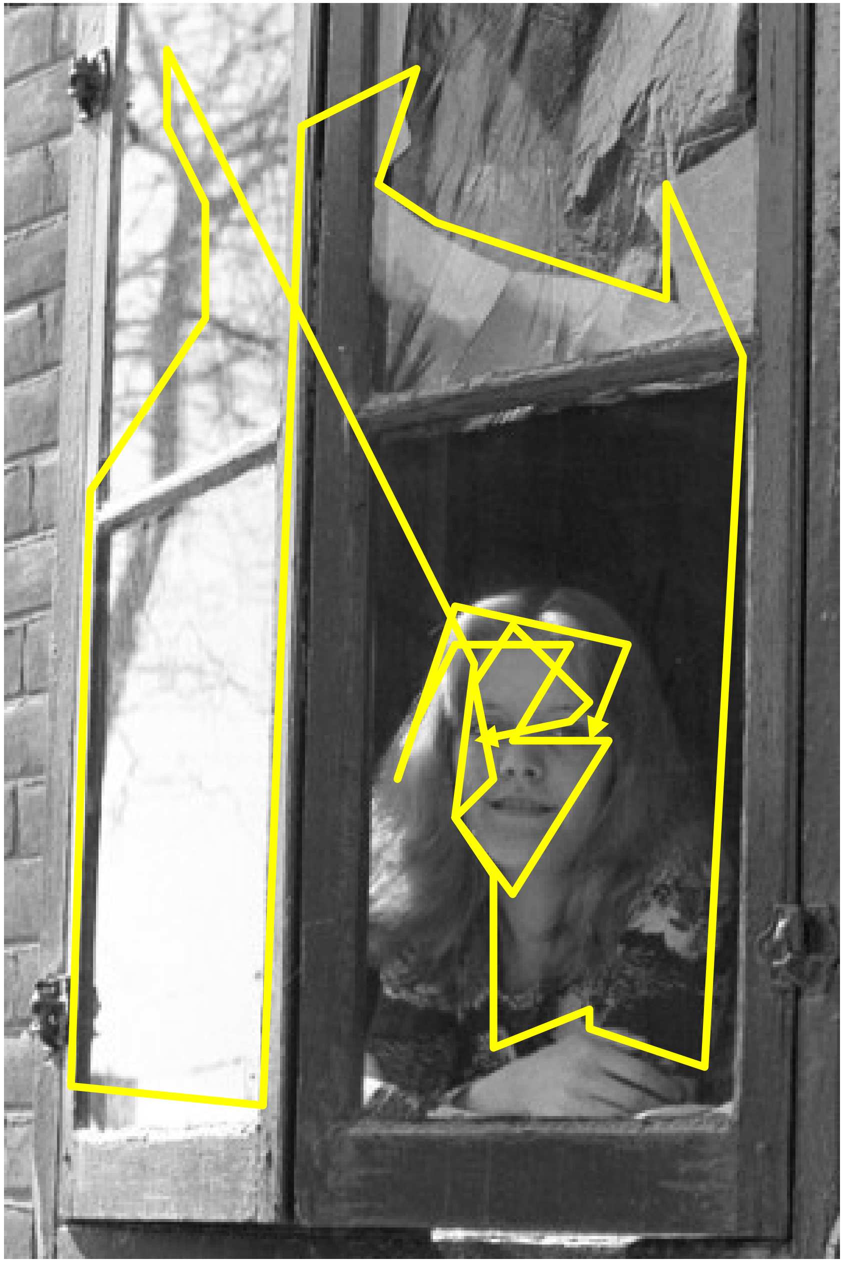
Creating eye movement

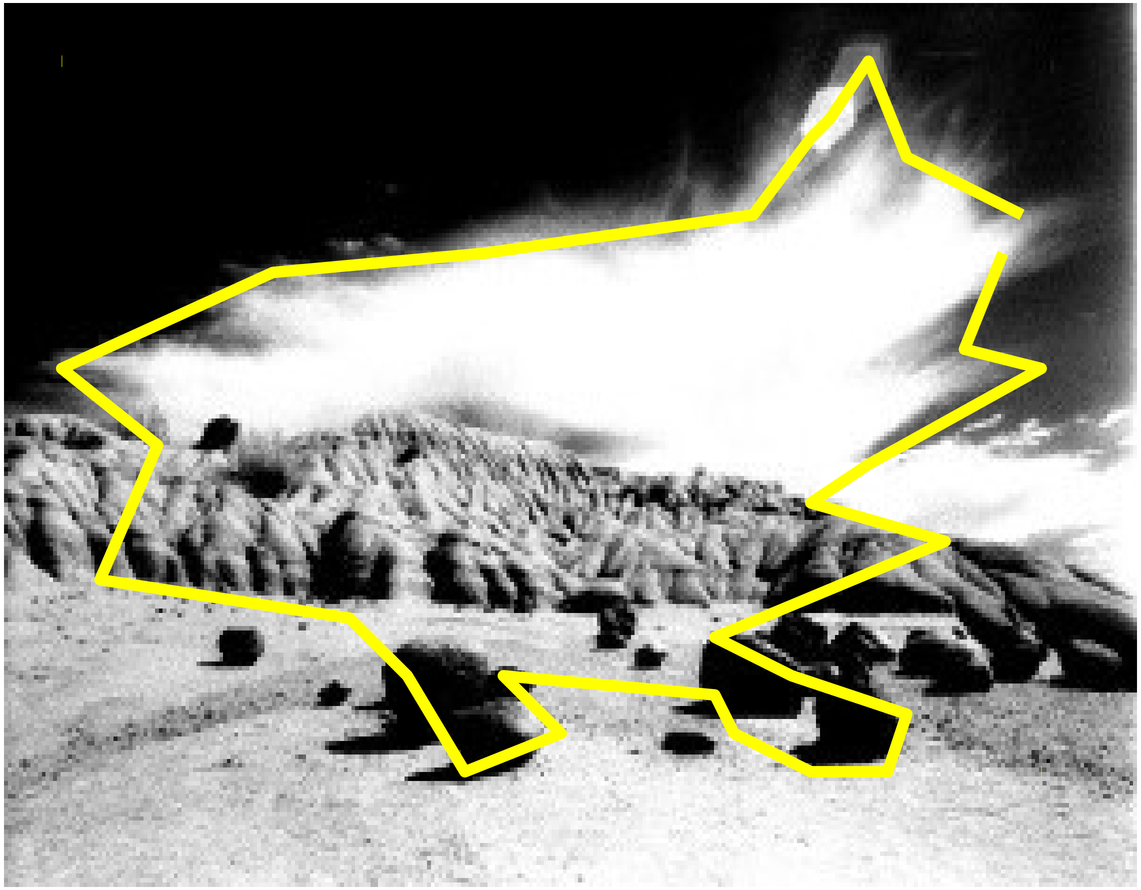
Contrast

- Contrast in form is implied in much of what we've seen thus far in the course. But what about content? Although this course does not delve into content much is certainly implied in most, if not all, of our photography. Images become powerful when contrast in form and content balance in an image. What is being contrasted in form and in content in the two images below? Again, note the eyes' preprocessing of visual information **before** the 'story' is inferred

Creating eye movement

Contrast





Creating eye movement

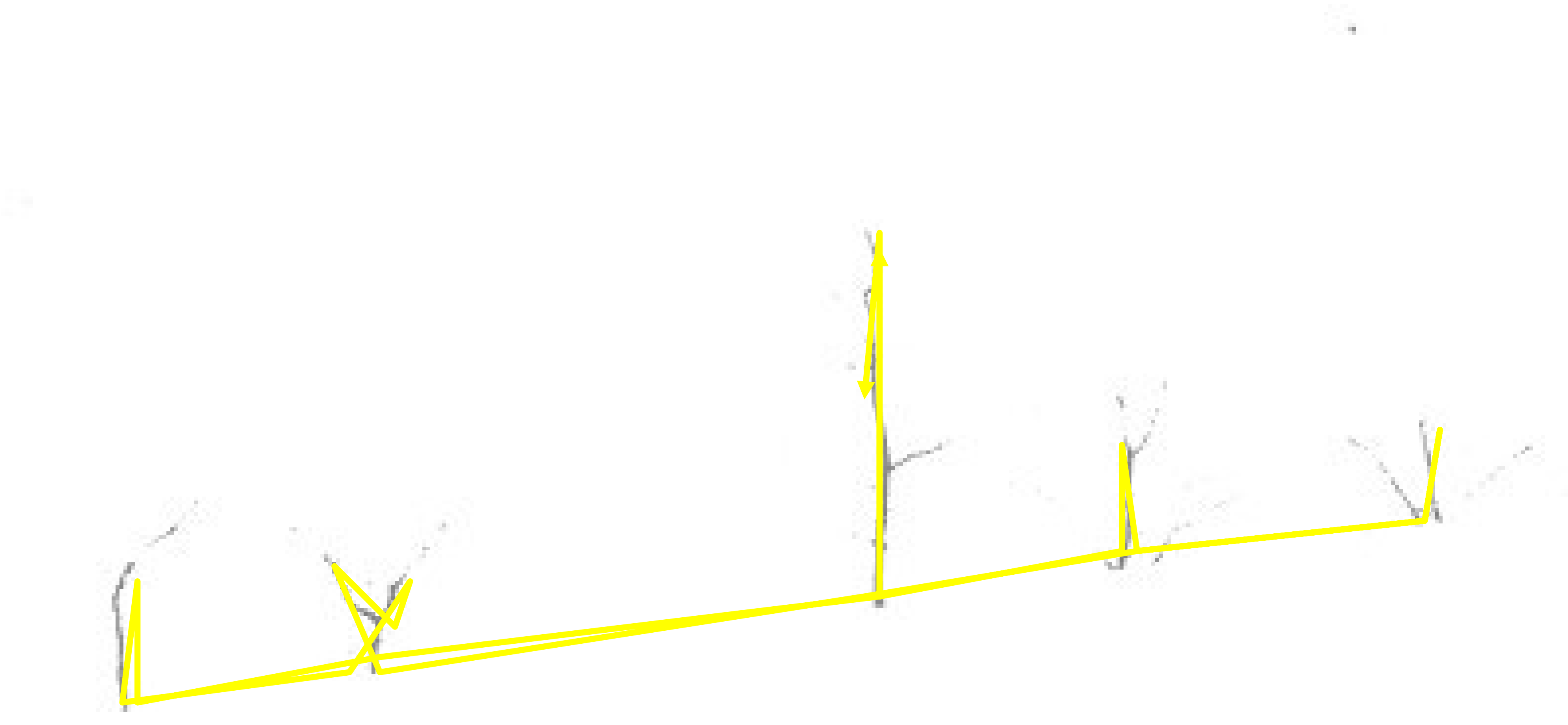
Gestalt Perception

- Gestalt theory, simply stated, is based on the idea that the whole (image) is greater than the sum of its parts. Gestalt theory has a number of 'Laws' , some of which we've seen as part of our investigation into Entropy and Information (Simplicity). We'll sideline the others (Proximity, Similarity, Common Fate, Good Continuation, Segregation) except for one, Closure, because it is relevant and very peculiar in our investigation of eye movement .

- The Law of Closure states that elements roughly arranged together complete an outline (even if it is not there). Can you see it in the image?

Creating eye movement

Gestalt Perception



Creating eye movement
Balance

- Balance is at the heart of composition, it almost demands that visual elements be organized such that eye movement in the image oscillates among two or more points of interest. Mutual opposition is the mechanism by which most balance is achieved and tools are the use of contrast both form (edges, lines) and content. As image entropy increases balance is compromised. Consider what would happen to balance and eye movement if the condensation trail was not in the image..

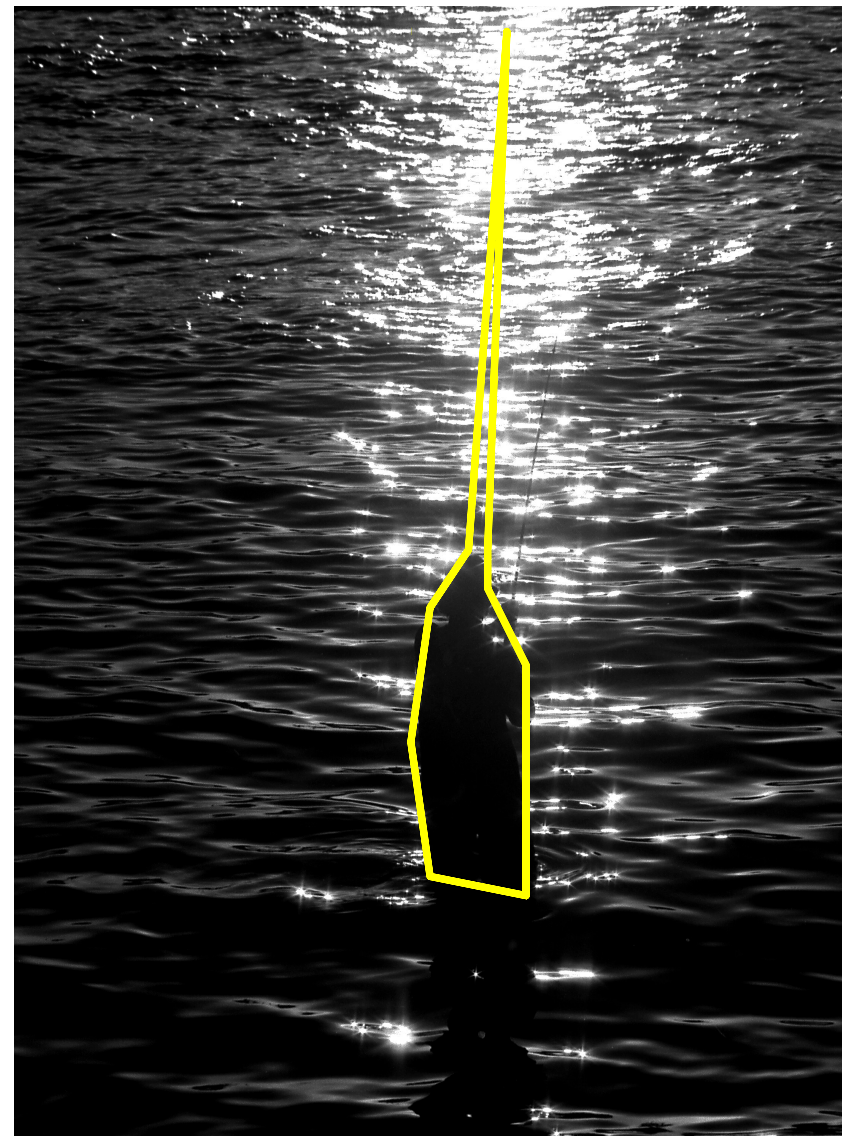
Creating eye movement

Order and minimizing entropy

- Repetition is a key element in creating order. Repeating pattern drive eye movement from one edge to the next repeating edge, ideally to some end element. Repeating patterns create rhythm and textures in design language

Creating eye movement

Order and minimizing entropy



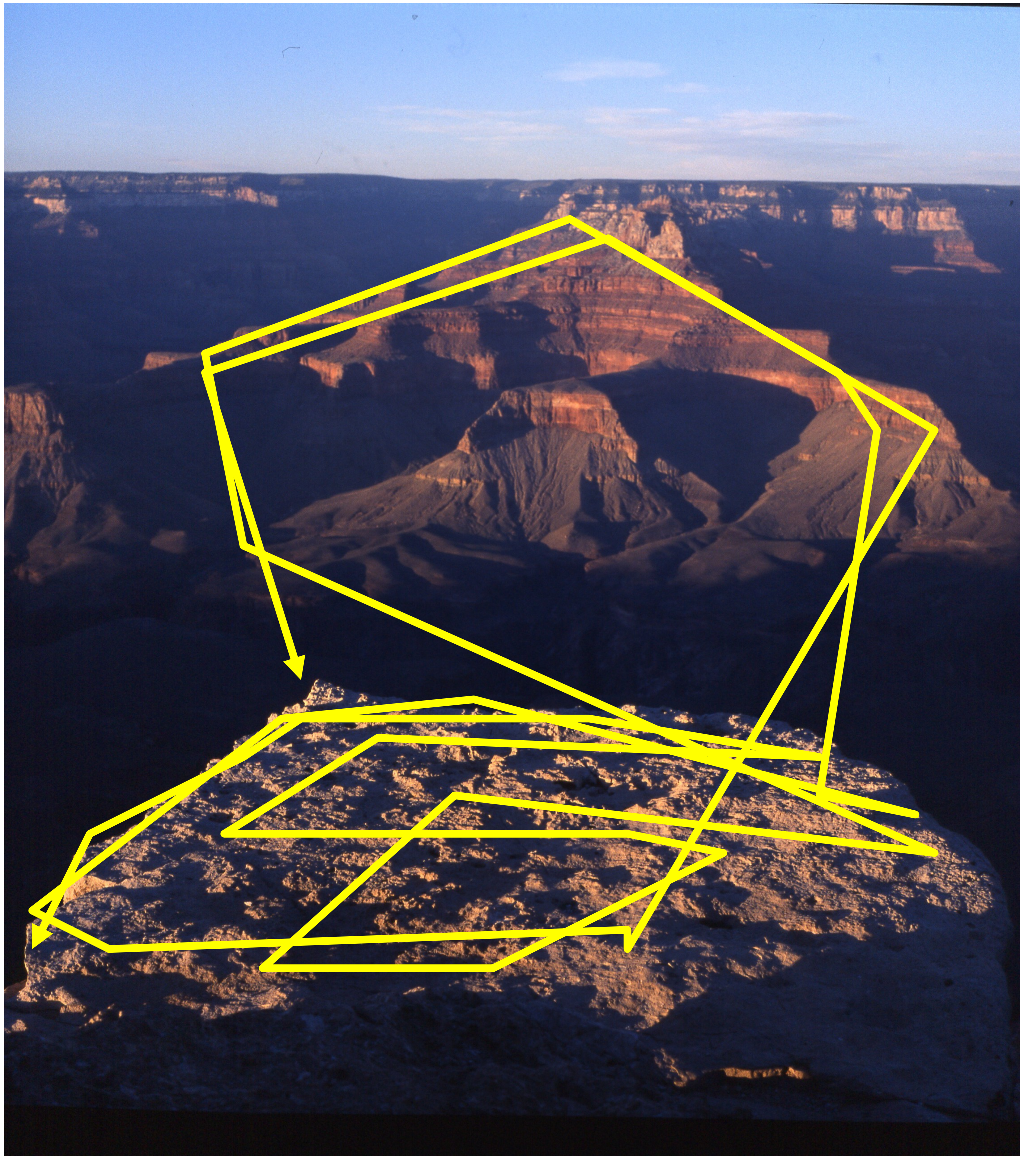
Creating eye movement

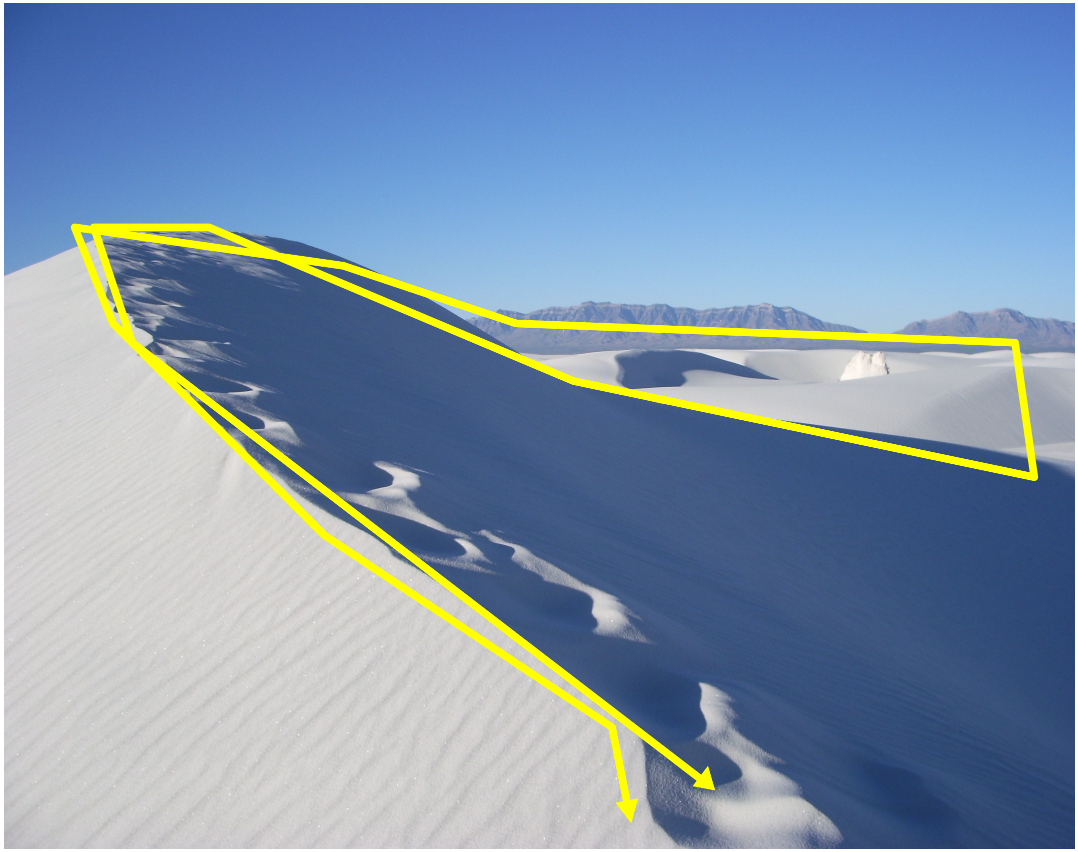
Depth and Perspective

- Generally photography portrays a 2D plane; eye movements are on this plane. Using elements to generate eye movement (our usual edges and lines) and strategic placement of the elements, we can generate the illusion of depth in the image.
- The principle (also common in 3D images and movies) involves placing a strong edge or line close to the frame border to ensure that the eye is anchored and then using that element to lead the eye away from that frame border. Generally the frame border has an object that is close to the photographer and that object leads to other elements that are further.

Creating eye movement

Depth and Perspective





Telling a story..

- A good picture, indeed, tells poignant story. As the photographer we may want to make a point of telling that story straight out but, at the very least, we want to make sure that the guiding elements (our lines and edges) lead to telling that story.
- Are the visual elements in this image effective in 'telling the story'? What is the balance/tension? What do you think the story is?

Telling a story..

Telling a story..

- Photography is one of the best mediums for making comments on our the social fabric. This image's story could be one of loneliness, marginalization, etc.. The 'story' and the interpretation could be very different indeed. But, there is no doubt that the human interest is the figure at the centre of the image and that the photographer has used an arsenal of edges, lines, patterns, textures to make that point.

Throwing out the rule book...

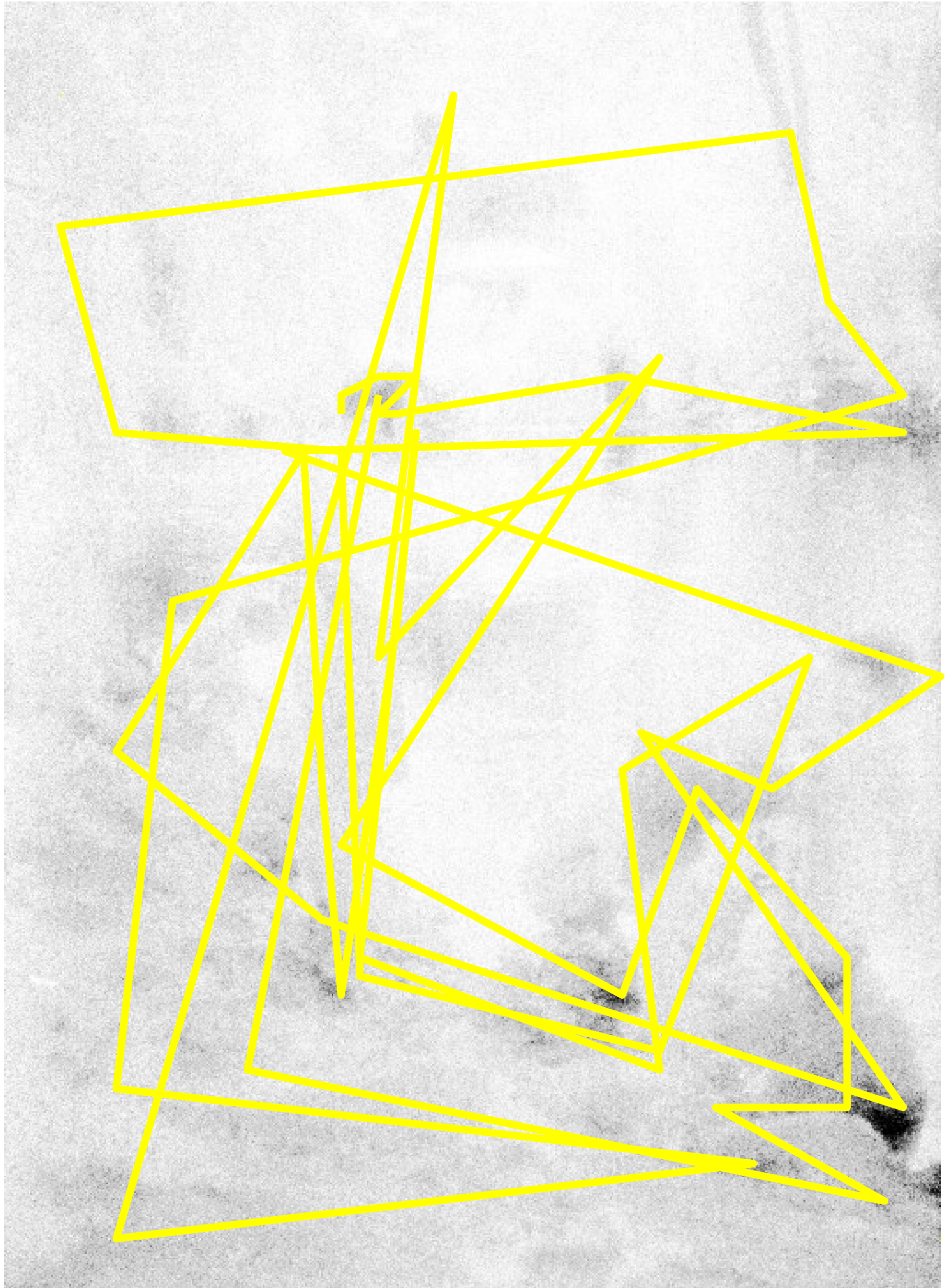
So, the eye is designed for seeing edges and lines, and physics tells us there is more information in order and edges, and that our eyes want to see order, and so on..

What would happen if we muted all the edges and lines in an image? If we challenged our audience perceptually by removing visual stimulant anchors?

A. The viewer may create their own reality ie. Top-down visual processing

B. if this image is entitled 'Four Persons by a Lake' (a code!) then...

Throwing out the rule book...

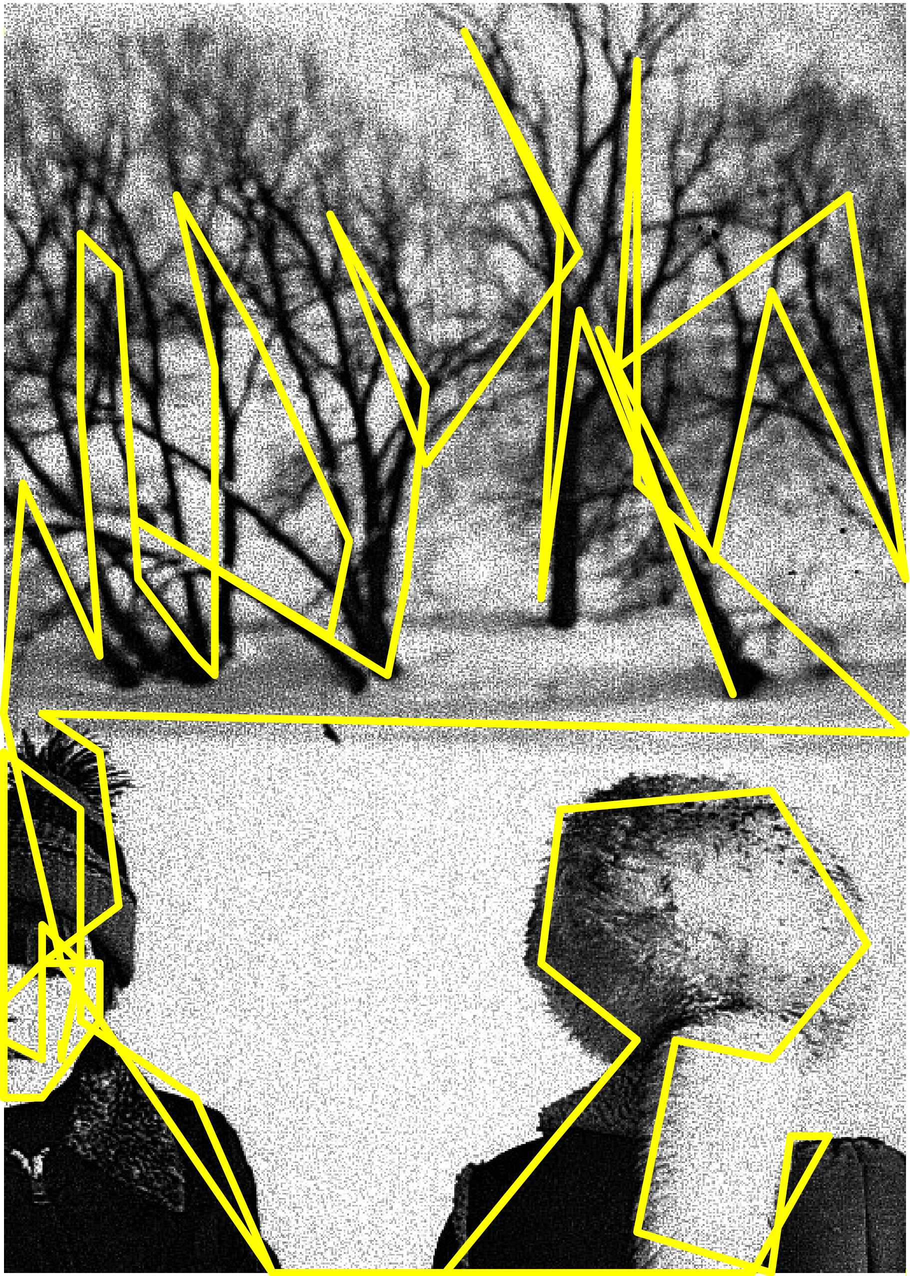


Throwing out the rule book...

Edges and lines are power elements in the visual field. The frame border, is one of the most powerful forms in an image, so much so that we devise strategies to prevent it from interfering with ***our*** edges.

In this image the border deliberately transects the boy's face; a power edge is used to create a story. Very unconventional but along with the other boy's back of head and the scene bisected horizontally in two, our sense of order is disturbed.

Throwing out the rule book...



From Theory to Practice..

Let's recap the most important points made in this course:

1. Our eyes like edges, lines and contrast. Understanding this fact and using it appropriately in our visual compositions will make our message more focused.
2. Our eyes are designed to search for order in our environment. Increasing order decodes randomness and conveys more information. Many visual design principles (repetition, balance, tension, etc.) are based on this natural predilection.
3. Simplify, reduce feature complexity. In photography, as opposed to painting, reducing image clutter can be challenging. Restricting image dynamic tension or interplay to two or three features should be a goal.
4. Using items 1-3, contain the viewer's attention in the frame

From Theory to Practice.. How?

The practice of a visual medium like photography can be made into an everyday activity, how can you make what you see more salient? How would you use photographic techniques such as zooming and cropping to reduce clutter, complexity and increase simplicity? Or maybe ensure that the edges and contrast are aligned and emphasized to focus your message through imaging software (GIMP, Photoshop, Photopaint)?

A basic example..

How would you crop this image? Remember, apply our course learning's!
What is what's superfluous, what's distracting what's essential?
What kind of eye motions are occurring?

A more subtle one..

And the final one..

...where a selective zoom may add the drama needed to make an image from “historically interesting” to “artistically interesting”. Does the zoomed image 'work' from the eye movement perspective? It certainly does not contain the eye, in fact all lines lead out of the frame, a large structure perhaps?

Automated Cropping?

We have investigated vision and noted that certain forms are more stimulating to the eye..

We have looked at information content in our environment and also noted that certain structures will contain more information than others..

We have looked at research techniques that track eye movement and fixation and that produce saliency maps and heat maps of stimulating and informative features of a scene..

Given the semblance of a process, can computers be programmmed to crop images to create a more pleasing aesthetic? If so, why, and what happens to the creative process?

Retargeting
Computer based cropping and resizing

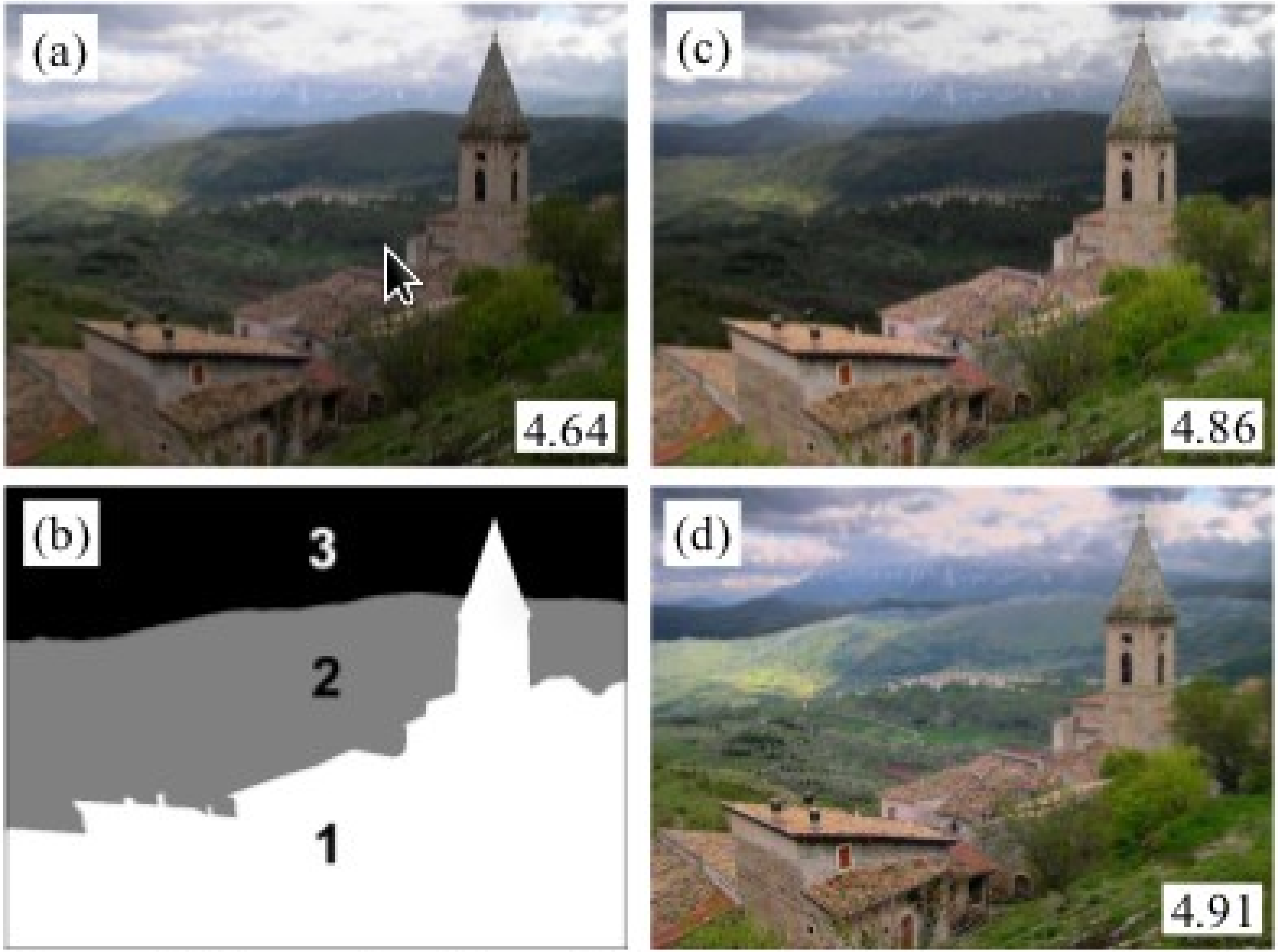
The technique combines typical image enhancement techniques such as contrast enhancement, color correction and balancing, and edge sharpening with computational visual attention models that simulate the human visual system to identify regions of interests (ROIs) in an image ie. salience or heat maps.

The main purpose of these techniques has been to adapt imagery and video to larger or smaller display formats such as from 4:3 to 16:9 TV or cellular devices such as iPhones.

It is implemented as a feature in Adobe Photoshop CS4, where it is called Content Aware Scaling and in the free GNU Image Manipulation Program as Liquid Rescaling. The technique is also known as Seam Carving (see Wiki)

http://liquidrescale.wikidot.com/en:start

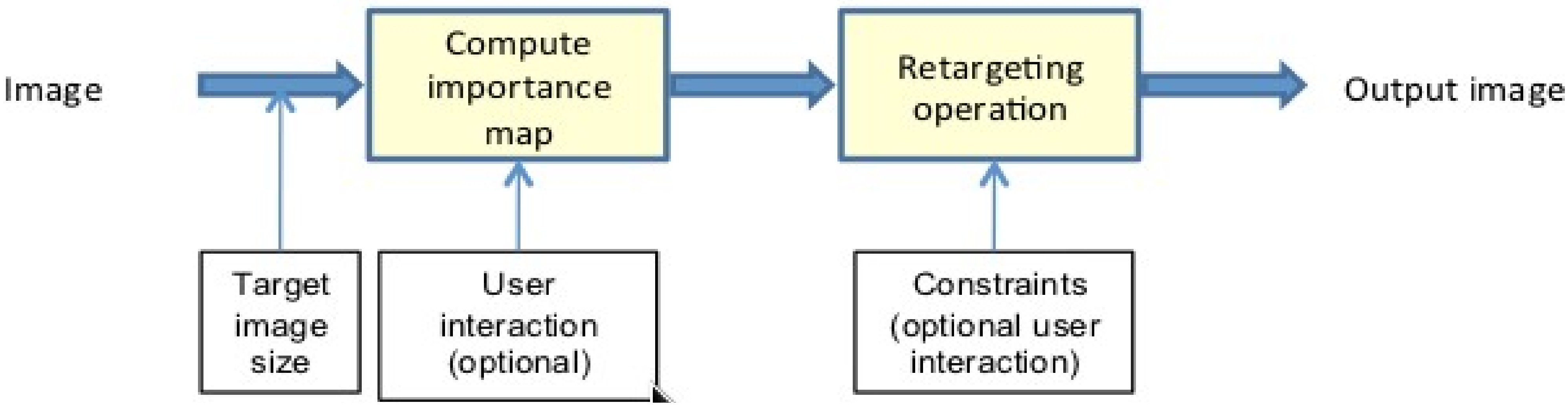
Retargeting techniques
Segmentation



Source: Saliency Retargeting: An Approach to Enhance Image Aesthetics by L. Wong K. Low

Retargeting techniques
Importance Map

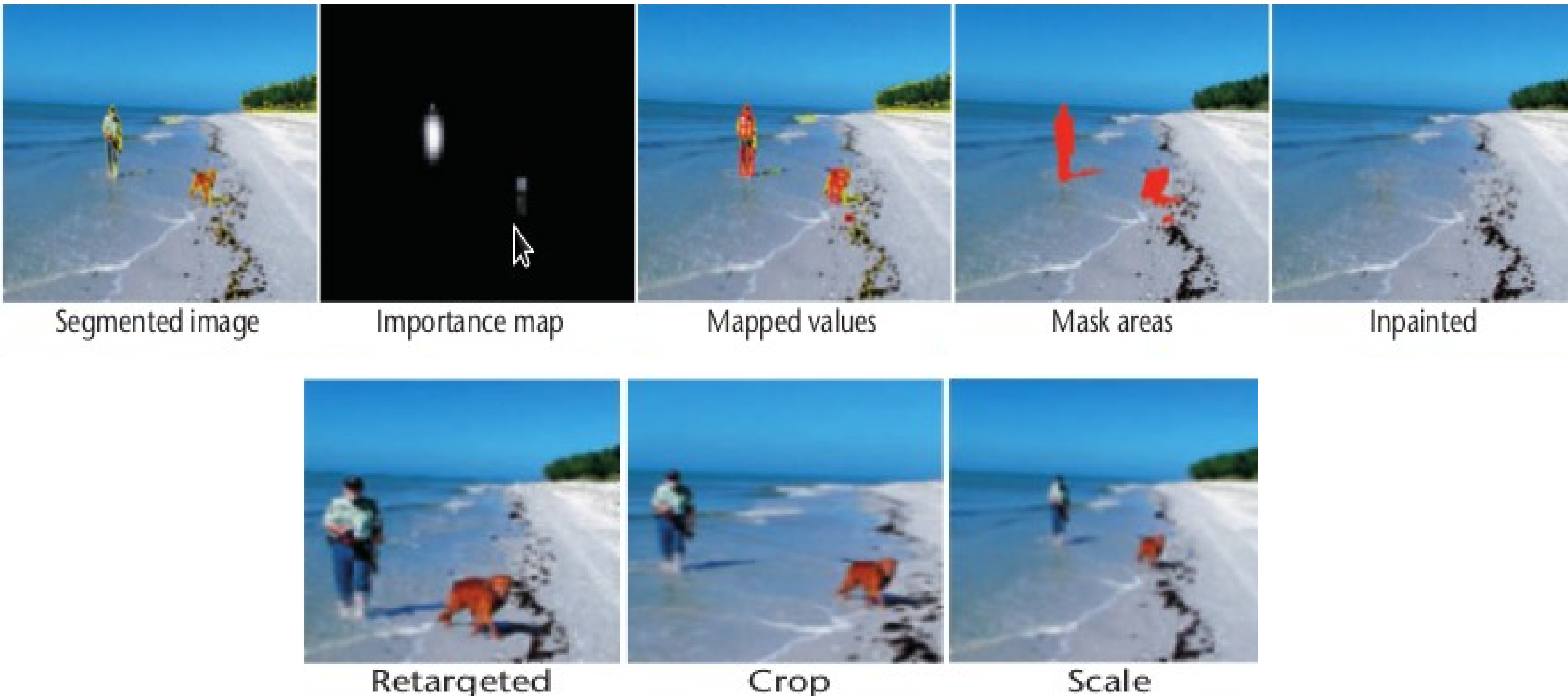
Most techniques typically start by computing an *importance map* (or salience map) which represents the relevance of every pixel, and then apply an operator that resizes the image while taking into account the importance map and additional constraints.



Source: A survey of image retargeting techniques
Daniel Vaqueroa , Matthew Turka , Kari Pullib , Marius Ticob , Natasha Gelfandb

Retargeting techniques

A complete example



Comparison between existing image resizing techniques and an automatic image retargeting method (from: Retargeting Images and Video for Preserving Information Saliency, V.Setlur,T. Lechner, M. Nienhaus, B. Gooch)

Wrapping Up..

- Preference..
- Our eye biology evolved as a function of survival, ensuring that we extract the most information and order from our environment.
- Real-time Information and order in the visual environment transposes to what the eye prefers to see in an image.
- Visual design principles are outcomes of these basic biological and physical determinants.

Wrapping Up..

- Statement #1: Our eye biology evolved as a function of ensuring that we extract the most visual information and order from our physical environment.
  - Does photography (and the visual arts in general) ***conform*** to this notion? If there was a notion of conformity in photographic composition would its basis be in biology and physics?
  - The sciences ground themselves in the scientific method; process replication of insight. Is there a scientific basis to photographic creation? How do principles of visual design substitute for 'process replication'?

Wrapping Up..

- Statement #2: Entropy in the universe is increasing overall. But pockets of increasing order are observed such as in nascent stellar clusters and with life on earth.
  - Humans (and living things in general), build order (minimize entropy). Is the process of photographing an extension this innate biology?
  - Humans die. The second law of thermodynamics prevails. Does photography mimic this universal truth? Does 'energy' or effort need to be expended to maintain photographic image order?

Wrapping Up..

- Statement #3: Information Theory states that information increases as the visual message becomes more complex. Without a code this complexity increases infinitely to eventual 'noise'.
  - What is the happy medium between an image that is too simple and lacks information and one that is too complex and undecipherable? What type of range is there between these two relative extremes. How do visual design principles enable a 'code' to decipher complexity?
  - To what extent is information ever present around us but is undecipherable (no code, no technology, etc..)?

Wrapping Up..

- Statement #4: The course has been based on the prerequisite to visual cognition is a bottom-up approach to visual stimulus (the micro-form lines and edges over macro-form of shapes associated with the top-down approach). Meaning, content, and significance have, for the most part, been avoided.
    - Can content ever *drive* form (does cognition precede behavior or stimulus)?
    - Does knowledge or memory of a image scene ever trump strict visual stimuli? (think people, family).

Applying our learnings..

- Using your knowledge of eye circuitry and information theory:
  1. Create image that has the lines, edges, textures etc. assembled to convey minimum image entropy.
  2. Create an image that conveys maximum entropy.
- Be prepared to speak to *how* each of the images' components achieve the two goals.
- Bring the images to next week's class on a USB drive.

Web Sites List

Retinal elements of the eye:

http://neuroscience.uth.tmc.edu/s2/chapter14.html

http://www.arn.org/docs/glicksman/eyw_041101.htm

http://www.youtube.com/watch?v=hgZFuq2S15A

Entropy in composition:

http://www.nici.kun.nl/~peterh/doc/sit.html

http://www.robertsmithson.com/essays/entropy.htm

http://www.naturephotographers.net/articles0808/ip0808-1.html

General Photo-composition:

http://www.colorpilot.com/comp_rules.html