

## Respostas do Teste de Análise de Imagens - 2004:

1. Diga com suas palavras o que é: (Valor total da questão: 1.2)

Filtragem passa alta (valor deste item até 0,3) - importante falar que apesar do nome ser no domínio da frequência pode ser usado na forma de máscaras direto no domínio do espaço, usando a ideia de projeto de filtros pela convolução inversa.

Transformada de Hough (valor deste item 0,3) - importante falar que pode ser usado para diversas equações de curvas paramétricas e na forma genérica.

Filtro Gaussiano (valor deste item 0,2) - importante falar de que usa o desvio padrão

Detecção de contornos (valor deste item 0,2) - importante falar que pode ser direcional não direcional, usando filtragem com máscaras ou não.

Textura (valor deste item 0,2) - importante falar dos textels

(depende do que o aluno escreveu com suas palavras receber o máximo de cada item nessa a correção, falando absurdos tem -0.1)

2. Calcule o histograma para a imagem "Degradê" abaixo. Essa imagem é de 100 x 100 pixels. E tem o mesmo tom em cada linha. Os tons variam de linha para linha a incrementos de 2 tons. Ou seja uma certa linha tem o mesmo tom da anterior mais 2.

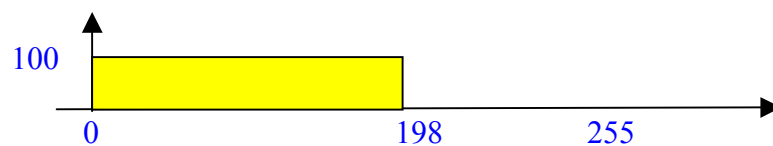
0 0 0 0 0 0 0 0 ....  
2 2 2 2 2 2 2 2 ....  
4 4 4 4 4 4 4 4....  
6 6 6 6 6 6 6 6....  
.  
.  
.

Sob ponto de vista da análise desse histograma, essa imagem apresenta um bom contraste? O que você recomendaria fazer quanto a isso? Faça essa sua recomendação nesta imagem. Como a imagem fica depois de processada? Houve alguma perda no processo? O que ocorre se você passar um filtro de média nesta imagem? E um filtro bem comum de detecção de lados verticais? (valor total 2)

A imagem tem 100 colunas. Logo terá 100 pixels de cada tom.

A imagem tem 100 linhas. Na linha 0 teve tom 0. Na linha 1, teve tom 2. Na linha 2, teve tom 4. Na linha 3, teve tom 6. Assim na linha 4, terá tom 8, e sucessivamente de modo que na 99 terá tom  $2 \times 99 = 198$ .

Assim o histograma dessa imagem irá do tom 0 ao tom 198 com 100 pixels em cada tom. De modo que terá uma envoltória com a aparência de um degraus em linha reta.



(0.4, se errou os valores numéricos 0.1)

Não. Esta imagem não usa toda sua possibilidade de contraste já que não há qualquer pixel com tons acima de 198: 200...a 255. (0.1)

Ela poderia ser melhorada usando funções para **aumentar o contraste** ou por **equalização**. (0,1)

Fazendo a sugestão na imagem ( 0.4)

A primeira forma é bem mais simples, bastando encontrar a função adequada para ampliar seu range de 198 para 255. Essa poderia ser obtida por regra de 3:

$$\text{tom\_atual} \times 255/188 = \text{tom\_novo} \quad \text{tom\_atual} \times 1.287878787879 = \text{tom\_novo}$$

A imagem ficará: (0.2)

```

0  0  0  0  0 ....
3  3  3  3  3 ....
5  5  5  5  5 ....
8  8  8  8  8 ....
10 10 10 10 10....
:  :  :
255 255 255 255 255.....

```

Como os tons são inteiros, ao ser ampliado o incremento de tom para o próximo nível que é constante de 2 na imagem original, a imagem processada teria incrementos não constantes. Isso não é perceptível por olhos humanos, mas em um processo de detecção de incrementos em visão de máquina poderia ser notado Assim essa é a única **perda** no processo, a constância dos incrementos tonais. (0.2)

Ao se passar um filtro de **média** nesta imagem nesta imagem, de tamanho impar , 3x3 ou 5x5 por exemplo, nada ocorre na imagem.(0.1) Ao se passar filtros de média de janela par: 2x2, 4x4 por exemplo o tom da imagem é alterado para um tom impar.(0.1) Se o filtro de média for 1D na horizontal nada é alterado. (0,1) . Se o filtro de média 1 D mas na vertical, depende de ser par ou impar como já comentado (0.1)

Um filtro bem comum de detecção de lados verticais é o definido pelas máscaras:

```

1 -1   ou   1 0 -1
1 -1         1 0 -1
              1 0 -1

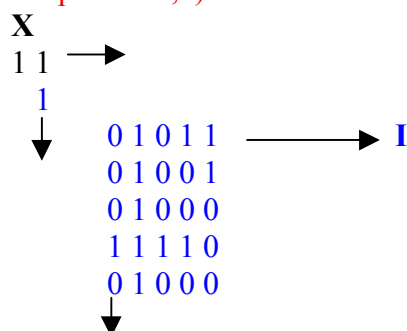
```

Esses filtros variam a imagem inteira ficar zerada. O que é um resultado esperado, já que na imagem não há qualquer lado para ser detectado. (0.2)

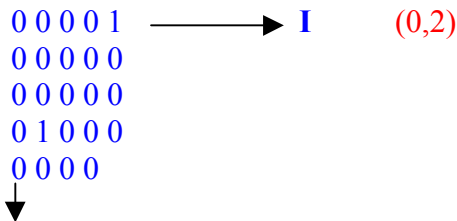
3 - Qual o resultado da abertura da imagem I abaixo pelo elemento X

X= {(0,-1) , (0,0) , (1,0) }

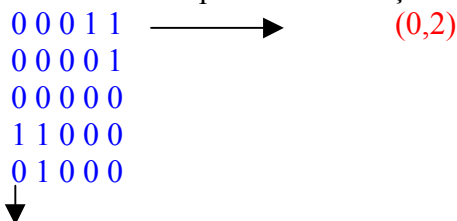
I= {(0,1), (0,3), (0,4), (1,1), (1,4) (2,1), (3,0) , (3,1) , (3,2) ( 3,3) , (4,1) } (valor total da questão 1,2)



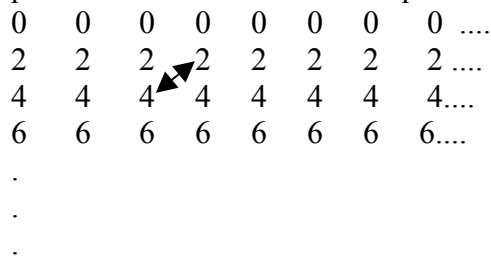
Abertura é a erosão seguida de dilatação  $(0,1)$ . A erosão busca as posições na imagem onde o elemento estruturante se encaixa completamente  $(0,1)$ . Na imagem dada essas posições são  $\{(3,1),(0,4)\}$   $(0,1)$  de modo que a imagem depois de erodida fica:



A dilatação pode ser obtida pela união das imagens resultantes da translação da imagem a ser dilatada por cada pixel do elemento estruturante  $(0,1)$ . A translação da imagem pelo pixel  $(0,-1)$  resulta em:  $\{(3,1-1),(0,4-1)\} = \{(3,0),(0,3)\}$   $(0,1)$ . A translação da imagem pelo pixel  $(0,0)$  não a modifica  $(0,1)$ . A translação da imagem pelo pixel  $(1,0)$  resulta em:  $\{(3+1,1),(0+1,4)\} = \{(4,1),(1,4)\}$   $(0,1)$ . A operação de união dos 3 conjunto resulta em  $\{(3,0),(0,3), (3,1),(0,4),(4,1),(1,4)\}$   $(0,1)$  e a imagem resultante da erosão e posterior dilatação ou sua abertura é:



4. Calcule os histogramas Soma e Diferença do Método de Unser? Para a imagem "Degradê" abaixo, a 45 graus e a distância de 1 pixel. A imagem é de 100 x100 pixels. E é a mesma usada na questão 2 anterior. (Valor total da questão: 1.5)

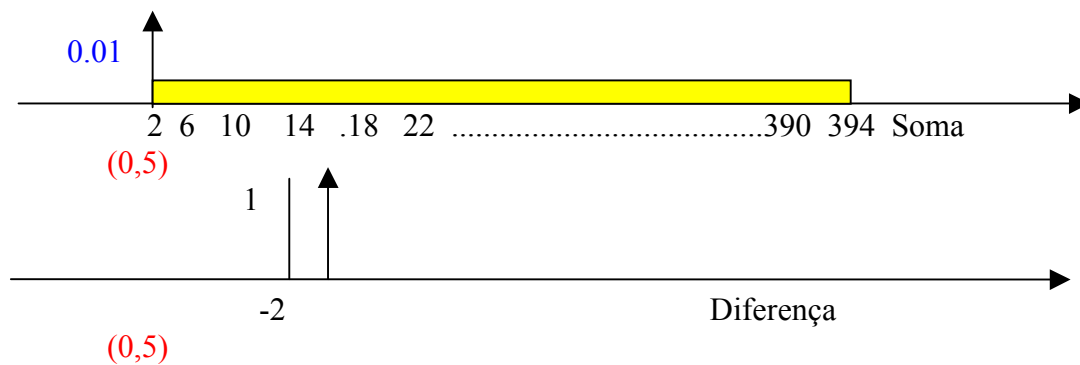


Continuando o degradê de aumento de 2 tons na linha 100 o valor do tom será:200, mas como a imagem é de 100 por 100 pixels, deve ir da linha 0 até a linha 99, ou seja até o tom de 198. Assim a última linha será:  
198 198 198 198 198 198 198 198 .....

As somas de vizinhos de 1 pixel desta imagem a 45 graus terão valores de 2 , 6, 10, 14 18 22 .... a .....  $194+196=390$   $196+198= 394$  e todos com a mesma probabilidade . Como há 100 linhas e 99x99 somas , cada uma ocorrendo 99 vezes, a probabilidade de cada uma das 99 é  $1/99= 0.01010101010101....(0,3)$

As diferenças de vizinhos de 1 pixel desta imagem a 45 graus terão valores de -2 em todas as linhas, ou seja esse valor terá 100% de probabilidade de ocorrer, ou seja 1.

Os histogramas de soma e diferença ficam: (0,2)



5. O que é Gradiente de uma imagem? O que é Laplaciano de uma imagem? Para que estas coisas são úteis? (Valor total questão:0.6)

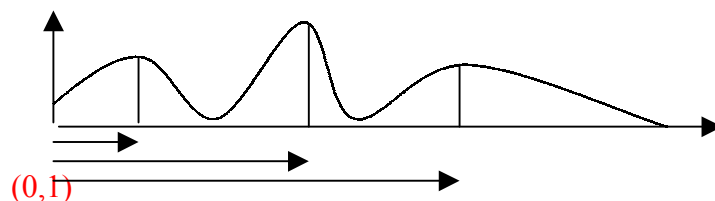
Gradiente de uma imagem é sua primeira derivada. (0,2)

Laplaciano de uma imagem é sua derivada segunda. (0,2)

Ambos são úteis para detecção de contorno ou aumento de contraste, realce ou ênfase na imagem. Podem ser a base de métodos de segmentação, e estarem associados a filtros passa alta. (0,2)

6. Faça um esboço de como seria a aparência de um histograma tri-modal e indique aproximadamente que valores teriam as modas. No seu histograma seria possível uma boa limiarização entre os objetos presentes na imagem? Se sim ou não diga o porque. (Valor total questão:0,5)

Um histograma tri-modal deveria ter 3 modas, ou seja e picos (0,1). Poderia ser algo como:



Neste histograma as modas seriam os valores na horizontal (0,1) correspondentes aos picos verticais. Ou seja os 3 valores indicados no gráfico (0,1)

Sim. Neste histograma como entre as normas há regiões com quase nenhum pixel é possível uma boa separação entre os 3 objetos, ou seja uma boa limiarização. (0,1)

7. Porque a Transformada de Fourier descreve a função no domínio da Frequência? Porque se diz que neste domínio a função tem uma parte Real e outra Imaginária? O que é amplitude e ângulo de fase ? O que é espectro de amplitude ou espectro de

Fourier de uma função? Onde geralmente se concentra a maior parte da informação no domínio da frequência? (Valor total questão:1.0)

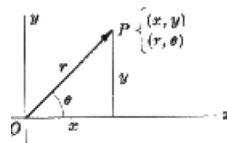
A Transformada de Fourier descreve a imagem no domínio da frequência porque  $\exp[.]$ , considerando as Identidades de Euler:

$$e^{i\theta} = \cos \theta + i \sin \theta, \quad e^{-i\theta} = \cos \theta - i \sin \theta$$

significa combinar senos e cossenos de  $[.]$  em uma só expressão: Já estas funções senos e cossenos são na realidade ondas. Nos fenômenos ondulatório a frequência é a quantidade de ondas que passam em uma região em um determinado tempo. Cada onda senoidal e cossenoidal se repete a cada 360 graus ou  $2\pi$ . Assim o que acaba caracterizando a frequência destas ondas são os valores que estão sendo multiplicados a  $2\pi$ , ou seja o  $u$  ou o  $v$  nas expressões da Transformada de Fourier: (0,3)

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

Neste domínio a função tem uma parte Real e outra Imaginária porque passa a ser descrita como um número complexo. Ou seja um número que tem duas partes: uma real e outra imaginária ( multiplicada por  $j$  ). Como o mostrado na figura abaixo.



A parte real é a representada como função ou série de cossenos e a imaginária como função ou série de cossenos. (0,2)

**Amplitude** de um número complexo é o seu módulo, o  $r$  no gráfico acima. (0,1) O **ângulo de fase** é o ângulo que o ponto, no plano complexo, faz com a horizontal, ou eixo real, o  $\theta$  no gráfico acima. (0,1) No caso da Transformada de Fourier essa **amplitude** e esse **angulo de fase** são dados por:

$$|F(u)| = [R^2(u) + I^2(u)]^{1/2}$$

$$\phi(u) = \tan^{-1} \left[ \frac{I(u)}{R(u)} \right]$$

Se for o caso de se calcular a transformada de Fourier de uma função de uma variável

**Espectro de amplitude ou espectro de Fourier de uma função** é o gráfico de sua amplitude *versus* cada frequência. Ou seja a descrição de qual é a amplitude  $|F(u)|$  para cada frequência  $u$ , em funções de uma variável ou o gráfico de  $|F(u,v)|$  para cada par de frequência  $(u, v)$ , para duas variáveis. Neste último caso o gráfico deve ser representado em 3D ou em 2D com tons de cinza. (0,2)

É na origem dos eixos de frequência onde geralmente a maior parte das informações se concentra. Principalmente em imagens.(0,1)

8. Porque em imagens quando é usada a codificação com perdas é importante avaliar a qualidade da compressão usada? Como o erro do processo pode ser medido? Dê exemplos claros de cada forma. ( Valor total questão: 0.5)

Principalmente pelo fato que só a **taxa de compressão** (bit rate) e o **tempo** do processo não medem a qualidade para o visualizador humano e o quanto foi suprimindo em diferentes níveis. (0.1)

Os erros podem e devem ser avaliados para uma melhor qualificação das técnicas usadas de forma **qualitativas ou quantitativas**. (0.1)

As formas **qualitativas** se baseiam nas diferenças entre as imagens antes da compressão e depois de comprimidas e restauradas. Neste caso é muito importante usar escalas amplificadoras adequadas. (0.1)

As formas de quantificar os erros geram números ou erros médio entre as imagens antes e depois. O mais comum é o **erro médio quadrático**, que faz um somatório entre a diferença de cada pixel das imagens elevado ao quadrado, depois desse número é extraída a raiz quadrada e divide-se, para ter uma média, pelo número de pixels somados:

$$e_{rms} = 1/N^2 (\sum_{x=i}^N \sum_{y=i}^N [e(x,y)]^2)^{0.5}$$

onde  $e(x,y) = (F(x,y) - A(x,y))$ ,  $F$  é a imagem final e  $A$  a imagem antes do processo.

Outros quantificações de erros comuns são a **relação sinal ruído** e a **relação sinal ruído de pico**. Nos **erros médios quanto maior o erro pior a imagem**, nas relações quanto maior o índice melhor a imagem. (0.2)

9- O que é *chain code*? Quais os dois tipos básicos? Em que parte da análise de imagens é geralmente usado? Desenhe uma pequena imagem e diga como seria seu *chain code*. (Valor total questão: 1,5)

É a codificação dos contornos por índices circulares na direção horária ou anti-horária variando de 0 a 3 ou de 0 a 7 de acordo com a posição do próximo ponto do contorno onde se localizar o ponto vizinho do atual. O ponto inicial não é relevante. Podem haver variações dessa numeração de modo que o código seja de 1 a 4 ou de 1 a 8 (0,3) Os dois tipos básicos são : 4 vizinhos e 8 vizinhos. 4 vizinhos considera a posição à direita, esquerda, acima e abaixo do atual. 8 vizinhos, incluem na outro as vizinhanças em diagonal (0,1) .

Essa codificação é usada na caracterização de contornos, vetorização de figuras, extração de medidas ( como: perímetro, área, largura e altura do retângulo envolvente) e reconhecimento de padrões (0,1) .

A resposta da codificação depende do desenho do aluno. (1,0)