

Processamento de Imagens

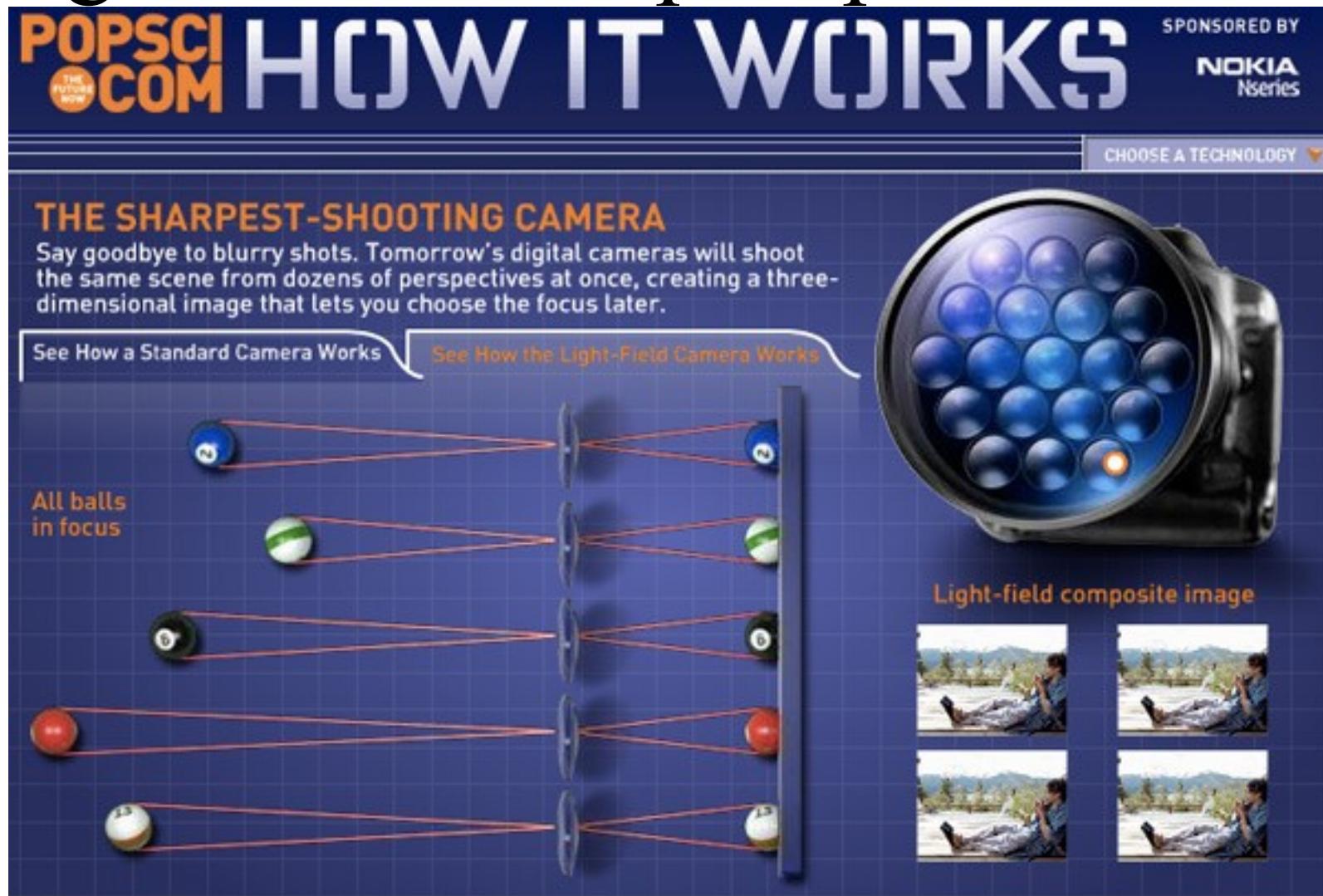
André Tavares da Silva

andre.silva@udesc.br

Roteiro

- O que é imagem
- Operações com imagens
- Convolução discreta
- Transformada de Fourier
- Transformada de Wavelets
- Morfologia matemática
- Extração de Características

Imagen com múltiplos planos de foco



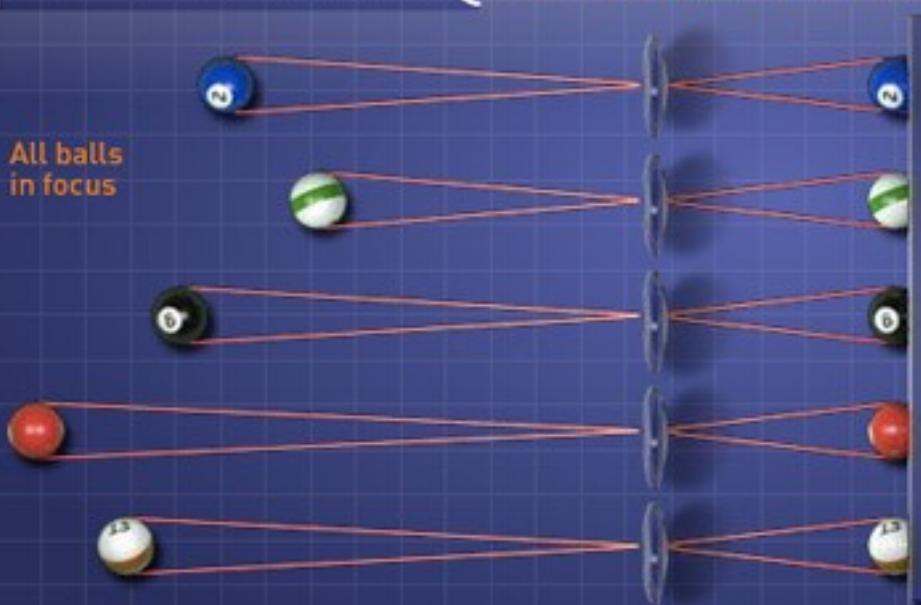
POPSCI.COM HOW IT WORKS SPONSORED BY NOKIA Nseries CHOOSE A TECHNOLOGY ▾

THE SHARPEST-SHOOTING CAMERA

Say goodbye to blurry shots. Tomorrow's digital cameras will shoot the same scene from dozens of perspectives at once, creating a three-dimensional image that lets you choose the focus later.

See How a Standard Camera Works See How the Light-Field Camera Works

All balls in focus



Light-field composite image



Imagen com múltiplos planos de foco



Imagen com múltiplos planos de foco



Imagen com múltiplos planos de foco (3^a geração da Adobe)



Imagen com múltiplos planos de foco



Patentes para geração de imagem

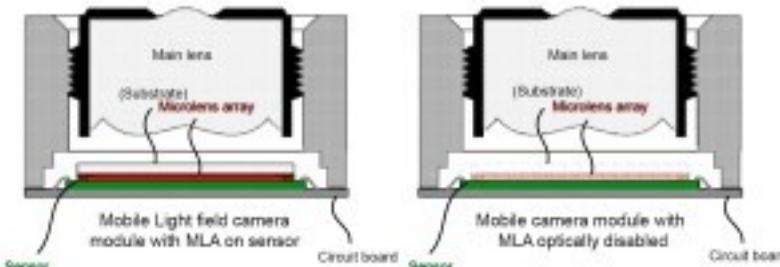
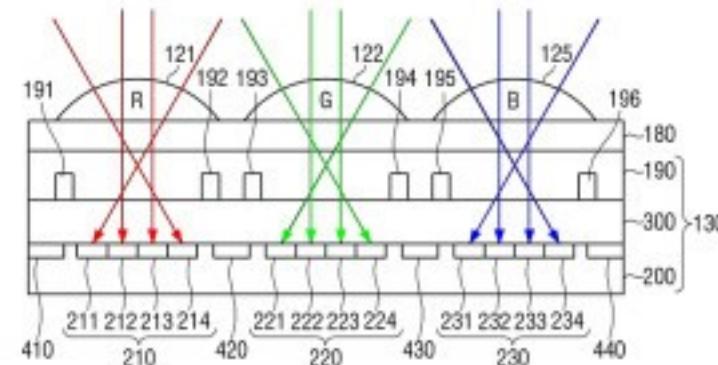


FIG. 4

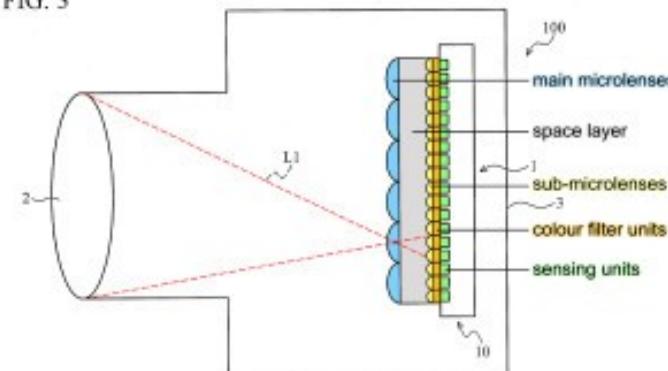
Bhat et al. 2014, modified by www.lightfield-forum.com

Lytro Patents Hybrid Camera for Conventional and Light Field Photography (2015)



Samsung Patents Modified Light Field Sensor with Monochrome Sub-Images (2014)

FIG. 3



Sony Patents Light Field Sensor with Full-Resolution 3D Stereo Output (2014)

Visera Patents Integrated Light Field Sensor Manufacturing Process (2014)

Imagen de microscopia óptica

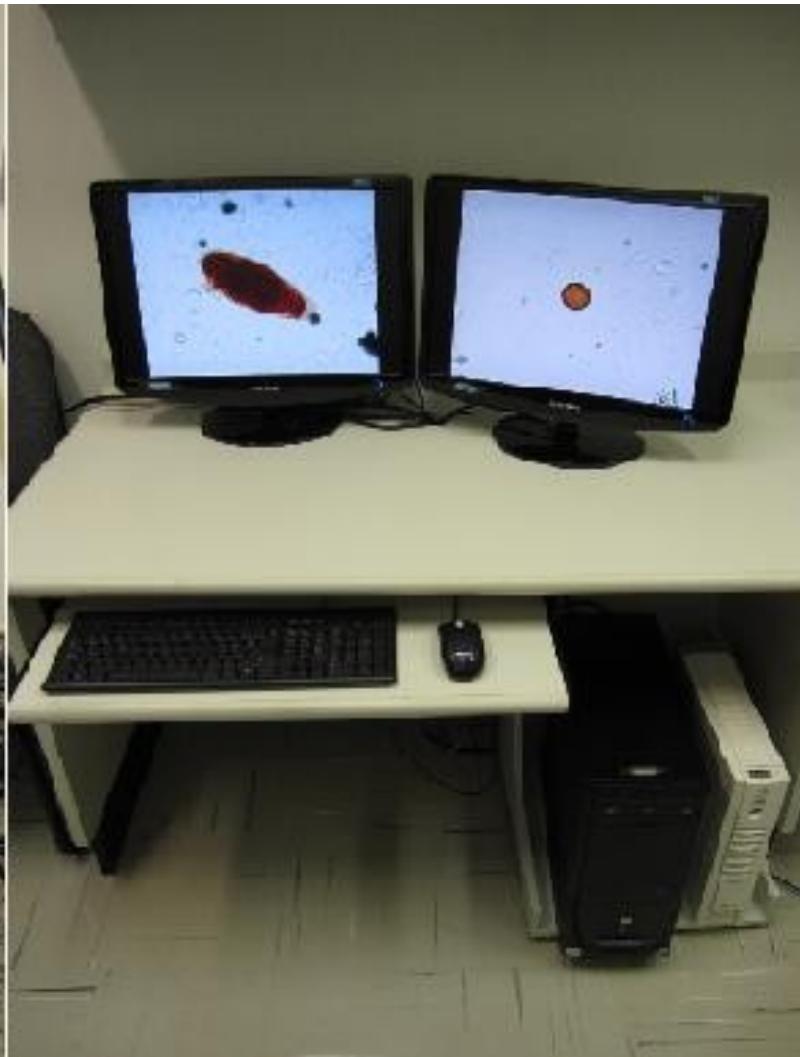


Imagen de sensoriamento remoto

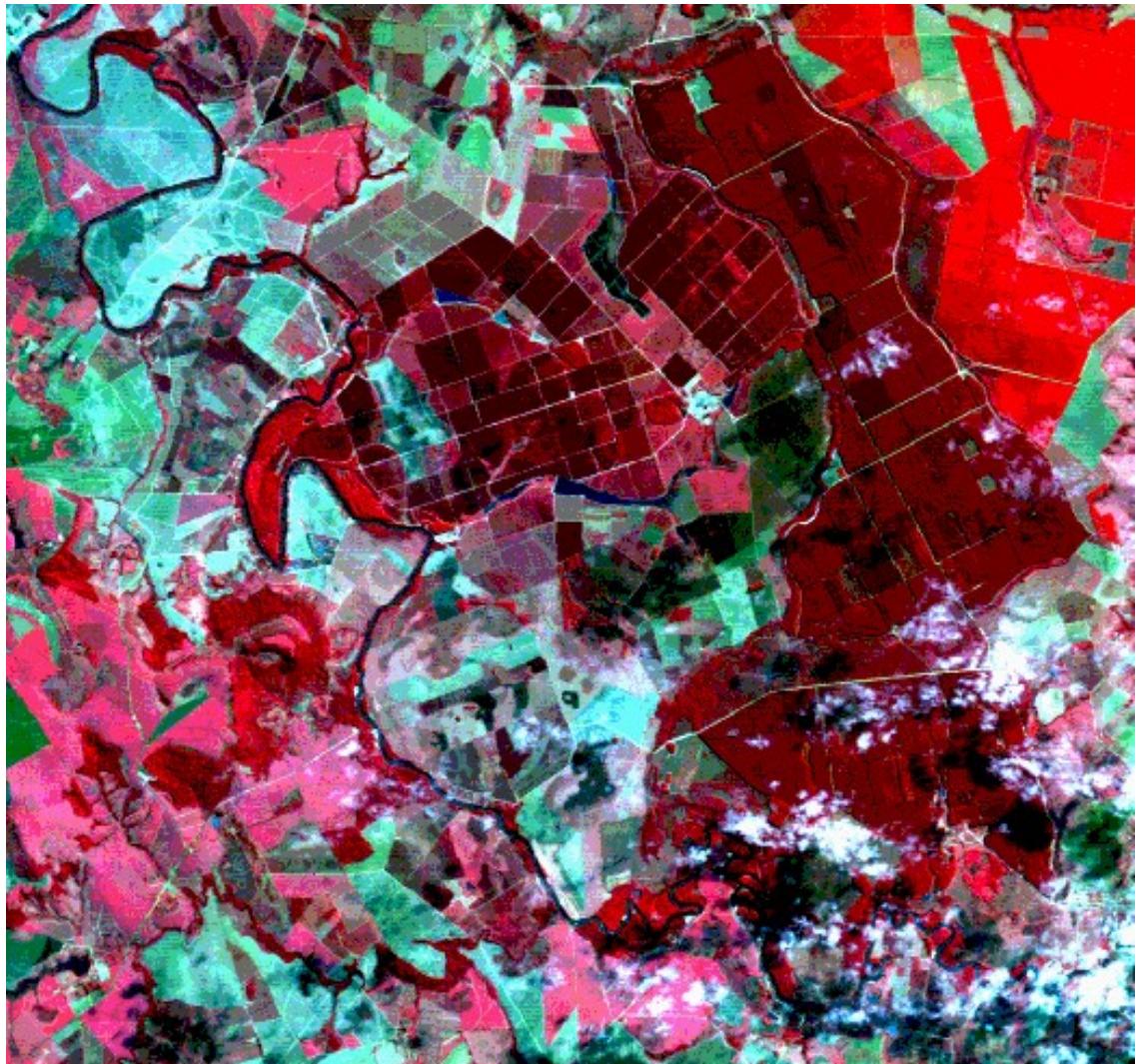
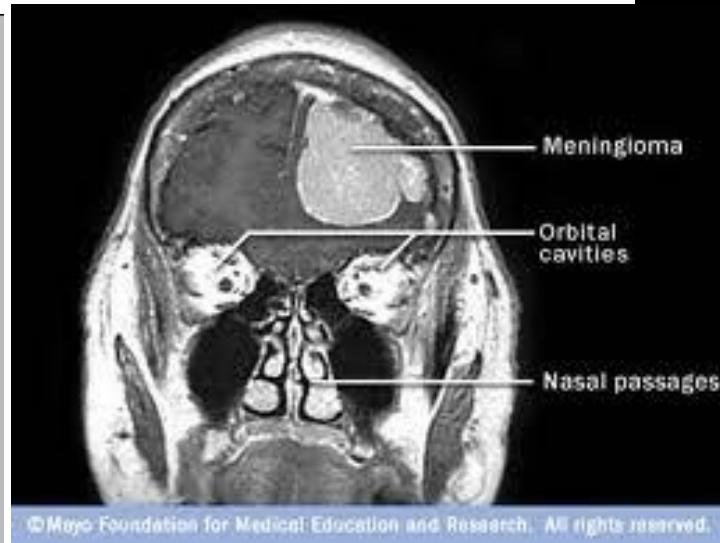
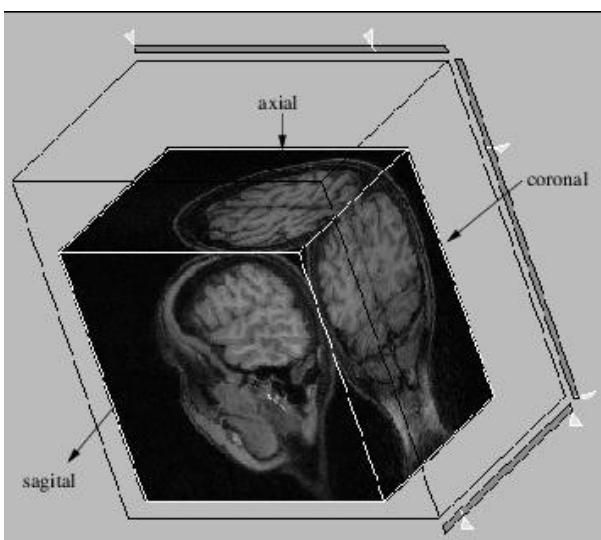


Imagen de sensoriamento remoto

Landsat-8 Bands	Wavelength (micrometers)	Resolution (meters)
Band 1 – Coastal aerosol	0.43 – 0.45	30
Band 2 – Blue	0.45 – 0.51	30
Band 3 – Green	0.53 – 0.59	30
Band 4 – Red	0.64 – 0.67	30
Band 5 – Near Infrared (NIR)	0.85 – 0.88	30
Band 6 – SWIR 1	1.57 – 1.65	30
Band 7 – SWIR 2	2.11 – 2.29	30
Band 8 – Panchromatic	0.50 – 0.68	15
Band 9 – Cirrus	1.36 – 1.38	30
Band 10 – Thermal Infrared (TIRS) 1	10.60 – 11.19	100
Band 11 – Thermal Infrared (TIRS) 2	11.50 – 12.51	100

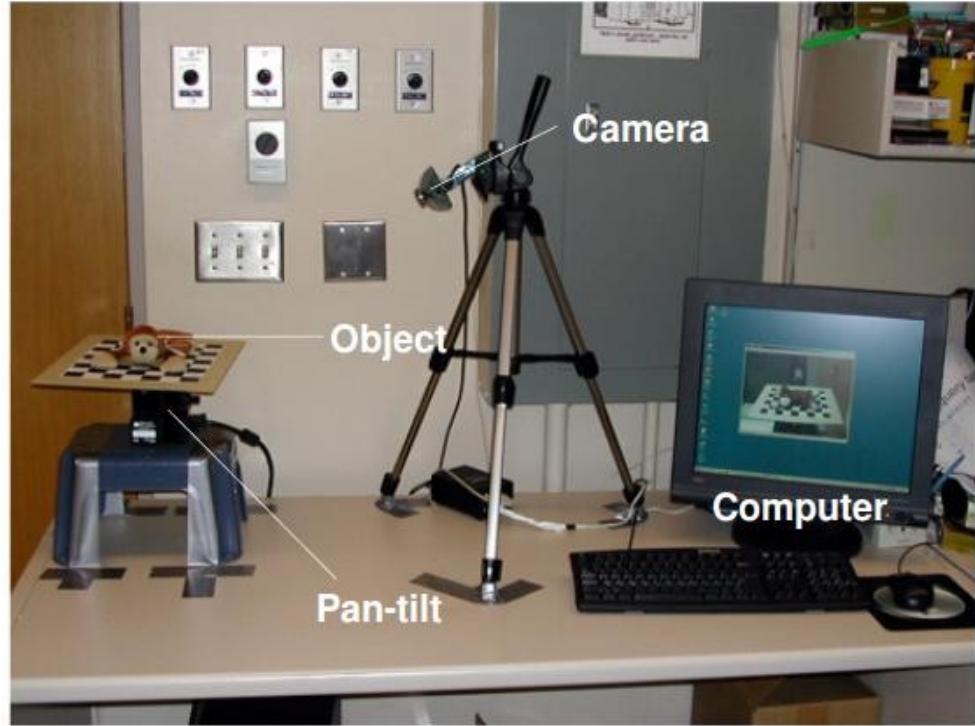
Tomografia / Ressonância Magnética



Tomografia / Ressonância Magnética

- A ressonância magnética mede o tempo de relaxação dos spins dos átomos de hidrogênio em (x,y,z) após desmagnetização de um campo magnético forte.
- Tecidos têm tempos de relaxação distintos, gerando *spels* (*voxels - volume elements*) com valores correspondentes.
- Tomografia de raios-X mede a atenuação do tecido à passagem de raios-X pelo corpo.

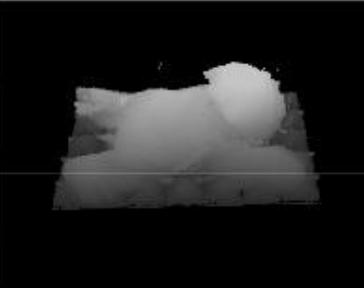
Image Based Rendering



Monkey

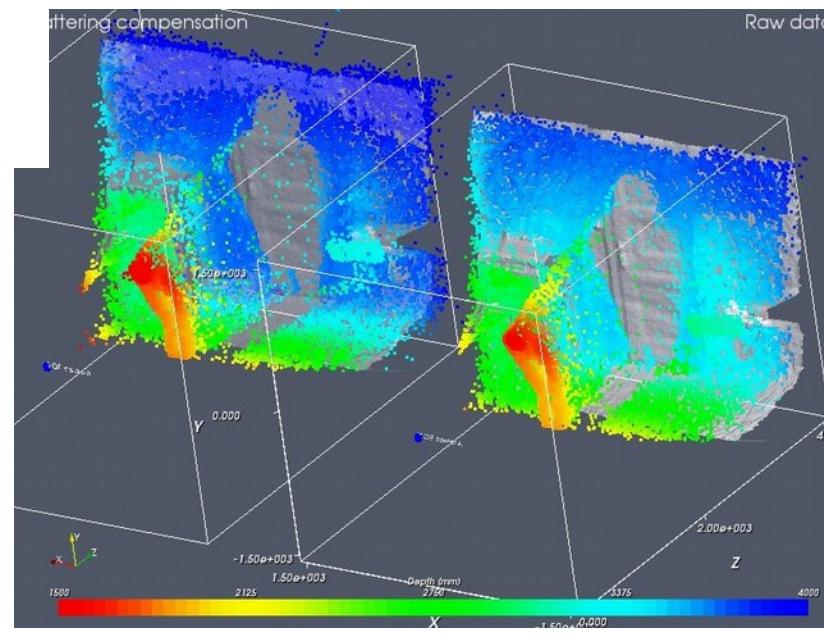


Rendered



Geometry

Sensores



Operações com imagens

- Filtros
 - Suavização
 - Realce
 - Transformações
- Operadores morfológicos
 - Abertura
 - Fechamento
 - Gradiente
- Extração de características
 - Descritores globais e locais

Convolução Discreta 2D

Filtro representado
por uma Matriz 3x3:

1	1	1
1	1	1
1	1	1

1/9

Imagem (5x5):

6	4	5	6	8
9	0	4	8	5
3	2	3	4	2
9	2	3	6	1
7	8	9	0	4

Algoritmo:

Para cada pixel da imagem

- Posicionar centro do filtro sobre o pixel
- Calcular média ponderada dos pixels vizinhos segundo os valores do filtro
- pixel correspondente na imagem final ganhará essa média

Exemplo no pixel (2,3):

6	4	5	6	8
9	0	4	8	5
3	2	3	4	2
9	2	3	6	1

$$(2+3+4+2+3+6+8+9+0) / 9 \approx 4$$

Considerações:

- Complexidade?
- Valores Negativos?
- O que fazer na borda?

Convolução

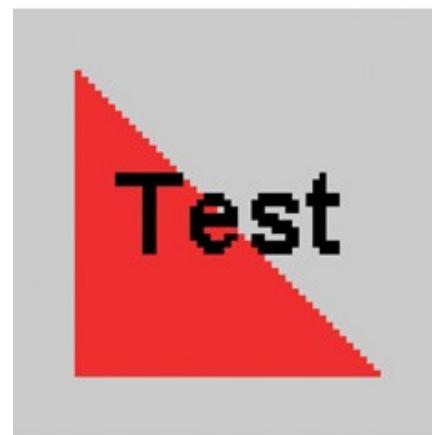
- Filtros (máscaras) comuns:
 - Média (passa baixa)
 - Laplaciana (passa alta)
 - Gaussiana
 - Sobel
 - Realce
 - Nitidez (*sharpen*)
 - Artísticos

Média ou Box

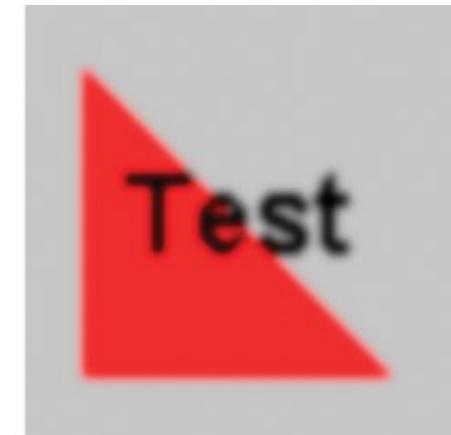
- O que acontece se fizermos uma convolução usando uma máscara como essa abaixo?

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Resultado



(a)



(b)

Resultado (máscara 5x5)

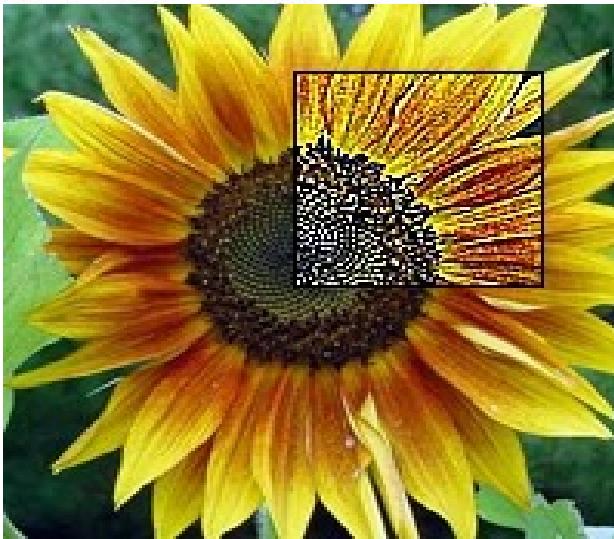


Nitidez (*sharpen*)

- O que acontece se fizermos uma convolução usando uma máscara como essa abaixo?

-1	-1	-1
-1	9	-1
-1	-1	-1

Resultado



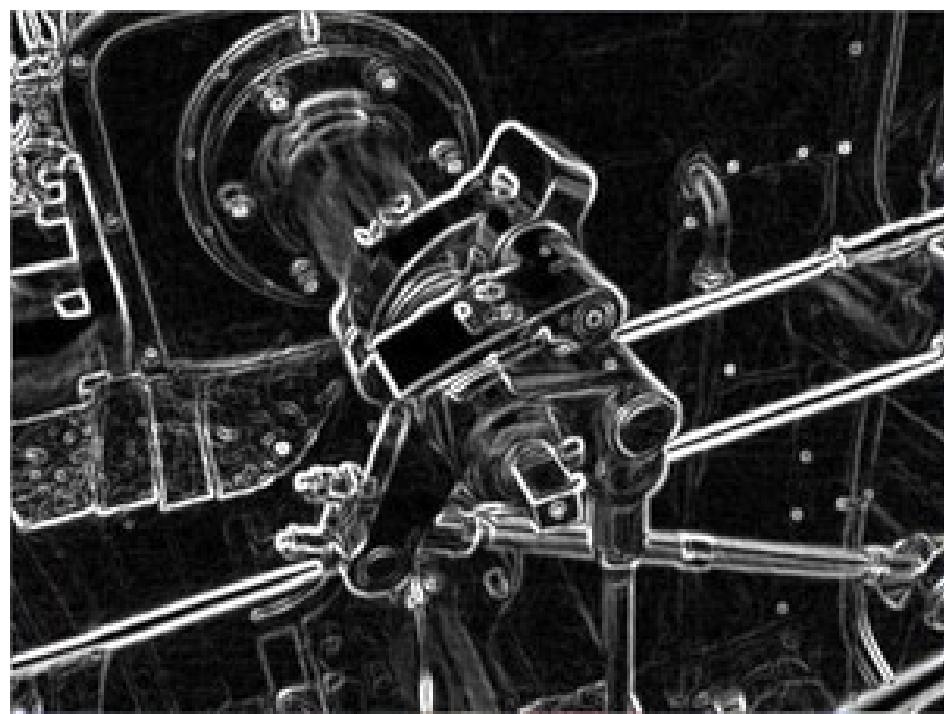
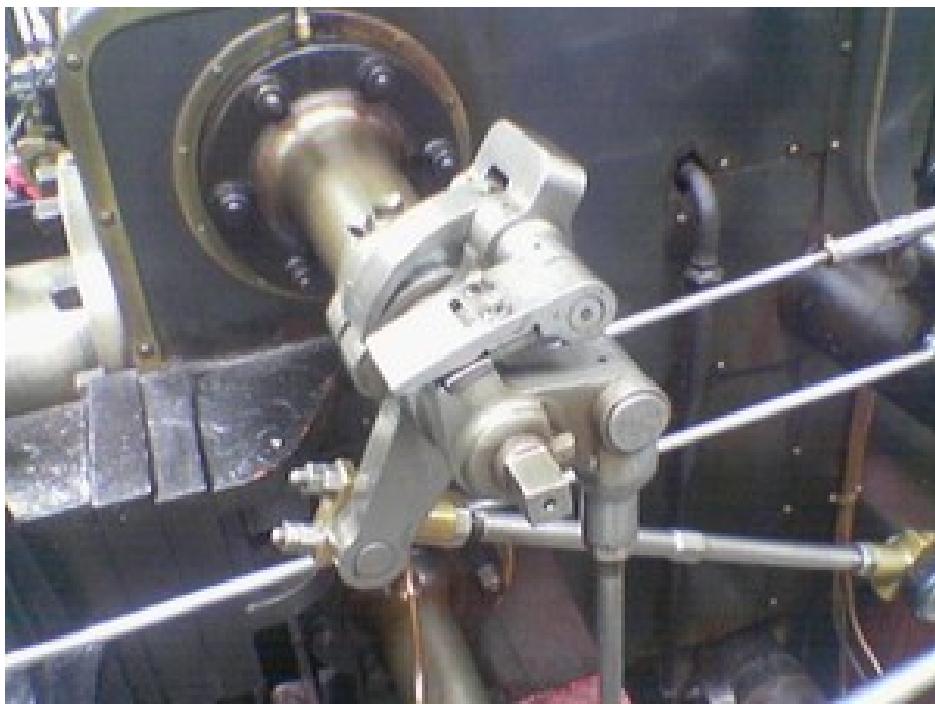
Resultado



Sobel

$$G_x = \begin{array}{|c|c|c|} \hline -1 & 0 & +1 \\ \hline -2 & 0 & +2 \\ \hline -1 & 0 & +1 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline -1 & 0 & 1 \\ \hline \end{array}$$

Sobel (borda)



Outros



Poprocket



Nashville



Gotham

Transformada de Fourier

- A teoria de Fourier diz que qualquer sinal, ou imagens, pode ser expresso como uma soma de uma série de sinusóides (senos e cossenos).
- No caso de imagens, essas são variações sinusoidais na intensidade luminosa.
- Em 1822, o matemático francês Jean-Baptiste Joseph Fourier mostrou que ondas sinusoidais podem ser usadas como bases para descrever, qualquer tipo de função.

Transformada de Fourier

- Muitas vezes ao invés de saber onde ou quando algum sinal é mais intenso precisa-se saber quanto frequentemente alguma intensidade ocorre, ou quanto por cento está acima de um certo valor, etc.
- Detalhes deste tipo são melhor respondido levando o sinal para o **domínio da frequência**.
- A transformada de Fourier (TF) faz essa transformação!

Transformada de Fourier

A **transformada de Fourier** $F(u)$, de uma função contínua $f(x)$ de uma variável real x pode ser definida como:

$$F(u) = \int_{-\infty}^{\infty} f(x) \exp[-j2\pi ux] dx \quad \text{onde } j = \sqrt{-1}$$

A partir de $F(u)$, pode-se obter $f(x)$ através da transformada **inversa** de Fourier:

$$f(x) = \int_{-\infty}^{\infty} F(u) \exp[j2\pi u x] du$$

Essas duas equações são chamadas de **par de transformada de Fourier** e podem existir se ambas forem integráveis e se $f(x)$ for contínua.

Transformada discreta de Fourier

- Pode-se representá-lo por amostras de N valores com intervalos uniformemente espaçados através da sequência: $\{f(0), f(1), f(2), \dots, f(N-1)\}$.
- Pode-se definir o par de **transformadas discretas de Fourier** como sendo uma soma finita de exponenciais complexas.
- O domínio da frequência também é considerando discreto e representado por
 - $u = (0, \Delta u, 2\Delta u, \dots, (N-1)\Delta u)$, onde $\Delta u = 1/N\Delta x$.
- A **transformada discreta de Fourier** se obtém através de:

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \exp[-j2\pi ux/N]$$

FT 2D

- Número de frequências corresponde ao número de *pixels* do domínio espacial, ou seja, a imagem do domínio espacial e Fourier são do mesmo tamanho.
- Em imagens (escala de cinza) a FT é calculada por:

$$F(k, l) = \frac{1}{N^2} \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} f(a, b) e^{-i2\pi(\frac{ka}{N} + \frac{lb}{N})}$$

$$f(a, b) = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} F(k, l) e^{i2\pi(\frac{ka}{N} + \frac{lb}{N})} \quad (\text{inversa})$$

FFT

- O cálculo da FT é feita em $O(n^2)$, mas existe um algoritmo chamado *Fast Fourier Transform* (FFT) que é realizado em $O(n \log n)$, baseado no método de dobramentos sucessivos.

Exemplo FT 2D

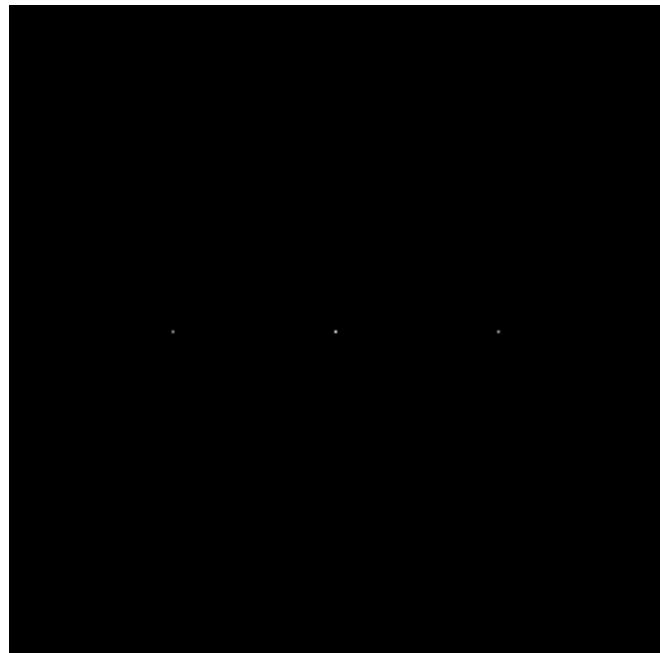
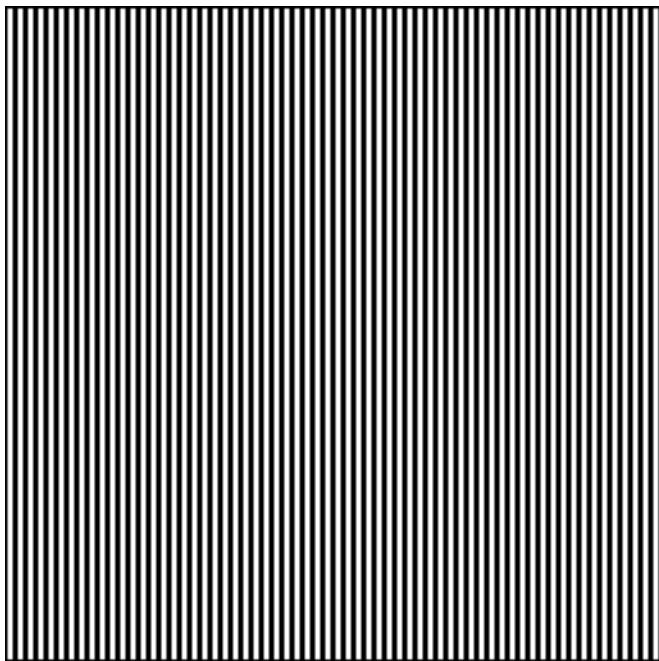


Imagen listras verticais
com 2 *pixels* de largura.

Filtros usando FFT

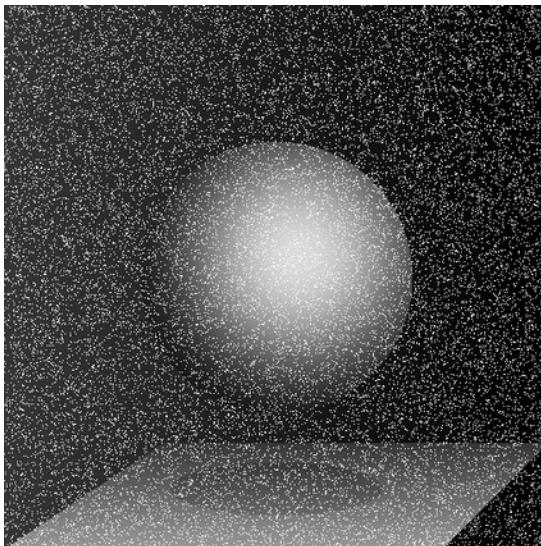
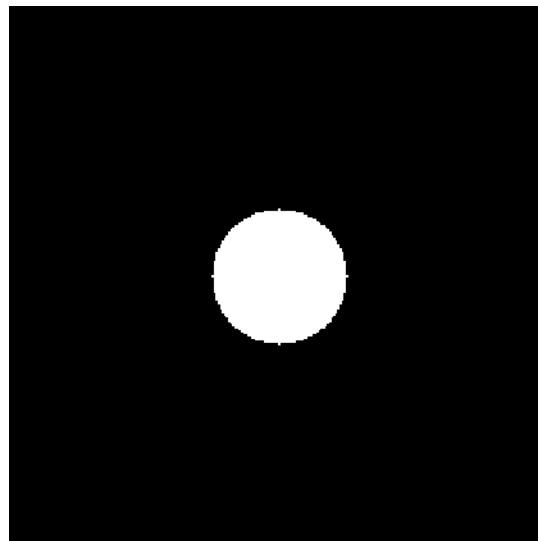
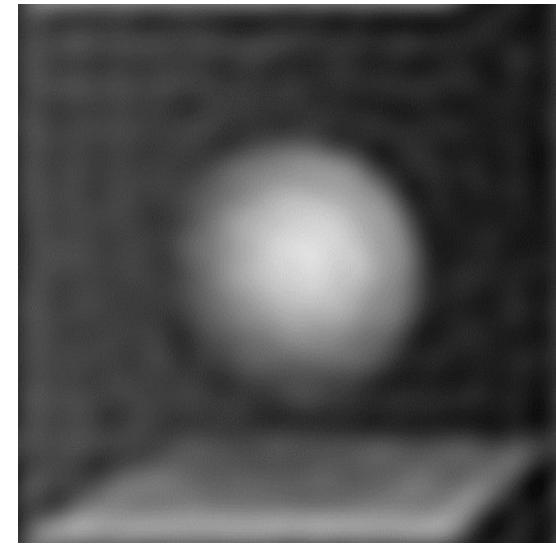


imagem original



máscara do filtro



reconstrução IFFT

FFT

- Fourier é ótima para sinais estacionários, cujo conteúdo não mudam no tempo.
- Neste caso, não é preciso saber "quando" um determinado componentes de frequência existe, já que todos os componentes de frequência existem em todos os momentos.
- Para contornar isso algumas propostas foram surgindo: técnicas que aplicam partições e multi resolução visando incluir informações espaciais junto com as de frequência
 - Windowed Fourier T
 - Short Time Fourier Transform (STFT).

Transformada de Wavelet

- A resolução ou detalhamento (da análise) no domínio da frequência (Fourier) diminui enquanto a resolução no tempo aumenta.
- Não é possível aumentar o detalhamento em um dos domínios sem diminuir no outro. Esta relação entre os domínios da frequência e do tempo é chamada de **princípio da incerteza**.
- Usando as wavelets (ondaletas), é possível escolher a melhor combinação dos detalhamentos para um objetivo estabelecido.

Transformada de Wavelet



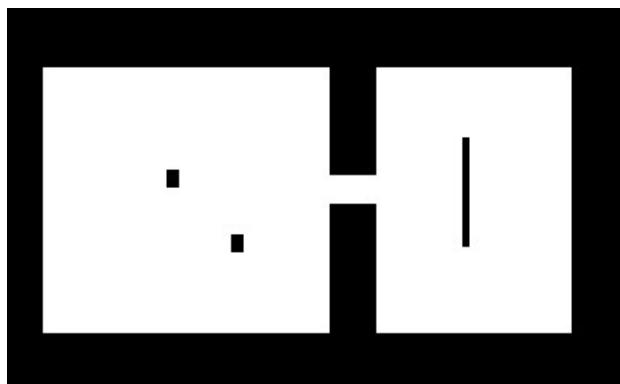
Morfologia Matemática

- A morfologia matemática foi criada por Georges Matheron e Jean Serra em 1964 para analisar a geometria dos objetos em imagens binárias, e depois foi estendida para imagens em tons de cinza.
- Transformações morfológicas essencialmente comparam uma imagem com outra menor (elemento estruturante), cuja geometria conhecida modifica os objetos da imagem.
- Duas operações básicas são dilatação e erosão. A primeira “engorda” os objetos de uma imagem binária (deixa mais clara a imagem cinza) enquanto a segunda “emagrece” os objetos (deixando mais escura). Outras operações podem ser feitas por combinação de ambas, como fechamento e abertura.

Morfologia Matemática



Original

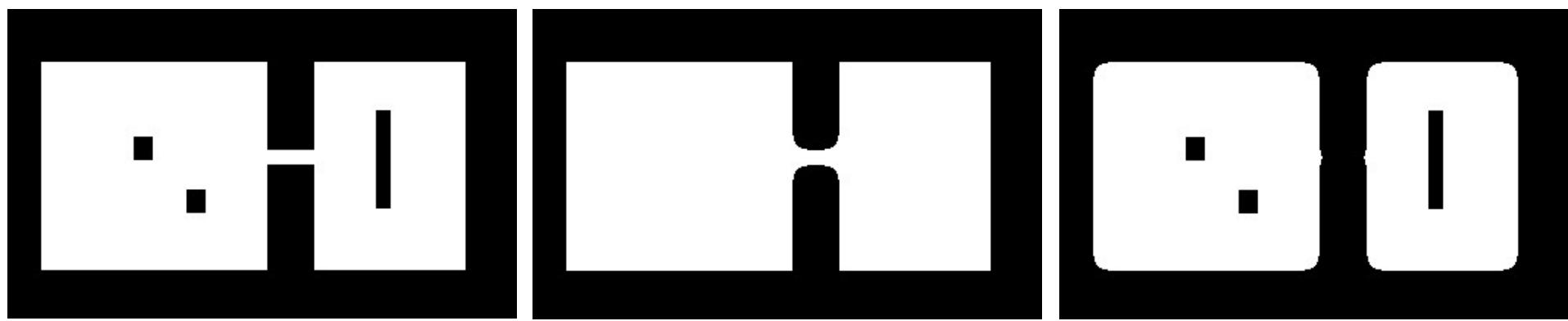


Dilatação



Erosão

Morfologia Matemática



Original

Fechamento

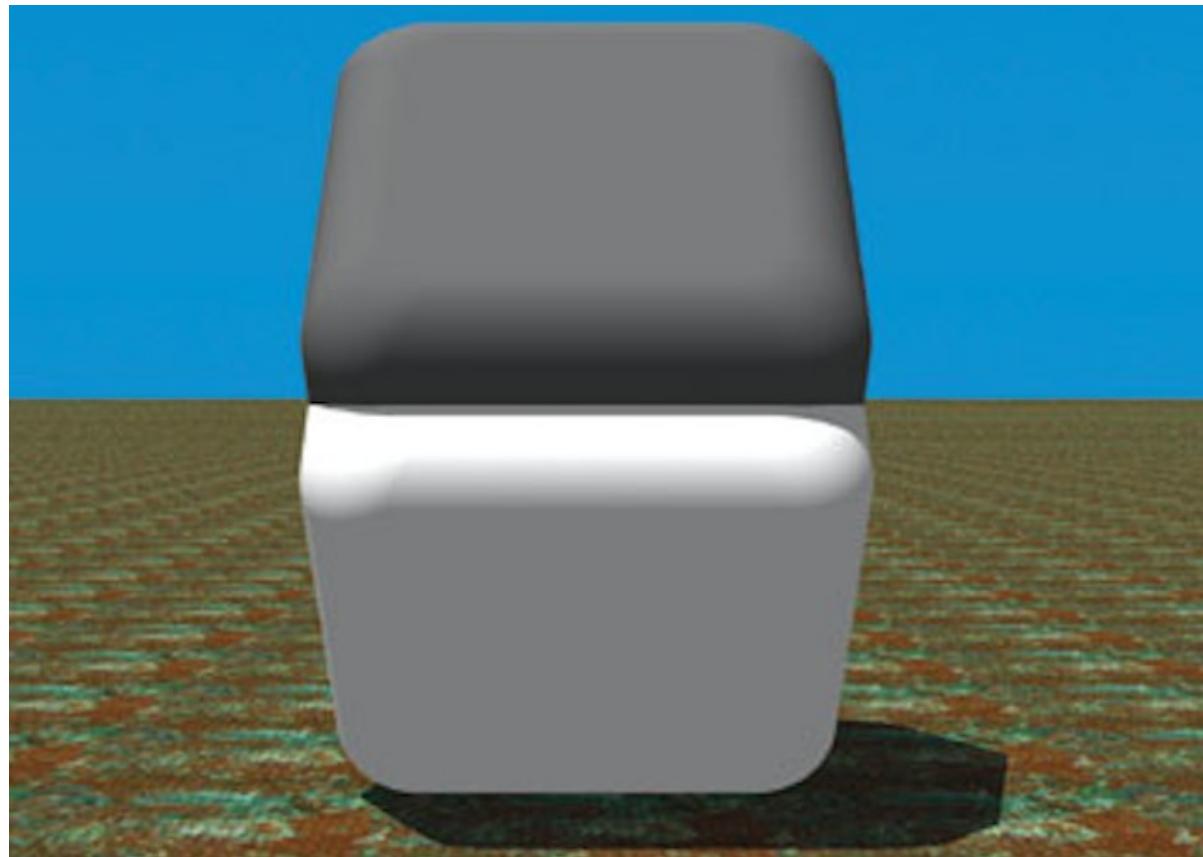
Abertura

Dilatação seguida de erosão = ?
Erosão seguida de dilatação = ?

Extração de Características

- A percepção visual dos objetos é subjetiva e por isso não existe uma única representação e nem mesmo uma melhor representação para uma dada característica.
- Como será apresentado a seguir, para cada característica existem várias representações distintas.

Extração de Características



Extração de Características

- A extração de características (descritores) é a base da recuperação de informação visual.
- Em imagens, estas características podem ser classificados como sendo de domínio geral ou de domínio específico.
 - O primeiro inclui características de cor, textura e forma, enquanto a última é dependente da aplicação, como por exemplo, classificação de impressão digital, placas de veículos, etc.

Domínio Geral x Específico

- As características de domínio específico é melhor apresentado na literatura que trata do padrão a ser reconhecido e envolve o conhecimento de características muito particulares ao problema em questão.
- Aqui serão apresentados alguns descritores com características mais gerais que podem ser utilizadas em diferentes aplicações.

Descritores Locais x Globais

- Descritores Locais
 - Descritores locais de imagem, calculados em regiões ao redor de pontos de interesse (na maioria das vezes vértices de contornos ou blobs) na imagem. Ex.: PCA-SIFT, SIFT, GLOH, SURF, DAISY
 - Eles são encontrados por detectores de pontos (Harris-Affine, Hessian-Affine, Fast Hessian, MSER, DoG) e possuem normalmente as seguintes informações: uma coordenada 2D na imagem, uma orientação, uma escala e características de uma região em torno do ponto.

Descritores Locais x Globais

- Descritores Globais
 - Abordagem tradicional de classificação de imagem.
 - Consideram a imagem como um todo, descrevendo imagens baseadas nas informações de cor, textura e/ou forma.

Descritores baseados em Cor

- A cor é provavelmente a característica mais utilizada para recuperação visual.
- Ela é relativamente robusta por apresentar independência do tamanho da imagem e da orientação da mesma.
- Uma vantagem do uso desse tipo de descritor é que as cores podem ser facilmente associadas a descrições textuais (nome da cor), facilitando a utilização em muitos sistemas CBIR.

Descritores baseados em Cor

- Os modelos de cor podem ser classificados em orientados ao hardware (RGB, CMY, YIQ) e orientados ao usuário (HSI, HSV, HSB, MTM, L^{*}u^{*}v^{*}, L^{*}a^{*}b^{*} e L^{*}C^{*}h^{*}).
- RGB é o espaço de cores mais utilizado para imagens digitais, mas não é uniforme: a percepção da diferença de cor não apresenta relação com a distância euclidiana no espaço de cores.
- HSI, HSV e HSB, apesar de serem modelos orientados ao usuário, apresentam o mesmo problema do RGB, por apresentarem uma escala não uniforme.

Descritores baseados em Cor

- Portanto, os outros espaços de cor (MTM, $L^*u^*v^*$, $L^*a^*b^*$ e $L^*C^*h^*$) são mais adequados para recuperação de informação visual.
- Apesar disso, a maioria dos classificadores utilizam RGB ou HSV para descrição das cores.

Descritores baseados em Textura

- Textura é uma propriedade presente em praticamente todas as estruturas, como nuvens, vegetação, paredes, cabelo e outros.
- Ela contém informação importante sobre o arranjo estrutural da superfície e sua relação com o ambiente.
- Por ser um padrão útil e importante em visão computacional e reconhecimento de padrão, é uma área de pesquisa rica em pesquisa das últimas décadas. Atualmente, são encontradas muitas pesquisas na área de recuperação de informação visual baseada em textura.

Descritores baseados em Textura

- Uma técnica muito popular para descrição de texturas é a utilização do máximo espectro de Fourier.
- Os padrões de textura estão normalmente associados a picos no espectro de Fourier da imagem, que podem ser detectados como máximos regionais na imagem de magnitude do espectro.
- O valor e a localização dos picos podem ser usados como características.
- As localizações descrevem a direção e a periodicidade da textura.

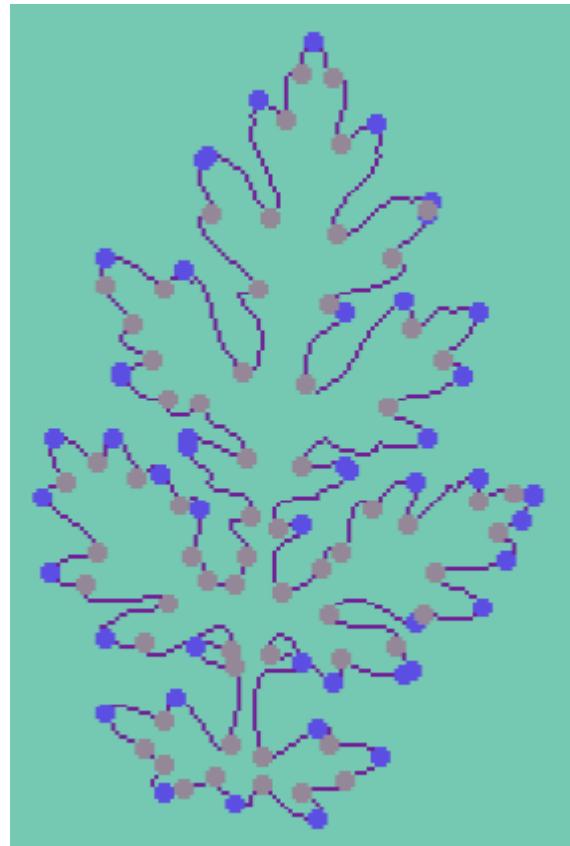
Descritores baseados em Forma

- Os primeiros descritores de forma se resumiam em definir informações simples como comprimento, perímetro, área e algumas relações entre essas medidas (regularidade e compacidade) de um conjunto de pixels que representam um objeto.
- Mais tarde, outros descritores mais complexos foram definidos, como os pontos de saliência.

Descritores baseados em Forma

- Pontos de saliência são aqueles de maior curvatura e são usados diretamente para definir as características de forma dos objetos.
- Um descritor mais simples do que as saliências de contorno, mas com maior eficácia em algumas aplicações, armazena as saliências de segmentos (Segment Saliences) como características.

Pontos de Saliência



Descritores baseados em Forma

- Outros descritores de forma tradicionais são os descritores de transformada de Fourier e invariantes de momento (*moment invariants*).
- A ideia principal de invariantes de momento é utilizar medidas baseadas na região que são invariantes às propriedades da forma, como rotação, escala e translação.

Considerações sobre imagens

- Normalmente trabalhamos com imagens digitais, que são uma transformação dos sinais contínuos reais (radiações eletromagnéticas) em uma matriz discreta.
- Essas imagens podem sofrer alterações como compressão com perda, filtragem, entre outros.
- Não faz parte do conteúdo deste curso estudar processos de manipulação de imagens como filtros para remoção de ruído, segmentação, detectores de contorno, etc.

Representação de imagem

- Modelo de cor
 - Existem diferentes espaços (modelos) de cor utilizados para representação de imagens (RGB, HSV,...).
 - Além dos espaços de cor, a quantidade de cores (tons) na imagem também é importante para alguns algoritmos de extração de características .
 - Desta forma, existem descritores que utilizam apenas imagem **binária**, outros tons de cinza ou coloridas.

Processamento de Imagens usando Grafos

- Teoria dos grafos é uma área bem estabelecida, rica em algoritmos eficientes e com várias aplicações nas ciências e engenharias.
- Neste sentido, soluções eficientes e eficazes para problemas de processamento de imagem podem ser obtidas pela simples redução do problema em um problema de teoria dos grafos com algoritmo conhecido.

Processamento de Imagens usando Grafos

- Esta estratégia também possibilita o desenvolvimento de variantes e novos algoritmos.
- Existem várias formas de modelar um grafo a partir de uma imagem, definindo os nós do grafo e uma relação de adjacência entre eles.
- Dependendo do problema, *pixels*, vértices de *pixels*, arestas de *pixels* ou regiões da imagem podem ser os nós.

Processamento de Imagens usando Grafos

- Como resultado, o problema de processamento de imagem é reduzido à aplicação de um algoritmo em grafo seguida de uma operação local sobre atributos do grafo resultante.
- Esta metodologia pode ser utilizada no projeto de operadores de filtragem de imagem, segmentação de imagem, representação e descrição de objetos de imagem, e classificação de padrões em imagem.

Imagen e Componentes Rotulados

- Muitas vezes, uma imagem possui vários objetos e eles precisam ser individualmente identificados com um rótulo (número).
- Uma descrição completa do objeto requer também a identificação de seus componentes.
- A segmentação realiza a primeira parte desta tarefa e a rotulação dos componentes pode ser obtida com uma busca em largura em um grafo, onde os pixels são os nós e os arcos são formados por *pixels vizinhos-4* (ou vizinhos-8).
- Uma vez rotulados, podemos descrever relações entre eles (i.e., outro grafo onde os objetos são os nós e os arcos representam estas relações).

Rotulação de Componentes

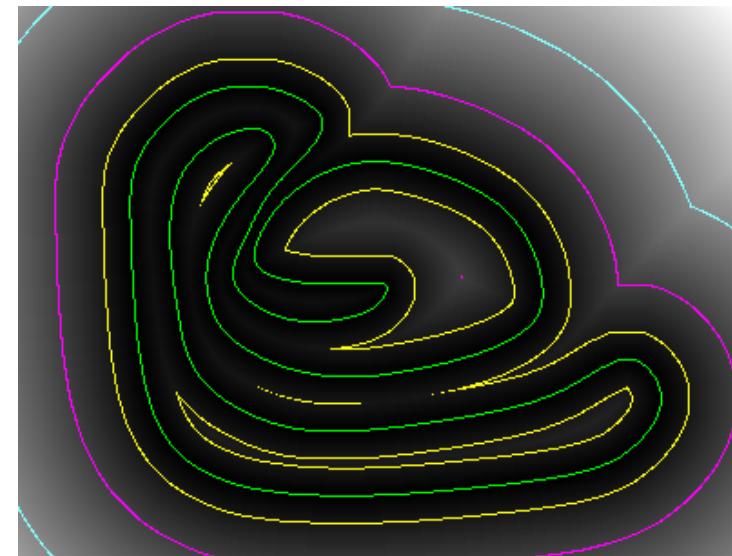
Imagen de 2 objetos rotulados
(saída da segmentação).

	1				1	1	1
1	1	1	1		2	2	2
1			1		2		2
1	1	1	1		2	2	2
1	1	1	1		2		2
					2	2	2

Elementos em Múltiplas Escalas

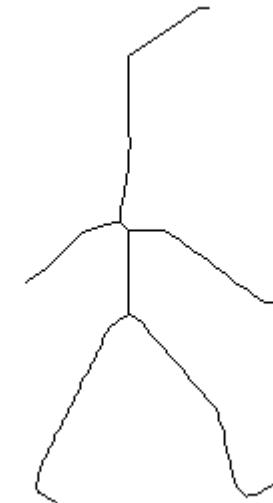
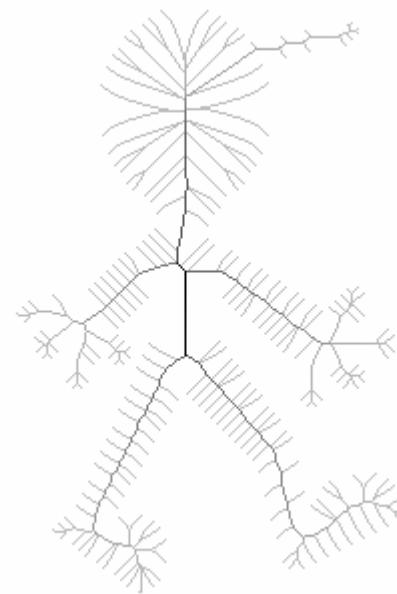
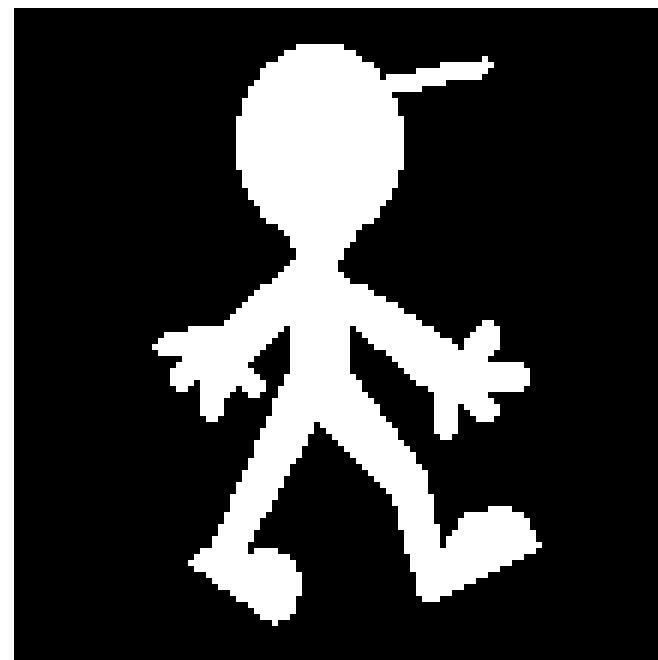
- Alguns elementos em imagens como contorno ou esqueletos podem ser definidos em diferentes escalas.
- No caso dos pontos de saliência, poderíamos ter para cada ponto um ângulo de abertura associado ao mesmo. Ao comparar duas imagens, podemos considerar somente ângulos acima de um limiar, por exemplo.

Contorno em Múltiplas Escalas



A TDE gera o conjunto de todas dilatações e erosões exatas da forma.

Esqueleto Multiescala



Limiar 1 a 120 (de 907)

Medidas de Avaliação de Descritores

- Ao avaliar um novo descriptor, os autores buscam medir:
 - Velocidade/complexidade do processamento da extração de características;
 - Separabilidade entre classes;
 - Percentual de acertos em uma classificação;
 - Gráfico Precisão-Revocação;
 - Exemplo individual.

Operação com Imagens

- O preprocessamento antes da etapa de extração de características é necessário para:
 - Melhorar condições de iluminação
 - Remoção de ruídos
 - Ajustar a posição e a escala dos objetos
 - Melhorar o foco

Ajuste de iluminação



Placa:
BRK 8558
Veículo:
Pajero 1995

Redução de ruídos



Importância da redução de ruídos

- Ruído geralmente se destaca das demais áreas da imagem:
 - Ele não está correlacionado à imagem
 - Muitos descritores identificam ele como algo significativo por diferir dos demais pontos
- A detecção de linhas e bordas é prejudicada pelo ruído pois o mesmo forma descontinuidades na imagem (Ex.: Transformada de Hough).
- Para a redução de ruídos pode-se aplicar diversos filtros, mas estes podem retirar detalhes importantes da imagem.

Ajustes de posição e escala

- As vezes a posição de um objeto na imagem prejudica a identificação
- A escala também afeta o reconhecimento: o mesmo objeto em escalas diferentes não necessariamente será reconhecido
- Estes problemas podem ser resolvidos utilizando segmentação e/ou descritores invariantes a posição e a escala
- Orientação do objeto no plano 3D é mais difícil de ser tratada

Ajustes de posição e escala



Influência da inclinação para a identificação de pessoas

Detalhes finais

- Processamento de imagens visa alterar ela de forma que ela possa ser melhor utilizada em sistemas de Visão Computacional.
- Para a parte de Síntese de Imagens, ela é útil para modificação em relação à iluminação ou modificação de outros aspectos visuais.
- No dia a dia, processamento de imagens são utilizados constantemente em sistemas como Instagram.