Fundamentos

Capítulo 2 - Gonzalez e Woods
Hitoshi

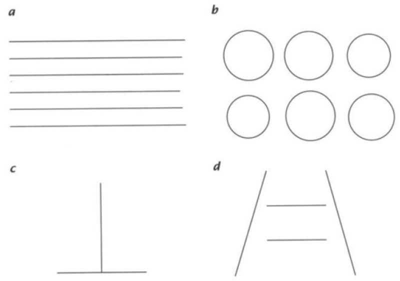
Sistema visual humano

- Vimos que o processo começa na aquisição do sinal e termina na interpretação
- Para que não haja falhas na interpretação, é necessário considerar também as características do sistema visual humano

Que características?

- Qual a diferença de intensidade que conseguimos distinguir
- Qual a resolução espacial do olho?
- Com que precisão nós estimamos e comparamos distâncias e áreas?
- Como nós percebemos cor?
- Que características nós usamos para detectar e distinguir objetos?

Imagens de teste



Testes de distância e área

a) linhas paralelas com até 5% de diferença no comprimento

b) Círculos com até 10% de diferença no raio

c) A linha vertical parece maior

d) perspectiva: a linha superior parece maior

Estrutura do olho humano

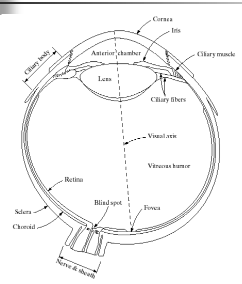


FIGURE 2.1 Simplified diagram of a cross section of the human eye.

- Forma é aproximadamente esférica
- Diâmetro médio: 20mm
- Possui 3 membranas:
-- Córnea e esclera
-- Coróide
-- Retina

Cones e bastonetes

- São células sensíveis a luz, contidas na retina
- Há três tipos de cones, localizados na fóvea, e são responsáveis pela visão fotópica (ou de luz intensa)
- Os bastonetes são responsáveis pela visão periférica (visão escotópica ou de baixa intensidade).

Distribuição dos cones e bastonetes

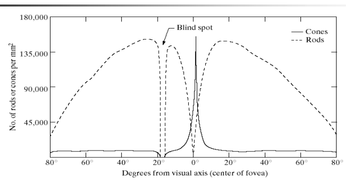
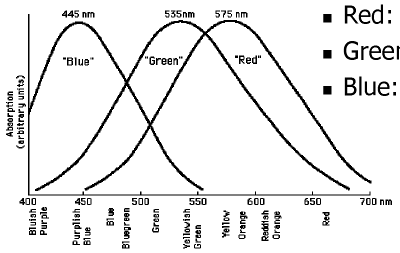


FIGURE 2.2 Distribution of rods and cones in the retina.

- Diâmetro da fóvea: 1.5mm
- cones: 6 a 7 milhões
- bastonetes: 120 milhões
- Ponto cego: nervo ótico se conecta com a retina

Cones



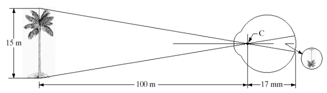
- Red: 64%
- Green: 32%
- Blue: 4%

Bastonetes

- 1000 vezes mais sensíveis à luz que os cones
  - capaz de reagir a um único fóton sob condições ideais
- Adaptação leva em torno de 10 minutos
- A sensitividade dos bastonetes é deslocada para um comprimento mais curto que os cones, mais sensível ao brilho de folhas verdes ao luar
- São mais sensíveis ao movimento

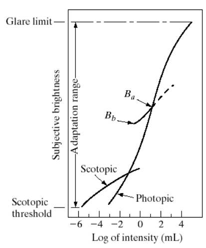
Formação da imagem

FIGURE 2.3 Graphical representation of the eye looking at a palm tree. Point C is the optical center of the lens.



- Comprimento focal: 14mm (perto) a 17mm (longe)
- 15 / 100 = h / 17 => h = 2.55mm

Adaptação à intensidade luminosa



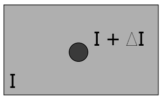
Intensidade subjetiva (percebida pelo sistema visual humano) é uma função logarítmica da intensidade luminosa.

A sensitividade do olho se adapta, por exemplo, em Ba

Bb = black

Sensibilidade ao contraste

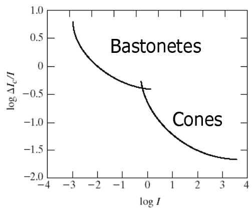
- A sensibilidade do olho à variações de contraste é muito importante
- I é uma intensidade luminosa uniforme
- ΔI é a alteração da iluminação, necessário para se notar uma diferença



Relação de Weber: $\Delta I/I$

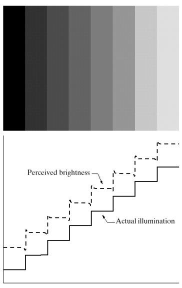
se a relação é alta, a discriminação é pobre

Discriminação de intensidade



- A discriminação de brilho é pobre quando a iluminação é baixa, e melhora muito quando a iluminação de fundo aumenta.
- Caso a iluminação seja constante ao invés de pulsar, o número de intensidades percebidas é bem menor (12 a 24)

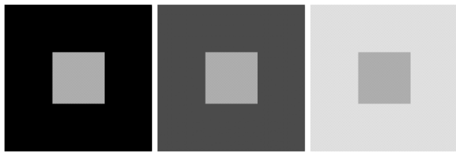
Bandas de Mach (1865)



a
b

FIGURE 2.7 (a) An example showing that perceived brightness is not a simple function of intensity. The relative vertical positions between the two profiles in (b) have no special significance; they were chosen for clarity.

Contraste simultâneo

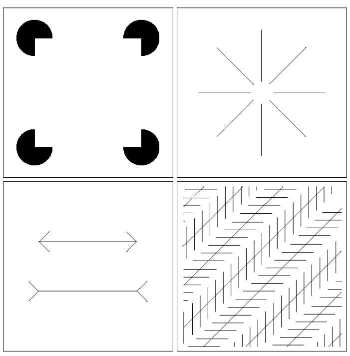


a b c

FIGURE 2.8 Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

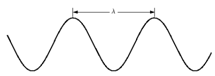
a b
c d

FIGURE 2.9 Some well-known optical illusions.



Luz e ondas eletromagnéticas

FIGURE 2.11 Graphical representation of one wavelength.



- Relação entre comprimento de onda */* e freqüência *f*
  / = *c* / *f*   onde *c* = velocidade da luz
- Energia *E* = *h f*   onde *h* = cte Planck

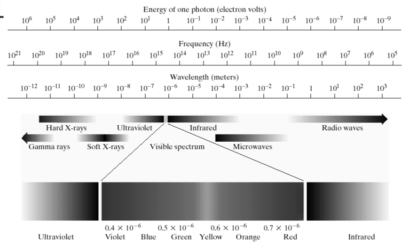
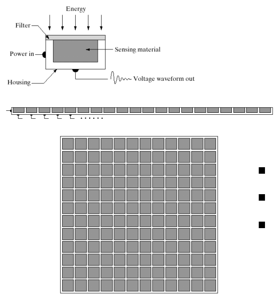


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

Aquisição de imagens

a
b
c
FIGURE 2.12
(a) Single imaging sensor.
(b) Line sensor.
(c) Array sensor.



- Único sensor
- Vetor (linha)
- Array (matriz)

Sinais

- Um sinal é uma função que carrega informação
- É comum que o valor de um sinal varie no espaço (tenha uma extensão) e no tempo (tenha uma duração).
  - Sinais espaço-temporais.

Sinais variantes no tempo

- Sinais que variam no tempo podem ser representados por

  f (t)
- exemplo: sinal de áudio
- esse sinal pode ser considerado como uma coleção de vários tons de freqüências diferentes que variam ao longo do tempo.

Sinais variantes no espaço

- Sinais também podem variar no espaço
- Uma imagem pode ser considerada como uma função de 2 dimensões espaciais:

  f(x,y)
- para imagens monocromáticas, o valor da função é a quantidade de luz naquele ponto
- Equipamentos médicos de tomografia (CAT) e ressonância magnética produzem imagens com 3 dimensões espaciais ***f(x,y,z)***

Sinais espaço temporais

- O que você imagina que seja um sinal da forma:

  f(x, y, t)
- onde x e y são coordenadas espaciais
- e t = tempo
- sinal que varia no espaço-tempo, como vídeo

Tipos de sinais

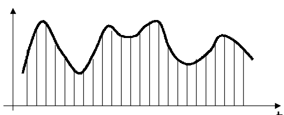
- a maior parte dos sinais que ocorrem naturalmente são funções definidas em um domínio contínuo
- Porém, para serem representadas em um computador, os sinais precisam ser definidos de forma discreta

Analógico x Digital

- a maior parte dos sinais naturais também variam continuamente, possuindo valores de precisão infinita
- No computador, a precisão deve ser finita, e assim a variação é discreta
- Sinais analógicos possuem domínio e variação contínuos
- Sinais digitais possuem domínio e variação discretos

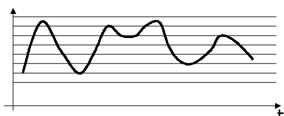
Amostragem

- amostragem = espaçamento entre valores discretos no domínio do sinal
- taxa de amostragem = quantas amostras são colhidas por unidade em cada dimensão? Exemplo: amostras por segundo.



Quantização

- Quantização = espaçamento entre valores discretos dos valores do sinal
- pode ser considerado como o número de bits usado para representar o valor de uma amostra.



Amostragem e quantização

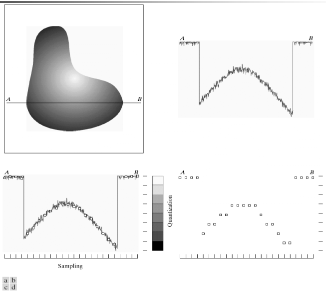
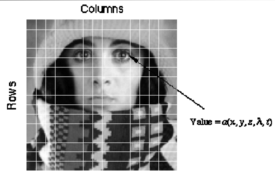


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

Representação de imagens digitais



- Uma imagem digital f(x,y) é uma imagem que foi digitalizada espacialmente e na intensidade de seus pixels
- O valor de um pixel é proporcional a intensidade luminosa (ou nível de cinza) daquele ponto

Exemplo

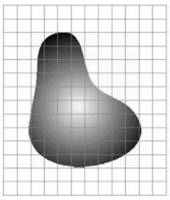


Image contínua projetada sobre o sensor

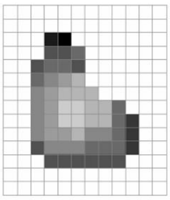


Imagem digitalizada resultado da amostragem e quantização

Função de intensidade de luz

- imagens são funções 2D de intensidade luminosa, f(x,y)
- A amplitude de f na coordenada espacial (x,y) fornece a intensidade (brilho) da imagem nesse ponto
- como a luz é uma forma de energia, ela deve ser maior que zero e finita

  $0 < f(x,y) <$ infinito

Iluminação e reflectância

- A f(x,y) pode ser caracterizada por duas componentes:
  - a quantidade de luz incidente na cena sendo observada => Iluminação i(x,y)

    $0 < i(x,y) <$ infinito
  - a quantidade de luz refletida pelos objetos na cena => reflectância r(x,y)

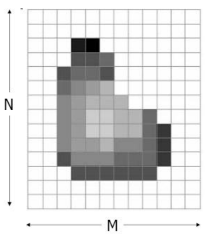
    $0 < r(x,y) < 1$

$f(x,y) = i(x,y) . r(x,y)$

Nível de cinza

- A intensidade de um pixel de uma imagem monocromática na posição (x,y) é chamada de nível de cinza (l) daquela posição
  - Lmin <= l <= Lmax
  - Lmin é não negativo e Lmax é finito
  - escala de cinza = [Lmin, Lmax]
  - usualmente, desloque a escala para [0, L]
  - 0 = preto e L = branco

Número de bits



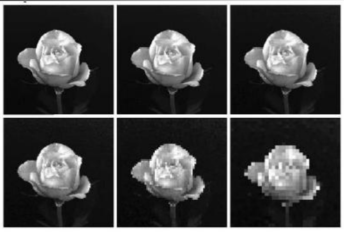
- O número de bits dos níveis de cinza é uma potência de 2
  - $L= 2^k$
- Número de bits necessário para armazenar uma imagem digital:

  $b= M \times N \times k$

Resolução

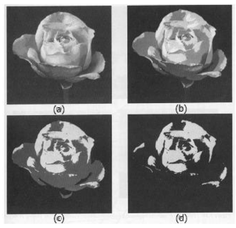
- Resolução (quantos detalhes da imagem você consegue observar?) depende da amostragem e dos níveis de cinza
- quanto maior a taxa de amostragem e a escala de cinza, melhor será a aproximação da imagem digital da imagem original
- quanto maior a amostragem, maior o tamanho da imagem (número de pixels)

Efeito de blocos



- imagens variam de resolução de 1024x1024 a 32x32
- Quando a resolução cai muito, o efeito de blocos aparece

Falsos contornos



- a) 16 níveis de cinza
- b) 8 níveis
- c) 4 níveis
- d) 2 níveis
- com poucos níveis de cinza, as áreas de pouco contraste são mais afetadas

Amostragem não uniforme

- Para um valor fixo de resolução espacial, a aparência da imagem pode ser melhorada através de taxas de amostragem adaptativas
  - amostragem fina: necessária nas vizinhanças de transições abruptas no nível de cinza
  - amostragem grossa: utilizada em regiões relativamente suaves, de pouco contraste

Exemplo

a b c

FIGURE 2.22 (a) Image with a low level of detail. (b) Image with a medium level of detail. (c) Image with a relatively large amount of detail. (Image (b) courtesy of the Massachusetts Institute of Technology.)

Exemplo

- a imagem de uma face sobre um fundo uniforme
  - fundo: pouco detalhe => amostragem grossa é suficiente
  - face: mais detalhes => amostragem fina
- se pudermos utilizar amostragem adaptativa, a qualidade da imagem é melhorada
- em particular, as regiões de alto contraste (contornos ou bordas) são importantes => transições abruptas entre objeto e fundo

Quantização não uniforme

- O número de níveis pode ser reduzido se eles forem espaçados de forma não uniforme
  - use poucos níveis de cinza nas vizinhanças das bordas. Por que?
    - pois nossa visão é relativamente fraca em observar sombreamentos nas vizinhanças de bordas
  - use mais níveis de cinza em regiões suaves para evitar o efeito de falsos contornos

Relações básicas entre pixels

- vizinhança de um pixel
- conectividade
- adjacência
- medidas de distância
- operações lógicas e aritméticas

Vizinhança de um pixel

- um pixel p na coordenada (x,y) tem:
  - $N_4(p)$: vizinhança-4 de p
    (x-1,y), (x+1,y), (x,y-1), (x,y+1)

	x	
x	p	x
	x	

  - $N_D(p)$: vizinhança diagonal-4 de p
    (x-1,y-1), (x-1,y+1), (x+1,y-1), (x+1,y+1)

x		x
	p	
x		x

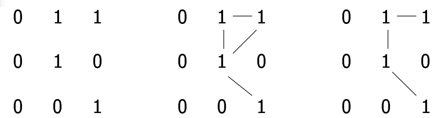
  - $N_8(p)$: vizinhança-8 de p
    $N_4(p) + N_D(p)$

x	x	x
x	p	x
x	x	x

Conectividade

- Seja V o conjunto de níveis de cinza utilizados na definição de conectividade
  - conectividade-4:
    - 2 pixels p e q com valores em V são conexos-4 se q está no conjunto $N_4(p)$
  - conectividade-8:
    - 2 pixels p e q com valores em V são conexos-4 se q está no conjunto $N_4(p)$
  - conectividade-m (misturada)
    - 2 pixels p e q com valores em V são m-conectados se
      - q está no conjunto $N_4(p)$ ou
      - q está no conjunto $N_D(p)$ e o conjunto $N_4(p) \cap N_4(q)$ é vazio
      - conjunto de pixels que são vizinhos-4 tanto de p quanto q e cujos valores estão em V

Exemplo



arranjo de pixels | conectividade-8 | conectividade-m

- conectividade-m elimina caminhos múltiplos possíveis na conectividade-8

Adjacência

- um pixel p é adjacente a um pixel q se eles são conexos
- duas regiões na imagem R1 e R2 são adjacentes se um pixel p em R1 for adjacente a um pixel q em R2

Percurso (path)

- um percurso do pixel p com coordenadas (x,y) ao pixel q com coordenadas (s,t) é uma seqüência distinta de pixels com coordenadas
  - (x_0,y_0), (x_1,y_1), ..., (x_n, y_n), onde
    - $(x_0,y_0) = (x,y)$
    - $(x_n, y_n) = (s,t)$
    - (xi, yi) é adjacente a (yi-1, yi-1)
- n é o comprimento do caminho
- pode-se definir 4, 8 ou m-percursos dependendo do tipo de adjacência

Exercício

- Considere as duas regiões S1 e S2

```
            S1              S2
    ---------------------------------
  0 | 0  0  0  0 | 0  0  1  1 | 0
  1 | 0  0  1  0 | 0  1  0  0 | 1
  1 | 0  0  1  0 | 1  1  0  0 | 0
  0 | 0  1  1  1 | 0  0  0  0 | 0
    ---------------------------------
  0   0  1   1  1  0  0  1  1  1
```

- para V={1}, determine se S1 e S2 são
  - 4-conexas
  - 8-conexas
  - m-conexas

Componentes conexos

- Seja S um subconjunto de pixels de uma imagem
- Se p e q estão em S, p está conectado a q se existir um percurso de p para q totalmente em S
- Componente conexo: conjunto de pixels em S que são conexos. Pode haver mais de um componente dentro de um dado S

Algoritmo de rotulação

- Percorra a imagem, pixel a pixel, da esquerda para a direita, de cima para baixo
- V é conhecido, por exemplo, V = {1} para imagens binárias
- Com base nos pixels já rotulados (a esquerda e acima), decida como rotular o pixel atual
- Na saída, a imagem rótulada significa que teremos, para cada pixel, um rótulo que corresponde ao objeto que pertence

Componentes conexos-4

	b	
d	p	

- para todos os pixels p
  - if p = 0 (background): sem ação
  - else // p = 1 (object): check b and d
    - if b=0 and d=0: crie um novo rótulo para p
    - if b=1 or d=1, use o mesmo rótulo para p
    - if b=1 and d=1
      - if b and d tem o mesmo rótulo => use o mesmo rótulo para p
      - rótulos diferentes => use um deles para p, e estabeleça equivalência entre os rótulos
  - Um segundo passo sobre a imagem é necessário para agrupar os rótulos equivalentes

Componentes conexos-8

a	b	c
d	p	

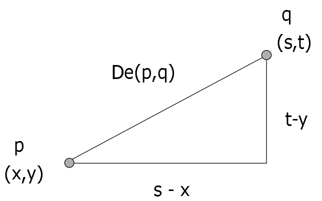
- para todos os pixels p
  - if p = 0 (background): sem ação
  - else // p = 1 (object): verifique a, b, c e d
    - if todos = 0: crie um novo rótulo para p
    - if somente um deles = 1, associe esse rótulo a p
    - if mais de um deles = 1, associe esse rótulo a p, e estabeleça equivalência entre rótulos

Medidas de distância

- Sejam os pixels p, q, e z com coordenadas (x,y), (s,t) e (u,v) respectivamente
- D é uma função de distância ou métrica se:
  - a) D(p,q) >= 0; D(p,q) = 0 iff p=q
  - b) D(p,q) = D(q,p)
  - c) D(p,z) <= D(p,q) + D(q,z)

Distância Euclidiana

- $D_e(p,q) = [(x-s)^2 + (y-t)^2]^{1/2}$



City block distance: D4

- $D4(p,q) = |x - s| + |y - t|$

		2		
	2	1	2	
2	1	0	1	2
	2	1	2	
		2		

- diamante centrado em (x,y)
- D4 = 1 são vizinhos-4 de (x,y)

Chessboard distance: D8

- D8 (p,q) = max (|x-s|+|y-t|)

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

- quadrado centrado em (x,y)

Distâncias D4 e D8

- Distância D4 (D8) entre os pixels p e q equivale ao comprimento do percurso-4 (percurso-8) entre os 2 pontos
- As distâncias D4 e D8 são medidas, não importando se há um percurso conexo entre os pontos (só depende das coordenadas)

Distância de conectividade-m

- A distância Dm entre 2 pontos é definida como o mínimo percurso-m entre os pontos.
  - a distância depende dos valores ao longo do percurso e dos valores de seus vizinhos.
  - Exemplo: Distância entre p0 e p4

	p3	p4
p1	p2	
p0		

	0	1
0	1	
1		

d = 2

	0	1
1	1	
1		

d = 3

	1	1
1	1	
1		

d=4

Operadores aritméticos

- Operadores aritméticos são largamente utilizados no PDI, realizados sobre toda a imagem, pixel a pixel, entre p e q:
  - adição: p + q
    - usado para calcular médias, para redução de ruído
  - subtração: p - q
    - usado para realce em imagens médicas
  - multiplicação: p x q
    - usado para correção de sombreamento resultante da não uniformidade na iluminação ou sensor
  - divisão: p / q

Operadores lógicos

- AND: p AND q
- OR: p OR q
- Complemento: not q
- operadores lógicos são utilizados apenas em imagens binárias, também pixel a pixel
- utilizados como máscaras, detecção de características e análise de forma

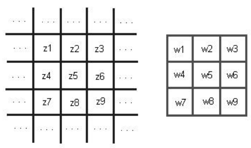
Operações sobre regiões

- Além de operações pixel a pixel, operações lógicas e aritméticas podem ser aplicadas também sobre regiões da imagem. Esse processamento se dá em geral na forma da aplicação de máscaras
- Máscaras são operadores sobre regiões. Calcula-se o valor resultante da máscara aplicada sobre uma região da imagem de entrada, e atribui-se esse valor a um pixel da imagem de saída

Exemplo

- Sejam os coeficientes da máscara $w_i = 1/9$
- O pixel da imagem de saída correspondente a z5 é calculado como:

  $Z = 1/9\ z1 + 1/9\ z2 + \ldots + 1/9\ z9 = \Sigma\ wi\ zi$



(a) Região na Imagem (b) Máscara

Definição da máscara

- A escolha dos coeficientes da máscara e sua aplicação para cada pixel da imagem encontra várias aplicações:
  - redução de ruído
  - afinamento de regiões
  - detecção de bordas
- A aplicação de máscaras é computacionalmente caro.

Transformações afins

- É uma classe de transformações que inclue:
  - rotação
  - translação
  - escala (uniforme e não uniforme)
  - reflexão
  - cizalhamento
- Propriedades:
  - todas as transformações preservam combinações afins entre pontos
    - $R = (1 - \alpha)P + \alpha Q \Rightarrow T(R) = (1-\alpha)T(P) + \alpha T(Q)$

Coordenadas homogêneas

- Para representar pontos e vetores em um espaço d-dimensional, utilizaremos vetores de comprimento d+1.
  - pontos tem a última coordenada = 1
  - vetores tem a última coordenada = 0
- Esse tipo de representação é denominado coordenadas homogêneas de um ponto ou vetor relativo a um frame de coordenadas F.

Observações

- Primeiro, um axioma, para facilitar nossa forma de notação:
  - 0 . P = vetor nulo
  - 1 . P = P
- Atenção: "vetor" pode ser usado como
  - **vetor livre**: entidade geométrica
  - **vetor de coordenada**: forma de representação que pode ser usada para vetores livres e para pontos.

Propriedades

A escolha de 1/0 para ponto/vetor não é arbitrária, ela possui algumas propriedades interessantes:

v = P - Q: a última coordenada se cancela

seja U e V pontos ou vetores. Após várias operações da forma U-V, U+V ou αU:

se a última coord = 0, o resultado é um vetor

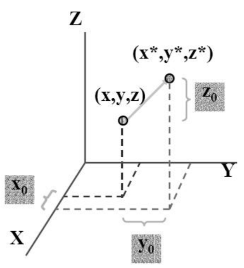
se a última coord = 1, o resultado é um ponto

caso contrário, não é uma operação afim válida.

Isso permite grande flexibilidade, como combinações do tipo:

centróide = (P+Q+R)/3

Translação



x* = x + x0

y* = y + y0

z* = z + z0

$$\begin{bmatrix} x^* \\ y^* \\ z^* \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Transformações: forma geral

v* = A v

A : matrix de transformação 4x4

v : vetor coluna contendo as coordenadas originais

v* : vetor coluna contendo o resultado da transformação

Matriz de translação

- T = $\begin{vmatrix} 1 & 0 & 0 & x0 \\ 0 & 1 & 0 & x1 \\ 0 & 0 & 1 & x2 \\ 0 & 0 & 0 & 1 \end{vmatrix}$

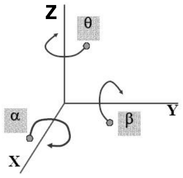
v* = T v

Escala

- escala não uniforme pelos fatores sx, sy, e sz, ao longo dos eixos x, y, z.
  - Caso particular: s = sx = sy = sz
- T = $\begin{vmatrix} sx & 0 & 0 & 0 \\ 0 & sy & 0 & 0 \\ 0 & 0 & sz & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$

Rotação

- rotação de um ponto ao redor do eixo z por um ângulo Θ no sentido horário



$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Rotação sobre x

- Rotação de um ponto ao redor do eixo x por um ângulo α

$$R_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Rotação sobre y

- Rotação de um ponto ao redor do eixo y por um ângulo β

$$R_\beta = \begin{bmatrix} \cos\beta & 0 & -\sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Concatenação e transformação inversa

- Várias transformações podem ser representadas por uma única matriz de transformação 4x4
  - exemplo: translação, escala e rotação ao redor do eixo z de um ponto V é dado por:

    $V^* = R_\theta (S (T\ V))$

    $= A\ V$

    ATENÇÃO: a ordem das matrizes é muito importante!