Universidade Estadual de Campinas Instituto de Computação

Introdução ao Processamento Digital de Imagem (MC920 / MO443)

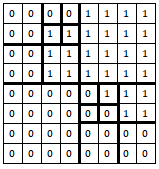
Professor: Hélio Pedrini

Lista II

1. Ache a representação *quadtree* do seguinte objeto:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Resposta:



1

1 (0, 1) (0, 0) 1

(0, 0) 1 (0, 0) (0, 1) 1 (0, 1) (0, 0) (0, 0)

(0, 0) (0, 0) (0, 1) (0, 1) (0, 0) (0, 1) (0, 0) (0, 0)

(1) (1)(0,1)(0,0)(1) (0,0)(1)(0,0)(0,1) (1)(0,1)(0,0)(0,0) (0,0)(0,0)(0,1)(0,1) (0,0)(0,1)(0,0)(0,0)

1 101001 0010001 1010000 00000101 00010000 (binary)

1A48D00510 (hexadecimal)

1. Quais os efeitos causados a representação *quadtree* de uma imagem após sofrer uma mudança de escala, translação ou rotação?

Resposta:

Mudança de escala: a representação continuará a mesma, apesar de estar suspeita a mudanças conforme ocorrer a perda de detalhes da imagem.

Mudança de translação ou rotação: a representação mudará completamente (*quadtree* é muito sensível a esse tipo de mudanças).

1. Suponha a imagem a seguir.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 21 | 21 | 21 | 95 | 169 | 243 | 243 | 243 |
| 21 | 21 | 21 | 95 | 169 | 243 | 243 | 243 |
| 21 | 21 | 21 | 95 | 169 | 243 | 243 | 243 |
| 21 | 21 | 21 | 95 | 169 | 243 | 243 | 243 |

* 1. calcule a entropia da imagem;

Resposta:

Imagem: *4x8 = 32*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Níveis de cinza (*i*) | 0 | 1 | 2 | 3 |
|  | 12 | 4 | 4 | 12 |
|  | 0.375 | 0.125 | 0.125 | 0.375 |

H = = **1.81**

* 1. calcule o código de Huffman para cada nível de cinza da imagem.

Resposta:

*1 0*

**0.375** 0.625

*1 0*

0.25 **0.375**

*1 0*

**0.125** **0.125**

**Huffman ~ Binary Search Tree -> Left Node: 1, Right Node: 0**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Fontes originais | | Redução de fontes | | | | |
| Símbolo | Probabilidade | Código | Etapa 1 | | Etapa 2 | |
|  | .375 | 1 | .375 | 1 | .625 | 0 |
|  | .375 | 00 | .375 | 00 | .375 | 1 |
|  | .125 | 011 | .250 | 01 |  |  |
|  | .125 | 010 |  |  |  |  |

**1.875** bits/símbolo

1. Dada uma fonte de cinco símbolos com probabilidades {0.55,0.15,0.15,0.10,0.05}, obtenha duas codificações de Huffman diferentes. Calcule o comprimento médio *L* de bits para cada um dos códigos.

Resposta:

**Primeira codificação:**

*1 0*

0.45  **0.55**

*1 0*

**0.15** 0.3

*1 0*

0.15 **0.15**

*1 0*

**.05 .10**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Fontes originais | | Redução de fontes | | | | | | |
| Símbolo | Probabilidade | Código | Etapa 1 | | Etapa 2 | | Etapa 3 | |
|  | .55 | 0 | .55 | 0 | .55 | 0 | .55 | 0 |
|  | .15 | 11 | .15 | 11 | .3 | 10 | .45 | 1 |
|  | .15 | 100 | .15 | 101 | .15 | 11 |  |  |
|  | .10 | 1011 | .15 | 100 |  |  |  |  |
|  | .05 | 1010 |  |  |  |  |  |  |

**1.900000001** bits/símbolo

**Segunda codificação:**

*1 0*

0.45  **0.55**

*1 0*

**0.15** 0.3

*1 0* *1 0*

**0.05** **0.10 0.15 0.15**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Fontes originais | | Redução de fontes | | | | | | |
| Símbolo | Probabilidade | Código | Etapa 1 | | Etapa 2 | | Etapa 3 | |
|  | .55 | 0 | .55 | 0 | .55 | 0 | .55 | 0 |
|  | .15 | 100 | .15 | 11 | .3 | 10 | .45 | 1 |
|  | .15 | 101 | .15 | 100 | .15 | 11 |  |  |
|  | .10 | 111 | .15 | 101 |  |  |  |  |
|  | .05 | 110 |  |  |  |  |  |  |

**1.9** bits/símbolo

1. Considere uma fonte com alfabeto Ω = {a,b,c}. A partir da sequência 1, 2, 2, 0, 2, 3, 4, 9, 10, 5, 9, construa o dicionário e decodifique essa sequência por meio do algoritmo LZW.

Resposta: bccacbcccccccccccaccc, d = {0: 'a', 1: 'b', 2: 'c', 3: 'bc', 4: 'cc', 5: 'ca', 6: 'ac', 7: 'cb', 8: 'bcc', 9: 'ccc', 10: 'cccc', 11: 'ccccc', 12: 'cac'}

|  |  |  |  |
| --- | --- | --- | --- |
| Input | Dictionary translation | Decode output | New dictionary entry |
| 1 | b | **b** | none |
| 2 | b c | b**c** | bc/3 |
| 2 | c c | bc**c** | cc/4 |
| 0 | c a | bcc**a** | ca/5 |
| 2 | a c | bcca**c** | ac/6 |
| 3 | c bc | bccac**b**c | cb/7 |
| 4 | bc cc | bccacbc**c**c | bcc/8 |
| 9 | ??? | ??? | ??? |

|  |  |  |  |
| --- | --- | --- | --- |
| Input | Dictionary translation | Decode output | New dictionary entry |
| 1 | b c | bc | bc/3 |
| 2 | c c | bcc | cc/4 |
| 2 | c a | bcca | ca/5 |
| 0 | a c | bccac | ac/6 |
| 2 | c bc | bccacbc | cb/7 |
| 3 | bc cc | bccacbccc | bcc/8 |
| 4 | cc ccc | bccacbcccccc | ccc/9 |
| 9 | ccc cccc | bccacbcccccccccc | cccc/10 |
| 9 | ccc cccc | bccacbcccccccccc | cccc/10 |
| 10 | cccc ca | bccacbccccccccccca | ccccc/11 |
| 5 | ca ccc | bccacbcccccccccccaccc | cac/12 |

1. Descreva os principais tipos de redundância que podem ser explorados pelos métodos de compressão de imagens.

Resposta:

* Redundância de codificação:
  + Mesma informação (ou símbolo) representada repetidas vezes;
  + Em geral, a redundância de codificação está presente quando os códigos atribuídos a um conjunto de eventos (tal como os níveis de cinza da imagem) não foram escolhidos de forma a explorar as probabilidades dos eventos;
* Redundância interpixel:
  + Explora a característica de que pixels vizinhos em uma imagem possuem alguma relação ou similaridade;
  + Contribuição visual de um único pixel é redundante - ela poderia ser prevista com base nos valores dos pixels adjacentes;
* Redundância psicovisual:
  + Há informações psicovisualmente redundantes - possuem menor importância relativa que outras no processamento visual e podem ser eliminadas sem prejudicar significativamente a percepção da imagem;
  + Redundância infere que dados presentes na imagem seriam eliminados de forma **irreversível**!

1. Explique as vantagens e desvantagens em se utilizar blocos de tamanhos diferentes de 8×8 pixels no cálculo da transformada DCT na padronização JPG.

Resposta:

* Vantagens:
  + Custo computacional menor em um bloco de tamanho para realizar operações como a transformada DCT;
  + Matrizes menores que não possuem informação suficiente e relevante para a compressão;
  + Aplicar a transformada na imagem inteira não é ideal devido ao conteúdo não uniforme da imagem;
* Desvantagens:
  + Imagem resultante terá tamanho maior do que se utilizados blocos maiores;

1. Qual a vantagem da ordenação ziguezague dos coeficientes DCT na padronização JPG?

Resposta:

A ordenação ziguezague dos coeficientes DCT tem o objetivo de facilitar a etapa de codificação por entropia. Em geral, é mais intuitivo que pixels em um bloco tenham mais em comum do que pixels em uma linha - ou seja, os pixels serão agrupados na ordenação de forma a terem mais características em comum, permitindo que a compressão seja mais bem-sucedida devido a redundância interpixel.

Outra Resposta:

Na matriz DCT, baixas frequências tem os maiores valores de energia. Como as baixas frequências estão nas linhas e colunas de índice baixo, ao percorrer em zigue-zague percorremos na ordem de menor energia esperada, esta ordem deve ter uma longa cauda de zeros o que otimiza a codificação em comprimento de corrida.

1. Descreva dois métodos de compressão de imagens com perdas.

Resposta:

Ambos os métodos partem da redundância psicovisual, a qual justifica as perdas durante a compressão.

Modulação Delta - o preditor supõe que o pixel atual seja igual ao anterior e, para cada pixel, retorna apenas a diferença entre este e o anterior. Na prática, recebe uma constante , o módulo máximo de diferença entre pixels adjacentes. O algoritmo consiste em limitar a diferença entre os pixels adjacentes em se a diferença for positiva e se a diferença for negativa..

Modulação de Código de Pulso Diferencial - o preditor supõe que o valor de um pixel depende de uma combinação linear de pixels anteriores e que o erro de quantização seja desprezível. O objetivo é minimizar o erro médio quadrático da predição do codificador.

1. Explique o princípio das técnicas preditivas de compressão de imagens. Descreva a principal diferença entre técnicas preditivas sem perdas e com perdas.

Resposta:

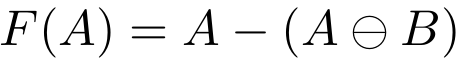
Técnicas preditivas de compressão são técnicas que envolvem uma etapa anterior à codificação dos símbolos, que explora a redundância interpixel na imagem.

As técnicas preditivas sem perdas utilizam conceitos de redundância como redundância de codificação e interpixel - basta representar a informação de uma maneira mais robusta para garantir a sua recuperação. Já as técnicas preditivas com perdas utilizam conceitos de redundância psicovisual, e se baseiam em alterar os valores dos pixels de forma que facilite a compressão da imagem, mantendo a essência da imagem e abstraindo os detalhes (de maneira imperceptível ou tolerada).

Traks:

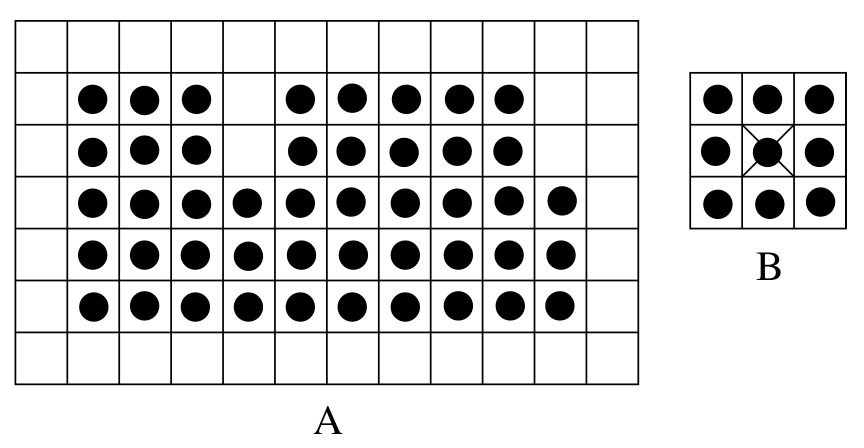
Também acho que faltou explicar o princípio. O princípio das técnicas preditivas é que é mais econômico guardar a diferença entre os pixels do que a intensidade dos pixels. Então numa compressão sem perda, você armazena todas as diferenças e recupera a imagem perfeitamente. Numa compressão com perdas, você estabelece limites que funcionam mais ou menos como: se a diferença entre os pixels for menor que X, não vale a pena armazenar isso e eles serão iguais. por isso tem erro. Ou como na compressão delta, se o erro for maior que X, colocamos um delta fixo e não armazenamos exatamente o erro, mas armazenamos esse delta.

1. A extração da borda *F* de um objeto *A* pode ser realizada através do operador morfológico

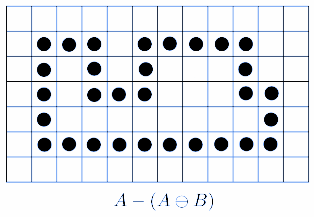
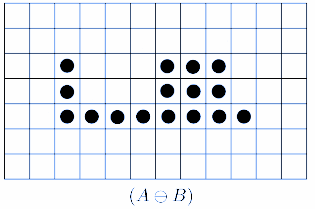


em que *B* e um elemento estruturante adequado.´

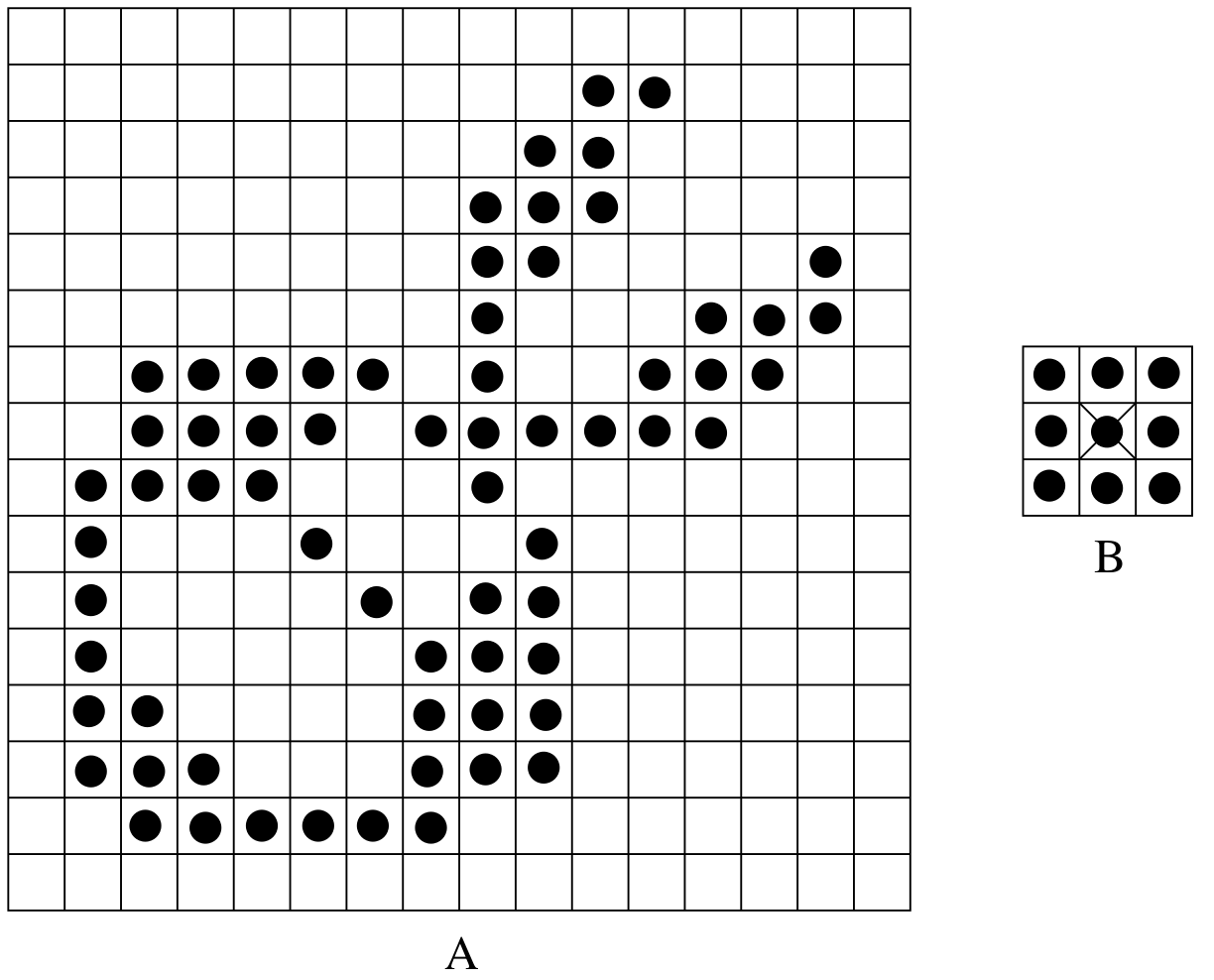
Aplique o operador acima para a imagem *A* a seguir, utilizando o elemento estruturante *B* de tamanho 3×3.



Resposta:

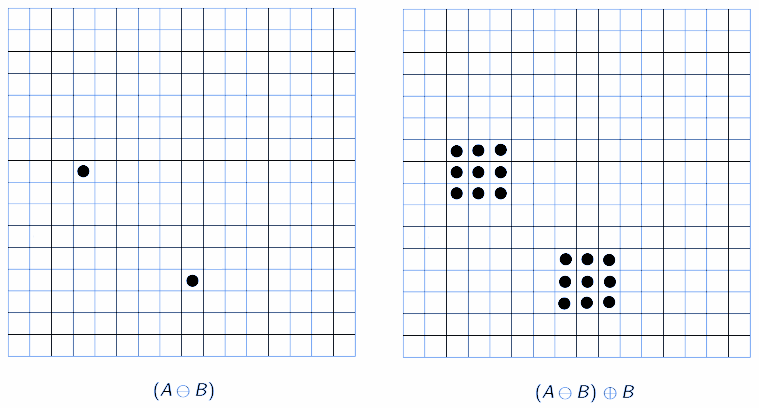


1. Aplique os operadores de abertura e fechamento morfológicos para a imagem a seguir.

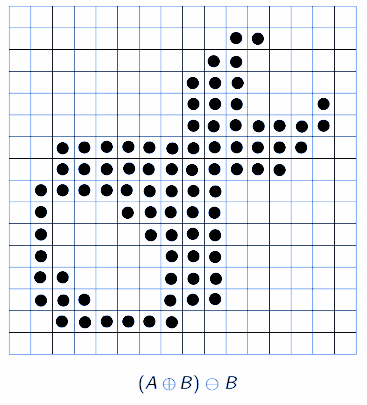


Resposta:

Abertura:



Fechamento:



1. Escreva um programa para calcular a área englobada por um contorno definido pelo código da cadeia. Considere que os pixels do contorno estão conectados por meio da vizinhança-4.

Resposta:

class Point {

public:  
 int x, y;

Point (int \_x, int \_y) : x(\_x), y(\_y) { }

};

int area(vector<int> chain\_code) {

vector<Point> polygon;

int area = 0;

/\* convert each vertex from chain code to a Point \*/

convert\_points(chain\_code, polygon);

for (int i = 0; i < polygon.size()-1; ++i) {

area += polygon[i].x\*polygon[i+1].y **-** polygon[i+1].x\*polygon[i].y;

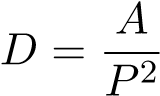
}

return area/2;

}

* [reference](http://mathworld.wolfram.com/PolygonArea.html)

1. Suponha o descritor *D* definido como



em que *A* é área (em pixels) do objeto *P* e seu correspondente perímetro ou comprimento de borda.

Comente:

* 1. *D* é máximo para um objeto igual a um círculo.

Resposta:

Se o objeto é um círculo, então . Sabendo que o círculo possui a menor(maior\*) distribuição de área **em relação ao seu perímetro**, será, portanto, máximo.

Traks:

O D de um quadrado é D = ¼. Ou seja, é maior que 1/4pi. Logo, D de um círculo não é MÍNIMO? #questionamento

D de um quadrado é (L^2)/(4L)^2 = 1/16 OBRIGADO AMIGO

* 1. *D* é sempre menor para um retângulo do que um quadrado, considerando que o lado menor do retângulo é igual ao lado do quadrado.

Resposta:

Para um retângulo, e, para um quadrado, temos . Sabendo que , e , temos que sempre será menor que e a afirmação é verdadeira, pois m > n > 0 ⇒ (m-n) > 0 ⇒ (m-n)2 > 0 ⇒ m2 - 2mn + n2 > 0 ⇒ m2 +2mn + n2 > 4mn ⇒ (m+n)2 > 4mn ⇒ (2m + 2n)2 > 16 mn ⇒ 1/16 > mn/(2m+2n)2 ⇒ DS > DR.

1. Descreva um conjunto de descritores que poderiam ser utilizados em um sistema de reconhecimento automático de placas de veículos a partir de imagens de vídeos.

Resposta:

Podemos utilizar:

1. Descritores de bordas para encontrar candidatos a polígonos;
2. Descritores de região para encontrar retângulos:
   1. Encontrar candidatos que satisfaçam relação de retangularidade adequada;
   2. Utilizar conceito de projeções horizontal e vertical para encontrar a orientação adequada;
   3. Utilizar mesmo conceito de B para reconhecimento das letras/números ou classificadores
3. Cite uma vantagem e uma desvantagem no uso de códigos da cadeia para representar bordas de objetos.

Resposta:

Vantagem: Ocupa menos espaço na memória para a representação do objeto, pode ser invariante quanto a escala, rotação e translação;

Desvantagem: Há um custo computacional para encontrar as bordas (a consulta não é constante), sensibilidade do código da cadeia a pequenas perturbações ao longo da borda devido a ruídos ou imperfeições na segmentação.

1. Represente a imagem binária mostrada a seguir pelo código de comprimento de corridas, iniciando-se com o comprimento das corridas de valor 1.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 |
| 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 |
| 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |

Figura 1: Imagem binária.

Resposta:

3, 4, 5, 4, 2, 1, 2, 1, 3

0, 4, 4, 3, 6, 1, 2, 1, 4

3, 7, 3, 4, 2, 2, 1, 1, 2

0, 2, 1, 2, 6, 8, 2, 4

5, 6, 4, 3, 2, 5

0, 1, 3, 2, 3, 1, 10, 3, 2

3, 10, 5, 2, 5

1. Qual o problema que a utilização da versão binária da unidade de textura, os padrões locais binários (LBP), reduz quando se efetua o cálculo das medidas?

Resposta:

O LBP reduz significativamente o número de entradas do espectro de textura, limitando o valor de cada pixel de **0** a **255**.

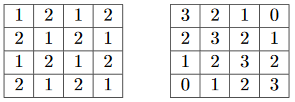
1. Por que os padrões locais binários (LBP) apresentam-se invariantes a transformações monotônicas aplicadas à imagem? Descreva as vantagens dessa característica.

Resposta:

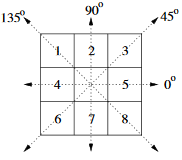
Porque os valores armazenados são relativos à intensidade entre os pixels e, portanto, o seu valor não sofre modificação desde que as relações entre as cores da imagem permaneça a mesma.

Isso é vantajoso porque permite a extração de características morfológicas da imagem, sem que seja necessário quantificá-la em relação às cores.

1. Utilizando calcule a matriz de co-ocorrência na orientação 0° para as imagens



Calcule o segundo momento angular para cada uma das matrizes de co-ocorrência. Essa medida pode ser utilizada para discriminar as texturas contidas nas duas imagens?

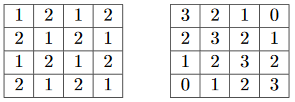
Resposta:

Segundo momento angular:

**(a)**

|  |  |  |
| --- | --- | --- |
|  | **1** | **2** |
| **1** | 0 | 12 |
| **2** | 12 | 0 |

**(b)**

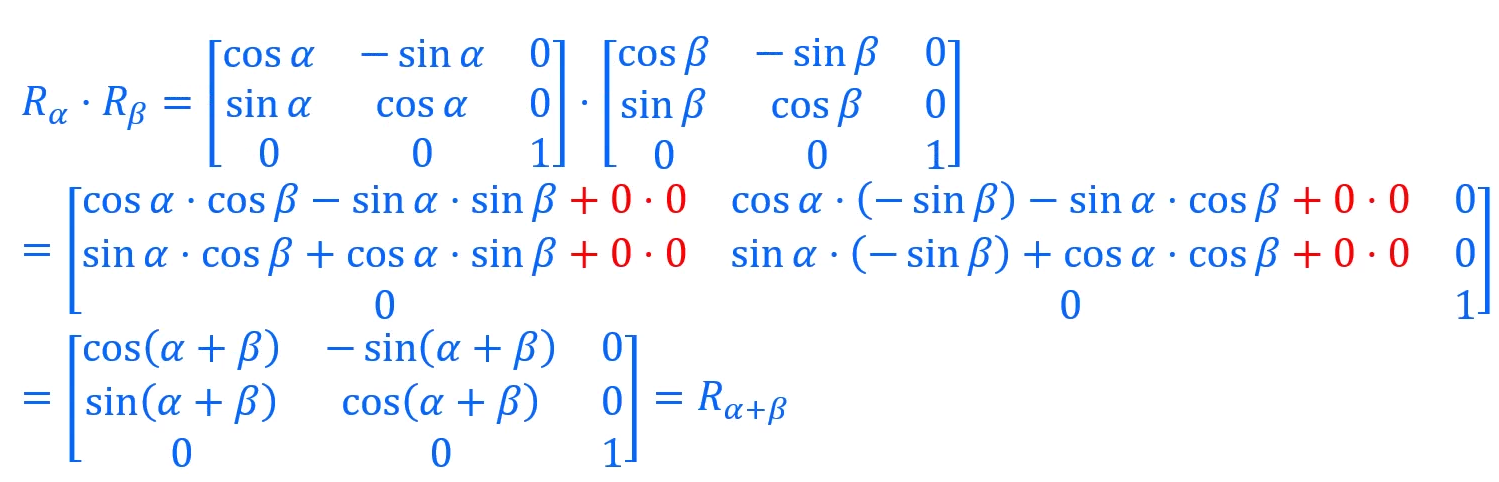


|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **0** | **1** | **2** | **3** |
| **0** | 0 | 2 | 0 | 0 |
| **1** | 2 | 0 | 4 | 0 |
| **2** | 0 | 4 | 0 | 6 |
| **3** | 0 | 0 | 6 | 0 |

*fsma* = 2 · (1/12)2 + 2 · (1/6)2 + 2 · (1/4)2 = 7/36 = 0.1944...

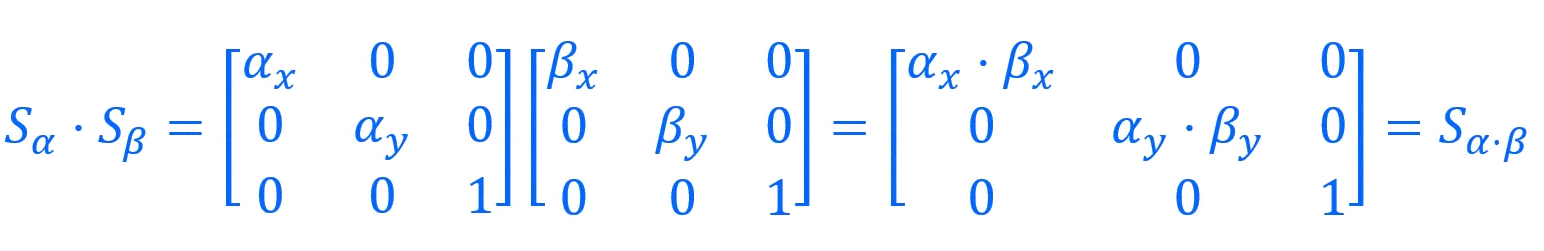
1. Mostre que duas rotações 2D sucessivas são aditivas.

Resposta:



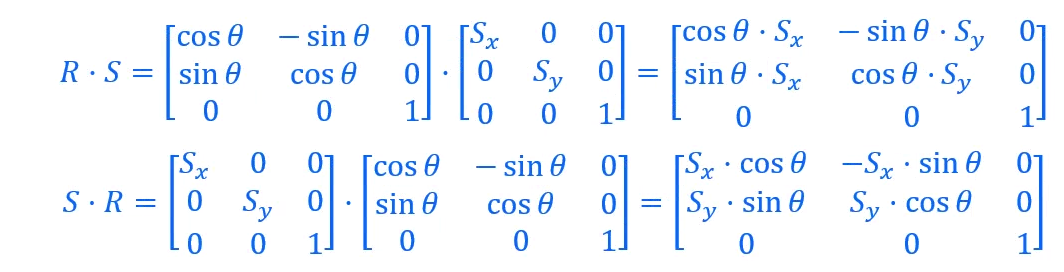
1. Mostre que duas escalas 2D sucessivas sao multiplicativas.

Resposta:



1. Mostre que a rotação e a escala 2D são comutativas se os fatores de escala *Sx* = *Sy*.

Resposta:



1. Descreva o conceito de transformada afim.

Resposta:

A transformada afim traduz transformações projetivas de coordenadas, permitindo a transformação de diversos pontos de uma mesma figura para uma determinada perspectiva através de uma matriz **M**.

As transformações afins generalizam as transformações de rotação, translação, escala, reflexão e cisalhamento. Elas preservam o paralelismo entre retas e planos, mas não preservam comprimentos, distâncias e áreas.

1. Descreva as principais diferenças entre projeção ortográfica e projeção perspectiva.

Resposta:

A transformação ortográfica realiza o mapeamento de pontos tridimensionais sobre o plano da imagem, tal que os pontos são projetados ao longo de linhas paralelas na imagem. Já a projeção perspectiva, apesar de também realizar transformação de pontos tridimensionais, altera o tamanho dos objetos de acordo com a distância do centro de projeção. Os efeitos de mudança de tamanho estão relacionados à percepção de profundidade do sistema visual humano - e, portanto, a forma dos objetos não é preservada.

1. Qual a vantagem da utilização de coordenadas homogêneas para a representação de transformações geométricas?

Resposta:

As coordenadas homogêneas possuem a vantagem de simplificar e unificar cálculos em matrizes. Por exemplo, elas permitem: realizar operações de translação e rotação simultaneamente; representação de pontos no infinito (o que permite eliminar casos especiais e tornar os cálculos mais simples e uniformes).