



Engenharia de Computação

PROCESSAMENTO DE IMAGENS

TRABALHO 1 - SÍNTESE

Guilherme Baccarin

21 de Junho de 2022

Considerações Iniciais	3
Entendendo o processamento digital de imagens	3
Relação entre processamento de imagens e computação gráfica	3
Natureza interdisciplinar do processamento de imagens	4
Organização do tutorial	4
Conceitos fundamentais	4
Natureza da luz	4
Estrutura do olho humano	4
Modelos cromáticos	5
Modelo de imagem digital	5
Amostragem e quantização	5
Sistema típico para processamento digital de imagens	6
Operações sobre imagens	6
Operações no domínio do espaço	6
Modificação histogrâmica	7
Inversão da escala de cinza	7
Expansão de contraste	7
Equalização histogrâmica	7
Filtragem espacial	7
Morfologia matemática	8
Segmentação	8
Limiarização (thresholding)	8
Segmentação orientada a regiões	8
Segmentação baseada em bordas	9
Extração de características e reconhecimento	9
Exemplos de aplicações	9
Segmentação de imagens	9
Reconhecimento de manuscritos	10
Classificação e recuperação de imagens por conteúdo	10
Considerações finais	11

Considerações Iniciais

Entendendo o processamento digital de imagens

O processamento digital de imagens é um conjunto de tarefas interconectadas, que se inicia pela captura de uma imagem, passa pelo processo de digitalização, pois essa imagem precisa ser representada de maneira adequada para tratamento computacional. As imagens podem ser representadas em mais de uma dimensão.

O primeiro passo efetivo desse processamento é o pré-processamento, onde filtra-se ruídos e é corrigido distorções geométricas.

Para a análise e identificação de objetos é requerido uma cadeia maior de processos. Extrai-se os atributos das imagens, seu background é separado, o que pode ou não ser uma tarefa simples, a depender da complexidade do objeto, mais complexos requerem técnicas sofisticadas como regularização e modelagem, que são técnicas que também podem ser aplicadas em restaurações e reconstruções.

A partir da forma geométrica do objeto, usa-se operadores morfológicos para analisar e modificar a forma e até mesmo para extrair informações do objeto para sua classificação. Essa classificação é uma tarefa de alto nível que visa reconhecer, verificar ou inferir a identidade de um objeto a partir de suas características.

Para objetos mais difíceis, usa-se a retroalimentação, ou, feedback entre as tarefas para ajustar alguns pontos e permitir que a classificação seja viável. Esse método é chamado de visão ativa. Em um cenário de agentes inteligentes, chamamos ciclos de ação-percepção.

Relação entre processamento de imagens e computação gráfica

Em geral ambas as áreas são tratadas como distintas, com o crescimento das áreas, há a resolução de problemas cada vez mais difíceis. A computação gráfica busca o fotorrealismo de cenas 3D geradas pelo computador e o processamento de imagens visa reconstruir uma cena 3D a partir de uma imagem real, advinda de câmera. Em suma, o processamento de imagens é um procedimento inverso ao da computação gráfica, mesmo que sejam sob o mesmo conhecimento.

Natureza interdisciplinar do processamento de imagens

A área de processamento de imagens tem fundamentos que vêm de várias ciências (física, computação, matemática, estatística, etc.) que são necessárias no processamento dessas imagens. Também é interligada à inteligência artificial e percepção visual, por exemplo. Por razões históricas, existem disciplinas que se desenvolveram parcialmente independentes, como a fotogrametria, astronomia e imageamento médico.

Organização do tutorial

Ele é estruturado em duas partes, sendo a primeira os fundamentos de processamento de imagem e a segunda os exemplos de aplicações.

Conceitos fundamentais

Natureza da luz

A luz apresenta um comportamento de onda que é caracterizado pela frequência e o comprimento dessa onda. Existe a luz visível, onde é sensível ao sistema visual humano, mas também existem os comprimentos de onda que não se enquadram na visualização humana. É possível que o olho humano perceba comprimentos de onda como cores diferentes, onde o único comprimento de onda é dado como monocromático e a sua cor, cor espectral pura.

Estrutura do olho humano

O olho humano possui um diâmetro médio entre 2 e 2,5cm, a radiação luminosa penetra na abertura frontal da íris (pupila), e de uma lente chamada cristalino e aí atinge a retina, que é a camada interna do globo ocular.

A retina tem dois tipos de fotossensores, os cones, que são sensíveis a cores e com alta resolução, e os bastonetes, que são insensíveis a cores e com baixa resolução. E eles que se encarregam de converter a energia luminosa em impulsos elétricos para o cérebro, para que possa existir a interpretação.

Toda a câmera fotográfica é uma câmara escura que recebe energia luminosa que visa uma imagem, e trata-se de uma extensão do corpo humano. Podemos comparar a pálpebra humana com o obturador da câmera, o diafragma da câmera, comparamos à íris humana, que controla a quantidade de luz que atravessa essa lente. A lente da câmera, comparada com o conjunto formado pelo cristalino, córnea, humor aquoso e humor vítreo, que focam a

luz e tornam as imagens nítidas. A coróide funciona como a câmara escura e a retina corresponde ao sensor da câmera fotográfica.

Modelos cromáticos

Objetos que emitem luz visível são percebidos como a soma das cores espectrais emitidas. Esse processo é chamado de aditivo. É uma combinação de componentes monocromáticos associados às sensações de verde, azul e vermelho, cores primárias, que formam todas as outras cores registradas pelo olho. Esse processo gerou a concepção chamada de RGB, e a combinação dessas cores, duas a duas e em igual intensidade, nos dão as cores secundárias (ciano, magenta e amarelo, CMY).

A cor oposta de uma cor secundária é a cor primária que não faz parte de sua composição. A cor branca é a combinação balanceada de vermelho, verde e azul ou, a combinação de qualquer cor secundária com sua cor oposta.

O processo de composição cromática é a absorção ou reflexão, de componentes verdes, vermelhos e azul, da radiação eletromagnética visível.

A formação de imagens em vídeo é um processo que integra verde, vermelho e azul. Dispositivos de impressão usam o sistema CMY, mas como o pigmento de tintas e cartuchos não produzem preto, é necessário usar um quarto pigmento, e esse sistema passa a se chamar CMYK.

Existem outros modelos em que a caracterização da cor se dá por outros atributos, como intensidade, matiz ou tonalidade (hue) e a saturação ou pureza.

Modelo de imagem digital

Uma imagem é uma função bidimensional contínua, como os computadores não são capazes de processar imagens contínuas, apenas arrays de números digitais é necessário apresentar imagens como arranjos bidimensionais de pontos. Cada ponto é denominado elemento de imagem, ou pixel. Em uma imagem digital colorida em RGB, por exemplo, um pixel pode ser visto como um vetor onde seus componentes representam as intensidades de vermelho, verde e azul de sua cor. Podem ser três imagens monocromáticas, cada uma em uma das cores, que combinadas, formam uma imagem colorida.

Amostragem e quantização

Para uma imagem ser armazenada ou até mesmo processada, é necessário que a sua discretização em nível de coordenadas espaciais e valores de brilho, esse processo de coordenadas espaciais é chamado de amostragem enquanto o processo de valores de brilho é chamado de quantização. Esses processos implicam a supressão de informação de uma imagem analógica, onde seu equivalente digital é uma aproximação.

Sistema típico para processamento digital de imagens

Na década de 80 surgiram diversas empresas especializadas no desenvolvimento de software para o processamento de imagens. Atualmente, ainda é de muito uso os sistemas de processamento de imagem, entretanto, tendências recentes voltam-se para a miniaturização de integração do hardware especializado nesse processamento de imagens a computadores de pequeno porte. Esse hardware é composto por componentes.

No que tange a aquisição, também conhecida como sensoramento de imagens digitais, é relevante dois elementos: dispositivo físico sensível à faixa de energia irradiada pelo alvo e o dispositivo conversor da saída, usualmente chamado como digitalizador.

No geral, hardware especializado é um digitalizador integrado a um hardware destinado à execução de operações primitivas, uma unidade lógico-aritmética para realizar as operações aritméticas e lógicas de imagens inteiras conforme são digitalizadas. Esse hardware é chamado de subsistema front-end, e seu diferencial é a velocidade de processamento.

Em nível de processamento, o computador é um hardware de uso geral que pode ser desde um PDA até um supercomputador. A depender da capacidade de processamento que se é exigida.

O armazenamento se torna um dos grandes desafios, pois é necessário cada vez mais espaço de armazenamento à medida que os sistemas se tornam mais aprimorados e os volumes, conseqüentemente, maiores. O armazenamento é composto por três níveis: armazenamento de curta duração, armazenamento online e arquivamento de imagens.

No âmbito de saída, normalmente encontramos duas alternativas: visualização de dados e impressão de dados. A primeira, exige monitor de vídeos coloridos que recebem dados de placas gráficas, enquanto o segundo usa dispositivos de impressão de pequeno, médio ou grande porte.

O software para processamento de imagens são módulos que se destinam a realizar tarefas específicas. E a conexão em rede de sistemas para processamento de imagens é cada vez mais necessária.

Operações sobre imagens

Operações no domínio do espaço

Essas são caracterizadas pela manipulação direta dos pixels de uma imagem. A imagem de saída é o resultado da combinação de uma ou mais imagens de entrada. Essas operações podem ser pontuais ou locais, na primeira, cada pixel da imagem de saída depende apenas do mesmo correspondente da imagem de entrada e qualquer operação pode ser interpretada como um mapeamento de pixels da imagem de entrada para uma imagem de saída, na segunda, o valor de saída em uma coordenada específica depende dos valores de entrada daquela mesma coordenada e da sua vizinhança.

Modificação histogrâmica

O realce de contraste visa melhorar a qualidade das imagens sob vista do olho humano. Ele é definido como a razão entre seus níveis de cinza médio. O histograma de uma imagem é a distribuição de seus níveis de cinza. Logo, o realce desse contraste é a redistribuição desses níveis de cinza.

Inversão da escala de cinza

A inversão da escala de cinza pode se tratar do negativo da imagem, após um registro fotográfico a partir de uma câmera convencional, existe a revelação do negativo do filme, que virará uma imagem positiva.

Expansão de contraste

O baixo contraste vem de um problema no processo de digitalização de uma imagem, essa redução de contraste de uma cena dificulta o seu entendimento, expandindo o contraste, podemos redistribuir os tons de cinza dos pixels de uma imagem e melhorar o discernimento de seus componentes.

Equalização histogrâmica

Um processo que visa o aumento da uniformidade da distribuição dos níveis de cinza de uma imagem, que visa realçar diferenças de tonalidade em uma imagem.

Filtragem espacial

Imagens apresentam bordas, as bordas, por sua vez, representam, alterações bruscas entre intervalos de níveis de cinza, sua representação gráfica é caracterizada por gradientes acentuados.

Os alvos que variam mais uniformemente com a distância, apresentam-se em forma de regiões homogêneas, correspondendo a feições de baixa frequência.

As técnicas de manipulação de contraste, de filtragem de imagem implicam em transformações pixel a pixel. A alteração de uma imagem filtrada não depende apenas do pixel de cinza da imagem original, mas também dos valores de sua vizinhança.

A filtragem espacial se fundamenta no uso de uma máscara e da imagem digital considerada. A máscara é um arranjo onde existem valores a serem aplicados sobre os pixels da imagem, é uma operação progressiva em forma de matriz (coluna a coluna, linha a linha).

Entre os filtros mais utilizados, estão o de média, mediana e moda. Todos eles com intuito de suavizar uma imagem, atenuando variações dos níveis de cinza.

Filtro da média de ordem: a sua suavização é em função do tamanho da vizinhança considerada, a cada interação da convolução da máscara de filtragem com a matriz original, extrai a média dos valores dos pixels em uma vizinhança.

Filtro da mediana de ordem: produz como pixel de saída a mediana dos valores dos pixels da imagem de entrada em uma vizinhança. A preservação da definição das bordas das regiões é superior ao do filtro de média.

Filtro de moda de ordem: produz como valor do pixel de saída a moda dos valores do pixel da imagem de entrada em uma vizinhança.

Enquanto esses filtros são para suavização de imagem, existem outras categorias, como os operadores de gradiente, que visam produzir a acentuação ou aguçamento de regiões de uma imagem.

Morfologia matemática

É uma modelagem destinada à análise da forma de um objeto digital. É baseado no conceito de que a imagem é um conjunto de pontos elementares (pixels ou voxels) que formam conjuntos bidimensionais ou tridimensionais, e a inter-relação entre eles formam a morfologia da imagem.

As suas operações básicas são a erosão, onde são removidos os pixels que não atendem a um padrão, a dilatação, onde uma área é alterada por um dado padrão. Elas são usadas para criação de transformações mais sofisticadas.

Segmentação

É a subdivisão da imagem em partes ou objetos constituintes. Algoritmos de segmentação permitem diferenciais de dois ou mais objetos, entre si ou entre si e o background. Na segmentação de imagens monocromáticas, o algoritmo é baseado na descontinuidade ou similaridade dos níveis de cinza. O primeiro, apresenta a imagem em zonas de mudanças bruscas de níveis de cinza e o segundo, consiste na limiarização e crescimento de regiões.

Limiarização (thresholding)

É a abordagem para a segmentação fundamentada na análise da similaridade de níveis de cinza que visa extrair objetos de interesse mediante a definição de um limiar que separa os agrupamentos de níveis de cinza de uma imagem.

Segmentação orientada a regiões

É fundamentada na similaridade dos níveis de cinza de uma imagem, o crescimento de regiões é um procedimento que agrupa pixels de uma imagem em regiões maiores. A variante mais simples é a agregação de pixels, onde define-se uma semente, um conjunto

de pontos similares em valor de cinza, onde as regiões crescem com a agregação de cada pixel à semente que tenha propriedades similares.

Segmentação baseada em bordas

A detecção de bordas possibilita a análise de descontinuidades dos níveis de cinza, e caracterizam contornos dos objetos nelas presentes. Pontos de borda são as posições dos pixels com variações abruptas de níveis de cinza. Esses pontos caracterizam as transições entre objetos diferentes.

Extração de características e reconhecimento

O objetivo é identificar objetos em cena a partir de um conjunto de medições. Cada objeto é um padrão e os valores medidos são suas características. Um conjunto de objetos similares com características similares são pertencentes à mesma classe de padrões. A forma de um objeto pode ser descrita em termos de suas bordas. Para realizar o reconhecimento de objeto, existem várias técnicas que são divididas em dois tipos principais:

Classificação baseada em aprendizagem supervisionada: /esses algoritmos podem ser paramétricos ou não paramétricos, onde o primeiro é treinado com uma grande quantidade de amostras rotuladas para que possa ser estimado o parâmetro de cada classe de padrão.

Classificação baseada em aprendizagem não-supervisionada: nela, o classificador particiona o conjunto de dados de entrada por um critério de similaridade, que resulta em um conjunto de clusters.

Exemplos de aplicações

Segmentação de imagens

Ela tem como principal objetivo a separação de objetos de interesse do background da imagem. Na segmentação por limiarização a escolha do limiar depende de características intrínsecas da imagem. Numa abordagem alternativa foi proposto um método de modelagem perceptiva que aprende a decisão humana na limiarização de uma rede de funções de base radial, uma máquina que permite aproximar a função que mapeia as características da imagem em limiares escolhidos por humanos. As decisões são armazenadas em uma tabela de 2 colunas, onde a primeira armazena o limiar escolhido e a segunda uma característica global da imagem.

Como resultado da comparação da modelagem perceptiva com três outros métodos automáticos de segmentação por limitação, foi verificado que as respostas humanas tinham alta relação com alguns dos métodos avaliados, demonstrando a viabilidade da abordagem. Uma abordagem para a segmentação dos blocos de endereço em envelopes postais baseada em histogramas e a operação de watershed: a partir então do histograma, é realizada um processo de agrupamento (operação de watershed) das regiões que correspondem a blocos de endereços e carimbos postais, selos e background de envelope.

O ponto de encontro entre duas áreas de watershed é uma borda, que é o resultado final da operação. Para evitar a supersegmentação, foi feita uma sequência de erosões morfológicas para reduzir o número de regiões a 3.

Reconhecimento de manuscritos

Os manuscritos tem várias formas, fontes e estilos. Então, foram apresentadas duas abordagens para o reconhecimento de palavras isoladas dos meses do ano, uma baseada em redes neurais e outra em modelos de markov escondidos. Em três estágios, primeiro foi a correção da inclinação geral dos caracteres, depois, detecção e correção de inclinações na linha de base da palavra inteira e por fim, utilização de um filtro para atenuação de imperfeições e falhas nos manuscritos.

Para o classificador neural, a imagem dos manuscritos foi dividida em 8 sub-regiões, correspondendo a aproximadamente o número médio de letras no conjunto de palavras a serem reconhecidas. Para cada sub-região, extrai-se 10 características perceptivas a partir da análise direcional dos pixels.

Para o classificador de modelos markov escondidos, a partir do histograma de projeção horizontal dos pixels da imagem, três zonas foram definidas. Ascendente, corpo e descendente. Um processo de segmentação variável, dependente das transições escuro-claro presentes na linha central da palavra e para cada segmento, é identificado características perceptíveis e características baseadas em deficiências na concavidade/convexidade dos traços dos segmentos.

Resultado: a melhor taxa de reconhecimento ocorreu para o classificador neural 81,8%, enquanto foi possível obter uma taxa de reconhecimento muito superior combinando 3 classificadores (um em modelos markov escondidos e 2 em redes neurais) foi possível encontrar um resultado de 90,4% de apuração.

Classificação e recuperação de imagens por conteúdo

Está relacionado fortemente com áreas de sistema de informação e bancos de dados. Então, é necessário incluir campos e operações sobre imagens em uma consulta de dados. Um sistema para classificação de imagens da web em duas classes semânticas, gráficos e fotografias foi apresentado. O sistema utilizou do método de classificação baseado em árvores de decisão. Houve a identificação de um conjunto de características adequadas à separação entre duas classes semânticas escolhidas. Essas características foram: existência de objetos reais, pequenas diferenças na proporção, poucas ocorrências de regiões com alta saturação de cores, presença de um grande número de cores. As características classificadas como marcantes de gráfico foram a presença de objetos artificiais com bordas bem definidas e presença de regiões cobertas com cores saturadas, grandes diferenças na proporção e tendência a serem menores em tamanho do que as fotografias.

Assim, foram definidas métricas sobre número de cores, cor predominante, vizinho distante, saturação, histograma de cores e do vizinho mais distante, proporção das dimensões e menor dimensão.

Na fase experimental foram definidos dois conjuntos de treinamento, e para cada conjunto, foram extraídas métricas. A aplicação do algoritmo gerou uma árvore de decisão para a classificação de cada conjunto. As taxas médias de classificação em imagens de teste, não utilizadas durante o treinamento, são de 97,3% para GIF e 93,9% para JPG.

Considerações finais

É fornecido uma visão geral da área de PDI, permitindo uma reciclagem ou um primeiro contato de profissionais dos diferentes setores da economia, e mesmo que não conseguindo incluir todos os tópicos relevantes, fornece o mínimo de detalhes de cada etapa de processamento em um sistema típico de PDI.