

Processamento de Imagens

SCC0251

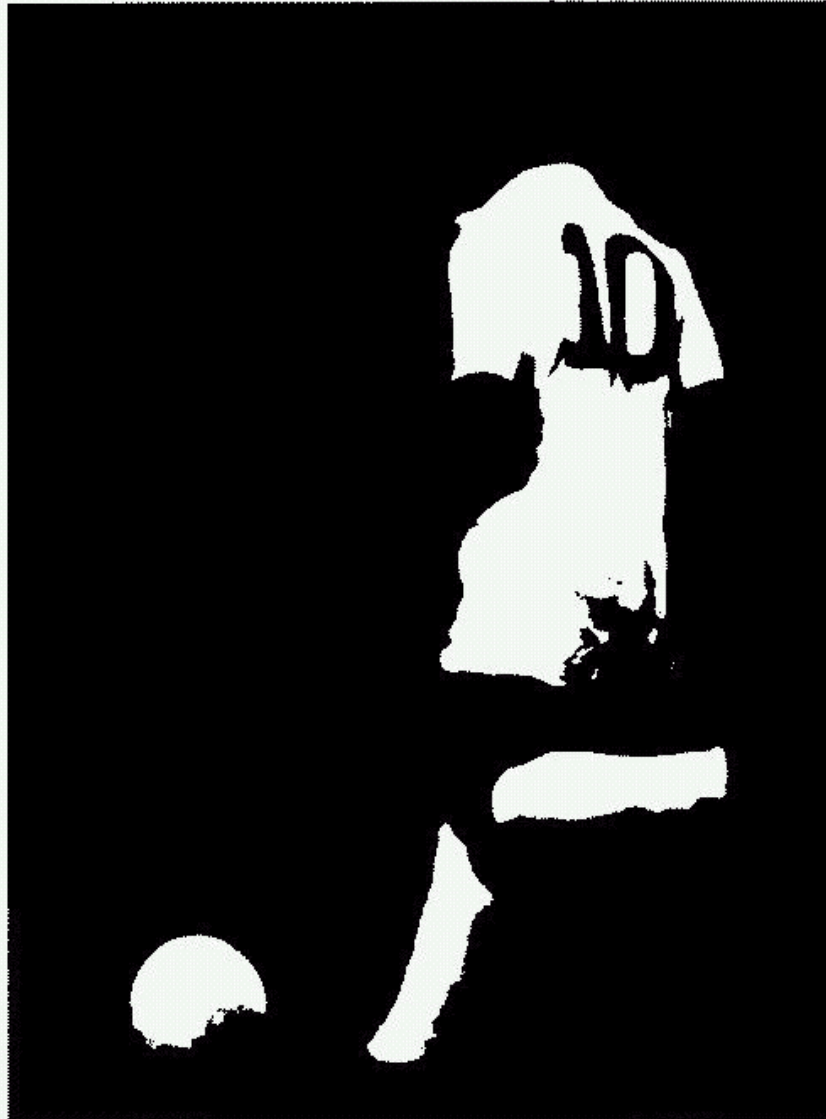
**Instituto de Ciências Matemáticas e de
Computação
USP**

Fundamentos de Visão Computacional e Processamento de Imagens

- *Computer Imaging*: Manipulação de imagem por computador
 - Aquisição e processamento de informação por computador
 - Sentido primário: visão
 - “Uma imagem vale por mil palavras”

QUINTA-FEIRA, 28 DE MAIO DE 1998

A SELEÇÃO NA FRANÇA



Domício Pinheiro/AE

Domício Pinheiro

Domício Pinheiro
Agência Estado

Oi! Eu sou Bill Gates, fundador da Microsoft. Voltei para apresentar o Upgrade do sistema operacional mais famoso do mundo...

Windows 2001[®]

LANÇADO ANTES MESMO QUE O WINDOWS 2000 TERMINASSE DE CARREGAR!



A **Microsoft** comunica que, para comemorar a chegada do novo milênio, as atualizações do Windows passam a ser **ANUAIS!!!** Sim: chega de esperar três anos para desfrutar dos bugs mais modernos que nossa tecnologia consegue criar!

SUPORTE INTEGRADO

Cada botão é um botão de Restart, por isso você não precisa se preocupar com as mensagens de erro NUNCA MAIS!

GAMES E APLICATIVOS ESPETACULARES

Além do clássico jogo de Paciência que ocupa 22 megas de HD, você vai se divertir com o hilário "Pra Que Diabos Serve Esse Ícone?". E entre os novos explicativos, há um corretor ortográfico certificadamente à porva de faias.

MANUAL SIMPLIFICADO

Agora com apenas 698 páginas - em cirílico! O novo Windows 2001 é tão intuitivo que precisa de apenas onze megas de arquivos Help!

MULTITAREFA

Consegue dar reboot e crash ao mesmo tempo. E agora ele não apenas dá pau; mostra uma tela de avisos e deixa você pressionar o OK!

MAIS VELOCIDADE

Se você quer que o Windows 2001 seja ainda mais rápido, use nosso exclusivo EMULADOR DE XT!

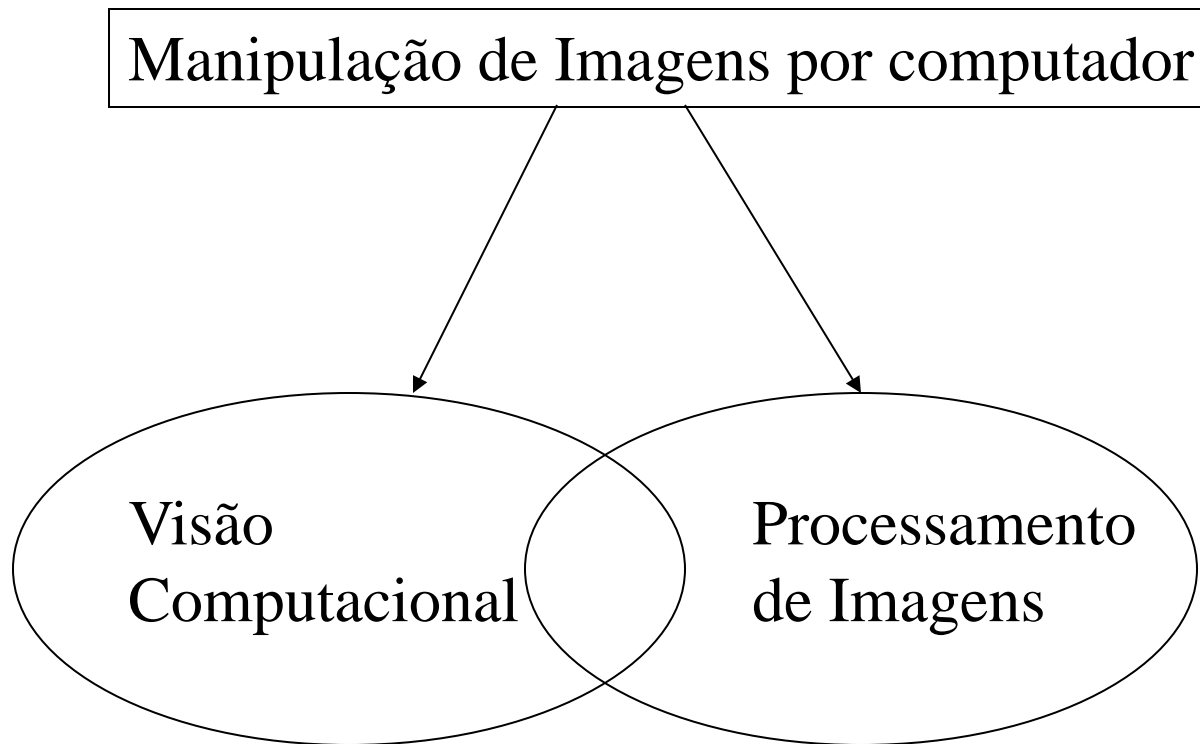
NOVOS PROGRAMAS

O Windows 2001 é 100% compatível com todos os produtos da Microsoft, exceto o software.

COMPATÍVEL COM O
BUG DO DÉCIMO MILÊNIO!
É PESSOAL, POR ENQUANTO
NINGUÉM TÁ LIGANDO, MAS EM
9999 A PARANÓIA VAI RECOMEÇAR!
COMPRE O UPGRADE POR APENAS
US\$ 99,95 E LEVE TRÊS BITS
TOTALMENTE GRÁTIS PARA
COMBATER O BUG DO
ANO 10.000!

TEXTO: SERGIO MARTORELLI FOTO: IRVINO SCHILD

Computer Imaging: duas áreas



Visão X Processamento

“Em aplicações de visão computacional as imagens processadas (saídas) são usadas pelo computador. Em processamento de imagens, são para consumo humano.”

“Historicamente, o processamento de imagens evoluiu a partir da Eng. Elétrica (Proc. Sinais). Visão computacional resultou de desenvolvimentos em Biologia + Exatas (Física, Matemática, C. Computação.”

Visão Computacional (VC)

- Aplicações que não requerem pessoas no ciclo visual, i.e.,
 - imagens são examinadas, manipuladas e `interpretadas` pelo computador
- O computador usa a informação visual diretamente
 - reconhecimento e inspeção de objetos
 - análise de imagens (*Image Analysis*)
 - Feature extraction
 - Pattern Recognition

Aplicações em VC

- Tarefas “tediosas” para seres humanos
 - ambiente hostil
 - alta taxa de processamento
 - acesso e uso de grandes banco de dados/imagens
- Encontrados em ambientes variados
 - controle de qualidade em sistemas de manufatura

Aplicações em VC (cont.)

– Ambiente médico

- detecção automática de tumores
- sistemas de auxílio a neuro-cirurgias
- identificação de impressões digitais
- identificação de impressão DNA (*DNA fingertips*)

– Outros

- Monitoramento de rodovias
- Sistemas de defesa (militares)
- Visão Robô (*vision-guided robot*)
- *Crowd control* (controle de multidão) em metrô

Processamento de Imagens (PI)

- Aplicações envolvem pessoas no ciclo visual, i.e., as imagens são examinadas e manipuladas por pessoas
- Necessário entender como funciona o sistema visual humano

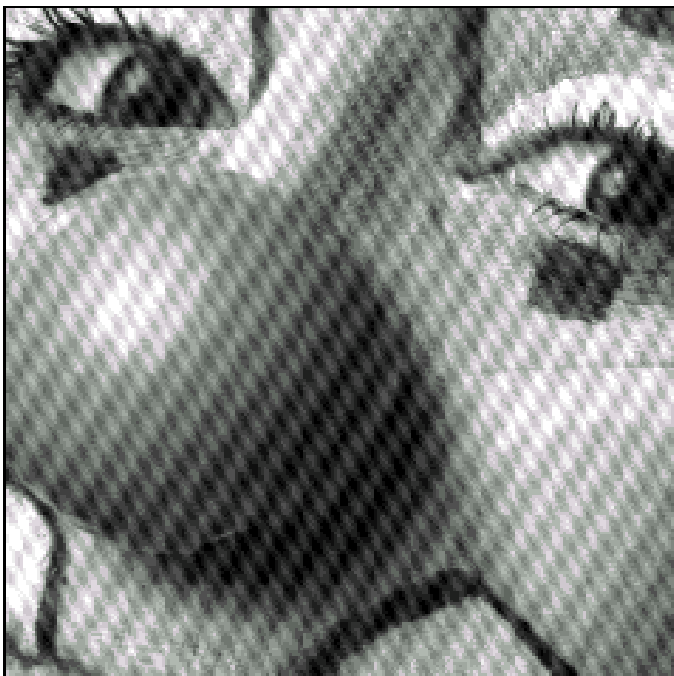
PI: principais tópicos

- Restauração (*restoration*)
- Realce (*enhancement*)
- Compressão (*compression*)

Restauração de Imagem

- Processo de recuperação da aparência original de uma imagem que possui algum grau conhecido (ou estimado) de degradação
- É possível desenvolver um modelo da distorção se soubermos algo sobre a degradação
- Modelo degradação: aplica processo inverso e obtém-se imagem restaurada

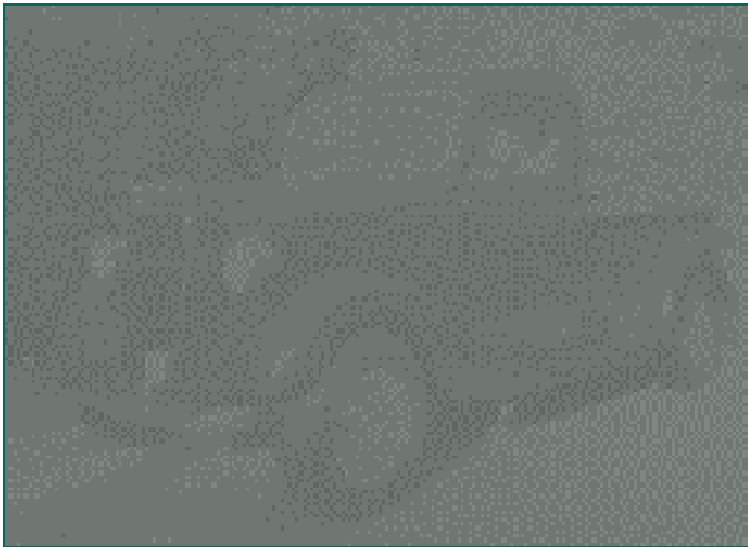
Exemplo de Restauração



Realce de Imagem

- “Melhorar” uma imagem visualmente, baseado na resposta do sistema visual humano.
- Solução ótima depende do problema (*problem/domain specific techniques*)
- Exemplo:
 - Expandir o contraste (*contrast stretching*)

Exemplo de realce de contraste



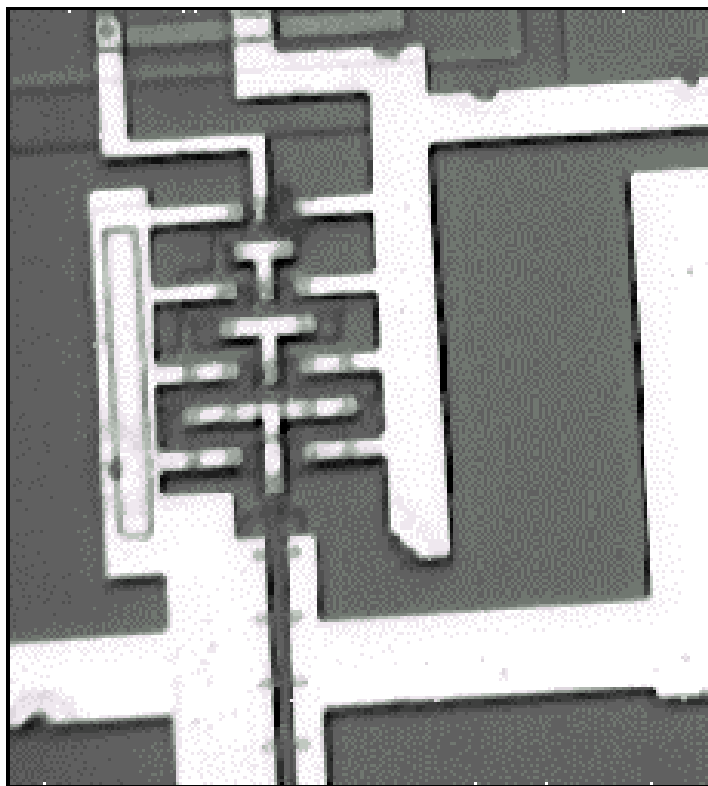
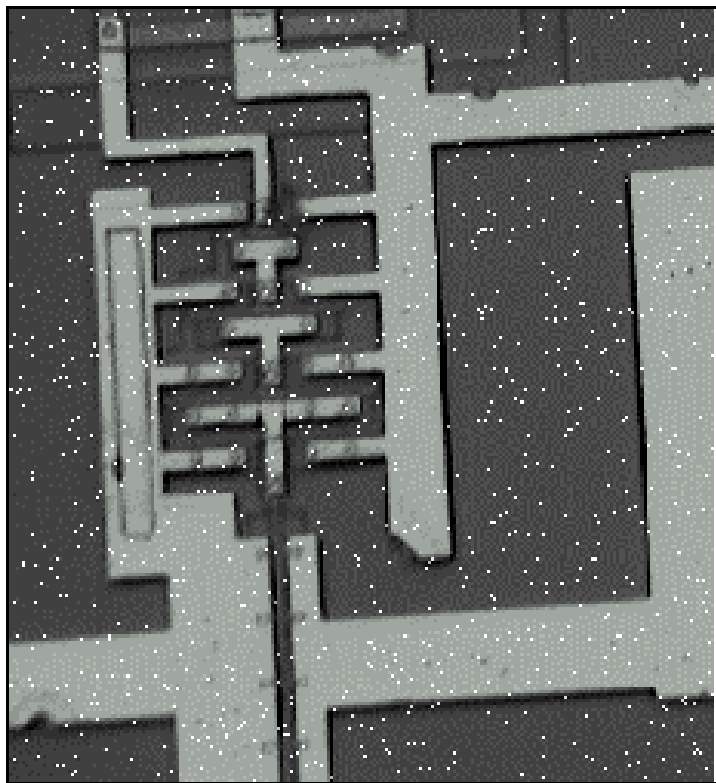
Típica operação do tipo $s = T(r)$, conhecida como *contrast stretching*.

Restauração X Realce

- Ambos levam a “melhora” na imagem

“Métodos de restauração procuram modelar a distorção e aplicar o processo reverso, enquanto os métodos de realce utilizam a resposta do sistema visual humano para “melhorar” a imagem visualmente.”

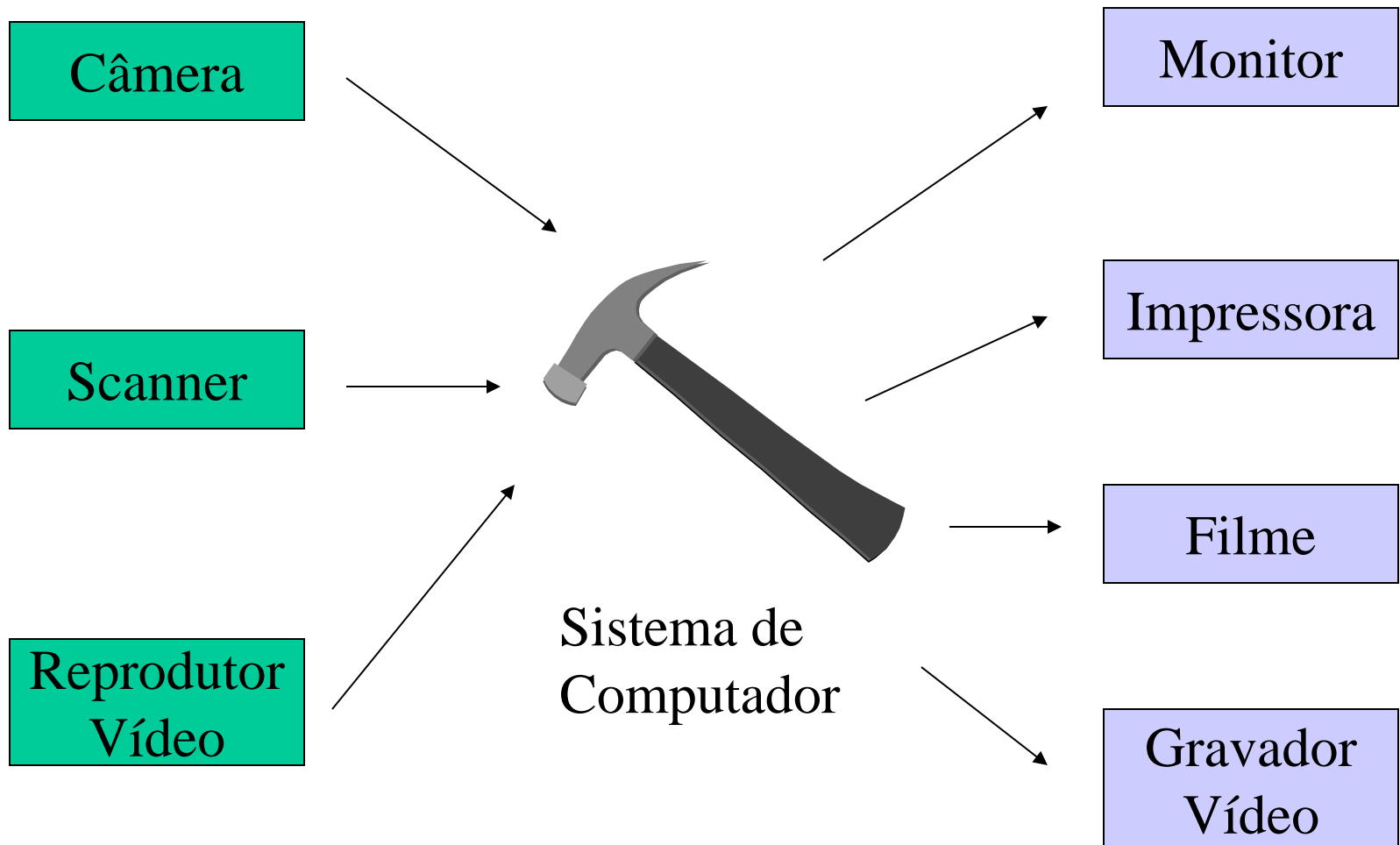
Realce: eliminação de ruído



Compressão de Imagem

- Redução da quantidade expressiva de dados necessária para representar uma imagem
- Eliminação do que é visualmente desnecessário
- Imagens apresentam redundância de dados
- Redução na ordem de 10 a 50 vezes

Sistemas de Imagem por computador



Digitalização

- Sinal de vídeo analógico é transformado em sinal digital através da discretização do sinal contínuo a uma taxa fixa
- Esse processo é feito muito rapidamente
 - Câmeras CCD (*charge-coupled device*)
- O resultado é um vetor bidimensional de dados, em que cada elemento é chamado pixel

Digitalização: discretização

- Amostragem (*sampling*)
 - discretização espacial do sinal
 - Imagem: uma matriz de pontos
- Quantização (*quantization*)
 - discretização da amplitude do sinal
 - Imagens: a faixa de valores que um `ponto' é capaz de armazenar

Quantization

- Exemplo
 - valores de 8 bits para representar voltagens de 0-10 V.
 - 8 bits: 256 valores distintos
 - $10 \text{ V} / 256 = 0.039 \text{ V} !$
 - $0 \text{ V} = 00000000$ $10\text{V} = 11111111$
 - Cada mudança de 0.039, indica um novo valor binário



8 bpp



4 bpp



2 bpp

False
Contour



1 bpp

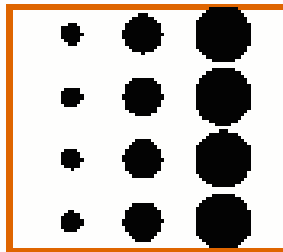
False
Contour

Como reduzir efeito da quantização ?

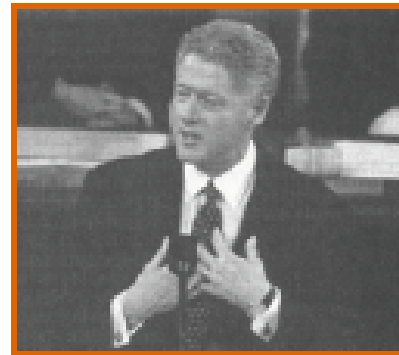
- *Halftoning*
- *Dithering*



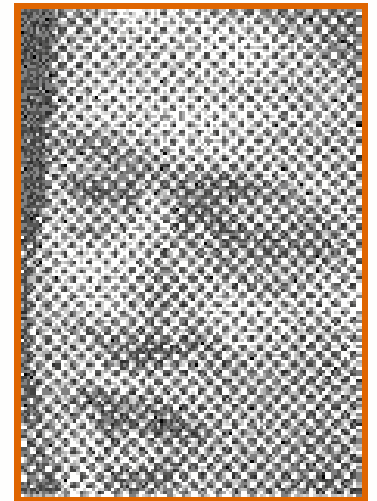
$I(x,y)$



$P(x,y)$

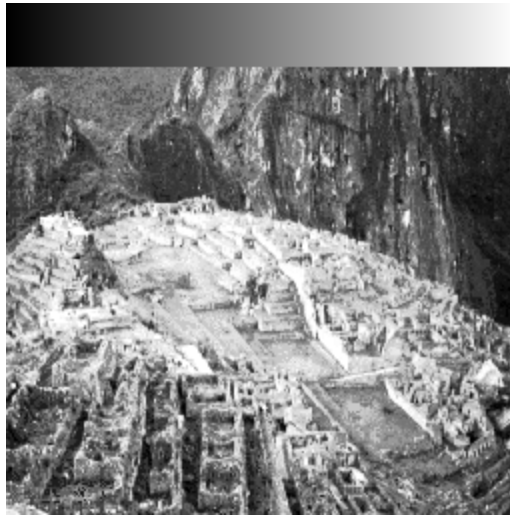


Newspaper Image



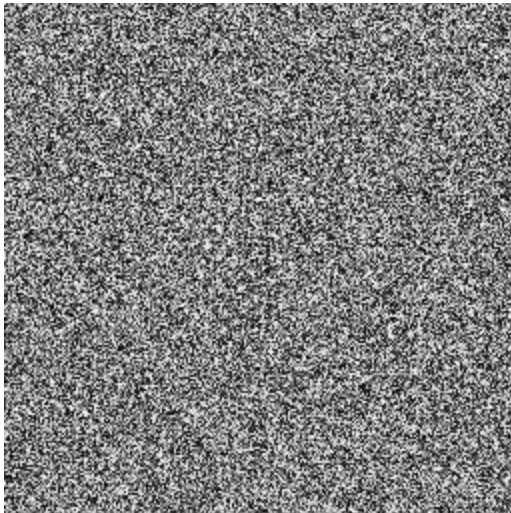
Halftoning

Dithering



original

Pixel > padrão: branco
Pixel < padrão: preto



Ruído uniforme



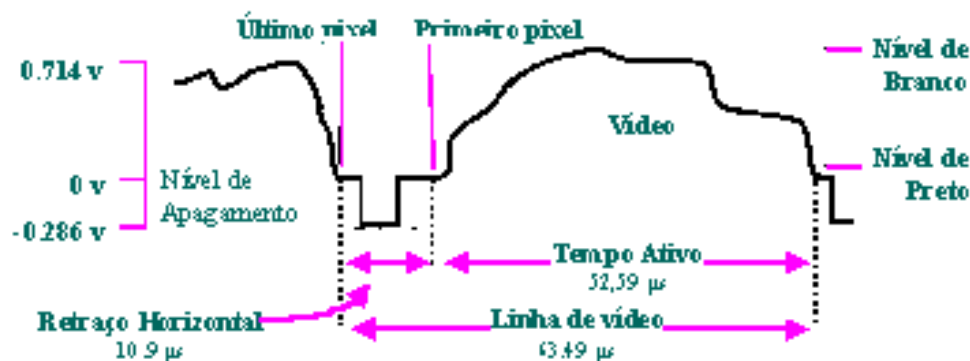
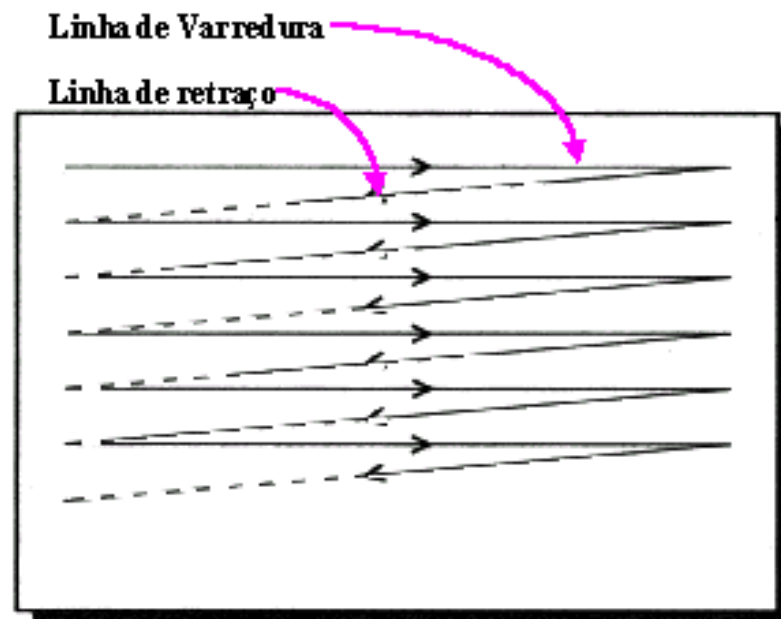
Imagem resultante (binária)

Halftoning/Dithering

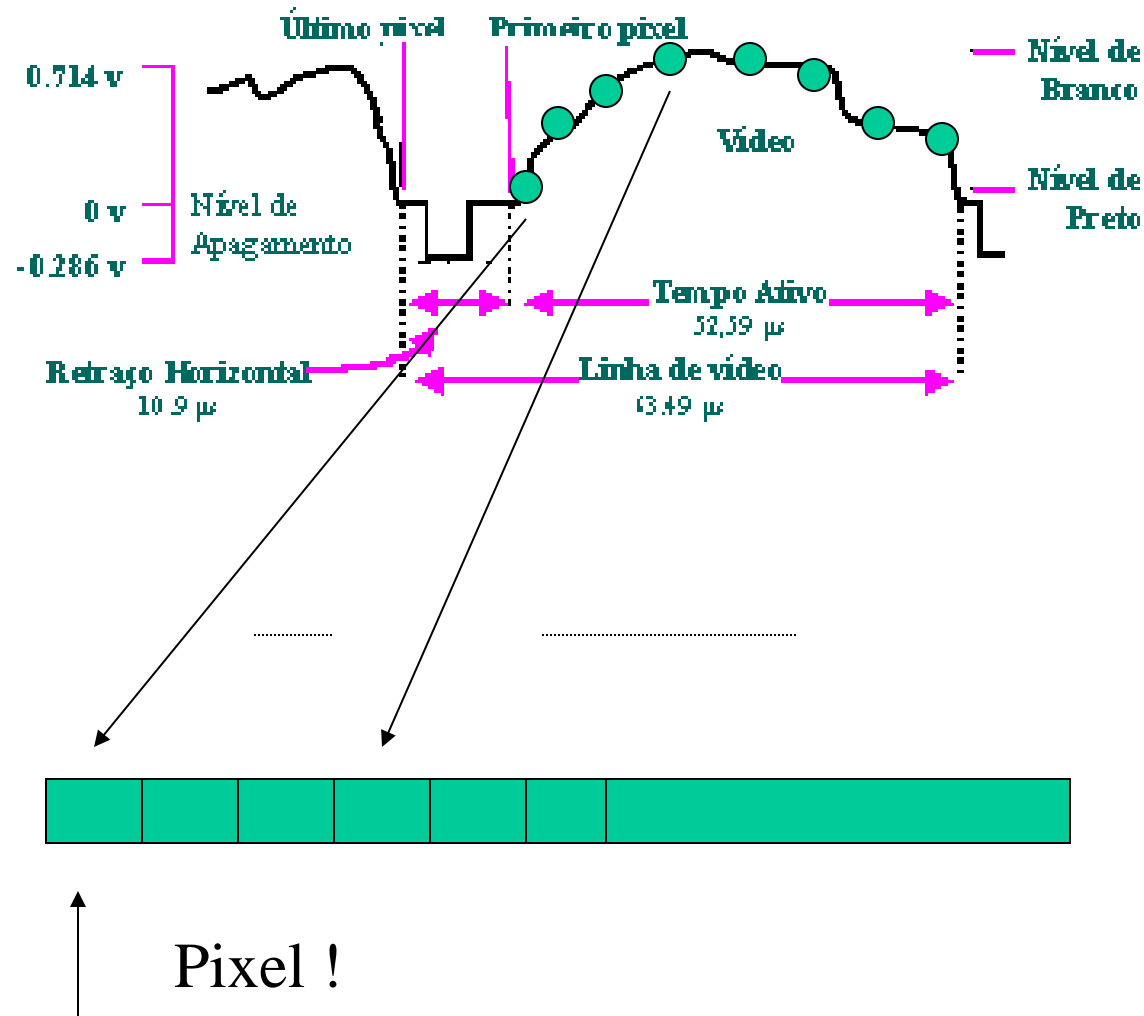
- <http://techterms.com/definition/halftone>
- <http://webstyleguide.com/wsg2/graphics/dither.html>
- <https://en.wikipedia.org/wiki/Dither>

● Digitalizadores de vídeo

- ☞ Resolução amarrada pelos padrões.
- ☞ Ruído mais elevado
- ☞ Permite captura de movimento
- ☞ Standard RS-170 (NTSC)
 - 525 linhas por quadro (485 ativas)
 - varredura entrelaçada 2:1
 - 30 quadros/s ou 60 campos/s
 - razão de aspectos 4:3
 - $485 \times 4/3 = 646.66$ pixels quadrados p/linha (640x480)
- ☞ Standard CCIR (PAL)
 - 625 linhas por frame (585 ativas)
 - varredura entrelaçada 2:1
 - 25 frames/s ou 50 campos/s
 - razão de aspectos 4:3
- ☞ HDTV
 - 1125 linhas por frame
 - varredura entrelaçada 2:1
 - 30 quadros/s ou 60 campos/s
 - razão de aspectos 16:9



Amostragem (*Sampling*)



Propriedades de uma imagem

- Vizinhança
 - 4-vizinhança: um pixel p tem 4 vizinhos
 - dois horizontais e dois verticais $N_4(p)$
 - 8-vizinhança: $N_8(p) = N_4(p) \cup N_d(p)$, em que $N_d(p)$ é o conjunto dos pixels na diagonal

Propriedades de uma imagem

- Conectividade
 - dois pixels estão conectados se são adjacentes segundo algum critério de vizinhança e se seus níveis de cinza satisfazem um critério de similaridade.
 - 4-conectado: p e q são similares e $q \in N_4(p)$
 - 8-conectado: p e q são similares e $q \in N_8(p)$

Propriedades: distâncias

- Sejam os pixels $p = (x,y)$, $q = (s,t)$, $z = (u,v)$. Uma função de distância D tem as propriedades:
 - $D(p,q) \geq 0$ ($D(p,q) = 0$, se e somente se $p = q$)
 - $D(p,q) = D(q,p)$
 - $D(p,z) \leq D(p,q) + D(q,z)$
- Distância Euclidiana $D(p,q) = \sqrt{(x-s)^2 + (y-t)^2}$
- Distância D_4 (city-block): $D(p,q) = |x-s| + |y-t|$
- Distância D_8 (Tabuleiro de xadrez)
 - $D(p,q) = \max(|x-s|, |y-t|)$

Operações aritméticas e lógicas

- Adição: $g = p + q$ (overflow ?)
- Subtração: $g = p - q$ (underflow)
- Correção por
 - corrigir escala: $g = 255 * (f - f_{\min}) / (f_{\max} - f_{\min})$
 - Truncamento: valor máx = 255 e valor min = 0

Operações lógicas

- AND
- OR
- XOR
- NOT
 - simples e bastante utilizadas.
 - <http://www.dai.ed.ac.uk/HIPR2/and.htm>

Alteração das dimensões

- Scaling
 - ampliar ou reduzir a imagem segundo um fator (igual para horizontal e vertical, ou não)
- Sizing (ou resizing)
 - diz-se o novo tamanho da imagem, ao invés de especificar o fator de ampliação/redução.

Ampliação: zoom in (um pixel se torna 4)

Redução: zoom out (4 pixels se tornam 1)

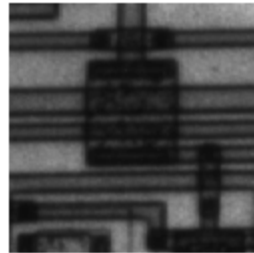
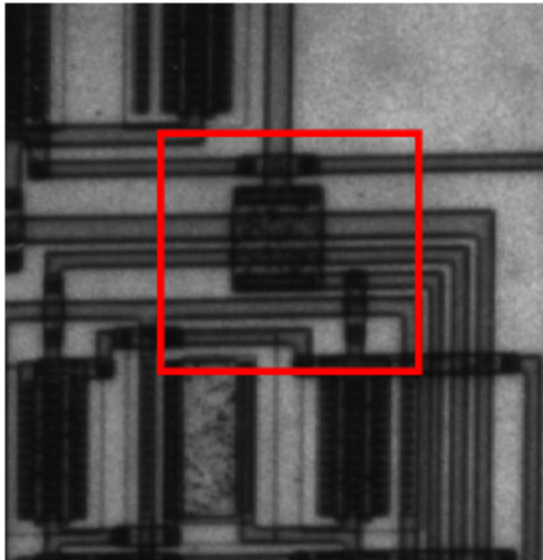
Transformações geométricas

- Seja o pixel de coordenada (x,y)
 - Translação: $(x',y') = (x+\Delta x, y + \Delta y)$
 - Rotação θ :
$$x' = x \cos (\theta) + y \sin (\theta)$$
$$y' = y \cos (\theta) - x \sin (\theta)$$
 - Espelhamento (*flip*): reflexão
 - *Warping*: projeção afim conforme um dado *template* ou imagem...

- <http://www.dai.ed.ac.uk/HIPR2/reflect.htm>

Crop, cut e paste

- *Cropping*: selecionar porção de uma imagem (*rubberband*)



Percepção Visual Humana

- Envolve componentes fisiológicos e psicológicos
- Por que estudá-lo?
 - Projetar algoritmos de compressão (reduzir qtd de informação, retendo informação visual)
 - algoritmos de realce de imagem (sabendo-se como funciona o sistema visual, pode-se aplicar técnicas que melhorem as imagens)

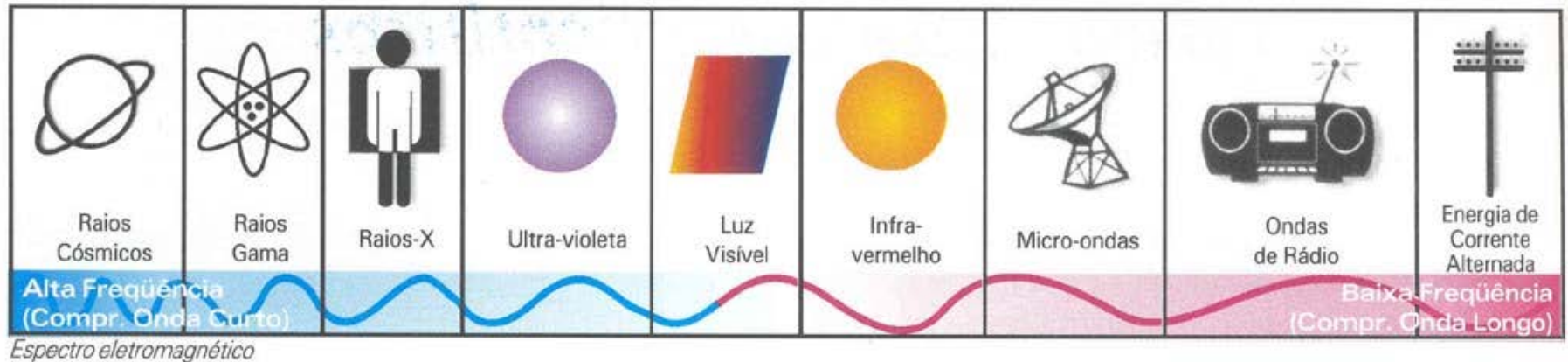
Sistema visual humano

- Energia luminosa focalizada pelas lentes do olho nos sensores da retina
- Estes sensores respondem à energia luminosa por uma reação eletroquímica que envia um sinal elétrico ao cérebro através do nervo óptico
- o cérebro usa esses sinais para criar padrões neurológicos que percebemos como imagens

Cores

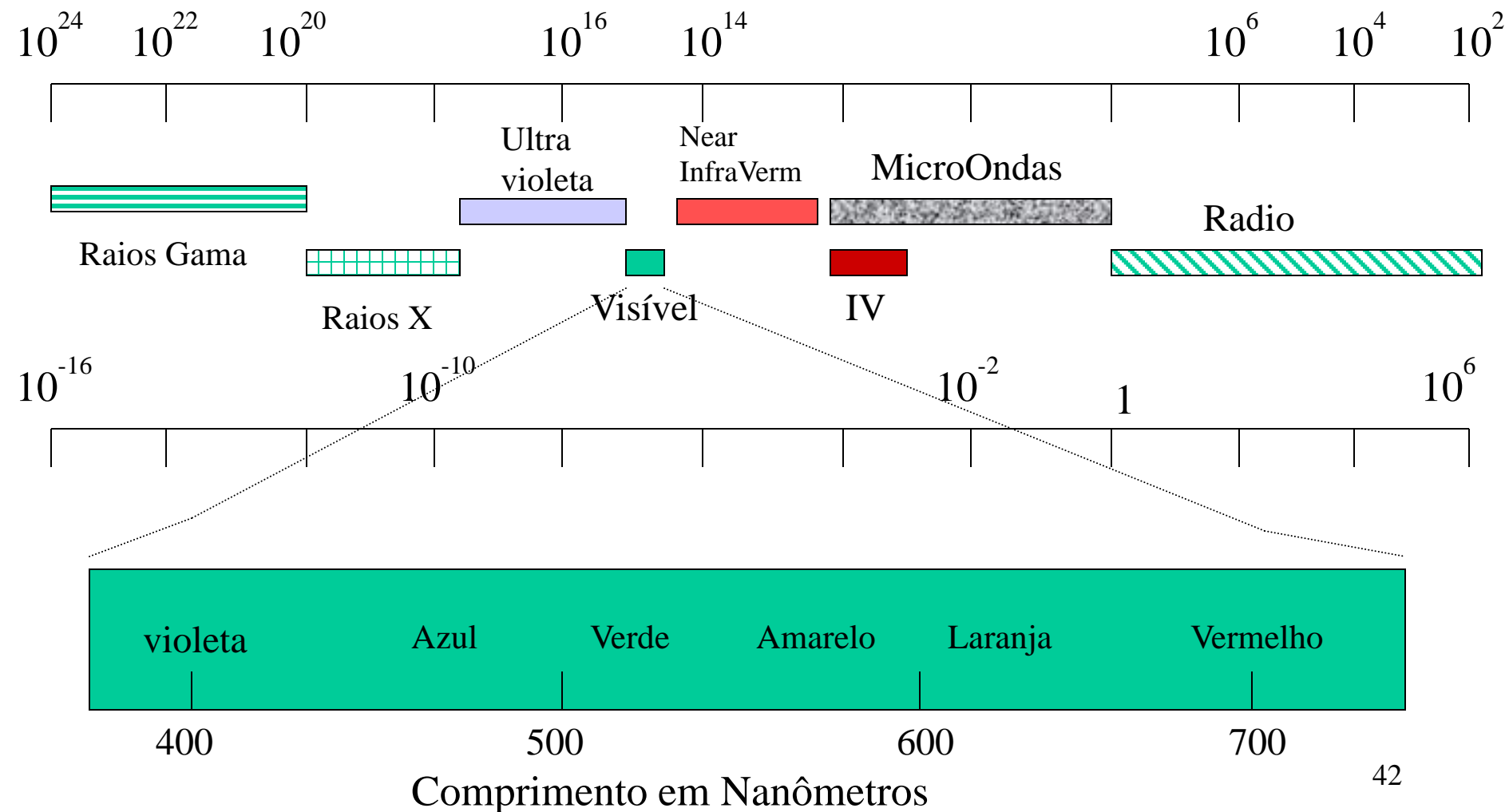
- Pode ser um poderoso descritor das propriedades de um objeto -> segmentação
- Humanos podem distinguir uma ampla variedade de nuances de cores, por outro lado, poucos tons de cinza distintos são perceptíveis (cerca de 100)
- *full color* ou pseudo-cor (falsa cor)

Luz “Visível”

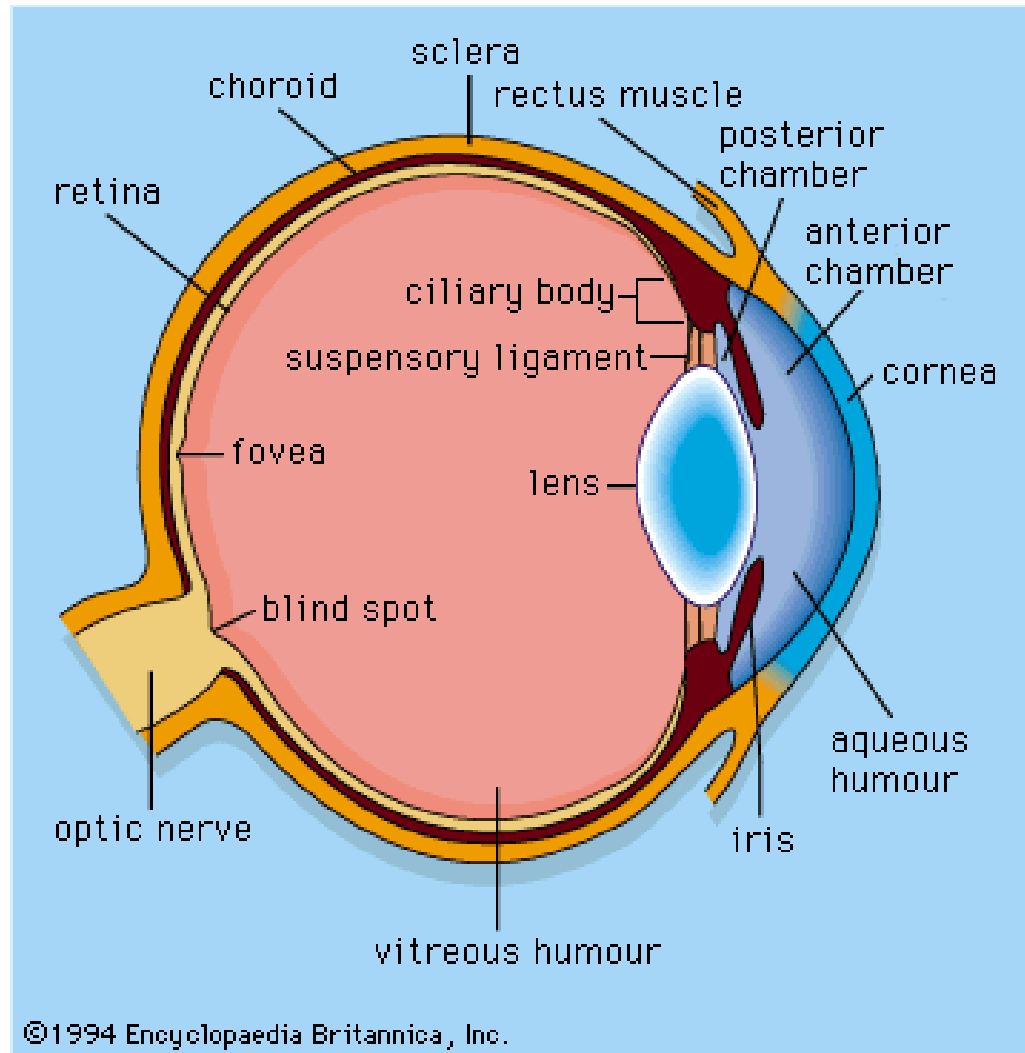


- Azul (400 - 500 nm)
- Verde (500-600 nm)
- Vermelho (600-700 nm)

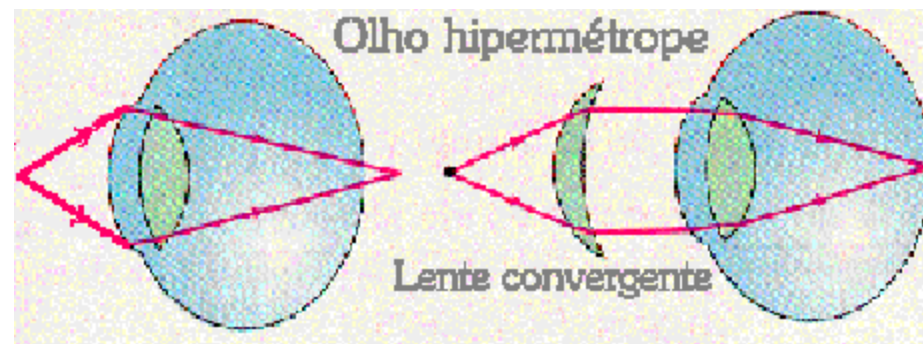
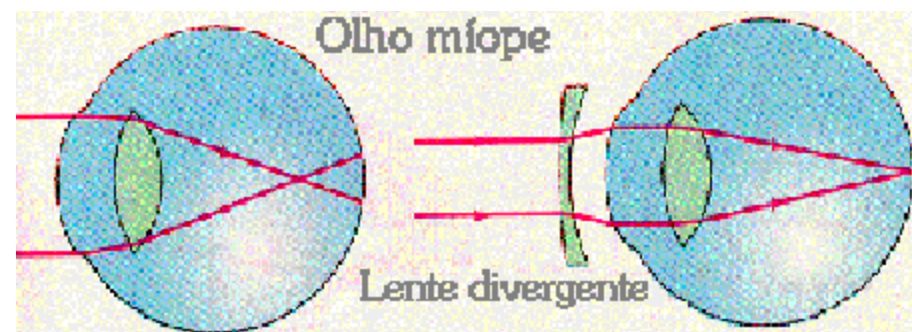
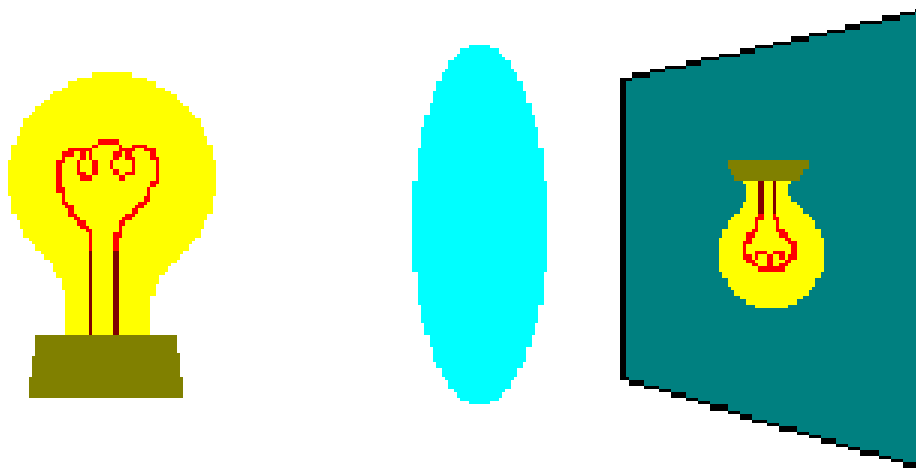
Espectro eletromagnético



Estrutura básica do olho



Formação da imagem



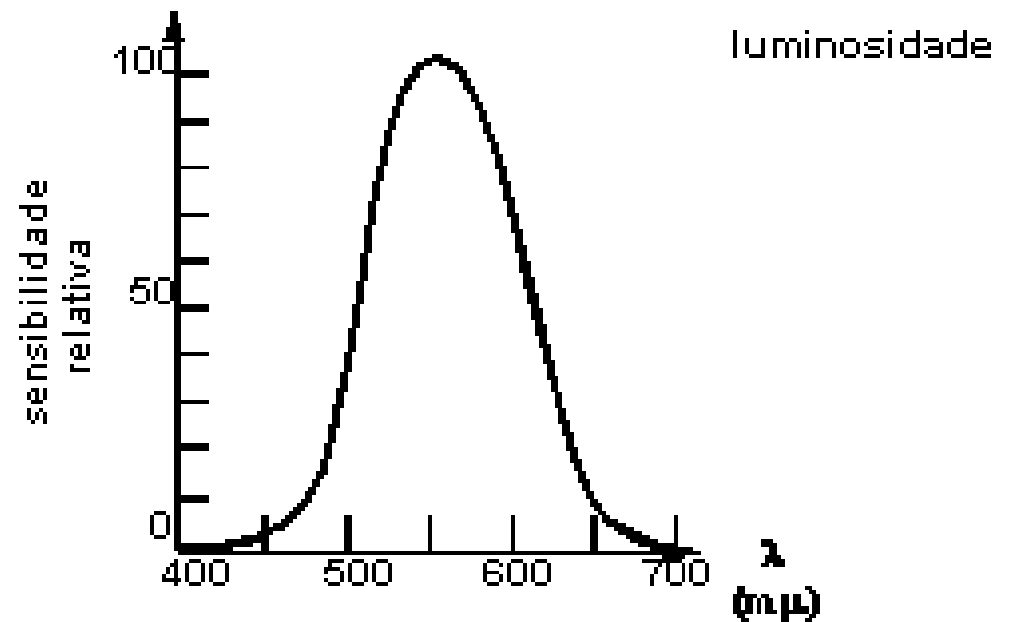
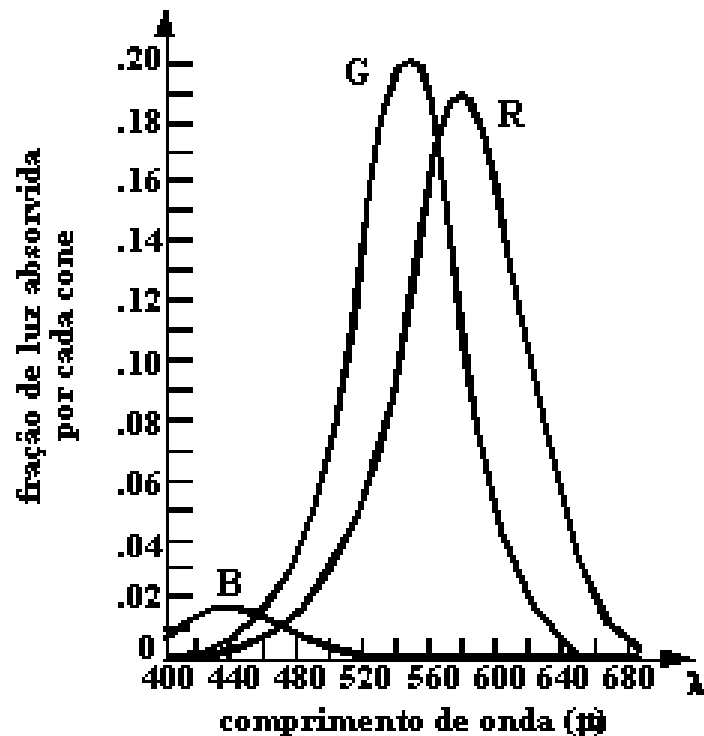
Características

- O olho é praticamente esférico: 20mm diâmetro
- córnea: transparente, cobre a superfície anterior
- esclera: opaca, dá continuidade à córnea
- retina: membrana mais interna onde a imagem é formada

A imagem formada deve-se a dois tipos de receptores de luz:

- cones, entre 6 e 7 milhões, localizados próximos a fóvea
 - são sensíveis à cor. Definem detalhes (visão diurna)
 - cada cone é conectado a um nervo óptico.
- bastonetes, entre 75 e 150 milhões, distribuídos por toda a retina
 - visão geral, panorâmica
 - vários bastonetes conectados a um único nervo
 - sensíveis a pouca iluminação (visão noturna)

Cones e Bastonetes



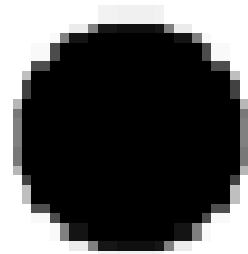
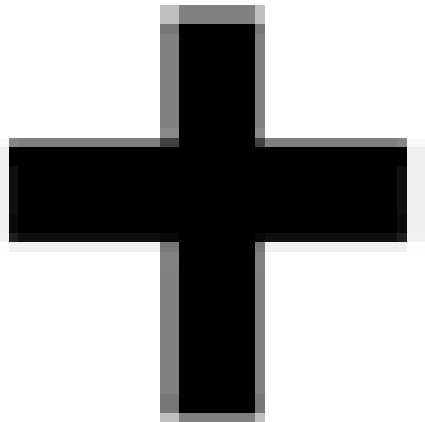
O olho humano percebe cor através do processo de tricomacia

Características (cont.)

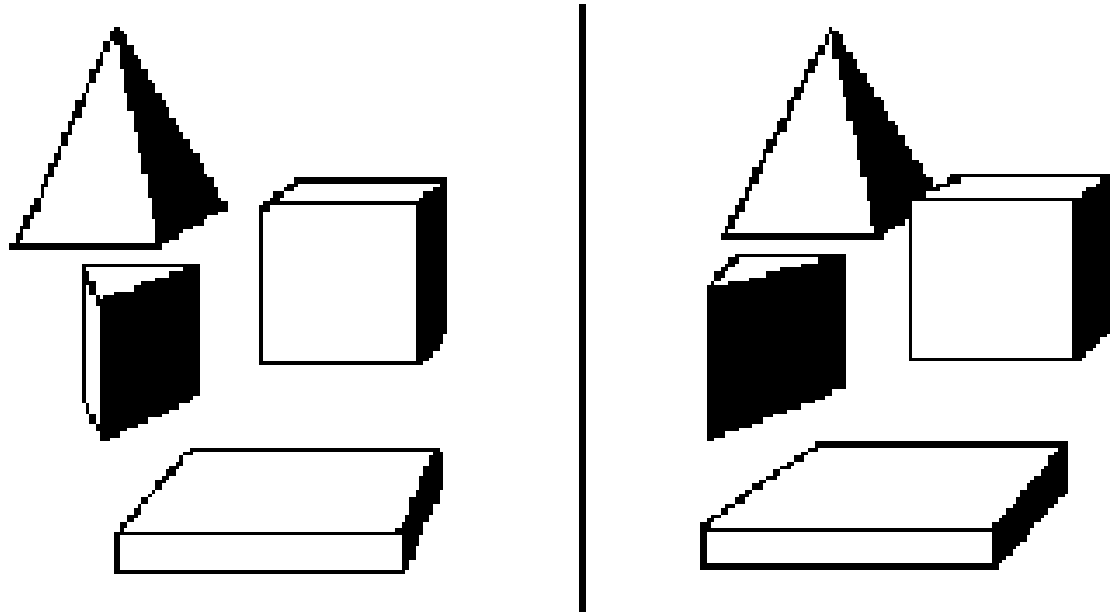
- as lentes (60-70% de água e 6% de gordura) são formadas por camadas de tecidos fibrosos. São flexíveis. O formato das lentes é controlado pelos *corpos ciliares*, músculos. Ficam mais grossas para focarem objetos próximos e mais relaxadas, achatadas, para focarem objetos distantes
- Ponto cego (*blind spot*): deslocado da região de maior

c

n



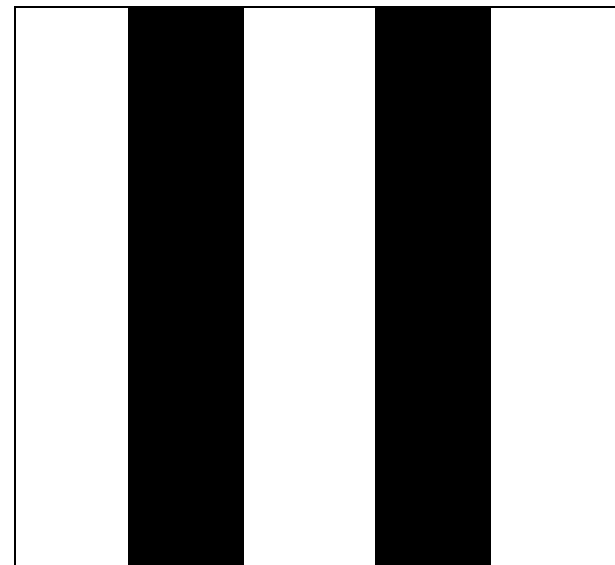
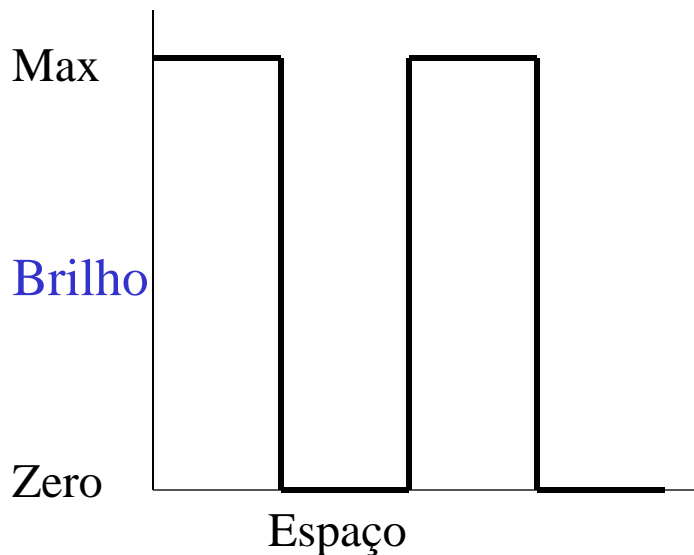
Visão Binocular



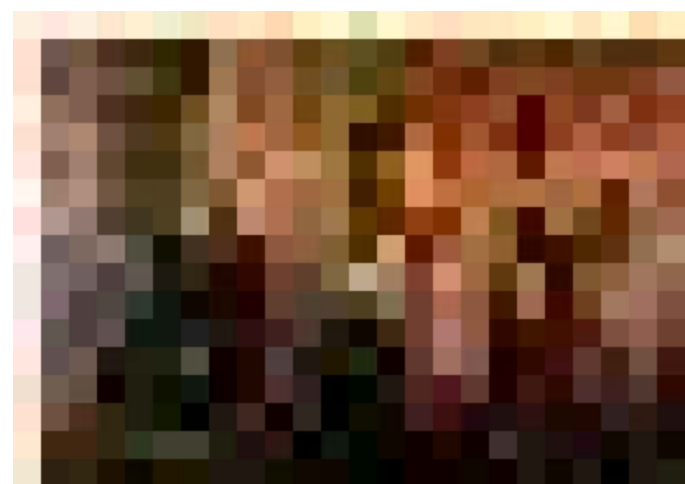
Imagens estereoscópicas

Resolução de frequência Espacial

- Resolução
 - habilidade em separar dois pixels adjacentes
- Frequência espacial
 - quão rapidamente o sinal muda no espaço



Resolução

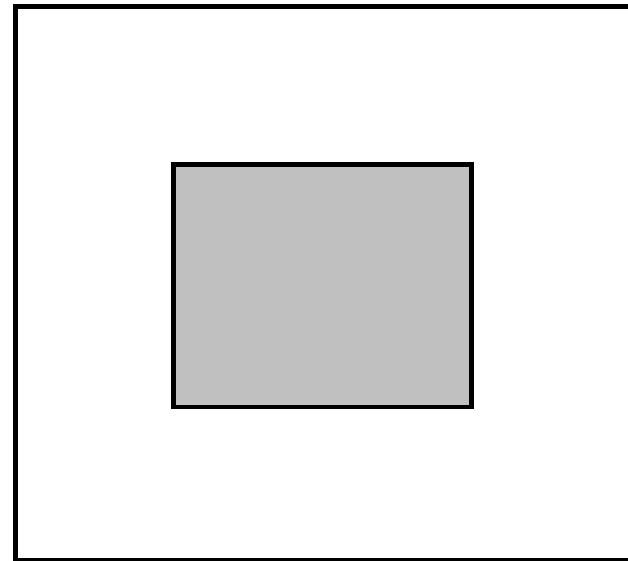
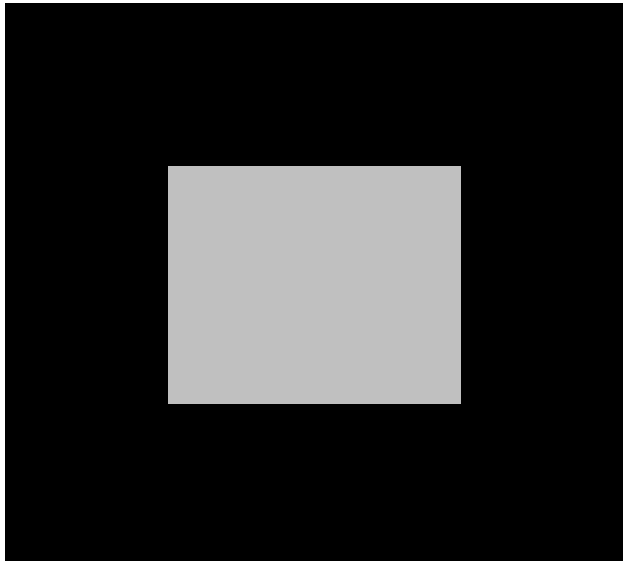


Topo, Esquerda (sentido horário): Original, 1/4, 1/8, 1/16 da original

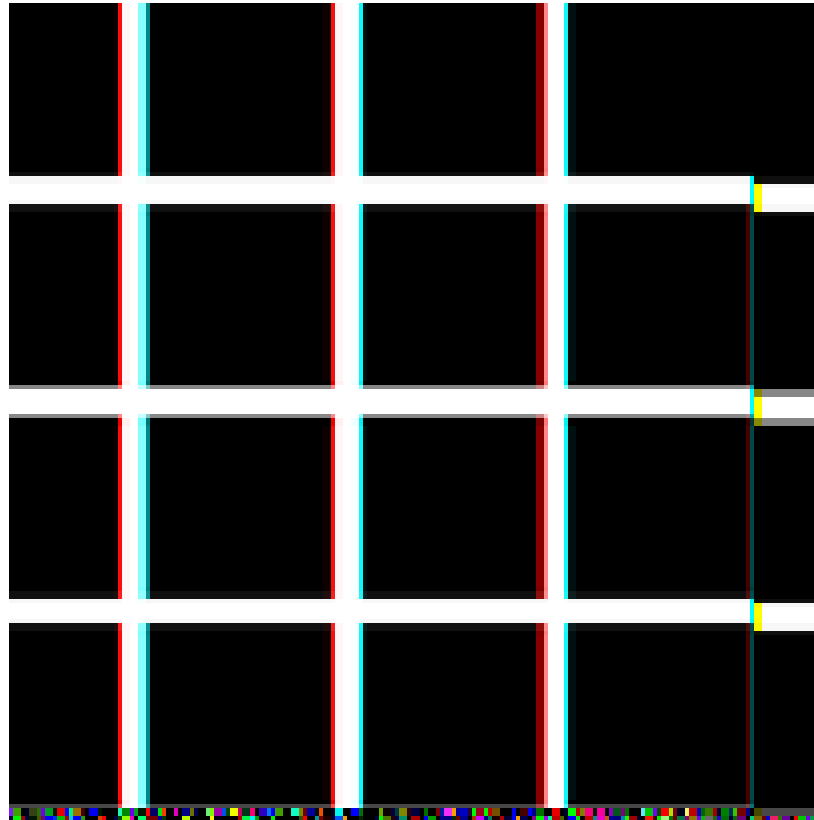
Adaptação ao brilho

- Espectro amplo de níveis de brilho
- Resposta depende da média de brilho observado e é limitado pelo limiar do escuro e o limite de clareza (*glare*)
- Adaptação não é instantânea (pupila)
- Distinguimos apenas 20 mudanças de brilho em uma área pequena de uma imagem complexa
- Imagem inteira: 100 níveis de cinza

Mudança do fundo e percepção de brilho



Contrastes



Manchas escuras nas junções dos quadrados ??

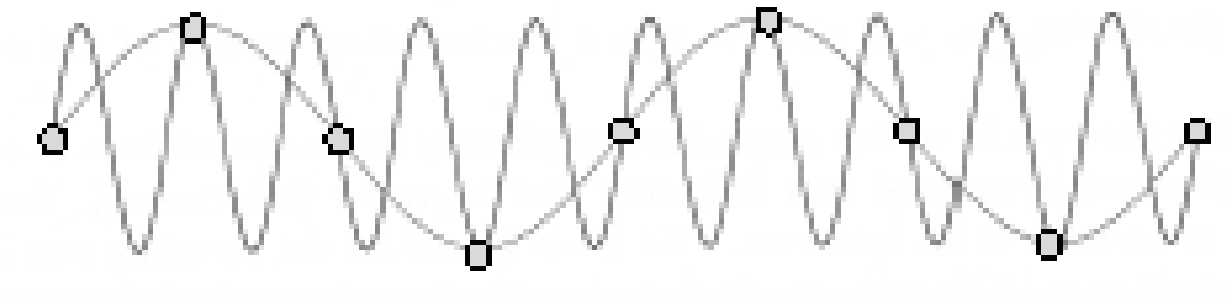
O fenômeno da Banda de Mach



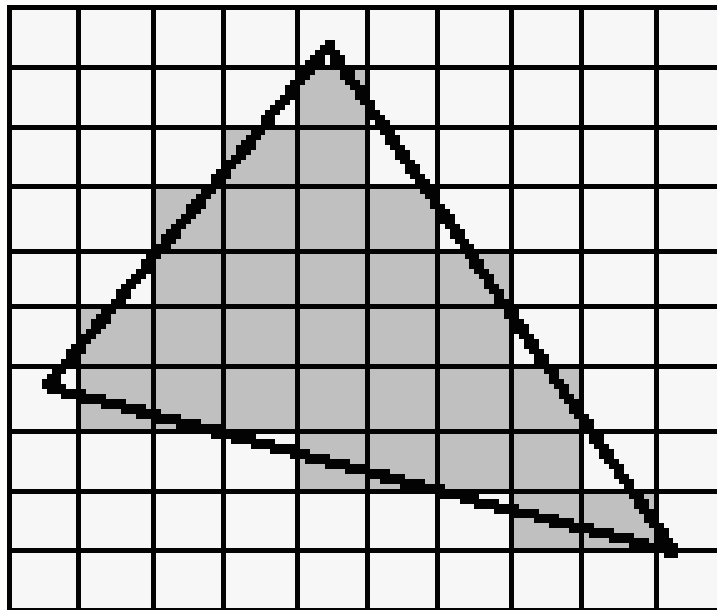
- Mudança brusca da intensidade
- Sistema visual realça (*overshoot*) as bordas (*edges*), criando um efeito escada

Aliasing

- Artefatos devido a sub-amostragem ou reconstrução ruim
 - Espacial
 - Temporal



Aliasing espacial



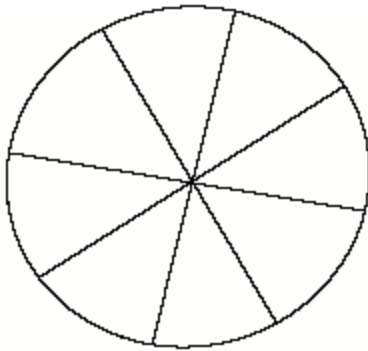
"Jaggies"

Jagged edges: efeito escada

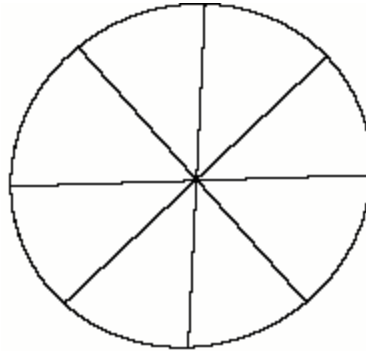
Aliasing Temporal

Efeitos de: *flickering* (monitor filmado na TV!)

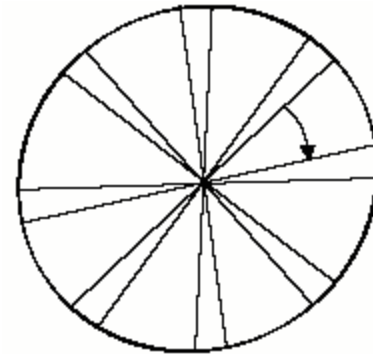
strobing (luz estroboscópica): roda do carro girando para trás na TV ou à noite !



t_1



t_2



strobing

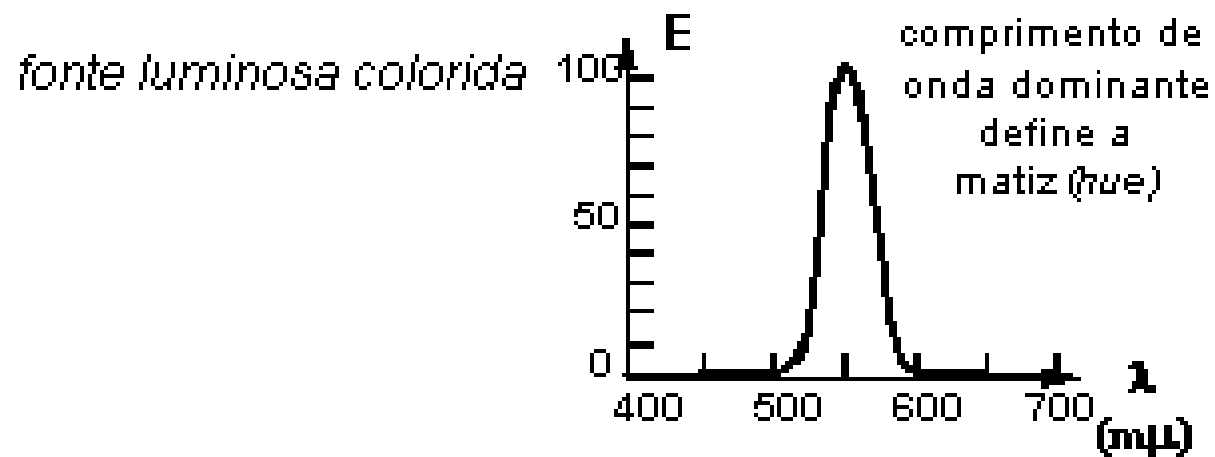
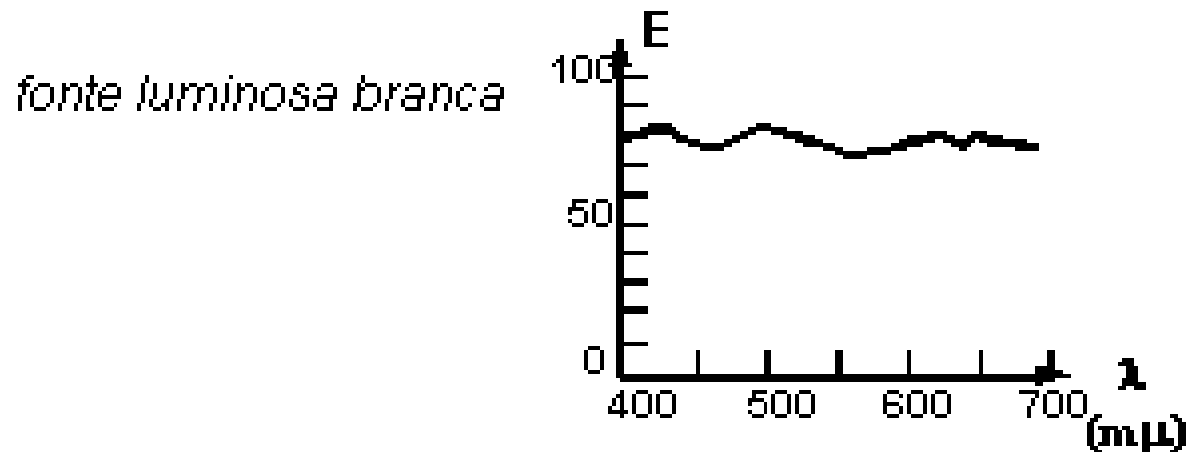
Representação de Imagem

- imagem = função bidimensional da intensidade de luz recebida

$I = f(x,y)$ taxa, nível de cinza proporcional ao brilho da imagem

- Binárias (0 - Preto, 1 - Branco)
- *Greyscale* (monocromáticas, one-colour)
- Coloridas...

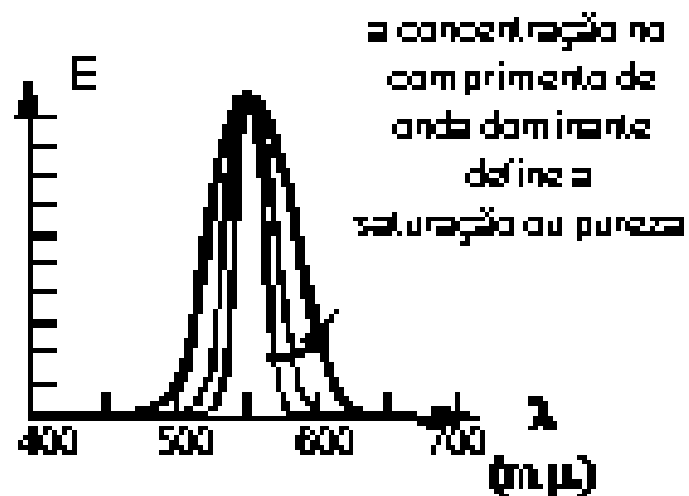
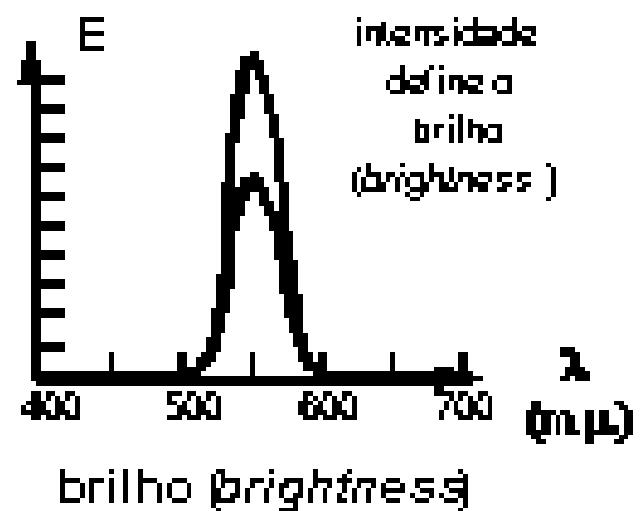
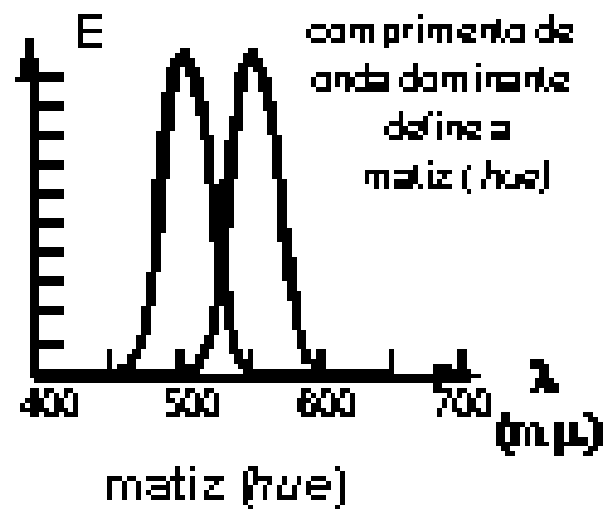
Cor



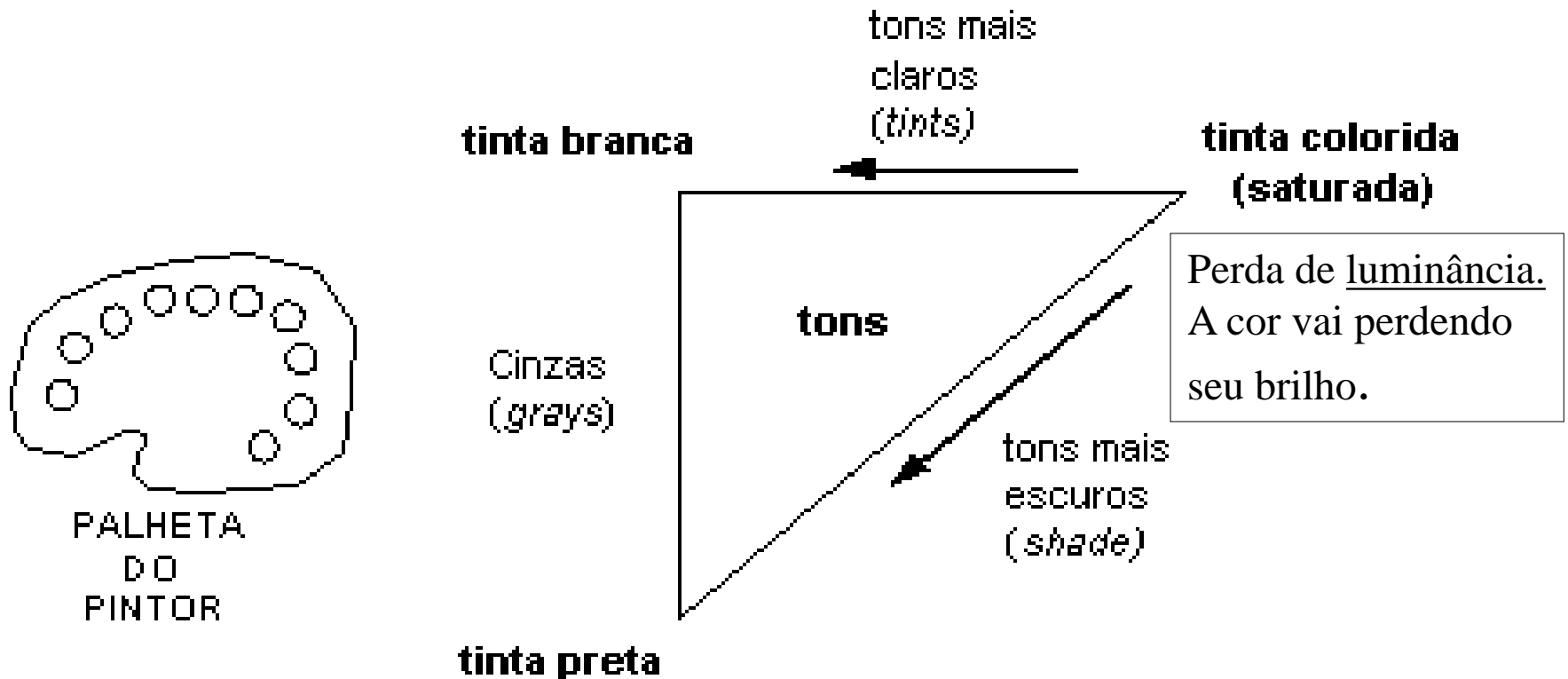
Cor

- Além do matiz, dois outros fatores caracterizam a cor
 - intensidade ou brilho (*brightness*): amplitude do comprimento de onda
 - saturação (pureza da cor): concentração em torno do comprimento de onda dominante

Cor

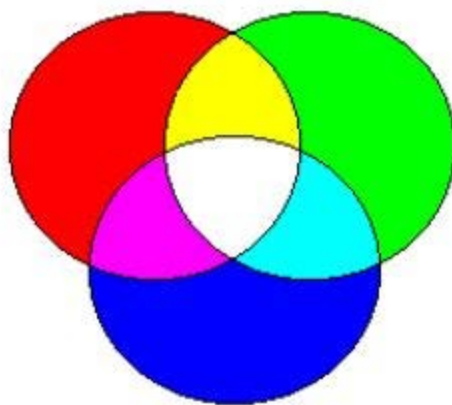
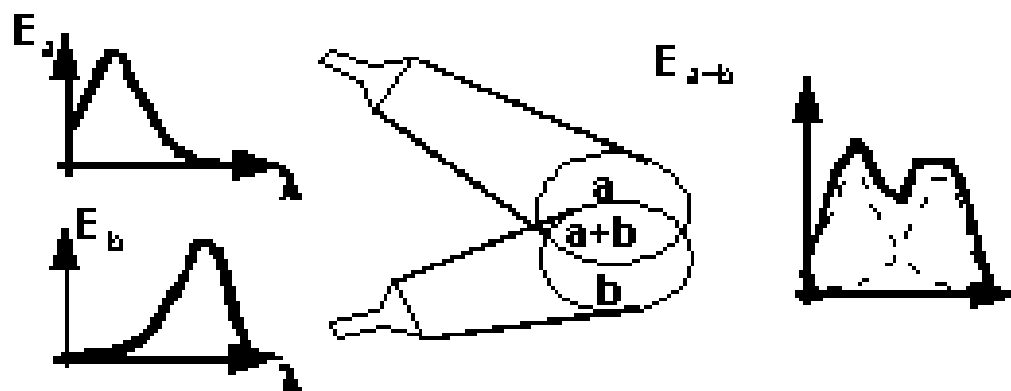


Cor na paleta de um pintor



Formação de cores por pigmentação

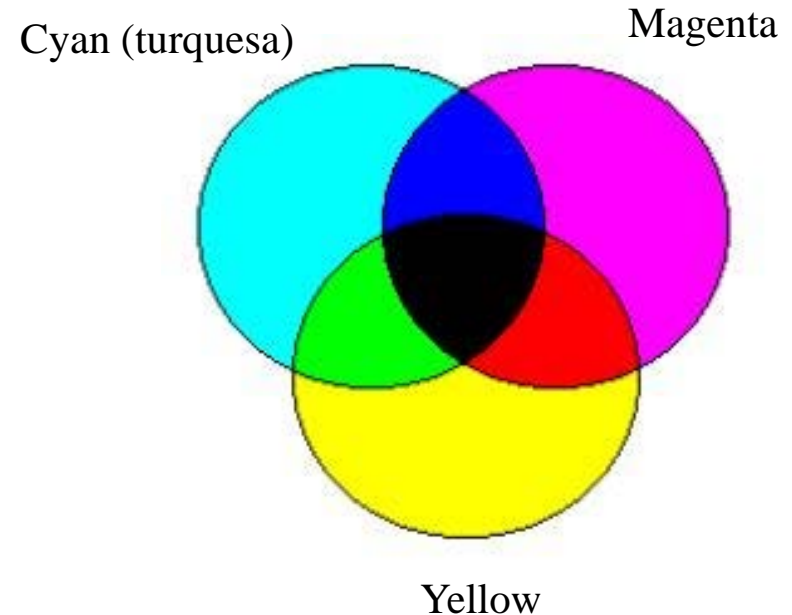
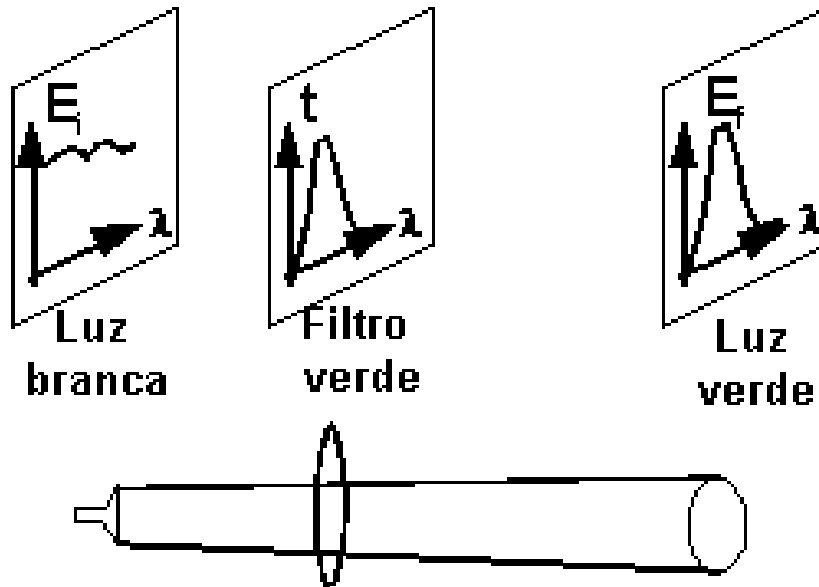
Cores por adição



Cores por adição: televisores

Cores por subtração

filtros ou corantes



Cores são geradas subtraindo-se o comprimento da onda dominante da luz branca, por isso, a cor resultante corresponde à luz que é refletida




Impressoras utilizam este sistema: <https://pt.wikipedia.org/wiki/CMYK>

Cor

- Luminosidade (*lightness*): associada à quantidade de luz refletida por um objeto
 - reflete mais -> mais branco
 - absorve mais -> mais escuro (preto)
 - depende do objeto (material)
- Brilho (*brightness*): depende da iluminação (emissor)

Luminosidade (preto-branco) e ***brilho*** (escuro-claro)

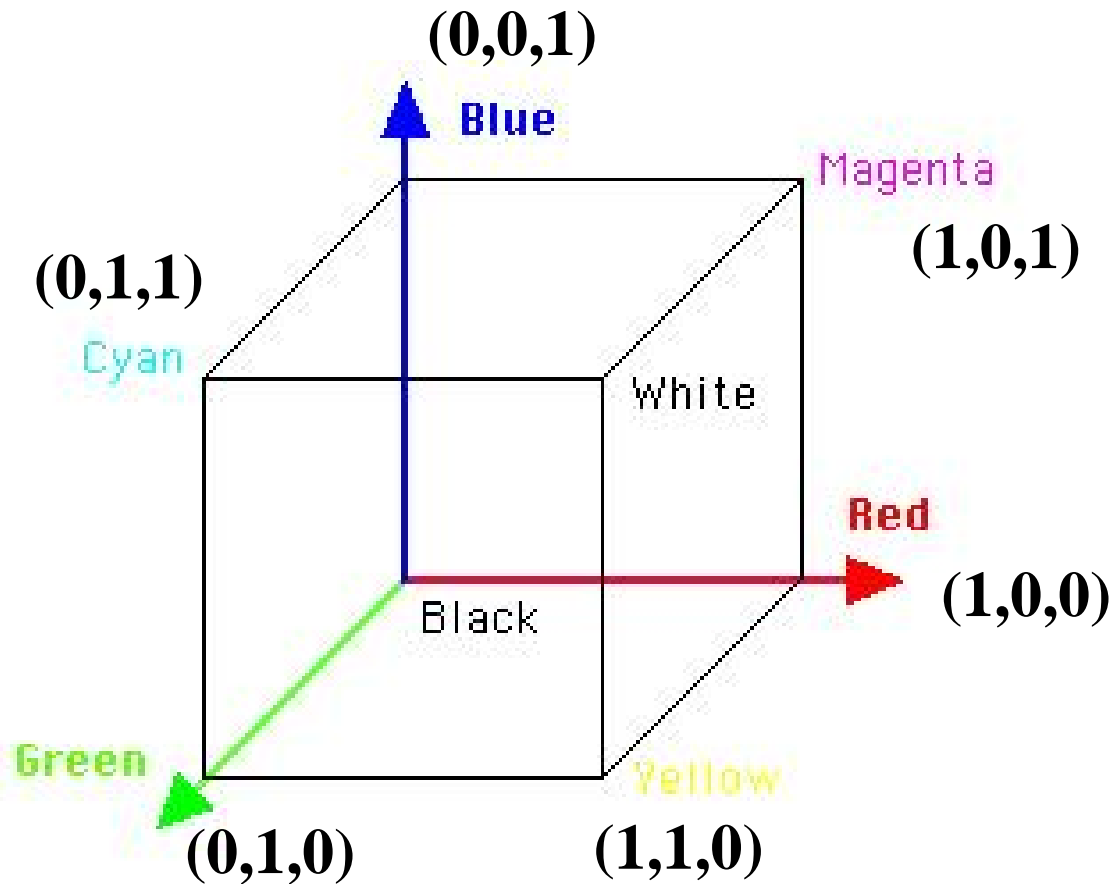
Imagens coloridas

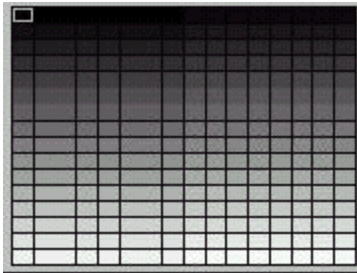
- Modeladas por 3 bandas monocromáticas de dados
- Cada banda  uma cor
- Dado  brilho do espectro de cada banda
- 8 bits por banda  24 bpp
- RGB nem sempre é “natural” ao homem

Imagens Coloridas

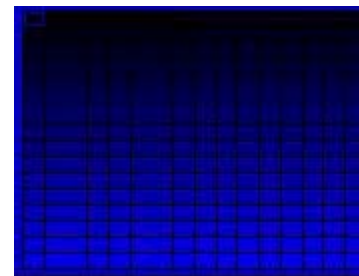
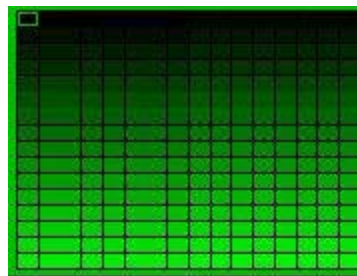
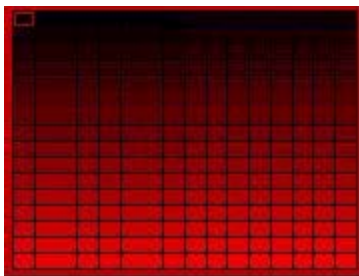
- *Teoria tricromática*: é possível formar uma cor arbitrária sobrepondo três cores básicas
- as 3 cores primárias (R, G, B) são projetadas numa região comum para reproduzirem a cor desejada
- quaisquer 3 cores podem ser primárias desde que elas sejam independentes entre si. Nenhuma pode ser mistura das outras duas.

Modelo RGB










Indexada: Pseudocolor



Indexada: True color: 3 LUT (look up tables)

- Indexado

o seu valor aponta para uma aquarela, uma tabela de cores
LUT (Look Up Table)

33	
34	
35	
36	
37	

Podem ser usadas uma ou mais tabelas

Ex: 3 tabelas x 256 tons x 8 bits (cada ton) = 16Mcores

- Não indexado (True Color)

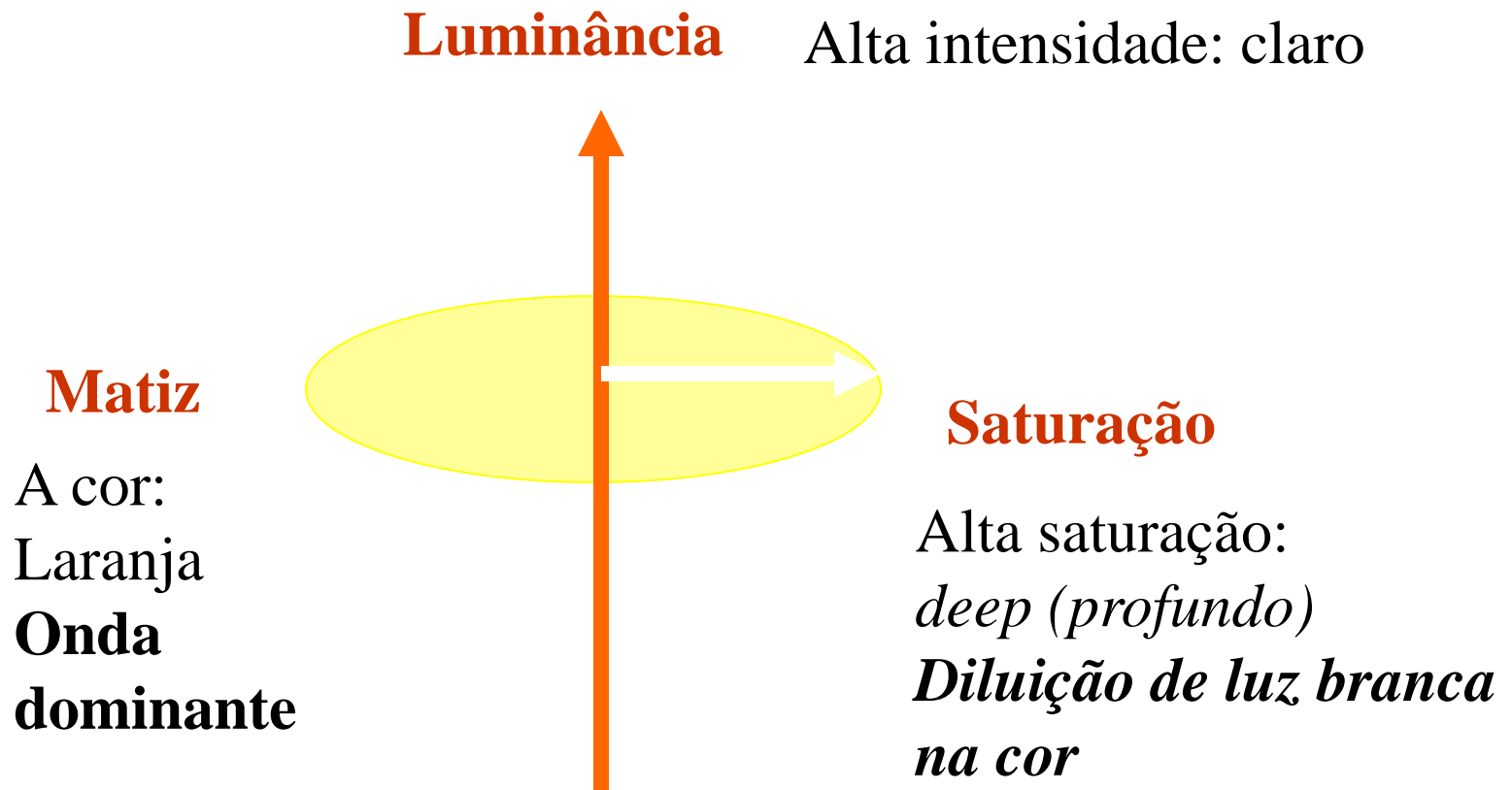
o valor do pixel representa o ponto (a cor)

1 byte =

alfa							
------	---	---	---	---	---	---	---

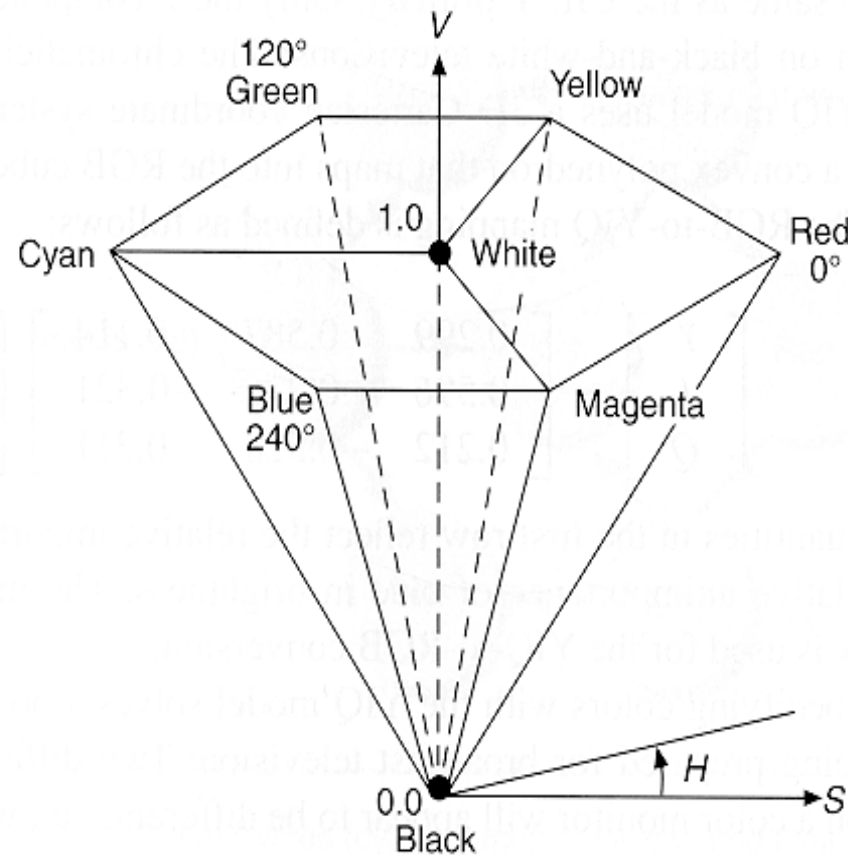
Modelo Perceptual

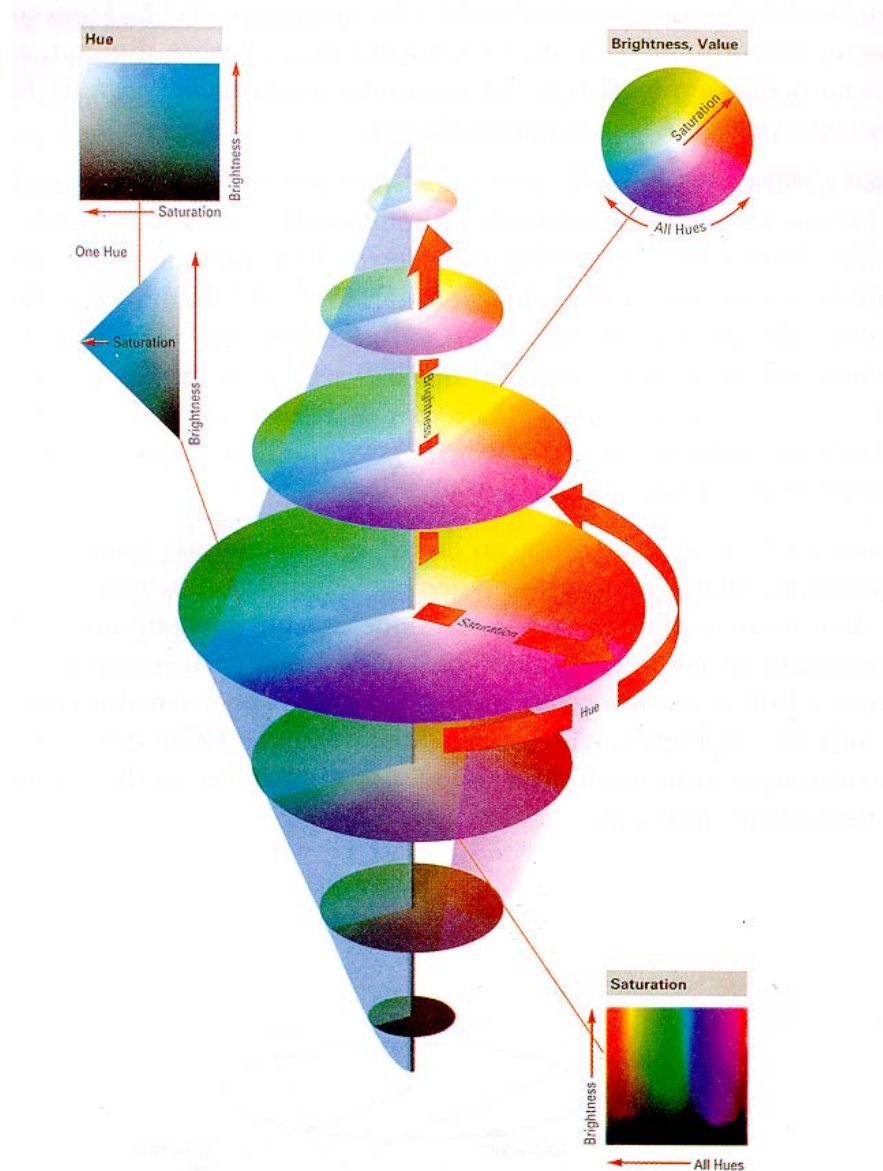
A deep, bright orange: um laranja claro e realçado (profundo)



RGB (A deep, bright orange): R-245, G-110, B-20

Modelo Perceptual HSV





Formatos de arquivos para imagens digitais

- Diversos formatos
 - requisitos diferentes, vários tipos de imagens
 - razões de mercado
 - software proprietário
 - falta de coordenação da indústria de imagens
- Computação Gráfica
 - imagens bitmap e imagens “vetores” (*vector images*)

Formatos

- Bitmap (*raster images*)
 - $I(r,c)$: pixel armazena o valor do brilho.
- Vector images
 - representação de linhas, curvas e formas
 - armazena os “pontos chaves”
 - rendering (renderização)

Imagens Bitmap

- BIN (*raw image*, sem cabeçalho)
- PPM (raw data + cabeçalho simples)
 - PBM(binário), PGM(grey-scale), PNM (ambos)
- TIFF (Tagged Image File format) 24 bpp + compressão (LZW, RLE, JPEG)
- GIF (Graphics Interchange Format) 8 bpp + compressão LZW.

Imagens Bitmap

- JPEG: Joint Photographic Experts Group
- JFIF: JPEG Interchange Format
 - utiliza compressão jpeg
- Sun raster (Sun)
- SGI (Silicon Graphics)