

Prova 2

Ieremies Vieira da Fonseca Romero

Slides

Cores

A formação de cores se dá por meio de dois processos:

aditivo energias dos fótons são combinados

subtrativo quando a luz passa por um meio que filtra certas frequências

As características das cores são *brilho*, *matiz* e a *saturação*.

brilho ou luminância representa a noção de intensidade da radiação (claro ou escuro)

matiz associada ao comprimento de onda (azul ou vermelho)

saturação pureza do matiz, grau de mistura do original com a luz branca (cores puras são completamente saturadas)

Matiz + saturação são chamados de *crominância*.

Modelos de cores

Sistemas de representação tri-dimensional das cores, usando as características acima. O espaço de cores possíveis num modelo é chamado de gamute.

Em modelos aditivos, a cor branca é a soma de todas as cores. Em modelos subtrativos, é a ausência (ausência da filtragem).

O modelo CMY, subtrativo, é usado em dispositivos de pigmentação e, para evitar o consumo excessivo de tinta, muitas vezes adiciona o preto, formando o CMYK. O modelo HSV (hue, saturation, value) forma uma pirâmide hexagonal, na qual a base é composta das 3 cores primárias e 3 secundárias e a altura é a luminância (sendo a base a mais clara). O modelo HSL forma dois cones de altura 1 total, onde uma das pontas é o preto e outra o branco. O meio, largo ficam as cores.

Textura

Finas :: interações aleatórias e grandes variações Ásperas :: interações melhor definidas e regiões homogêneas

Extração procura retirar um quantidade de dados representativa e simplificada, enquanto a seleção visa reduzir o número de medidas

Matriz de coocorência

Matriz definida a partir de uma relação entre pixels (por exemplo, visinho ao lado), onde a posição i,j da matriz é a quantidade de pares na relação que possuem uma transição do nível de cinza i para j . Utilizamos esta matriz na forma normalizada (dividindo cada elemento pelo número total de transições).

Para generalizar, podemos definir a relação entre pixels a partir da distância e ângulo (zero graus é direita) entre eles. Qualquer alteração nessas medidas altera de forma significativa na matriz, a qual não tem mais informações espaciais sobre os pixels originais.

Segundo momento angular (energia) expressa a uniformidade de uma textura $f_{sma} = \sum \sum p_{i,j}^2$. Em texturas ásperas, poucos elementos da coocorencia normalizada apresentam valores diferentes de zero e, quando ocorrem, são próximos de um e o segundo momento angular apresenta valor próximo a 1.

Entropia quando uma imagem não é uniforme, as entradas $p_{i,j}$ apresentam valores próximos a zero e $f_{ent} = - \sum \sum p_{i,j} \log(p_{i,j})$ em valores altos não normalizados.

Contraste diferença entre tons de cinza, baixo ocorre quando há pequena diferença entre os níveis em uma região contínua $f_{con} = \sum \sum (i-j)^2 p_{i,j}$.

Heterogeneidade

Correlação

Homogeneidade

Matrizes de comprimento de corridas de cinza

Matriz dada por $P(i,j|\theta)$ contém a quantidade de corridas do mesmo nível de cinza i e comprimento j na direção θ . Uma corrida de tamanho 4 e 21 de tamanho -> fina. Várias corridas de tamanho grande -> espessa

Função de autocorrelação

Finas -> primitivas pequenas -> frequências espaciais altas A função de auto correlação descreve as interações espaciais entre as primitivas

Unidade de textura

Definida a partir da relação do pixel central com seus 8 vizinhos. Para cada um deles, olhamos se ele é menor, igual ou maior que o valor do pixel central. Associamos 0, 1, 2 respectivamente e definimos a unidade como o polinômio $a_1 3^0 + a_2 3^1 \dots$ um número na base 3, Perceba que a ordem importa!

LBP Padrões locais binários usa a mesma ideia só que zeros e uns

Registro

Transformações geométricas: transformação espacial (reorganização dos pixels no plano) + interpolação de intensidade.

Mapeamento direto: da original para a transformada, podendo mais de um pixel cair no mesmo lugar Mapeamento indireto: usa-se a inversa, aplicando-a transformada, mas que faz pixels da resultante serem mapeados ao mesmo da original

Coordenadas homogêneas: para permitir que as transformações espaciais possam ser realizadas por meio de multiplicação de matrizes e que possa haver combinação delas $(x,y,z) \rightarrow (Wx, Wy, Wz, W)$.

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ t_{31} & t_{32} & t_{33} & t_{34} \\ t_{41} & t_{42} & t_{43} & t_{44} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

em que é assumido o valor $W = 1$ nessa representação em coordenadas homogêneas.

Transformações afim

generalizam transf. de rotação, translação, escala, reflexão e cisalhamento. Preservam o paralelismo e a proporção entre volumes, áreas e comprimentos.

$$\begin{aligned} X' &= aX + bY + cZ + j \\ Y' &= dX + eY + fZ + k \\ Z' &= gX + hY + iZ + l \end{aligned}$$

na forma matricial utilizando coordena:

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c & j \\ d & e & f & k \\ g & h & i & l \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- Mudança de escala

$$X' = S_x X$$

$$Y' = S_y Y$$

$$Z' = S_z Z$$

coordenadas do novo ponto.

variação de escala pode ser representada

$$S = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Translação

$$X' = X + t_x$$

$$Y' = Y + t_y$$

$$Z' = Z + t_z$$

translação não pode ser representada

$$T = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Rotação Em 2d é dado pela matriz $\cos -\sin$ $\sin \cos$

Projeções

Cada um dos pontos que formam uma cena no espaço tridimensional possa ser projetado no plano de imagem.

Ortográfica pontos são projetados ao longo de linhas paralelas na imagem, projetamos em um dos plano.

A matriz é a identidade com coeficiente zero no plano a ser projetado.

Perspectiva tamanho dos objetos reduz conforme a distância. Lembre-se de ótica do EM. O centro da lente fica no eixo z, a uma distância f focal da origem.

$$\frac{X'}{f} = -\frac{X}{Z-f} \quad \frac{Y'}{f} = -\frac{Y}{Z-f}$$

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ W \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/f & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} W X' \\ W Y' \\ W \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ i & j & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

$$X' = \frac{aX + bY + c}{iX + jY + 1}$$

$$Y' = \frac{dX + eY + f}{iX + jY + 1}$$

Interpolação

Técnicas de registro

Iterativo, correlação de fase (fourier)

Compressão

Sem perda imagens cujos dados são de difícil aquisição.

Com perda nem toda informação é recuperada, mas tá tudo be,

Em geral, as técnicas se baseiam na redução de redundâncias. Redundância é medido como a parte que foi jogada fora na compressão. Uma compressão de 10:1, joga 90% fora. Informação baseia-se na capacidade de obter significado.

Redundância de codificação Seja \bar{L} o comprimento médio de bits para representação de um pixel, uma imagem de $M \times N$ possui codificação ótima com $MN\bar{L}$

bits. Um código é determinado ótimo se seu comprimento mínimo é \bar{L} .

Redundância interpixel Pixels vizinhos possuem valores próximos. Armazenar seus valores absolutos pode gastar mais espaço que só armazenar a diferença entre eles.

Codificação por comprimento de corrida

(valor, quantidade_{deleconsecutivas}) para cada linha

Redundância psicovisual Podemos remover algumas informações que o olho humano tende a não dar atenção.

Teoria da informação

A informação obtida a partir de um evento aleatório é dado por $I(E) = \log_b \frac{1}{P(E)} = -\log_b P(E)$. Se um evento sempre/nunca ocorre, não há nenhuma informação a ser obtida. Quanto mais improvável, maior a quantidade de informação a é necessário para comunicar o evento.

Entropia é $E = -\sum p_i \log_2 p_i$ A eficiência da codificação pode ser definida como $n = E/\bar{L}$

Métodos de compressão

- Sem perdas
 - Huffman Podemos usar códigos de tamanho variável desde que nenhum código seja prefixo de outro de tamanho maior. Para determinar os códigos, utilizamos a técnica de redução de fonte:

- * ordenamos os símbolos por probabilidade
- * a cada passo, combinamos os dois com menor probabilidade, somando-as.

Fontes Originais		Redução de Fontes			
Símbolo	Probabilidade	1	2	3	4
S_2	0.40	0.40	0.40	0.40	0.60
S_6	0.30	0.30	0.30	0.30	0.40
S_1	0.10	0.10	0.20	0.30	
S_4	0.10	0.10	0.10		
S_3	0.06	0.10			
S_5	0.04				

- * Depois, retornamos, da direita à esquerda atribuindo códigos às probabilidades. Adiciona-se um bit a cada símbolo previamente agrupado

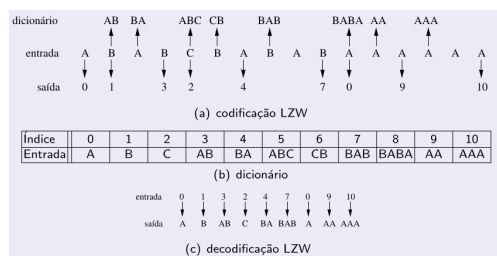
Fontes Originais		Redução de Fontes							
Símbolo	Probabilidade	Código	1	2	3	4			
S_2	0.40	1	0.40	1	0.40	1	0.40	1	0.60
S_6	0.30	00	0.30	00	0.30	00	0.30	00	0.40
S_1	0.10	011	0.10	011	0.20	010	0.30	01	
S_4	0.10	0100	0.10	0100	0.10	011			
S_3	0.06	01010	0.10	0101					
S_5	0.04	01011							

As vezes, resolver esse processo pode ser computacionalmente complexo ou gerar código proibitivamente longos. Uma solução para isso é aplicar à apenas os m símbolos mais frequentes enquanto o resto usa um prefixo livre e tamanho fixo.

- Shannon-Fano Divisão e conquista, cada divisão um fica com 0 e outro com 1.
- Dicionário LZ78: vou adicionando símbolos à minha palavra enquanto a palavra resultante ainda aparecer no dicionário. Quando isso deixar de ser verdade, eu uso o código da maior palavra que consegui e começo de novo.

LZW: Iniciamos com um dicionário com todas as palavras de um símbolo. A cada passo, adicionamos o próximo símbolo c a nossa palavra I . Se $I + c$ existe no dicionário, passamos para o próximo. Se não, utilizamos o último símbolo que tínhamos para I , criamos uma palavra para $I + c$ e recomeçamos com $I = c$.

Dessa forma, não é necessário transmitir o dicionário. No processo de decodificação, vamos descobrindo as palavras conforme decodificamos.



- Comprimento de corrida Ou eu uso ($onde_{começa}$, $quanto_{dura}$) para cada corrida ou, começando de uma corrida preta, alterno em quanto dura a corrida.

Por árvore binária: eu divido a linha na metade e marco se as partes são inteiras brancas/pretas. Caso contrário, ramifico.

- Planos de bits Caso sua imagem não seja binária, não tema. Podemos utilizar os planos de bits que correspondem ao mapeamento de cada i -ésimo bit em cada valor.
- Preditiva sem perdas Redundância interpixel Ao invés de salvar a variação entre os pixels, utilizamos uma função que tenta adivinhar qual o próximo pixel e que na verdade salvo o erro dessa função. Assim, na hora de decodificar, tentamos prever o pixel e adicionamos o erro salvo. Quanto mais precisa for a predição, menor os valores de erro, menor o espaço.

- Com perdas

- Preditiva com perdas Mesmo processo do sem perdas, mas agora tendemos a jogar fora algumas informações do erro para que este não se acumule. A função quantizadora pode, por exemplo, considerar erros muito pequenos como zero. Para evitar que esse “arredondamento” acumule-se, utilizamos-no também no cálculo do erro.
- Modulação delta O preditor preve que o pixel será igual e o erro só pode ser \pm o delta
- Modulação Código de Pulso Diferencial Assume-se que o erro devido a quantização é irrelevante e utiliza-se um preditor mais sofisticado, de forma a minimizar o erro médio quadrático.
- Transformada Utiliza-se janelas pequenas de 8×8 ou 16×16 nas quais aplicamos transformadas como a de Fourier ou discreta do cosseno para descobrir os coeficientes que descrevem aquela região. Podemos descartar os coeficientes que descrevem o menor número de informações a fim de reduzir o espaço de armazenamento e guardar o resto para ser decodificado.

Não podemos usar grandes janelas pois estas não possuem uniformidade o que causaria muitos coeficientes.

Padronização JPEG

Ordenamos por zigue-zague para facilitar a codificação por entropia Realizar as operações em matrizes maiores possui um custo computacional elevado quando tratamos de transformadas, mas, até um certo ponto, há vantagem já que o valor médio, chamado DC e cada janela, é muito similar em janelas vizinhas, o que torna muito útil técnicas preditivas.

Lista 2

1. Extração de borda por operador morfológico
2. Efeitos causados à representação quadtree após mudança em escala, translação ou rotação?
3. Entropia / código de huffman

Compressão Entropia é dada pela soma $-\sum p(s_i) \log_2 p(s_i)$ Código Huffman usa a redução

de fontes na frequência que esses valores apreecem. COmbinando 95 com 169, depois o resultante com 21. 243 código 0, 21 código 10, 95 código 110 e 168 código 111.

4. Codificações de Huffman

Compressão

simbol	prob	1	2	3
a	0.55	0.55	0.55	0.55
b	0.15	0.15	0.30	0.45
c	0.15	0.15	0.15	
d	0.10	0.15		
e	0.05			

a = 0 b = 10 c = 110 d = 1110 e = 1111 média = 0.55
+ 0.30 + 0.45 + 0.4 + 0.2 = 1.9

a = 0 b = 100 c = 101 d = 110 e = 111 média = 0.55
+ 0.45 + 0.45 + 0.3 + 0.15 = 1.9

se eu fiz certo, eles possuem o mesmo comprimento médio

5. Construa dicionário e LZW

Compressão a b c bc cc ca ac cb bcc ccc cccc ccccc 1 2 2
0 2 3 4 9 10 5 9 b c c a c bc cc ccc cccc ca ccc
a b c 1 2 2 0 2 3 4 9 10 5 9 b

6. Cores subtrativos / aditivos

Cores Modelos de cores subtrativos são baseados na adição de filtros para remoção de frequências, como CMY utilizando em toners, apesar de nesse caso ser necessário adicionar o K de black para diminuir o uso de cores. Nesse modelo, preto é a presença de todos os filtros.

7. Vantagens da codificação aritmética para compressão

Compressão Ele é extremamente eficiente em tamanho da representação, apesar de conter problemas de precisão no desempenho da decodificação.

8. Redundância em compressão

Compressão Redundância de coficação, interpixel e interpretação humana. A primeira se dá a um uso excessivo de pixels para representar os símbolos, a segunda pelo fato de pixels próximos terem valores próximos (podemos então usar compressões preditivas) e o terceiro se dá pelo fato do olho humana não conseguir distinguir todas as informações presentes e dar mais valores a algumas.

9. Vantagens e desvantagens de usar blocos de tamanhos diferentes no cálculo da transformada discreta do cosseno no JPEG

Compressão Blocos maiores possuem mais informações mas são bem mais custosos de computar. Blocos menores são mais fáceis de realizar a DCT mas abstraem menos da informação.

10. Vantagem da ordenação zig-zag do JPEG?

Compressão Facilita a codificação por entropia dos coeficientes AC da transformada discreta do cosseno dentro das janelas 8x8.

11. Compressão com perda

Compressão Predição com perdas na qual fazemos uma quantização do erro em pról de reduzir o tamanho necessário deste. Por transformada, na qual dividimos a imagem em partes, aplicamos uma transformada e descartamos coeficientes que menos adicionam informações.

12. Técnicas preditivas de compressão de imagens. Descreva a principal diferença entre técnicas preditivas sem e com perdas.

Compressão Técnicas preditivas sem perdas armazenam o valor exato do erro, enquanto técnicas com perda utilizam a quantização ou modularização do valor do erro para economizar espaço em detrimento de um pouco de qualidade.

13. Código de comprimento de corridas

Compressão 3 4 4 4 2 1 2 1 3 ... lembrar que transformamos a imagem num grande vetor 1D

14. Versão binária da unidade de textura, padrões locais binários.

Textura Reduz o número de entradas no espectro de textura, o que permite uma representação mais sucinta.

15. Padrões locais binários demonstram ser invariantes a tranformações monotônicas aplicadas à imagem. Quais as vantagens?

Textura O fato da transformação ser monotônica faz com que as comparações entre os valores dos pixels não mude,

o que, pela definição de LBP, não modifica a representação. Assim, temos que, apesar de uma transformação alterar os valores da imagem, conseguimos manter a nossa representação do que é ainda a mesma estrutura de textura.

Podemos, sem medo de incubir o custo de recomputar o LBP fazer alterações e garantir que texturas que soferam apenas transformações monotônicas manteram sua unidade de textura.

16. Matriz de concorrência, momento angular, discrminação de texturas.

Textura

x	0	1	2	3
0				
1			6	
2		6		
3				

$$\text{Momento angular} = 1/12^2 + 1/12^2 = 1/72$$

x	0	1	2	3
0		1		
1	1		2	
2		2		3
3			3	

$$\text{Segundo Momento Angular} = (2 + 8 + 18)/12^2 = 14/72$$

Essa medida pode ser usada sim, SMA maiores, representam níveis de energia maiores e portanto texturas ásperas

SEGUNDO MOMENTO ANGULAR ALTO -> ÁSPERA

17. Rotações 2d são aditivas

Ou seja, rotação por α_1 e posterior rotação por α_2 é igual a rotação por $\alpha = \alpha_1 + \alpha_2$.

Só expandir as definições de rotação como $x = x \cos - y \sin$ e $y = x \sin + y \cos$. No final tem que fazer a regrinha de trigonometria.

18. Escalas 2d são multiplicativas

São multiplicativas pq na conta dos fatores fica uma multiplicação entre eles vezes o valor original

19. Mostre que a rotação e escala são comutativas se os fatores de escala $S_x = S_y$

Mesma ideia das anteriores.

20. Transformada afim

As transformadas afins generalizam transformações como rotação, translação, escala enquanto mantém o paralelismo e as proporções entre volumes, áreas e comprimentos entre objetos da imagem. Podem ser representadas na forma matricial com coordenada homogênea de forma que a última linha seja 0 0 0 1.

21. Projeção ortográfica e projeção perspectiva.

Registro Projeção ortográfica possui um centro de projeção no infinito, mantém as retas paralelas e as dimensões intactas. Já a projeção de perspectiva modifica a dimensão dos objetos baseado em suas distâncias ao foco da lente (distância focal)

23. Coordenadas homogêneas para representação de transformações geométricas

Registro Permite que modelemos as transformações via matrizes e utilizemos as operações matriciais para combinar transformadas.

24. Descreve três técnicas de registro de imagem.

Registro