

Aspectos de Percepção Visual

Carlos Alexandre Mello

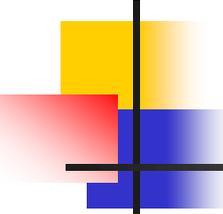
Pós-Graduação em Ciência da Computação

História

- *“What is real? How do you define real? If you’re talking about what you can feel, what you can smell, what you can taste and see, then real is simply electrical signals interpreted by your brain. This is the world that you know.”*

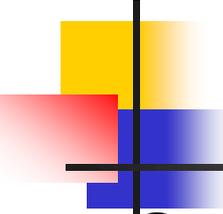


Morpheus,
Matrix, 1999



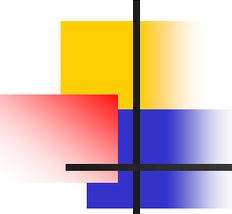
História

- Essas questões vêm sendo levantadas pela humanidade há muito tempo
- Até hoje só temos suposições sobre o que é real ou o que achamos que é real
- A percepção em seu sentido mais amplo varia de espécie para espécie e até dentro de uma mesma espécie



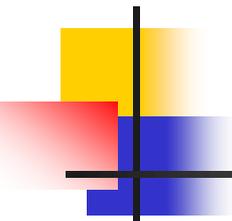
Sensação e Percepção

- Sensação e percepção são dois estágios no processo no qual nós construímos uma representação interna do ambiente
- Sensação: energia física é convertida em sinais neurais
- Percepção: sinais são selecionados, organizados e interpretados
 - Cocktail-Party problem
- Nosso objetivo, como organismos biológicos é responder, interagir e adaptar ao ambiente, extrair o significado dos objetos que existem e dos eventos que ocorrem ao nosso redor



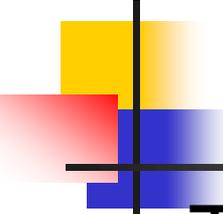
Percepção

- Percepção não depende apenas de energia e eventos que mudem o ambiente, depende também de características do perceptor
- Adaptação dos sentidos a estímulos
 - Ex.: Cheiro diferente ao entrarmos em um ambiente – logo, nos adaptamos e não percebemos mais



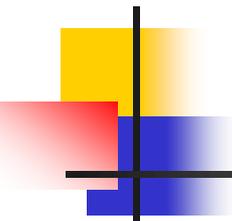
O que é Percepção?

- Transdutores são estruturas que mudam a energia de uma forma para outra
- Transdutores transformam informações sobre o mundo (luz, som, pressão, composição química) em sinais neurais que possam ser interpretados pelo cérebro
- Percepção lida com a interpretação desses sinais



Sensação \Rightarrow Atenção \Rightarrow Percepção

- 1) Transdução: receptores sensoriais traduzem estímulos físicos em sinais neurais
- 2) Sinapses transmitem essa informação ao cérebro
- 3) Códigos neurais/mentais rudimentares, representando características do ambiente, são construídos no cérebro/mente
- 4) Características (propriedades e partes) são combinadas (agrupadas) para formar objetos (inteiros); pode requerer atenção



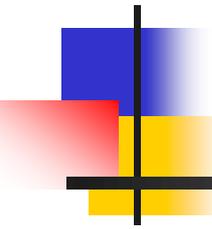
Sensação \Rightarrow Atenção \Rightarrow Percepção

- Da forma descrita anteriormente, a progressão do processamento da informação acontece de forma *bottom-up*
- Podemos também identificar influências do processamento *top-down*
 - Nesse caso, memória e conhecimento geral fazem parte da percepção consciente
 - Exemplo:

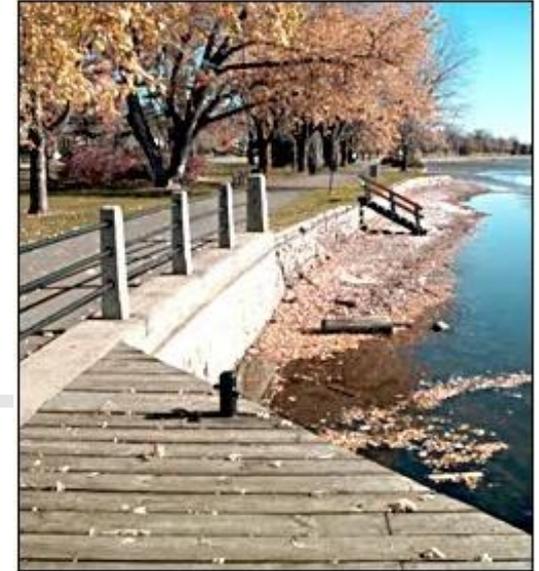
Sensação \Rightarrow Atenção \Rightarrow Percepção



A compreensão correta das palavras depende do nosso conhecimento prévio sobre a linguagem

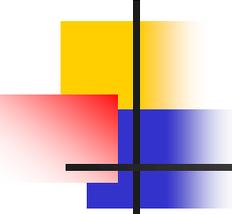


Percepção de Profundidade



Carlos Alexandre Mello

Pós-Graduação em Ciência da Computação



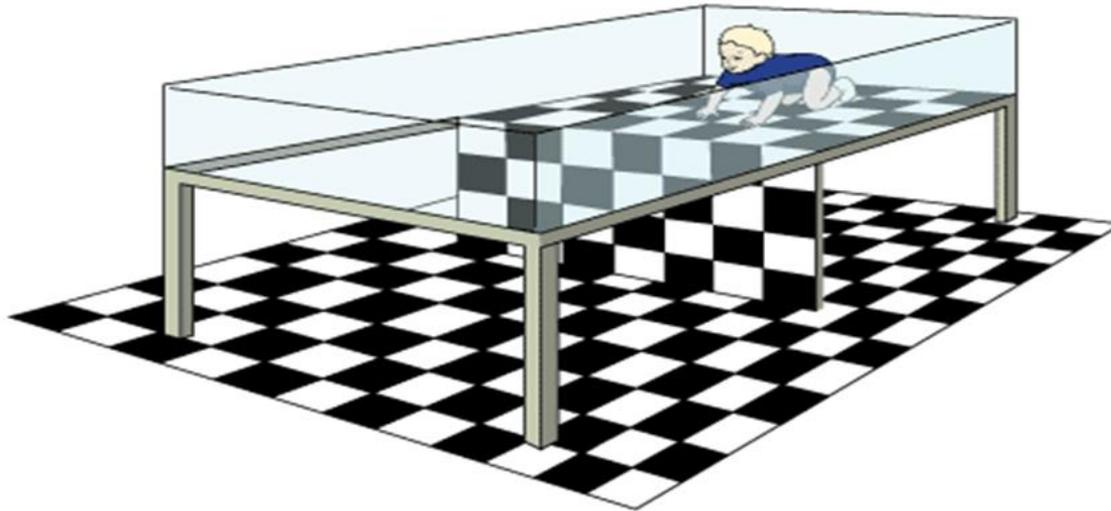
Percepção de Profundidade

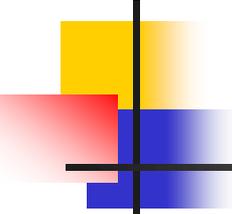
- Percepção de profundidade
 - Habilidade de perceber objetos tridimensionalmente
 - Experimento: The Visual Cliff

Percepção de Profundidade

■ Visual Cliff

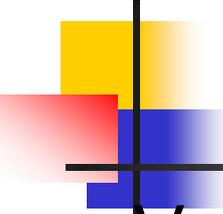
- <http://www.youtube.com/watch?v=p6cqNhHrMJA>





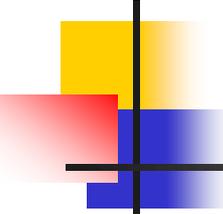
Percepção de Profundidade

- Percepção de profundidade
 - Variáveis fisiológicas ou psicológicas
 - Fisiológicas (ou internas) envolvem o sistema visual
 - Psicológicas (ou externas): aprendidas com a experiência
 - Variáveis monoculares ou binoculares
 - Binoculares dependem dos dois olhos
 - As monoculares podem ser apresentadas em uma representação bidimensional como uma imagem



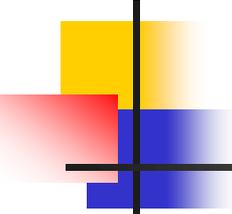
Percepção de Profundidade

- Variáveis Fisiológicas (ou internas)
 - Acomodação
 - Convergência
 - Disparidade binocular
 - Paralaxe de movimento
- Variáveis Psicológicas (ou externas)
 - Tamanho da imagem na retina
 - Perspectiva linear
 - Perspectiva aérea
 - Oclusão
 - Sombras e sombreamento
 - Gradiente de texturas



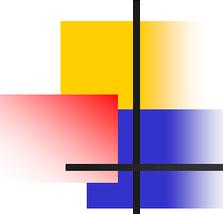
Percepção de Profundidade

- Variáveis monoculares
 - Variáveis de perspectiva
 - Tamanho
 - Oclusão
 - Profundidade de foco
 - Sombra
 - Forma a partir de movimento
- Variáveis binoculares
 - Convergência
 - Profundidade estereoscópica



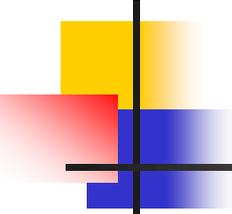
Percepção de Profundidade

- Existem quatro variáveis fisiológicas e seis psicológicas
- Essas variáveis contêm informação as quais, quando adicionadas à imagem 2D projetada na retina, nos permite relacionar os objetos da imagem ao espaço 3D



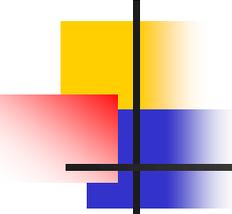
Percepção de Profundidade

- Variáveis fisiológicas:
 - Acomodação: o ajuste do comprimento focal das lentes;
 - Convergência: o ângulo feito pelos dois eixos de visão de um par de olhos;
 - Disparidade binocular: a diferença entre as imagens do mesmo objeto projetadas nas retinas;
 - Paralaxe de movimento: o resultado de mudanças de posição de um objeto no espaço devido tanto ao movimento do objeto quanto ao movimento da cabeça do observador.



Percepção de Profundidade

- **Acomodação e Convergência**
 - Acomodação e convergência estão associadas com os músculos dos olhos e interagem uma com a outra na percepção de profundidade
 - Acomodação é considerada uma variável de profundidade monocular já que é disponível mesmo com um único olho
 - Essa variável é efetiva apenas quando combinada com outras variáveis binoculares e para uma distância visual de menos de 2 metros

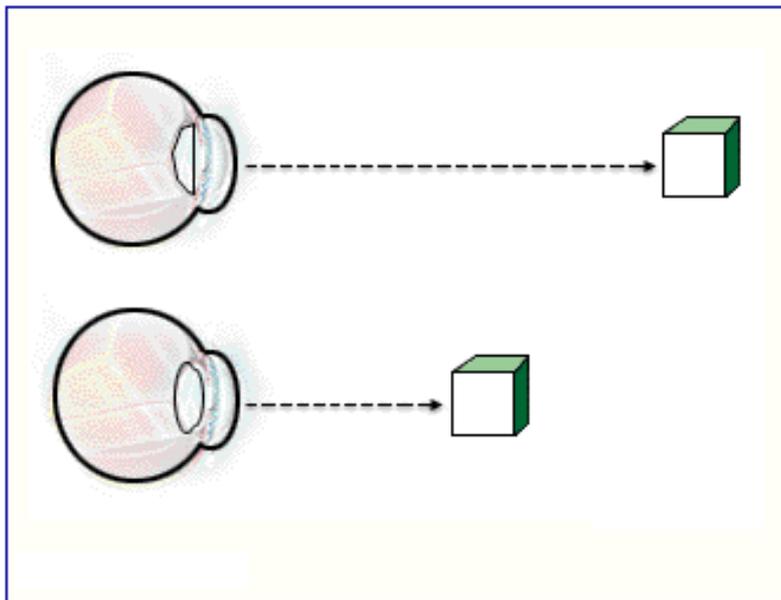


Percepção de Profundidade

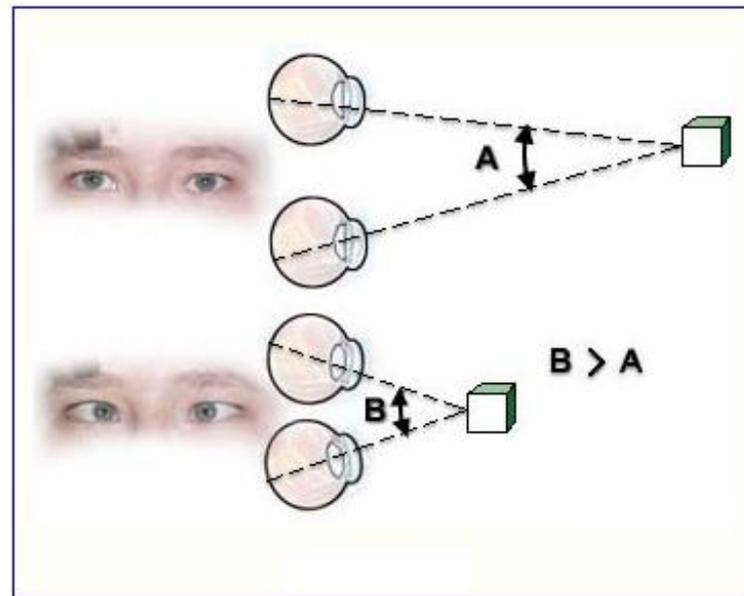
- **Acomodação e Convergência**
 - Quando vemos objetos perto, nossos olhos **convergem** tal que uma única imagem é formada em ambas as retinas
 - **Mudanças na tensão muscular** são detectadas e interpretadas pelo cérebro e usadas para determinar a profundidade e distância de um objeto
 - Nosso cérebro detecta que quanto mais próximo o objeto, maior a convergência

Percepção de Profundidade

■ Acomodação e Convergência



Acomodação



Convergência

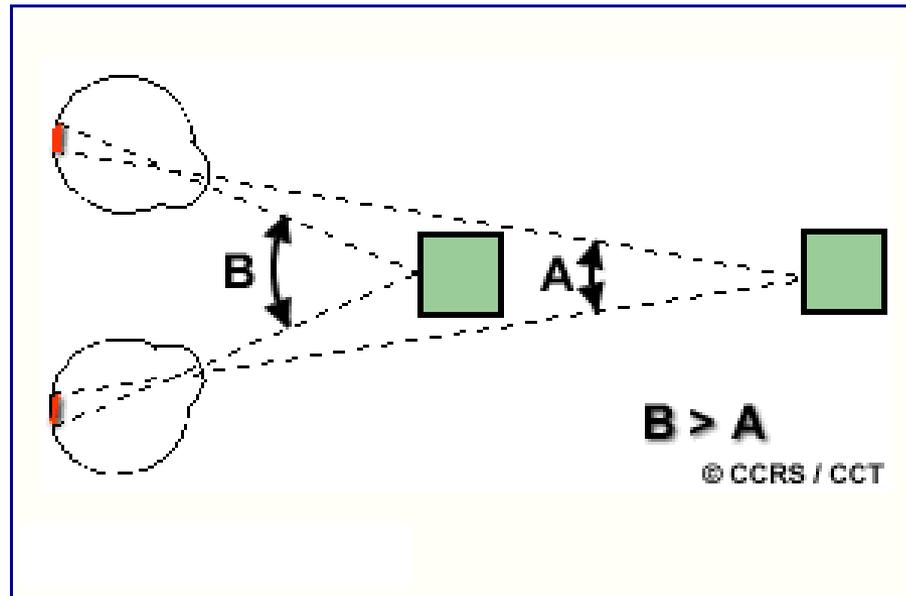
Percepção de Profundidade

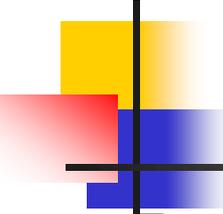
■ Disparidade Binocular

- Essa é considerada a mais importante variável na percepção de distância em distâncias visuais médias
- A paralaxe binocular é a diferença entre as imagens de um mesmo objeto projetadas na retina
- O grau da disparidade entre as duas imagens depende do ângulo de convergência (ângulo paralático – perpendicular ao horizonte)
 - Esse é o ângulo formado pelos eixos ópticos de cada olho convergindo em um objeto
 - O ângulo paralático está relacionado à distância do objeto aos olhos
 - Em grandes distâncias, o ângulo paralático diminui e a percepção de profundidade se torna mais difícil
 - O menor o ângulo paralático que uma pessoa é capaz de discernir é 3 arcos de segundos

Percepção de Profundidade

- Disparidade Binocular

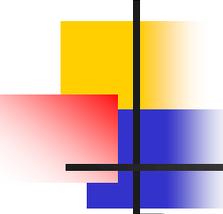




Percepção de Profundidade

■ Paralaxe de Movimento Visual

- É uma função da taxa a qual a imagem de um objeto se move pela retina
- Objetos distantes vão aparecer mais devagar em comparação a objetos mais próximos mesmo que os dois se movam à mesma velocidade
- A paralaxe de movimento também pode ser causada pelo movimento da cabeça do observador
- Objetos mais próximos do observador vão aparentar serem mais rápidos que aqueles mais distantes
- Essa é uma importante variável para aqueles que só enxergam por um olho

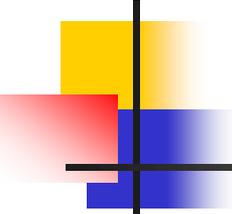


Percepção de Profundidade

■ Paralaxe de Movimento Visual

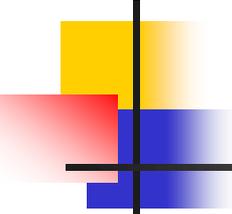
■ Experimento simples:

- Coloque seus dois dedos indicadores em linha, um afastado 15 cm de seus olhos e o outro 30 cm
- Feche seu olho esquerdo e observe a cena apenas com o direito
- Fixe o olhar no dedo mais afastado e movimente sua cabeça para os lados
 - Observe que o dedo mais próximo aparenta se mover em um movimento contrário ao da sua cabeça
- Fixe o olhar no dedo mais próximo e movimente sua cabeça para os lados
 - Observe que o dedo mais afastado aparenta se mover na mesma direção da sua cabeça (ou aparenta estar parado)



Percepção de Profundidade

- Variáveis Psicológicas
 - Tamanho da imagem na retina
 - Perspectiva linear
 - Perspectiva aérea
 - Oclusão
 - Sombras e sombreamento
 - Gradiente de texturas



Percepção de Profundidade

- Variáveis psicológicas são variáveis aprendidas, além do mais, elas são auxiliadas pela experiência
- Quando combinadas, essas variáveis melhoram bastante a percepção de profundidade

Percepção de Profundidade

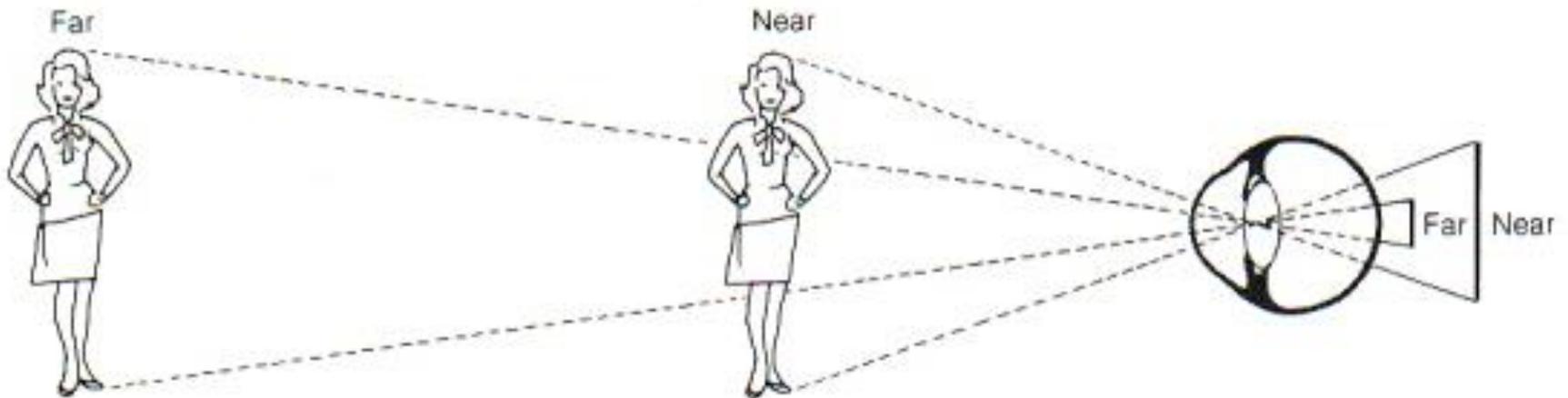
■ Variáveis Psicológicas

- Tamanho da imagem na retina: quanto maior a imagem de um objeto mais próximo ele parece



Percepção de Profundidade

- Variáveis Psicológicas
 - Tamanho da imagem na retina



Percepção de Profundidade

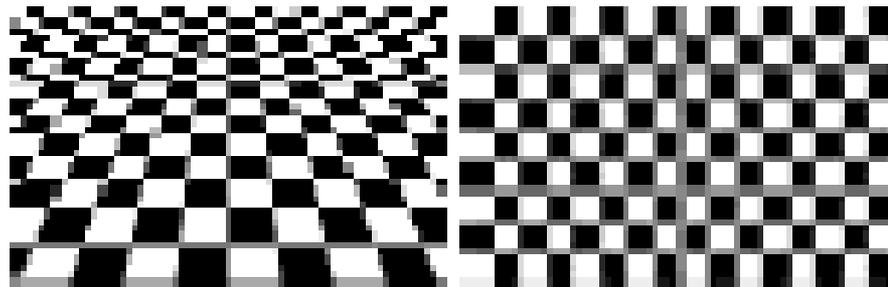
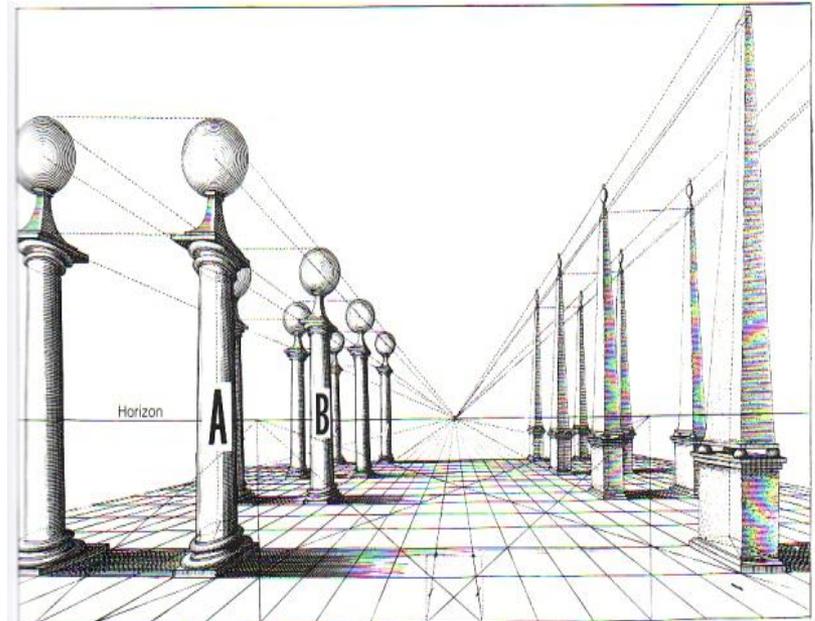
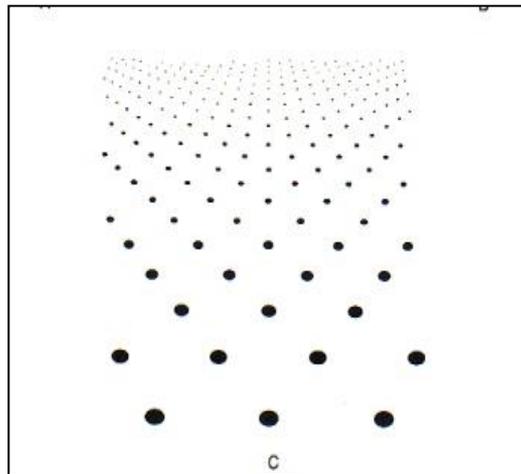
■ Variáveis Psicológicas

- Perspectiva linear: redução gradual do tamanho da imagem à medida que a distância ao objeto aumenta



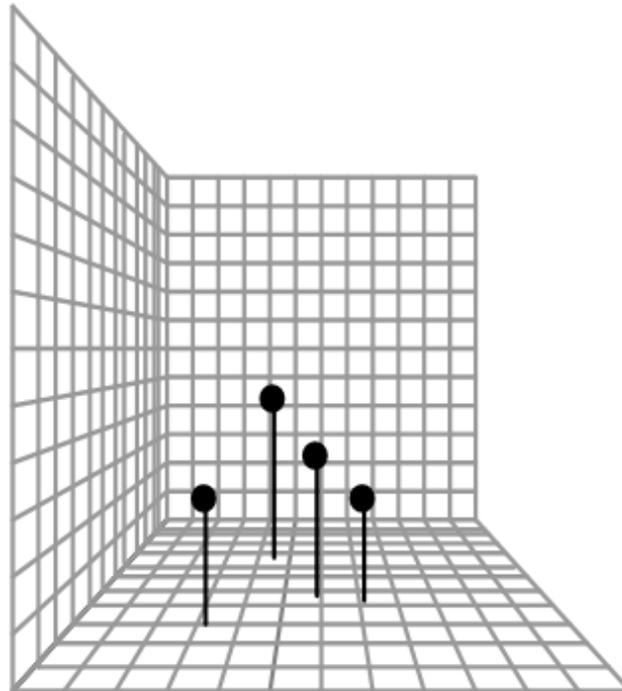
Percepção de Profundidade

- Variáveis Psicológicas
 - Perspectiva linear



Percepção de Profundidade

- Variáveis Psicológicas
 - Perspectiva linear



Percepção de Profundidade

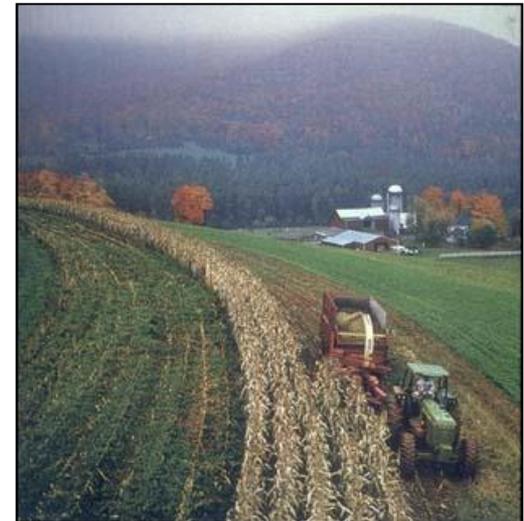
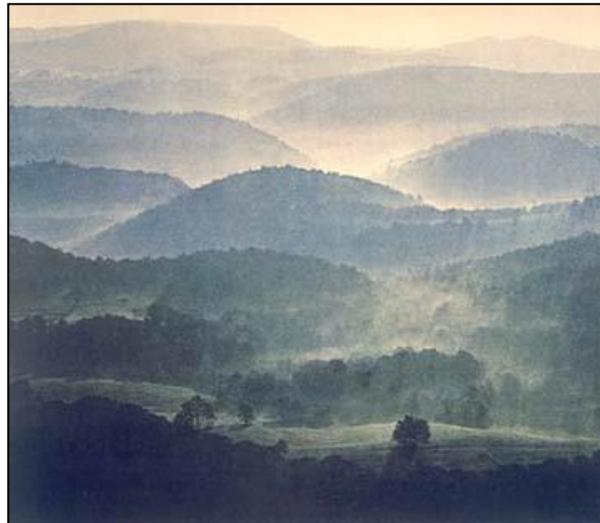
- Variáveis Psicológicas
 - Perspectiva linear



Percepção de Profundidade

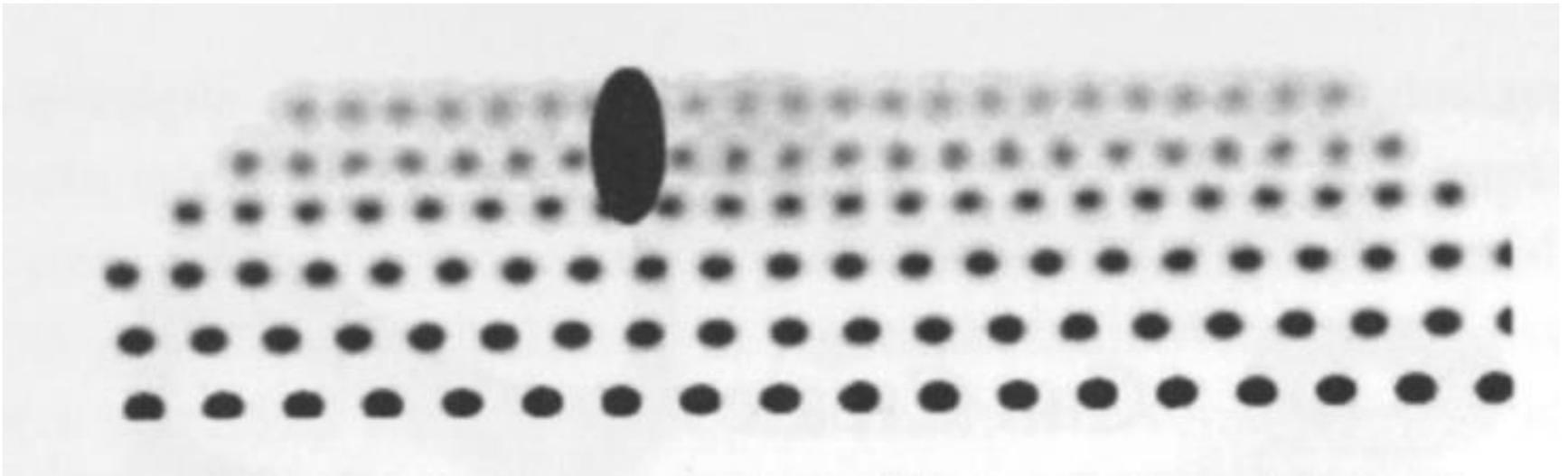
■ Variáveis Psicológicas

- Profundidade de foco: nebulosidade de objetos distantes



Percepção de Profundidade

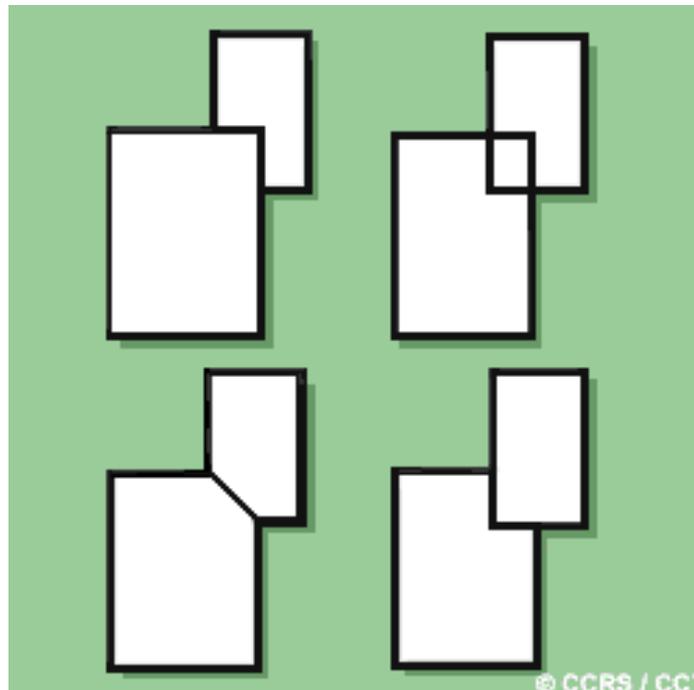
- Variáveis Psicológicas
 - Profundidade de Foco
 - Depende do ponto de fixação do usuário



Percepção de Profundidade

■ Variáveis Psicológicas

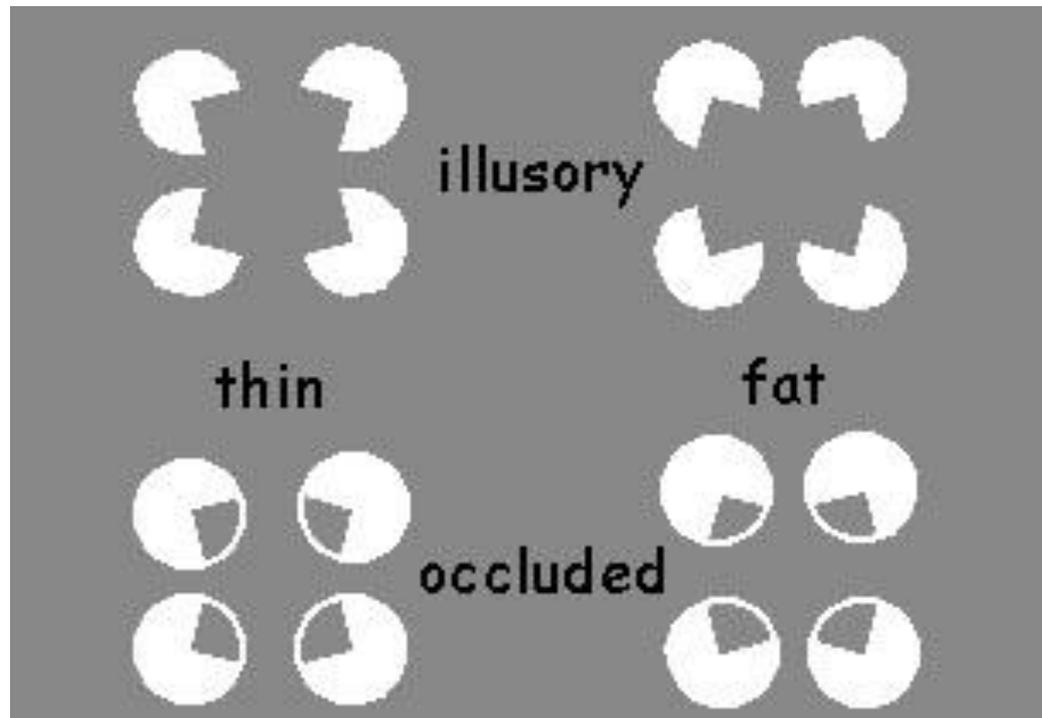
- Oclusão: o objeto que está mais próximo esconde parte de outro objeto mais distante



Percepção de Profundidade

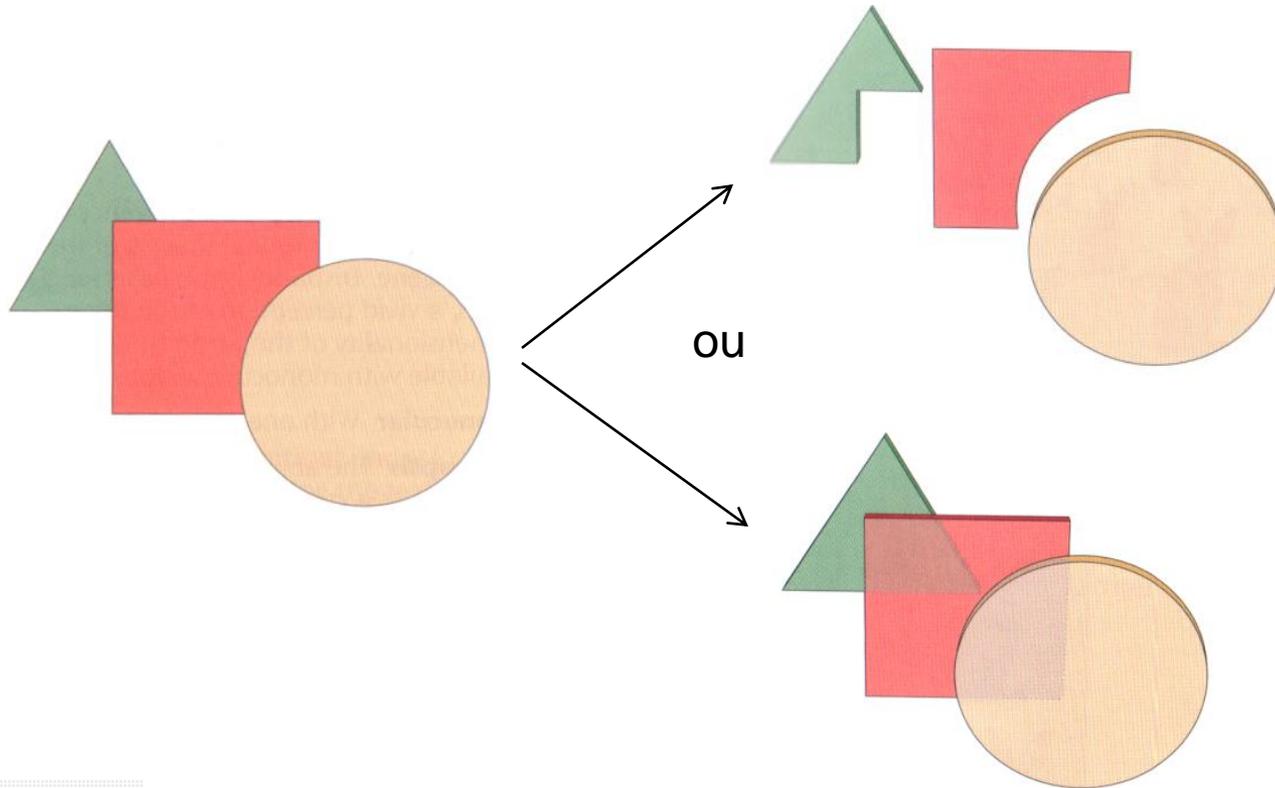
■ Variáveis Psicológicas

- Oclusão também é responsável pela ilusão de contornos



Percepção de Profundidade

- Variáveis Psicológicas
 - Problema: existe oclusão?



Percepção de Profundidade

- Variáveis Psicológicas

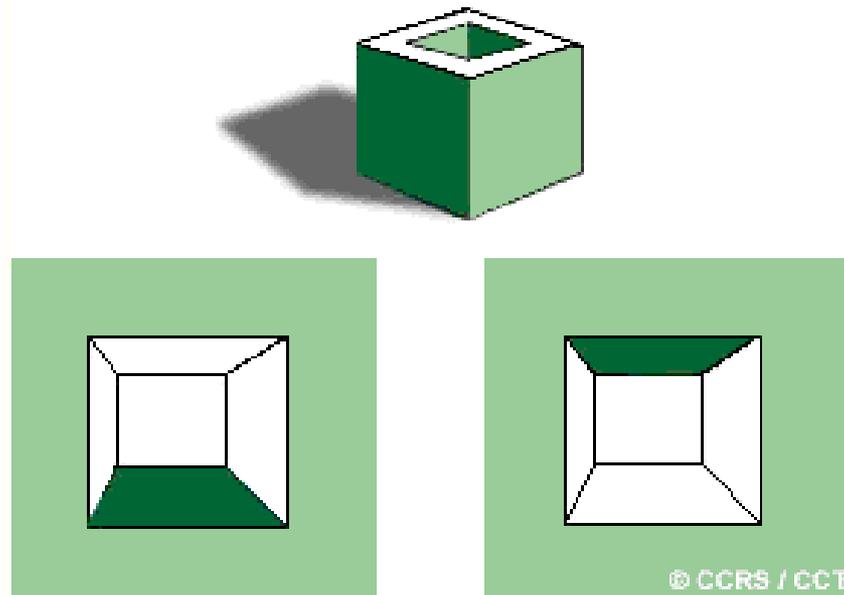
- Problema: Em alguns casos, assumimos que há



Percepção de Profundidade

■ Variáveis Psicológicas

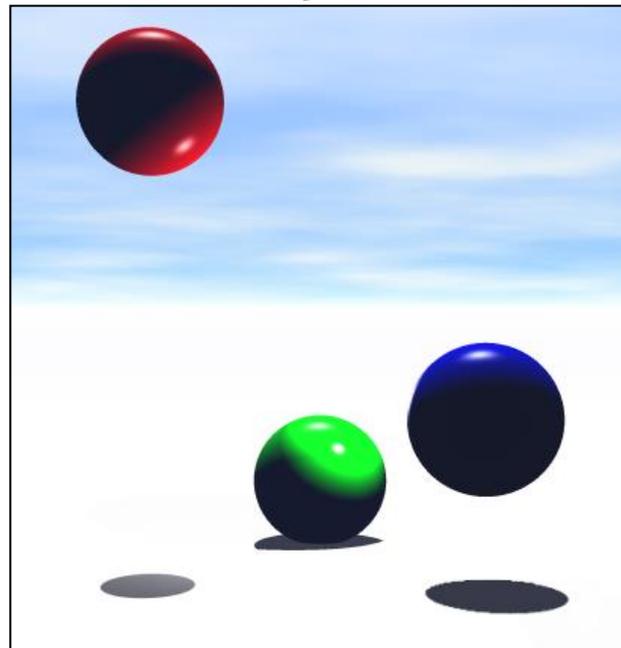
- Sombras e sombreamento: a impressão de convexidade ou concavidade baseado no fato que a maior parte da iluminação vem de cima



Percepção de Profundidade

■ Variáveis Psicológicas

- Sombras: Dá uma dica sobre a altura em que o objeto se encontra
- Forma indireta de percepção de profundidade



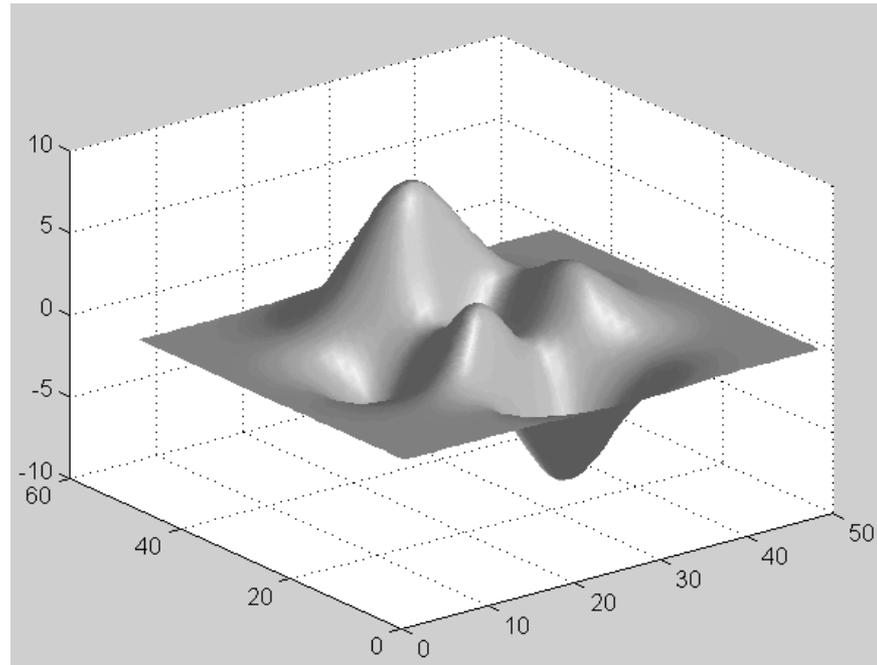
Percepção de Profundidade

- Variáveis Psicológicas
 - Sombras: As sombras sobre as montanhas e árvores dá uma noção de profundidade



Percepção de Profundidade

- Variáveis Psicológicas
 - Forma a partir de sombras (*shape from shading*)



Percepção de Profundidade

■ Variáveis Psicológicas

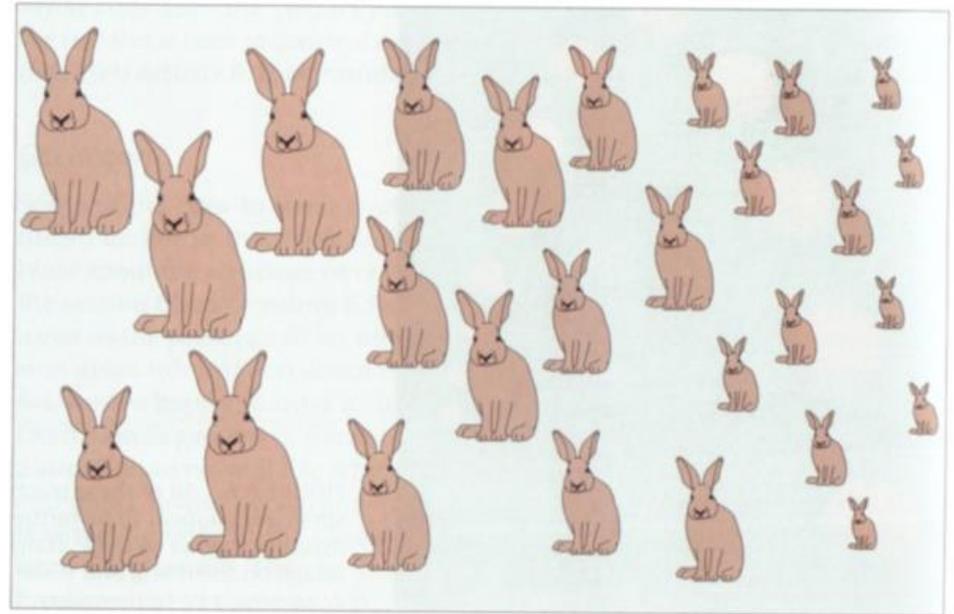
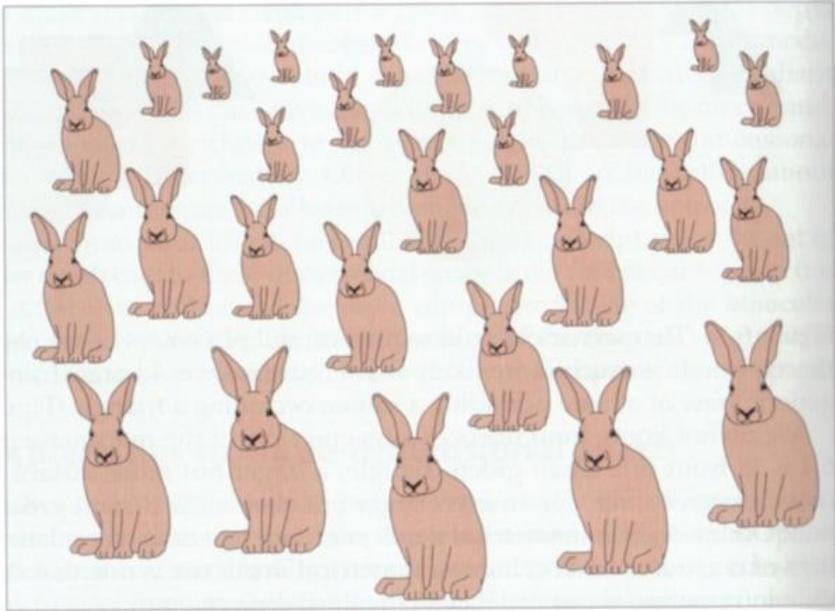
- Gradiente de texturas: a quantidade de detalhes de uma cena diminui à medida que a distância aumenta
- O *foreground* é mais detalhado que o *background*



Percepção de Profundidade

- Variáveis Psicológicas

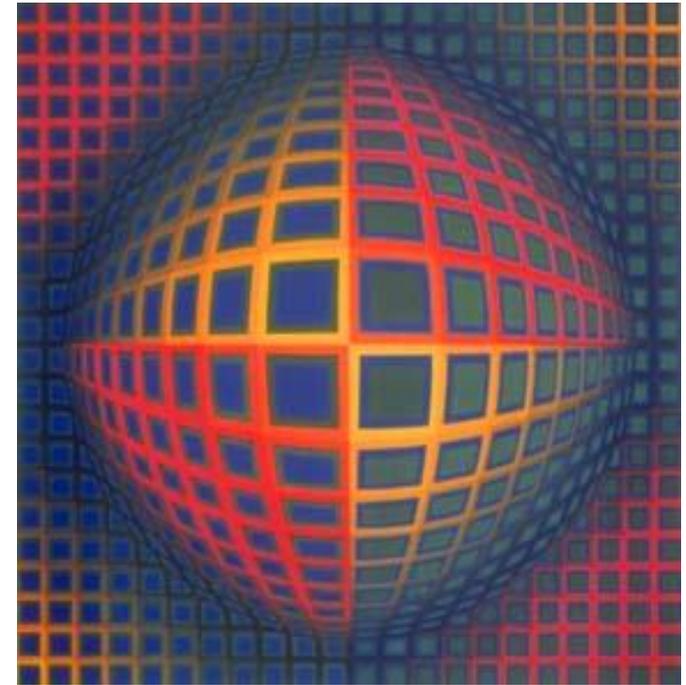
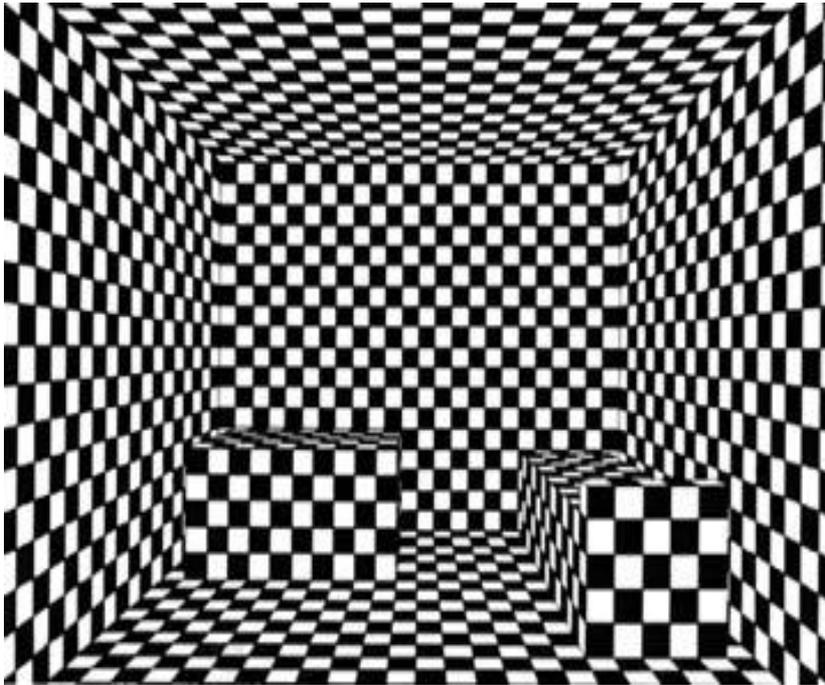
- Gradiente de texturas de coelho:

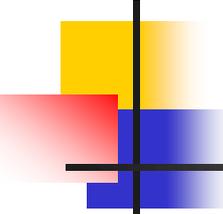


Qual dessas imagens dá melhor noção de profundidade?

Percepção de Profundidade

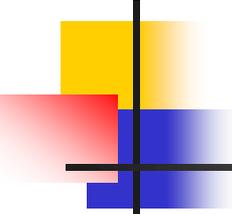
- Variáveis Psicológicas
 - Gradiente de texturas





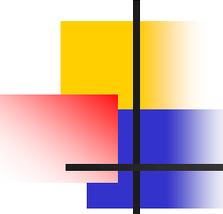
Constância Perceptuais

- Princípios de percepção aprendidos com a experiência
- Isso significa que quando objetos mudam em forma, tamanho ou brilho, nós ainda os vemos *como eles são* – algo ‘constante’ e **familiar**
- Nossa **percepção** estabelece que mesmo que as imagens na retina estejam alteradas, os objetos ainda são familiares



Constância Perceptuais

- Tamanho
- Brilho
- Forma
- Orientação



Constância de Tamanho

- Mantemos o tamanho percebido de um objeto mesmo que o tamanho da imagem na retina se altere devido à distância do objeto
- Se olharmos um caminhão fora de nossa casa, sabemos que o caminhão não é pequeno e, à medida que ele se afasta, sabemos que ele não está encolhendo
 - Isso é óbvio para adultos, mas crianças também percebem isso....

Constância de Tamanho

- + Perspectiva linear

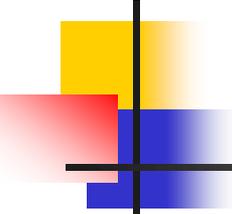
Qual a cena mais natural?



Constância de Tamanho

- Qual o maior balão?



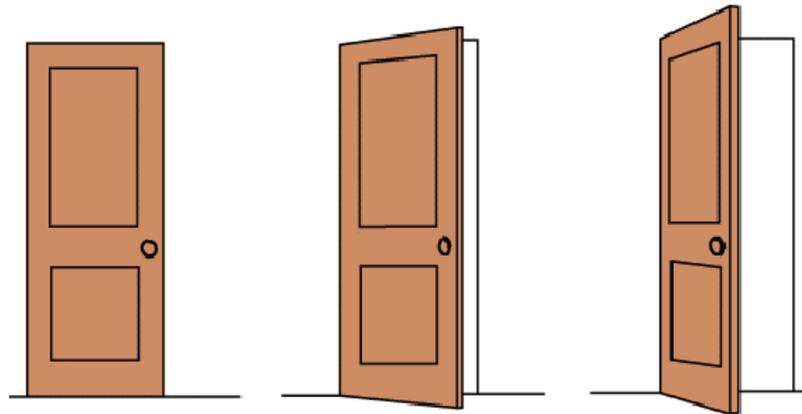


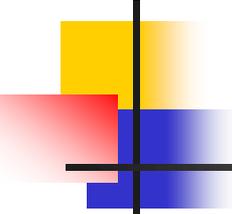
Constância de Brilho

- Mantemos o brilho percebido de um objeto apesar das mudanças na quantidade de luz recebida pela retina
 - Por exemplo, quando você vê uma camisa branca num quarto escuro, ela continua sendo percebida como branca embora seja vista como cinza

Constância de Forma

- A percepção da forma de um objeto permanece mesmo que o objeto seja observado por um ângulo diferente
 - Uma porta fechada, meio aberta ou aberta ainda é percebida como uma porta com forma retangular



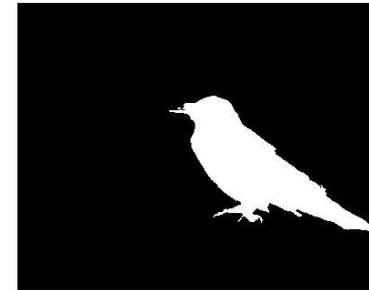
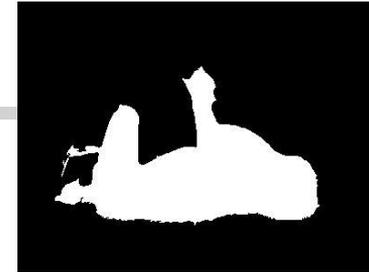


Constância de Orientação

- Essa é nossa tendência de manter a localização de um objeto no espaço constante – e percebemos o mundo como **de baixo para cima (vertical)**

Atenção Visual

Detecção de Regiões Salientes



Rafael Galvão de Mesquita

rgm@cin.ufpe.br

Carlos Alexandre Mello

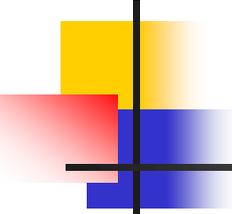
cabm@cin.ufpe.br

Carlos Alexandre Mello – cabm@cin.ufpe.br

Detecção de regiões salientes

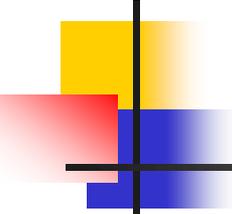
- Aplicações
- Recuperação de imagens baseada em conteúdo





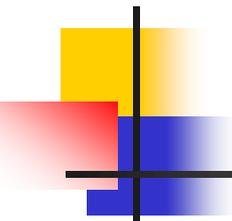
Detecção de regiões salientes

- Aplicações
- Segmentação
 - Questão subjetiva!
 - O que segmentar?
 - Possível abordagem
 - Segmentação baseada nas regiões mais “importantes”



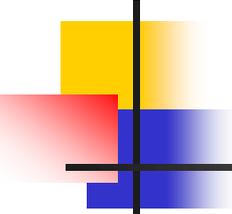
Detecção de regiões salientes

- Aplicações
- Compressão de imagens
 - Concentrar maior perda de informação em regiões menos importantes
 - Menor perda em regiões mais importantes



Detecção de regiões salientes

- Aplicações
- Busca visual/Reconhecimento de objetos
 - Priorização de locais da cena eleitos como mais importantes
 - Controle do nível de detalhe com que cada região deve ser processada
 - Áreas classificadas como background devem ser descartadas, ou processadas com um menor nível de detalhe
 - Inspiração no sistema visual humano
 - Atenção visual



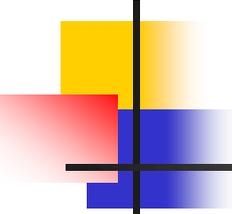
Atenção Visual

- Recebemos diversos estímulos em todos os sentidos o tempo inteiro
- É preciso selecionar o que nos interessa
 - Atenção Seletiva
 - Conceito que existe para cada um dos sentidos
 - Por exemplo, em uma sala cheia de pessoas falando, podemos nos concentrar na voz da pessoa que estiver falando diretamente conosco

Atenção Visual

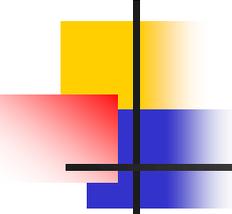
- Comparável com um foco de luz em um quarto escuro
 - Movendo o foco de luz pelo quarto, podemos ter uma impressão do conteúdo da cena da mesma forma que obtemos detalhes de uma cena quando movemos os olhos por ela
 - Dada a quantidade de estímulos que recebemos, nosso cérebro precisa priorizar alguns deles





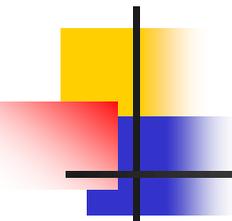
Atenção Visual

- Problema semelhante acontece com sistemas de visão computacional
 - Dentre os milhares/milhões de pixels de uma cena, quais realmente interessam
- Assim, como os conceitos de atenção seletiva do ser humano podem ser adaptados para sistemas computacionais?
 - Aplicações, em geral, em sistemas real-time



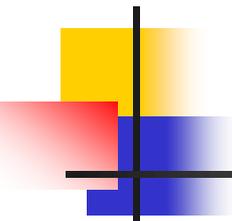
Atenção Visual

- O que é Atenção Visual?
 - Temos a impressão que retemos uma rica interpretação do nosso mundo visual e que mudanças no nosso ambiente vão atrair nossa atenção
 - Mas, de fato, apenas uma pequena parte da cena é analisada em detalhes a cada momento



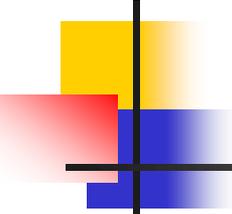
Atenção Visual

- O que é Atenção Visual?
 - A região que está em atenção que, não necessariamente, é a região na qual os olhos estão fixados
 - As outras regiões que estão fora de atenção são ignoradas
 - Efeito chamado de “change blindness”
 - Uma pessoa vendo uma cena não percebe (grandes) mudanças na cena
 - <https://www.youtube.com/watch?v=uO8wpm9HSB0>
 - <https://www.youtube.com/watch?v=0grANlx7y2E>



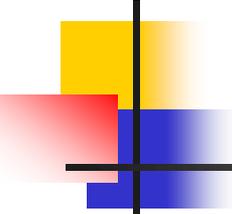
Atenção Visual

- As pessoas são capazes de automaticamente prestar atenção a regiões de interesse ao seu redor e varrer uma cena mudando rapidamente o foco de atenção
- A ordem na qual uma cena é investigada é determinada por mecanismos de atenção seletiva



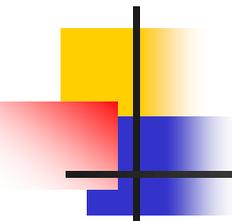
Atenção Visual

- Atenção define a habilidade mental em selecionar estímulos, respostas, memórias ou pensamentos que são relevantes entre muitos outros que são irrelevantes
 - Corbetta, 1990
- Atenção pode significar alerta, mas, para nós se refere a seletividade perceptual



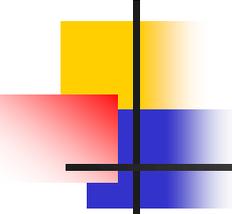
Covert vs Overt Attention

- Direcionar o foco de nossa atenção a alguma região de interesse é associado ao movimento dos olhos
 - *Overt Attention*
- Também somos capazes de ter atenção a localizações periféricas de interesse sem mover nossos olhos
 - *Covert Attention*



Bottom-Up vs Top-Down Attention

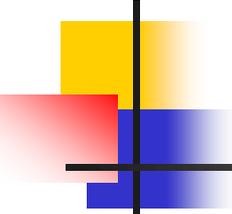
- Existem dois principais fatores que atraem a atenção:
 - Fatores *Bottom-Up*
 - Derivados apenas da cena visual
 - Regiões de interesse que atraem nossa atenção de uma forma *bottom-up* são chamadas *salientes*
 - Fatores *Top-Down*
 - Guiados por fatores cognitivos como conhecimento, expectativa, objetivos
 - Se você está procurando uma caneta vermelha em sua mesa, objetos vermelhos chamarão sua atenção



Visual Search and Pop-Up Effect

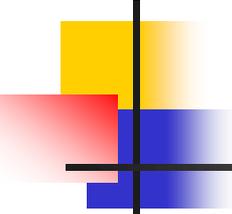
- *Visual Search*

- Questão principal: dado um alvo e uma imagem teste, existe alguma instância do alvo na imagem?
- Executamos *visual search* todo o tempo em nossa vida
 - *Unbounded visual search*
 - *Bounded visual search*



Visual Search and Pop-Up Effect

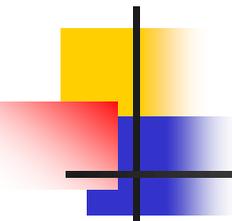
- *Unbounded Visual Search*
 - O alvo não é conhecido
 - Nenhuma otimização é permitida
 - Reconhecimento apenas *top-down*
 - *Data-direct vision* (Marr)
 - Problema NP-Completo
- *Bounded Visual Search*
 - O alvo é conhecido
 - Reconhecimento com conhecimento
 - Problema com complexidade linear



Visual Search and Pop-Up Effect

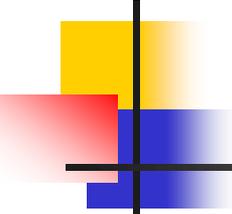
■ *Efficient search*

- Algumas propriedades particulares as quais formas de objetos são compostos aparecem no momento que os vemos, enquanto outras só aparecem após contemplação
- Efeito conhecido como *Pop-Up*
 - O *pop-up effect* ocorre apenas quando as distrações são homogêneas
 - A *busca* é menos eficiente quando o alvo é definido por diversas características
 - Características pré-atentivas também importam



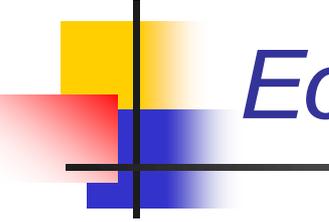
Eccentricity Effect

- A diagramação física da retina
 - Alta resolução no centro
 - Baixa resolução na periferia
 - Faz com que alvos em localizações periféricas sejam mais difíceis de detectar



Eccentricity Effect

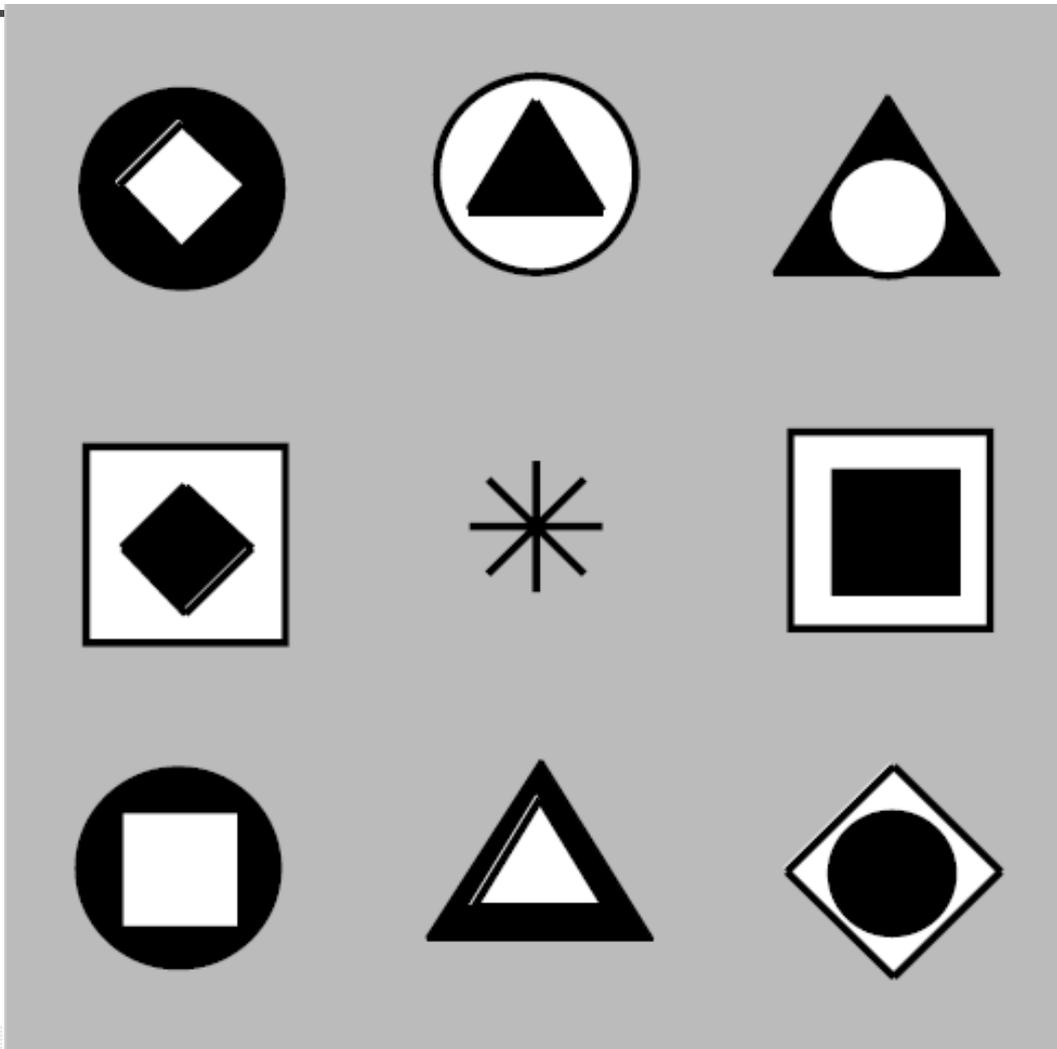
- Use a próxima imagem para se fixar no centro....
- Na figura seguinte, olhe para o ponto central apenas e procure localizar um círculo preto circunscrevendo um quadrado branco...

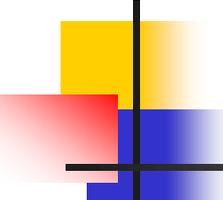


Eccentricity Effect



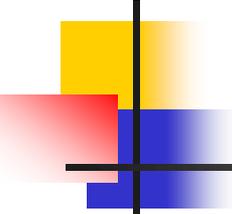
Eccentricity Effect





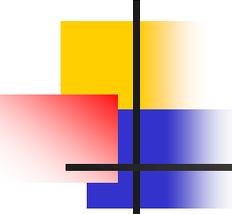
Atenção Visual

- Mecanismos Bottom-Up vs. Top-Down
 - Bottom-up
 - Derivado apenas de características da cena
 - Estímulos de baixo nível
 - Não há influência do observador
 - Processo exógeno
 - Top-down
 - Guiados por fatores cognitivos como conhecimento, expectativa ou objetivos
 - Detectado por estímulos de alto nível
 - Processo endógeno



Atenção Visual

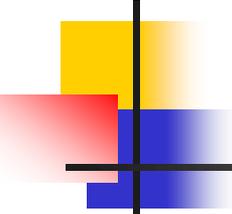
- **Processamento pré-atentivo**
 - Habilidade do sistema visual humano de baixo nível identificar rapidamente certas propriedades visuais básicas
 - Acontece antes que a atenção seja direcionada a algum objeto ou um ponto focal



Atenção Visual

- Processamento pré-atentivo
 - Procure um “3” aqui

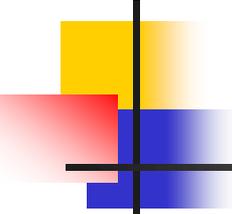
85689726984689762689764358922659865986554897689269898
02462996874026557627986789045679232769285460986772098
90834579802790759047098279085790847729087590827908754
98709856749068975786259845690243790472190790709811450
85689726984689762689764458922659865986554897689269898



Atenção Visual

- Processamento pré-atentivo
 - Não foi necessário pedir que buscassem pelo “3” ...

85689726984689762689764358922659865986554897689269898
02462996874026557627986789045679232769285460986772098
90834579802790759047098279085790847729087590827908754
98709856749068975786259845690243790472190790709811450
85689726984689762689764458922659865986554897689269898

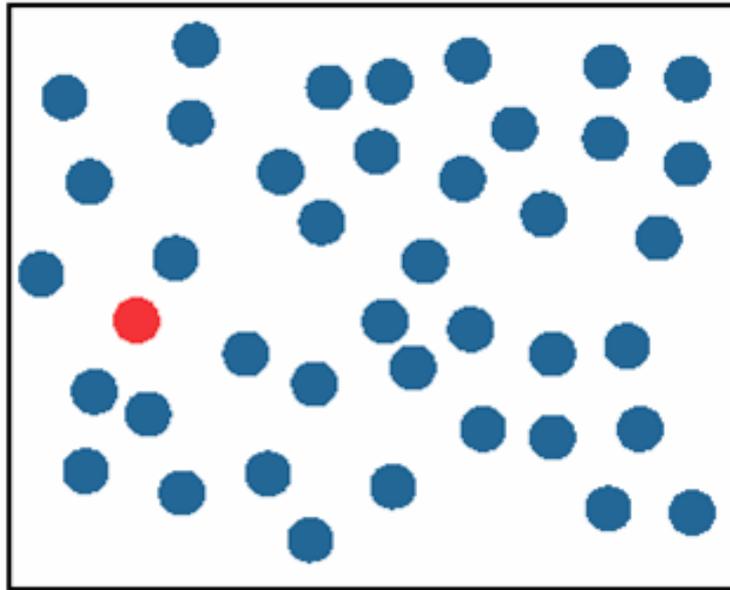


Atenção Visual

- **Processamento pré-atentivo**
 - Se o processamento for pré-atentivo, o tempo levado será o mesmo independente da quantidade de distrações
 - Em geral, leva 10 microssegundos
 - Se > 40 mseg = Não-pré-atentivo

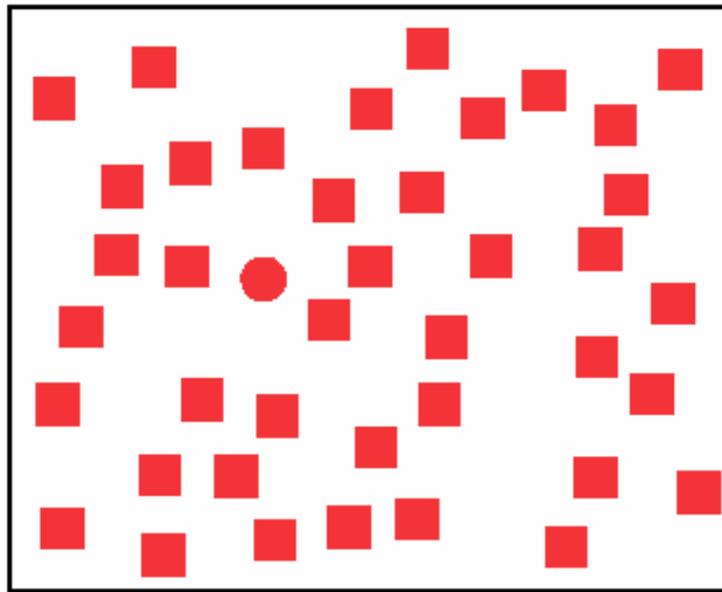
Atenção Visual

- Processamento pré-atentivo (Teste 1)
 - Encontre o círculo vermelho



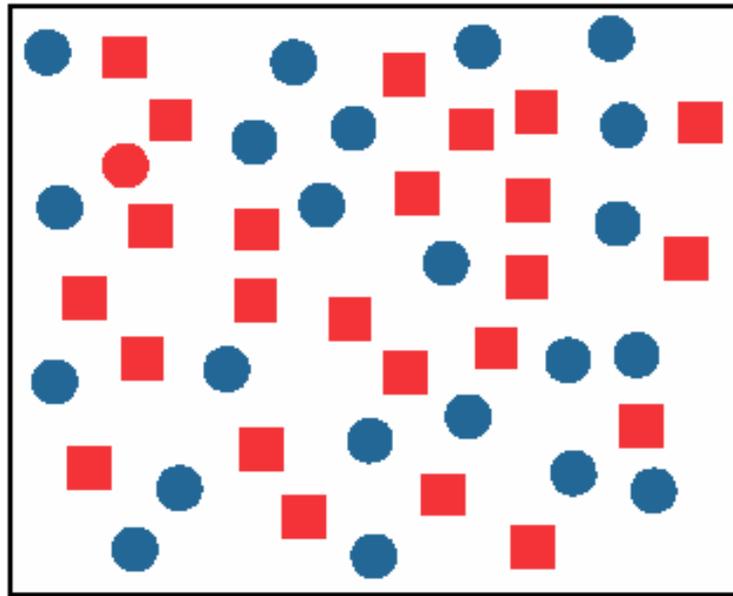
Atenção Visual

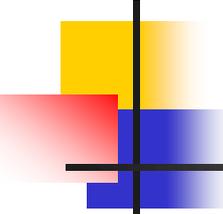
- Processamento pré-atentivo (Teste 2)
 - Encontre o círculo vermelho



Atenção Visual

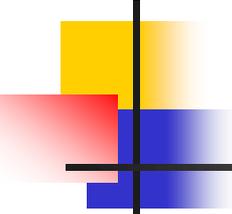
- Processamento pré-atentivo (Teste 3)
 - Encontre o círculo vermelho





Atenção Visual

- SVH vs. Sistemas Computacionais
 - Analogia com visão computacional: Dentre os milhares/milhões de pixels em uma imagem, quais realmente interessam?
 - Conceitos de atenção seletiva do ser humano podem ser adaptados para sistemas computacionais

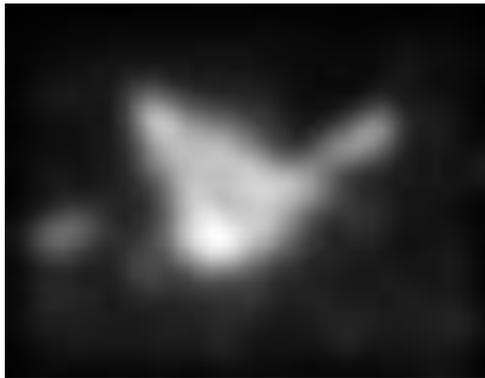


Detecção de regiões salientes

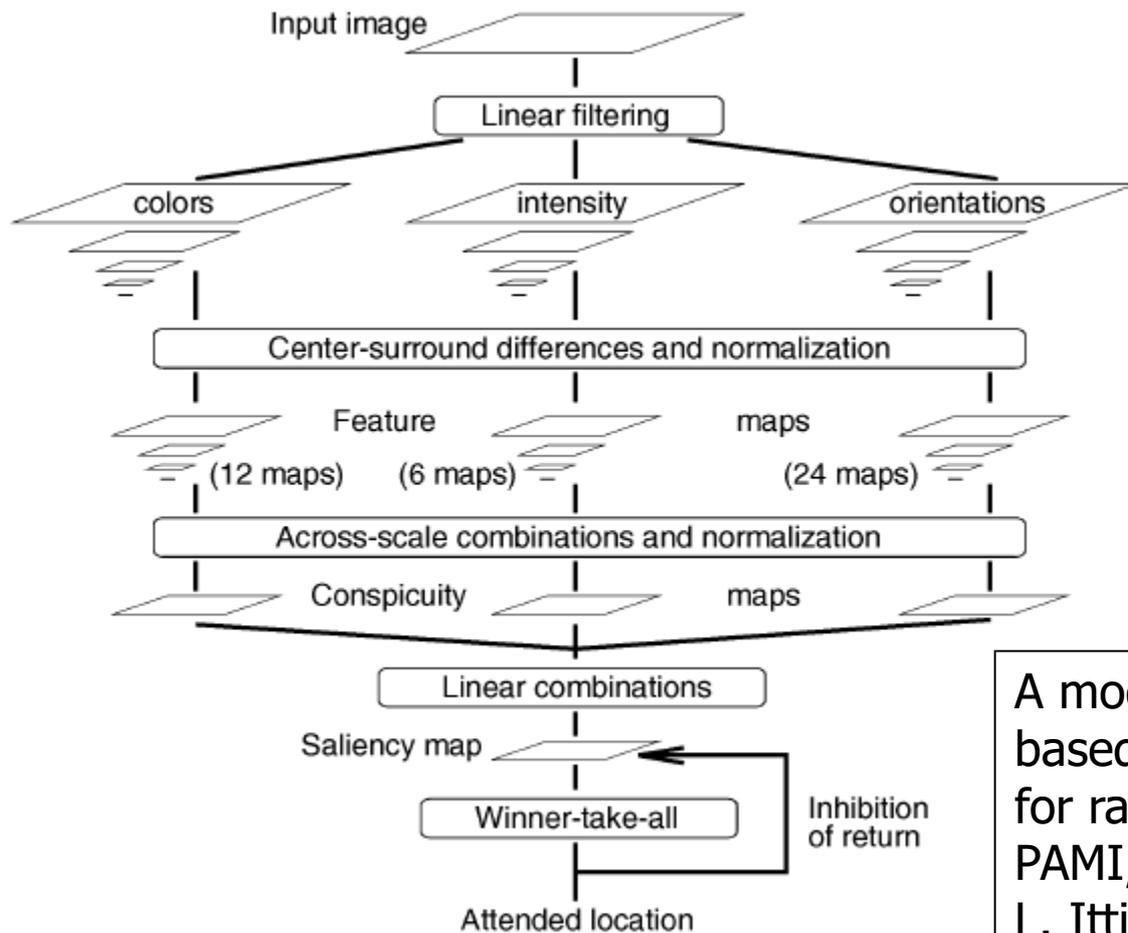
- Como as regiões salientes de uma imagem são representadas?
 - Mapas de saliência!
 - Imagem em tons de cinza
 - Quanto maior o valor de um determinado píxel, mais saliente ele é

Detecção de regiões salientes

- Mapas de saliência!

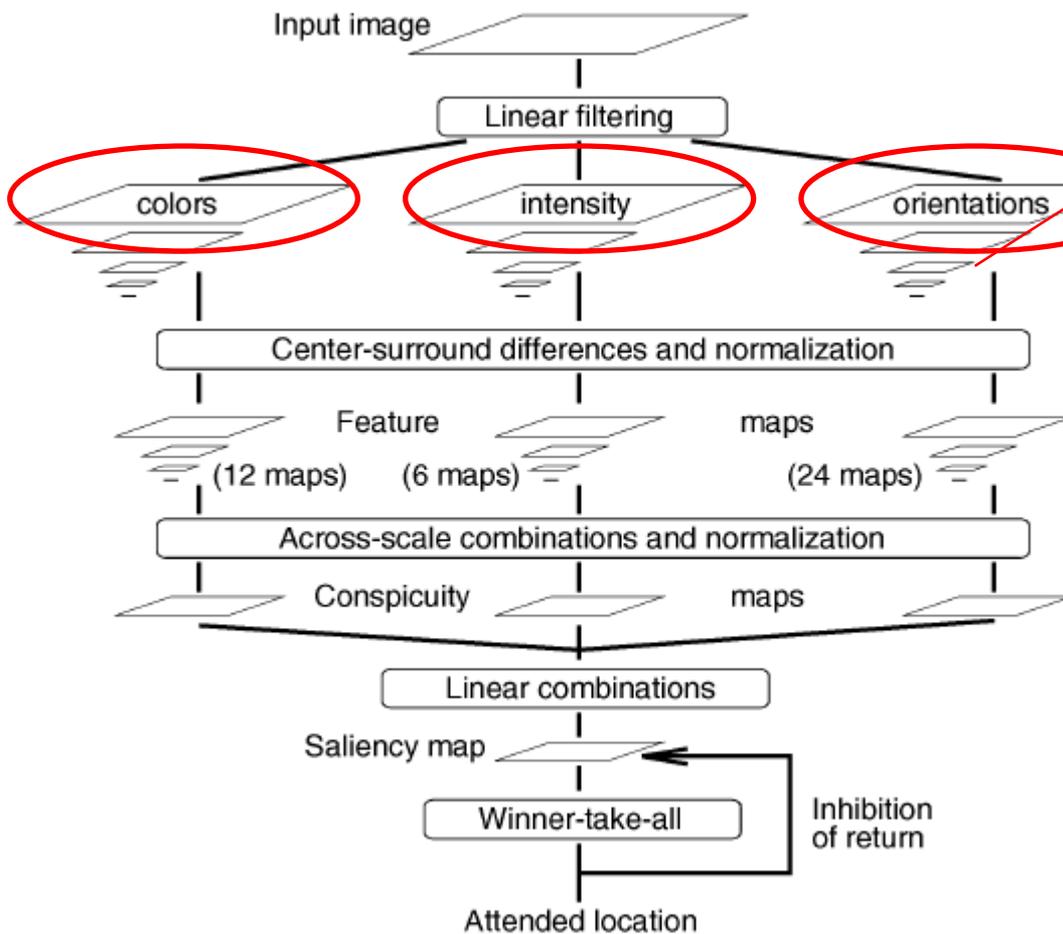


Itti & Koch

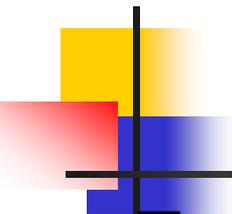


A model of saliency-based visual attention for rapid scene analysis, PAMI, 1998
L. Itti & C. KOCH

Itti & Koch



Pirâmide Gaussiana:
Sequência de
filtragens passa-baixa
seguida de
subamostragem.



Itti & Koch

■ Extração de características de baixo nível da cena

■ Intensidade

- $I = (r+g+b)/3$

■ Cor

- $R = r - (g+b)/2$

- $G = g - (r+b)/2$

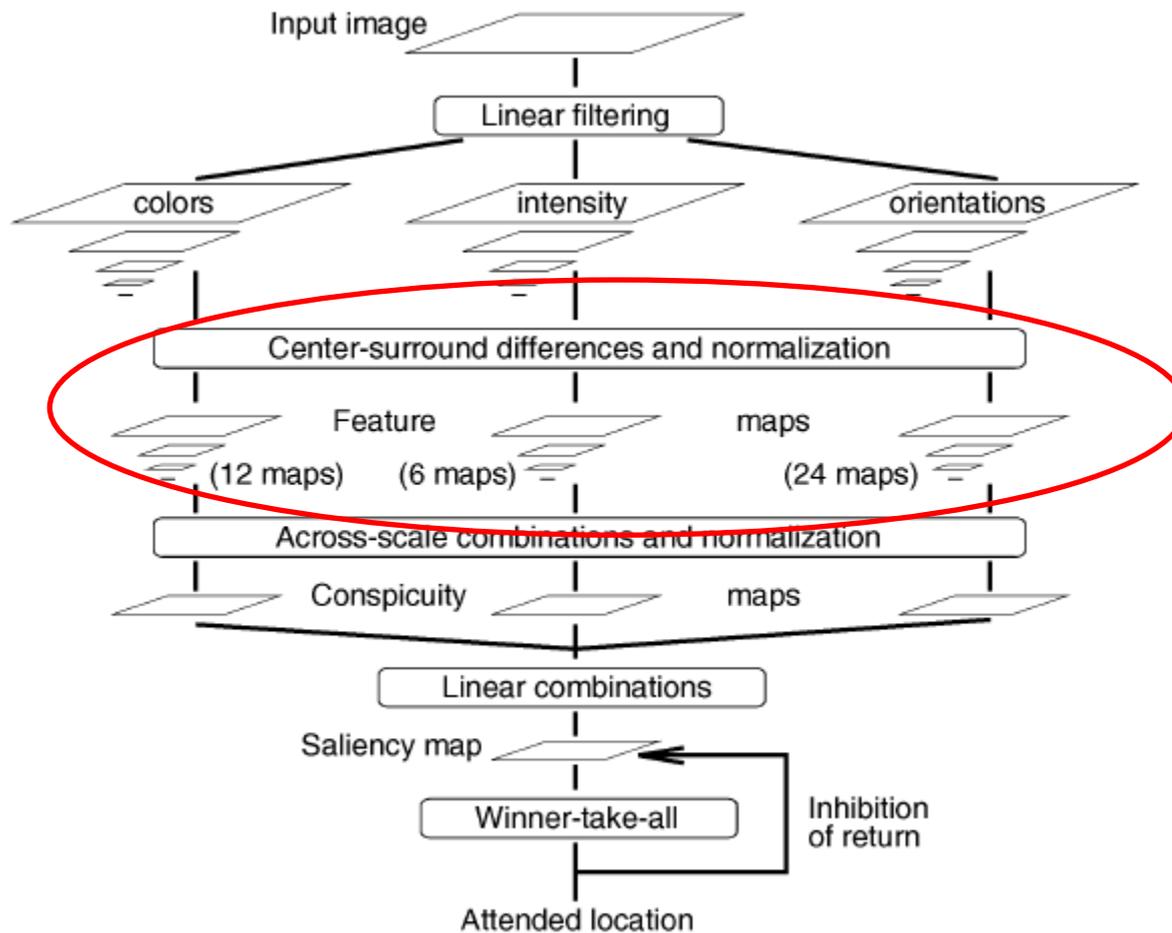
- $B = b - (r+g)/2$

- $Y = (r+g)/2 - |r-g|/2 - b$

■ Orientação

- Mais adiante....

Itti & Koch



Itti & Koch

- Obtenção mapas de características
 - Center-surround operation
 - Semelhante aos campos receptivos visuais
 - Detecção de regiões que se destaquem da sua vizinhança
 - Implementação: diferença entre escalas finas e grosseiras
 - Escala: uso de pirâmides gaussianas



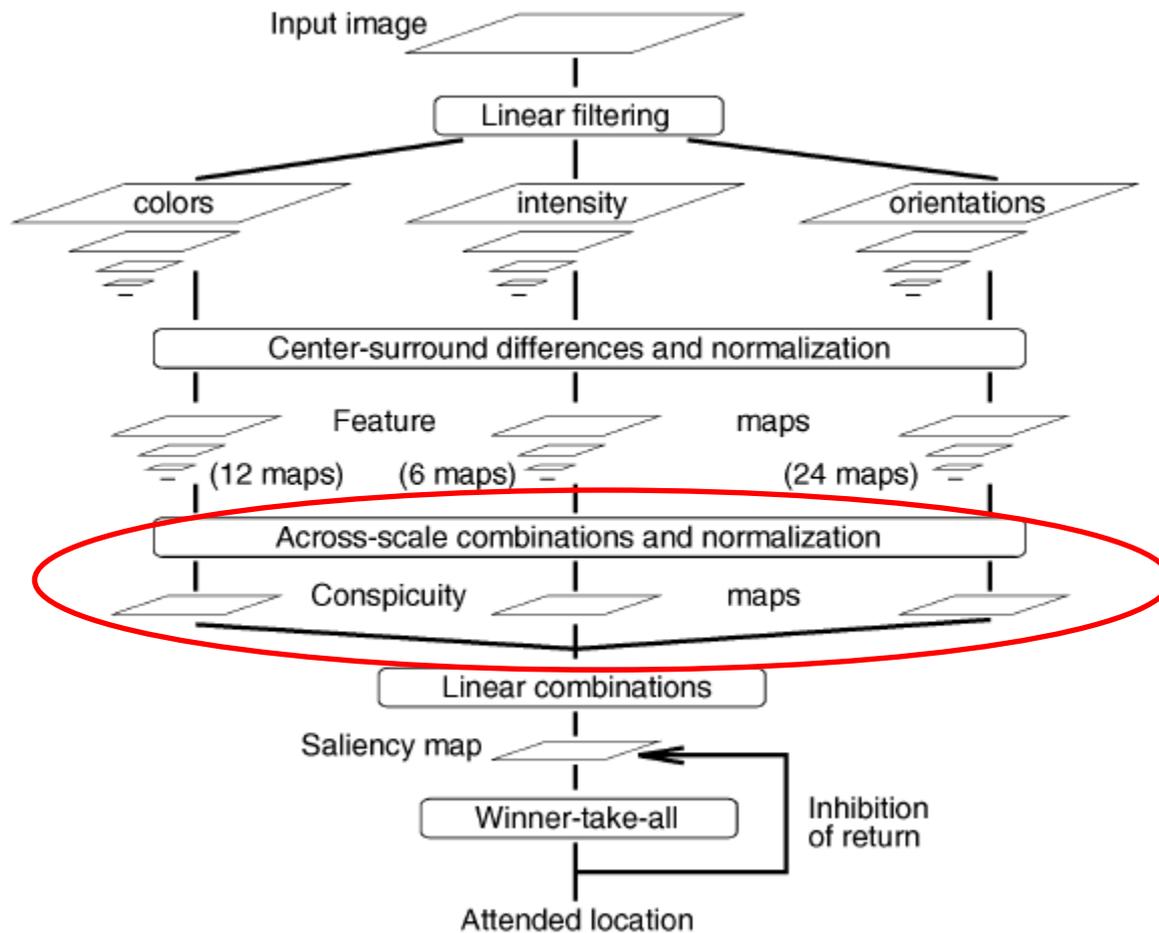
- Obtenção mapas de características
 - Center-surround operation
 - Diferença entre escalas θ
 - Interpolação para escala mais fina, e subtração pixel a pixel
 - Diferença entre escalas finas e grosseiras
 - Center: $c \in \{2,3,4\}$
 - Surround: $s = c + \delta, \delta \in \{3,4\}$
 - Intensidade
 - Seis escalas
 - $I(c, s) = |I(c)\theta I(s)|$

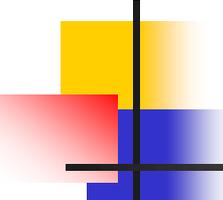
- Obtenção mapas de características
 - Center-surround operation
 - Cor
 - “Color double opponent system” (cortex)
 - Centro dos campos receptivos
 - Neuronio são excitados por uma cor (vermelho, por exemplo), e inibidos por outra cor (verde, por exemplo)
 - Na vizinhança dos campos receptivos o oposto ocorre
 - Oponencia vermelho-verde
 - $RG(c, s) = |(R(c) - G(c))\theta(G(s) - R(s))|$
 - Oponencia azul-amarelo
 - $BY(c, s) = |(B(c) - Y(c))\theta(Y(s) - B(s))|$

- Obtenção mapas de características
 - Center-surround operation
 - Orientação
 - Obtenção da informação de orientação
 - Pirâmides de Gabor $O(\sigma, \theta)$
 - Filtro de Gabor: filtros sensíveis à orientação [3][4]
 - Construída a partir de I
 - $\sigma \in [0 \dots 8], \theta \in \{0, 45, 90, 135\}$
 - Mapas de características de orientação
 - $O(c, s, \theta) = |O(c, \theta) \ominus O(s, \theta)|$

- Obtenção mapas de características
 - 42 mapas
 - Center: $c \in \{2,3,4\}$
 - Surround: $s = c + \delta, \delta \in \{3,4\}$
 - 6 para intensidade
 - 12 para cor
 - 24 para orientação

Itti & Koch





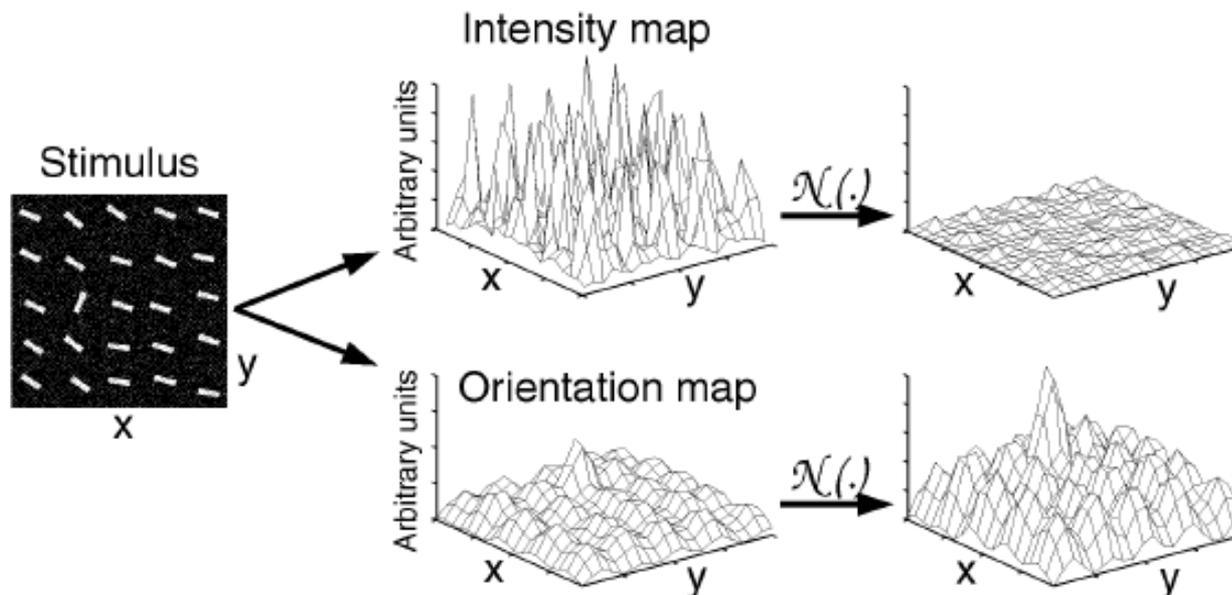
Itti & Koch

Mapas de conspicuidade

- Operador de normalização $N(.)$
 - Promove mapas em que um pequeno numero de picos de atividades (locais conspícuos) estejam presentes
 - Suprime mapas que contenham uma grande quantidade de picos comparáveis
- Implementação $N(.)$
 - Normaliza todos os valores no mapa
 - Busca ponto de máximos global (M) e a média de máximos local (m)
 - Multiplica o mapa por $(M - m)^2$
 - É possível mensurar o quão diferente o local mais ativo é em relação à média
 - Quando a diferença é grande, o local mais ativo se destaca e o mapa é promovido
 - Quando a diferença é pequena o mapa é suprimido

Itti & Koch

- Mapas de conspicuidade
 - Operador de normalização $N(\cdot)$



- Mapas de conspicuidade

- 3 mapas

- Intensidade, cor e orientação

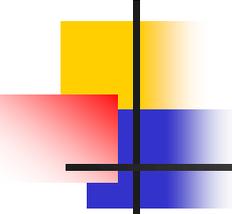
- Combinação dos mapas de características em mapas de conspicuidade

- Redução de cada mapa para a escala 4, seguido de adição pixe

$$\bar{I} = \bigoplus_{c=2}^4 \bigoplus_{s=c+3}^{c+4} \mathcal{N}(I(c, s))$$

$$\bar{C} = \bigoplus_{c=2}^4 \bigoplus_{s=c+3}^{c+4} \left[\mathcal{N}(\mathcal{R}\mathcal{G}(c, s)) + \mathcal{N}(\mathcal{B}\mathcal{Y}(c, s)) \right].$$

$$\bar{O} = \sum_{\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}} \mathcal{N}\left(\bigoplus_{c=2}^4 \bigoplus_{s=c+3}^{c+4} \mathcal{N}(O(c, s, \theta)) \right).$$



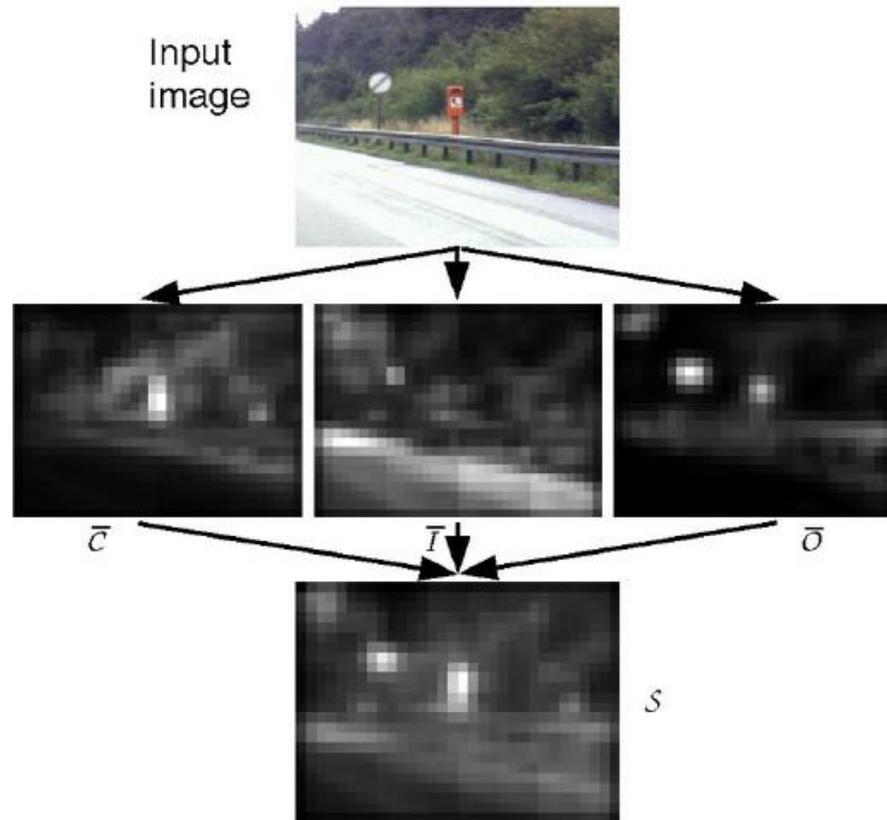
Itti & Koch

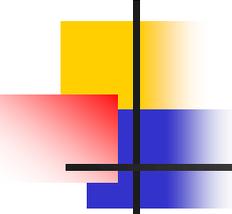
- Mapas de saliências
 - Mapas de conspicuidade são normalizados e somados novamente
 - Geração do mapa de saliências final

$$S = \frac{1}{3} (\mathcal{N}(\bar{I}) + \mathcal{N}(\bar{C}) + \mathcal{N}(\bar{O})).$$

Itti & Koch

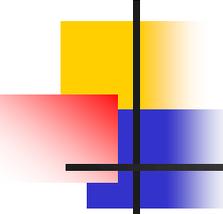
■ Mapas de Saliências





Itti & Koch

- Toolboxes para MATLAB:
 - <http://www.saliencytoolbox.net/download.html>
 - C++
 - <http://ilab.usc.edu/toolkit/>



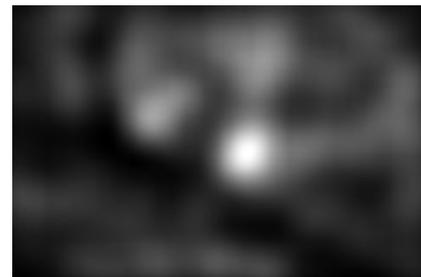
Outros métodos

- CA, COV, FES, GR, MC, PCA, SEG, SeR, SIM, SR
 - Disponíveis em:
<http://mmcheng.net/salobjbenchmark/>

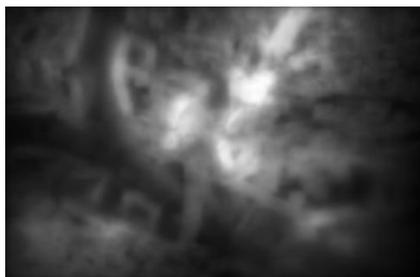
Outros métodos - Exemplo



SIM



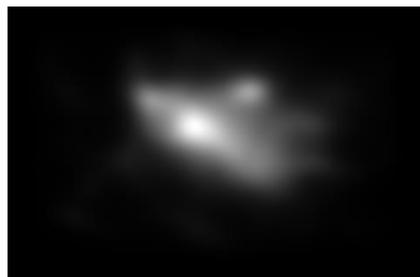
SR



CA



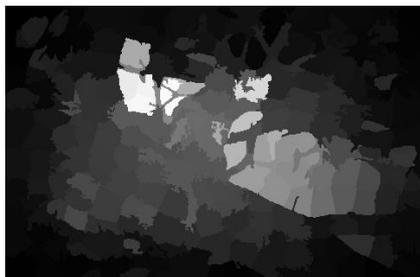
COV



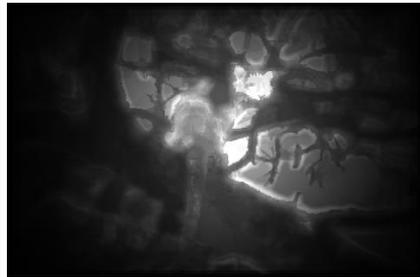
FES



GR



MC



PCA



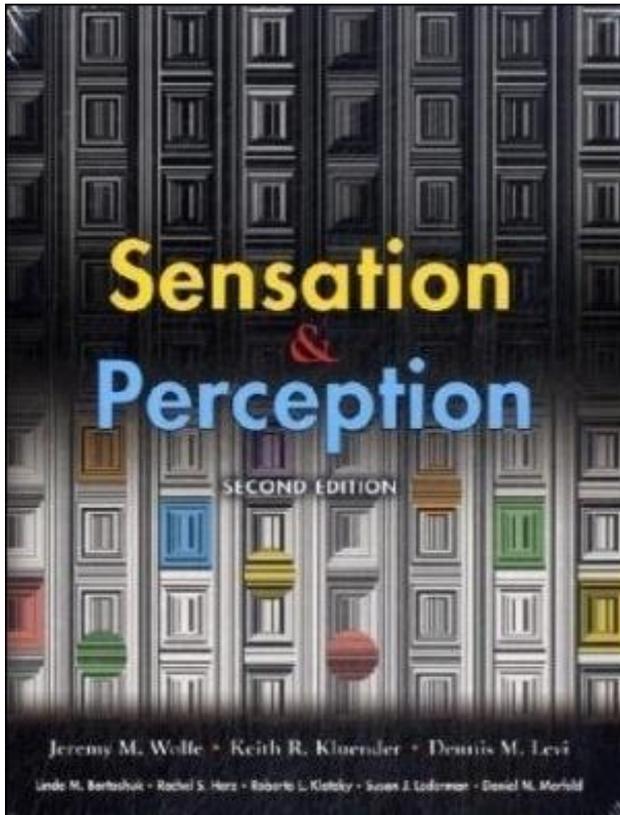
SEG



SeR

Percepção

■ Referência



Sensation and Perception,
Jeremy Wolfe, Keith Kluender, Dennis Levi
Sinauer Associates, 2009