

MO443 - Processamento de Imagem Digital

Prof. Helio

Ieremies Vieira da Fonseca Romero

Helio Pedrini's Homepage

Slides [50%]

Introdução

Fundamentos

Amostragem discretiza o espaço e quantização diz o número de níveis de cinza.

Precisamos de uma amostra a cada $1/2B$ onde $2B$ é o tamanho da banda $[-B, B]$ (limite de nyquist). Caso não seja usado uma frequência de banda suficientemente larga, sinais com frequência alta são reconstruídos com frequências baixas causando **aliasing**.

Resolução espacial tem a ver com o quanto de espaço real um pixel representa.

Realce

Filtragem em frequência

Fundamentos

Sinal :: representação de um fenômeno a partir da variação de características físicas ao longo do tempo ou espaço Tipos de sinais:

contínuos estados são definidos ininterruptamente

discretos estados definidos em instantes específicos

analógicos estados podem assumir qualquer valor real ou complexo

digitais estados podem assumir um subconjunto de valores discretos enumeráveis

Sinais analógicos podem ser convertidos em sinais digitais a partir da amostragem, quantização e codificação. Na amostragem, o espaço (tempo) de amostragem é dividido em partes discretas. Na quantização há a aproximação do valor de amplitude ao conjunto finito de possibilidades. Codificação corresponde à correlação de tais níveis a um código.

Transformadas de imagens

Operações que alteram o espaço de representação de uma imagem para outro domínio de forma que:

- a informação presente na imagem seja preservada no domínio da transformada
- a transformada seja reversível

Um exemplo destas é a transformada de Fourier

Amostragem

representação de um sinal contínuo por um número finito de valores

Teorema da amostragem é possível reconstruir um sinal (de banda limitada) contínuo a partir de suas amostras se a frequência de amostragem exceder duas vezes a largura da banda.

Transformada discreta de Fourier (DFT) converte um conjunto finito de N amostras uniformemente espaçadas numa função de senoidais complexas.

Convolução

To entendendo nada, as parece ser a soma dos sinais, um vindo da esquerda pra direita e outro da direita para esquerda.

Segmentação

Particionar o conjunto de dados em estruturas com conteúdo semântico -> após determinarmos os objetos, caracterizamo-nos por meio de suas propriedades.

Os métodos se baseiam nas descontinuidades (mudanças abruptas) e similaridades.

Detectar pontos pode ocorrer por meio da máscara

-1	-1	-1
-1	8	-1
-1	-1	-1

Para segmentos de reta, fazemos algo similar só que com um traço em diversas direções.

Detecção de bordas

Podemos fazê-la a partir do entendimento da mudança brusca nos níveis de cinza, observando a primeira e segunda derivadas.

Assim, podemos também utilizar o conceito de gradiente, dada a natureza \mathbb{Z}^2 das imagens. A direção do gradiente, quando sua magnitude é suficientemente grande, indica a orientação da borda do objeto. Podemos obter uma aproximação suficientemente descente da magnitude do gradiente somando a derivada em cada direção ou pegando a maior delas. Já do ângulo, tiramos o $\arctan(\frac{g_x}{g_y})$. Existem diversas máscaras que aproximam o gradiente.

Laplaciano

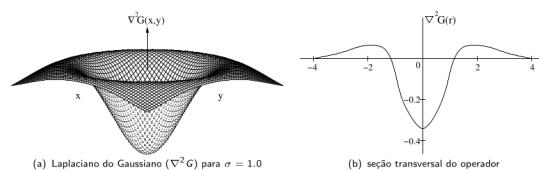
∇^2 Operador que pode ser utilizado no gradiente e é aproximado por uma máscara 3x3 na qual o pixel

central é positivo e todos os outros negativos. Idealmente, os mais próximos possuem valores ainda maiores.

Laplaciano do gaussiano e diferença do gaussiano

As derivadas segundas são sensíveis a ruídos. Podemos reduzi-los aplicando um filtro gaussiano e, posteriormente, detectamos bordas pelo cruzamento em zero da segunda derivada e um pico significativo na primeira.

Resolvendo a equação do laplaciano do gaussiano, temos uma função que parece um chapéu mexicano invertido.



Há evidências que o olho faz algo similar.

Ele pode ser aproximado pela convolução de uma máscara que é a diferença entre duas funções Gaussianas com valores de σ significativamente diferentes.

Canny

- Gaussiano
- computa-se a magnitude e direção do gradiente a partir das derivadas parciais
- borda é pontos cuja magnitude seja localmente máxima na direção do gradiente (supressão não máxima)
- Usa-se um limiar para remover os fragmentos espúrios

Boie e Cox

Similar ao Canny, mas usam o cruzamento do zero na derivada segunda.

Transformada de Hugh

Técnica para determinar se uma curva passa por um certo conjunto de pontos. Para retas, usamos o fato que a reta deve possuir os mesmos coeficientes angulares e lineares, ou fazemos isso usando coordenadas polares, onde ρ é a distância da origem à reta e θ é o ângulo. O espaço (ρ, θ) chamado espaço de Hough, pontos colineares no espaço (x,y) correspondem a curvas senoidais que se interceptam.

Para implementá-lo, discretizamos o espaço (ρ, θ) e cada célula do espaço é considerada uma célula de acumulação. Testamos várias combinações e somamos baseado na quantidade de pontos colineares para tais combinações. Picos são usados para determinar segmentos reais na imagem.

Útil para várias retas, já que apresentaram-se na forma de diversos picos e é insensível a ruídos e descontinuidades, já que estes representaram apenas uma redução do máximo local. O problema é na detecção de formas com mais parâmetros, o que aumenta a computação necessária -> podemos reduzir isso usando o gradiente das curvas.

O mesmo é feito para circunferências. Já para formas em geral, utilizamos a ideia de manter uma tabela de distância e ângulo de cada ponto da forma e utilizamos a ideia de acumulação da mesma forma.

Limiarização

Classificação dos pixels baseado num limiar que os separa em objeto e fundo. No caso global, toda a imagem obedece o mesmo valor, no caso local, este é calculado baseado numa janela da imagem, o que é severamente influenciado pelo tamanho da janela.

No método global, uma forma é tentada minimizar o erro da distribuição da gaussianas que definem o histograma dos pixels.

Segmentação de regiões

1. Cada pixel deve pertencer a uma região da imagem
2. Pixels satisfazem critérios de conectividade
3. Regiões devem ser disjuntas

4. Pixels de uma certa região devem satisfazer uma propriedade
5. Regiões adjacentes devem diferir em tal propriedade

Técnicas:

Crescimento de regiões Agrega pixels com propriedades similares e, portanto, precisa de pixels sementes de onde começar

Segmentação de regiões começamos de uma grande região com a imagem inteira e vamos dividindo (quadtree).

Representação e descrição

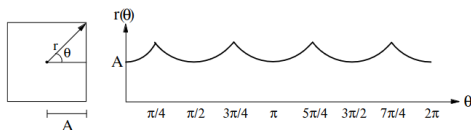
Um objeto pode ser descrito por suas características externas (bordas) ou internas (pixels que compõem). Para descrevê-lo, precisamos de uma forma que seja pouco afetado pela presença de ruído.

Esquemas de representação

- Freeman - código da cadeia cada número representa a direção do próximo pixel da borda (zero é sempre pra direita, crescendo no sentido horário). Ele varia conforme o pixel inicial, mas podemos normalizar encontrando o menor número inteiro que o represente. Para torná-lo invariante com a mudança de escala, a borda pode ser reamostrada em uma grade de espalhamento conveniente. Diferença no código da cadeia é a soma das diferenças entre cada dígito e o próximo (mod vizinhança). Uma mesma borda rotacionada possui a mesma diferença no código da cadeia independente da orientação, o que caracteriza o mesmo objeto. Vantagens: redução do espaço de armazenamento Desvantagem: sensibilidade a pequenas perturbações ao longo da borda.
- Polígonos Representamos a forma a partir de um polígono, dado uma precisão. Em cortonos fechados, podemos torna-la exata se a quantidade de vértices for igual ao número de pontos na

borda. Isso pode ser caro, mas tem certas técnicas para aproximar em tempo aceitável. Uma forma de fazer a borda é dividir em segmentos, toda vez que a distância perpendicular exceder o erro aceitável, esse ponto torna-se um novo vértice.

- Assinatura Função de onda da distância do centro a borda



Invariante em relação a translação, mas precisamos determinar um ponto inicial para ser indiferente a rotação ou normalizar os vetores para obter a invariância em relação à escala.

- Fecho convexo Menor polígono convexo que abrange o objeto
- Esqueleto de um objeto
 - ponto médio: conjuntos de pontos que equidistam de bordas (baseado nas maiores esferas)
 - transformada de distância: mapa da distância entre cada ponto interior, valores mais altos formam o esqueleto
 - diagrama de voroni: para uma forma de n vértices, repartimos o plano em n subconjuntos tais que os pontos interiores daquele conjunto estão mais próximos do que de qualquer outro ponto.

Técnicas de afinamento

Técnicas de afinamentos devem se atentar a não torna-lo desconexo ou causa erosão excessiva.

- Zhang e Suen: um pixel é removido se
 - tem mais que um e menos que 7 vizinhos,
 - se o número de transições de branco para preto na vizinhança ordenada é 1

- se ao menos um dos pixels cima direita baixo é fundo
- se ao menos um dos pixels direita baixo esquerda é fundo

Depois repetimos o processo, só que agora nas últimas duas instruções usamos baixo esquerda cima e esquerda cima direita. Repetimos tudo até não haver mais remoções.

- Holt Refina o método de Zhang: passos de zhang removendo serrilhamento a partir da percepção de que pixels que formam uma escada pode ser removidos sem quebrar a conectividade do objeto.
- Stentiford e Mortimer Quatro máscaras 3x3 com 0 1 1 no meio em cada uma das direções. Um ponto final é um ponto que só possui um pixel preto como vizinho, este não pode ser removido.

Descritores de borda

Diâmetro a maior distância entre dois pontos da borda.

Perímetro o número de pixels na borda.

Curvatura a medida do angulo de dois segmentos de reta da borda (\arctg).

Energia de deformação

$$\text{Energia} = \frac{1}{L} \sum_{\rho=1}^L \kappa^2(\rho)$$

Descritores de região

Área a soma dos pixels 1 ou pela soma abaixo



/home/ieremies/proj/mest/disc/mo443/_20230509_144318screenshot.png

Compacidade o quadrado do perímetro dividido pela área, o círculo possui a menor compacidade

Projeções horizontal e vertical para um pixel (x,y) a projeção horizontal é a soma dos pixels na linha y (resp para vertical)

Propriedades topológicas O número de buracos ou componentes conexas são propriedades invariantes das formas. O número de Euler é dado como $E = C - H$ (conexas - buracos) também é uma propriedade topológica

eixo maior o comprimento do maior segmento de reta (resp menor).

excentricidade o comprimento do maior dividido pelo menor.

Retângulo envolvente o retângulo de menor dimensões que envolve o objeto bidimensional e cujos lados são paralelos aos eixos.

Convexidade mede o grau com que o um objeto difere de um objeto convexo, pode ser dada pelo perímetro convexo dividido pelo perímetro real (o valor será 1 para os objetos convexos).

Retangularidade a área do objeto dividido pela área do retângulo envolvente.

Solidez mede a densidade de um objeto, área do objeto dividido pela área do fecho convexo (1 para sólidos e menor que 1 para objetos com bordas irregulares ou buracos).

- Momento

Sintáticos ou relacionais

...

Morfologia matemática

Construção de operadores para descrição de objetos.

Fundamentos matemáticos

Teoria de conjuntos $\{(x,y) | f(x,y) = 1\} \in \mathbb{Z}^2$, ou seja, conjunto dos pontos 1. Assim definimos propriamente união, intersecção, translação ($A + p = \{a + p | a \in A\}$), reflexão ($\{-a | a \in A\}$), complemento e diferença. Operador morfológico: mapeamento entre um conjunto A que define uma imagem e um conjunto B, chamado de elemento estruturante.

Adição de Minkowski (dilatação) \oplus para cada ponto da imagem, carimbamos o B. (comutativa, associativa).

Subtração de Minkowski (erosão) \ominus para cada ponto da imagem, se o B cabe inteiro, a origem é marcada.

É importante marcar que, para as duas operações, os elementos da primeira imagem não estão necessariamente contidos, quando dizemos “carimbar”, referimos a colocar na imagem resultante que inicialmente é vazia.

As propriedades de associatividade e distributividade podem nos ajudar a reduzir os elementos estruturantes à componentes menores.

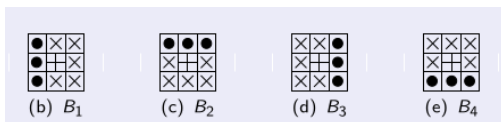
Abertura (círculo vazio) de A por B é denotada por $(A \ominus B) \oplus B \rightarrow$ partes estreitas são removidas

Fechamento (círculo preenchido) de A por B é denotada $(A \oplus B) \ominus B \rightarrow$ buracos são preenchidos

Acerto-ou-erro de A por B_1 e B_2 é definida por $A \circledast (B_1, B_2) = (A \ominus B) \cap (A^c \ominus B_2)$, ou seja, o resultante possui todos os pontos nos quais B_1 coincide com A e não coincide nenhum com B_2

Usaremos tais operadores para extrair características de objetos

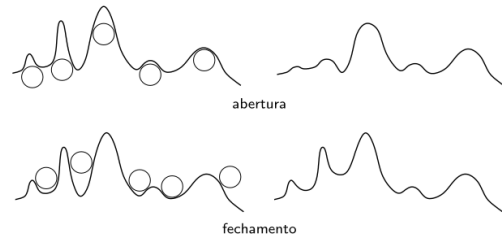
- Extração de borda $E(A) = A - (A \ominus B)$, já que a erosão vai deixar só os pixels de dentro (gradiente interno), é como se eu tivesse marcando a borda por dentro. Ou $E(A) = (A \oplus B) - A$, já que a dilatação vai crescer um pouco o objeto (gradiente externo), é como se eu estivesse marcando a borda por fora. Ou $E(A) = (A \oplus B) - (A \ominus B)$, que é a soma dos dois anteriores (gradiente morfológico).
- Preenchimento de regiões Fazemos a dilatação a partir de um ponto semente dentro do objeto. A cada iteração dilatamos mais, mas também tiramos os elementos que pertencem à borda, para que não ultrapassemos-na.
- Extração de componentes conexos Fazemos a dilatação a partir de uma semente e retiramos todos que não estão em A . Assim, quando não há mais mudança, é porque percorremos todos os pixels conectados destes componente.
- Fecho convexo Sejam B_1, B_2, B_3 e B_4 , o processo consiste em realizar a o acerto-ou-erro para B_1 afim de formar D_1 e assim respectivamente para os demais. O fecho convexo é a união dos D_i .



- Afinação e espessamento Afinação :: $A \otimes B = A - (A \circledast B)$, em que $B = (B_1, B_2)$ distintos. Espessamento :: $A \otimes B = A \cup (A \circledast B)$, em que $B = (B_1, B_2)$ distintos.
- Extração do esqueleto
- Imagens monocromáticas Para imagens monocromáticas, operamos com valores de intensidade. Na dilatação, aplica-se a translação do

elemento estruturante sobre todas as posições da imagem e para cada uma os valores do elemento estruturantes tomando-se o máximo. A erosão é similar, tomando-se o mínimo. Vale mencionar que colocamos o valor final na origem.

A abertura e fechamento podem ser modificados a partir disso. A visão geométrica na qual ordenamos o os pixels e a altura é a intensidade.



- Realce de contraste

top-hat a diferença entre a imagem e sua abertura

bottom-hat a diferença seu fechamento e imagem.

Podemos realçar o contraste da imagem somando a ela o top-hat e subtraindo depois o bottom-hat. O borrimento pode ser reduzido a partir da substituição do pixel original pelo pixel correspondente na abertura ou fechamento, o que estiver mais próximo do original.

- Granulometria Estimativa da distribuição dos tamanhos dos objetos. Podemos fazê-lo contando o número de componentes conexos após cada operação de abertura (progressivamente removendo os elementos em ordem crescente de tamanho).
- Atenuação de ruído Podemos atenuar o ruído sal e pimenta com operações sucetivas de abertura e fechamento.

Lista 1

1. Podemos diferentes diagramas para cada cor componente da imagem (RGB) ou estabelecer alguma relação entre os valores das 3 cores e um

- número, como a soma ou composição de um inteiro.
2. Filtro da média suaviza a imagem e elimina altas frequências, já a mediana exarceba diferenças entre regiões.
 3. As 4 características principais do filtro gaussiano são suficientes para mostrar que ele é bom para suavizar (separabilidade, simétrico, alguma coisa do desvio padrão, controle).
 4. -
 5. Detectar mudanças bruscas, não sofrer grandes alterações com ruído e detectar mudanças na direção desejada, evitar descontinuidades, fina e contínua, evitar fragmentos expúrios, eficiente.
 6. Profundidade: quantidade de bits para cada pixels. Resolução espacial: quantidade de detalhes do ambiente.
 7. está conectado, 9 componentes na vizinhança-4 e 4 na vizinhança-8
 8. O gradiente indica a direção e frequência com a qual a intensidade da imagem está mudando. Valores maiores indicam mudanças bruscas, o que indica bordas no sentido perpendicular a direção do gradiente.
 9. ??????
 10. -
 11. -
 12. -
 13. A entropia é uma medida da quantidade de informações contidas na imagem que se baseia na probabilidade de cada valor de intensidade aparecer. Quanto mais espalhada em cada valor possível, maior a entropia.
 14. Tais operações são utilizadas na erosão e abertura de objetos, bem como detecção de bordas e preenchimento.
 15. A amostragem define as linhas e colunas e está relacionado
 16. Por ser derivado de segunda ordem ele é suscetivo a ruído.
 17. ????
 18. A(s) bin(s) com maior quantidade
 19. -
 20. Dois vetores que multiplicadas dão a matriz que queremos. Se for possível, temos que é separável
- Um filtro passa alta retorna zero numa região homogênea Na convolução o Kernel tem que sofrer uma rotação de 180 (??)

Provas

P1

P2

Trabalhos

Trabalho 1

Trabalho 2

Trabalho 3