

Introdução à visão computacional

Augusto de Holanda B. M. Tavares

Universidade Federal da Paraíba
Centro de Informática
Departamento de Sistemas de Computação

25 de junho de 2024

- ▶ A visão computacional é uma disciplina que estuda os métodos de **reconstrução** e **interpretação** de **cenas em três dimensões** a partir de **imagens em duas dimensões** de acordo com as **propriedades das estruturas** presentes em uma determinada cena.

- ▶ Coloca-se uma pergunta: o conceito de visão computacional colocado anteriormente pode ser estendido para a visão humana?
- ▶ O sistema visual humano opera em quantas dimensões?
- ▶ Quais são as suas limitações?
- ▶ O ser humano é capaz de fornecer interpretações e reconstruir uma cena a partir da sua visão?

Visão computacional

- Descreva a cena ao lado.



- ▶ Desde que uma cena se encontre dentro dos contextos com os quais um indivíduo possui algum grau de experiência, em geral este é capaz de realizar com facilidade um conjunto de tarefas de alta complexidade:
 - ▶ Classificação: atribuir a uma cena ou a objetos dentro desta o pertencimento a uma determinada categoria.
 - ▶ Localização: determinar a posição de objetos dentro da cena.
 - ▶ Segmentação: separar uma cena ou objetos dentro desta em partes componentes.
 - ▶ Atribuição semântica: definir um significado para a cena de acordo com uma percepção geral do que está representado nesta.
 - ▶ Contextualização: definir o papel de objetos na cena em termos de situações específicas.

- ▶ Sendo assim, é possível considerar que os problemas mais complexos de visão computacional estão resolvidos de maneira trivial para o caso do ser humano.
- ▶ No entanto, e o que ocorre no caso não humano?
- ▶ Especificamente, considere os sistemas automatizados que desejam realizar interação humano máquina/ambiente máquina.
- ▶ Caso a interação não seja dependente de dados visuais, não há problema.
- ▶ E nos casos em que a informação visual é crítica para a aplicação?

Visão computacional

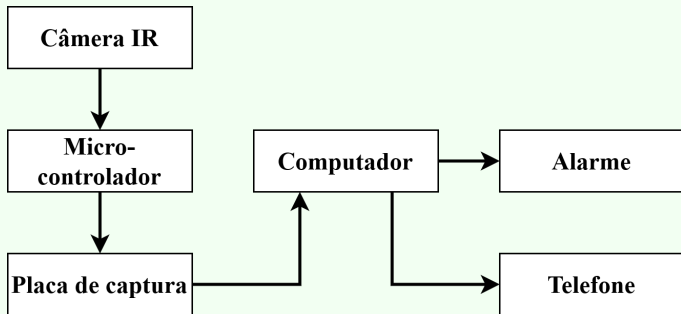
- ▶ Quando a informação visual é crítica, o sistema deve ser capaz de realizar parte ou todas as tarefas listadas anteriormente (classificação, localização, segmentação).
- ▶ Considerando os sistemas baseados em computadores, o problema evidente é que eles não possuem sequer a capacidade de enxergar, muito menos de interpretar imagens.
- ▶ O primeiro problema pode ser "resolvido" com a inclusão de sensores, na forma de câmeras, mas a interpretação exige um esforço adicional.
- ▶ Assumindo que uma imagem é fornecida em bom estado em um formato compatível com a máquina, como, por exemplo, esta pode identificar a presença de um objeto?
- ▶ Em outras palavras, como um computador pode fazer o que um ser humano faz ao se deparar com uma imagem?

- ▶ Assim, o objetivo final da visão computacional é modelar, replicar e eventualmente **exceder** a visão humana utilizando uma combinação de *software* e *hardware* em diferentes níveis. Para isto é necessário combinar os conhecimentos de ciência da computação, engenharia elétrica, matemática, fisiologia, biologia e ciência cognitiva, entre outros.

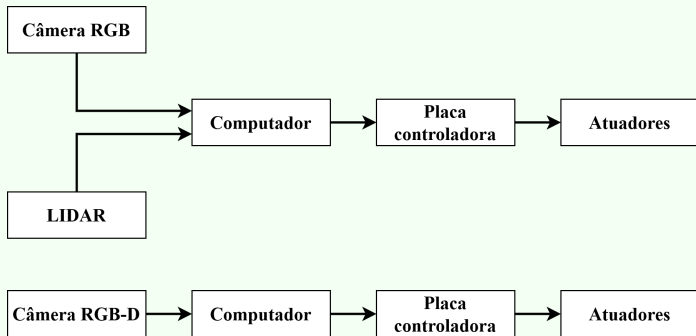
- ▶ Como é possível exceder o sistema visual humano?
- ▶ Limitações do sistema visual humano:
 - ▶ Limitação do espectro visível: o olho humano é sensível a uma faixa de comprimentos de onda estreita do espectro eletromagnético.
 - ▶ Memória: a medida que o tempo passa, há uma modificação acerca da percepção dos detalhes contidos em uma determinada cena.
 - ▶ Ilusões de ótica: existem determinados padrões visuais que interagem de maneira negativa com o sistema visual humano.

- ▶ Possíveis vantagens de um sistema visual computacional:
 - ▶ Espectro visível arbitrário: com a escolha apropriada de sensor e as técnicas de processamento corretas, a informação luminosa pode ser observada em qualquer faixa do espectro eletromagnético dentro dos limites tecnológicos.
 - ▶ Memória: contanto que o arquivo correspondente a uma determinada imagem ainda esteja armazenado, este pode ser recuperado e reinterpreteado exatamente da mesma maneira que quando a imagem foi capturada inicialmente.
 - ▶ Correção de ilusões de ótica: as aberrações e peculiaridades de um sistema de captação de imagens podem ser modeladas e corrigidas.

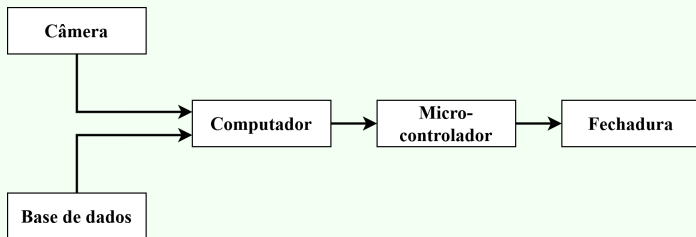
Exemplo



Exemplo



Exemplo



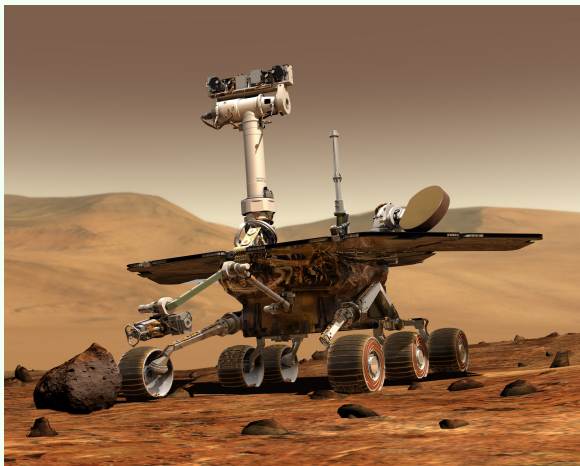
Exemplo

- ▶ Robótica.
- ▶ Medicina.
- ▶ Segurança.
- ▶ Transporte.
- ▶ Automação industrial.
- ▶ Bases de dados de imagem/vídeo.
- ▶ Interface humano-computador.

Exemplo

- ▶ Em robótica:
 - ▶ Determinar a localização do robô de maneira automática.
 - ▶ Desvio de obstáculos.
 - ▶ Servo-controle visual.
 - ▶ Montagem.
 - ▶ Manipulação.
 - ▶ Interação humano-robô.

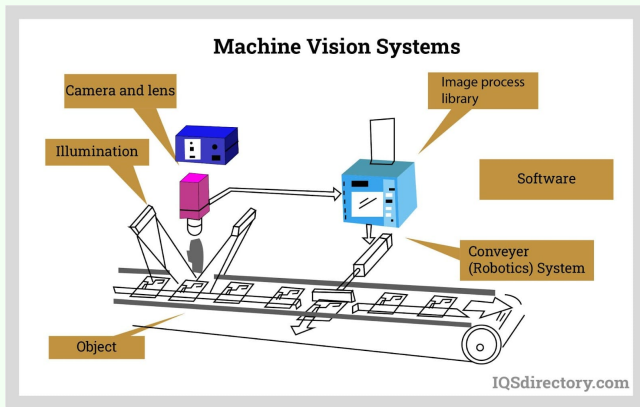
Exemplo



Exemplo

- ▶ Em automação industrial:
 - ▶ Inspeção industrial.
 - ▶ Montagem.
 - ▶ Leitura de códigos de identificação.
 - ▶ Organização de objetos.
 - ▶ Leitura de documentos.

Exemplo

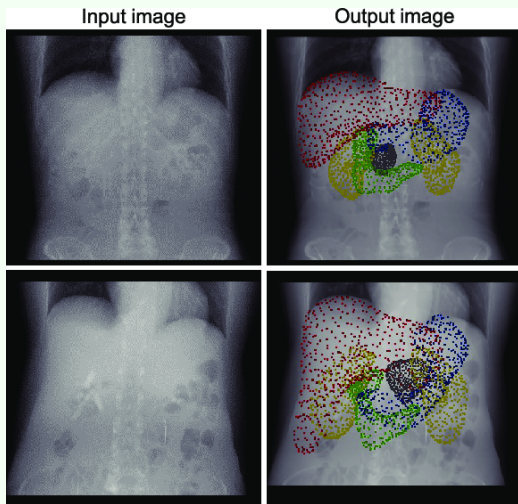


► Seleção de tomates

Exemplo

- ▶ Em medicina:
 - ▶ Classificação e detecção.
 - ▶ Segmentação.
 - ▶ Reconstrução de órgãos (ressonâncias, ultrassom).
 - ▶ Cirurgia robótica guiada por visão.

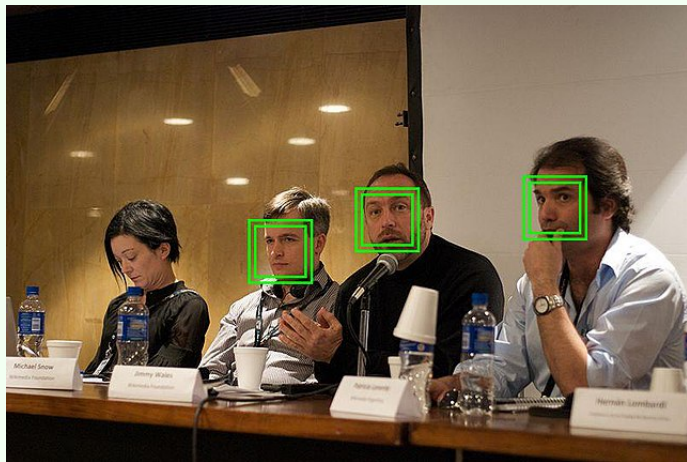
Exemplo



Exemplo

- ▶ Em segurança:
 - ▶ Biometria (reconhecimento ocular, impressão digital, reconhecimento facial).
 - ▶ Vigia (detecção de movimentos e comportamentos).
- ▶ Em transporte:
 - ▶ Veículos autônomos.
 - ▶ Segurança no trânsito.

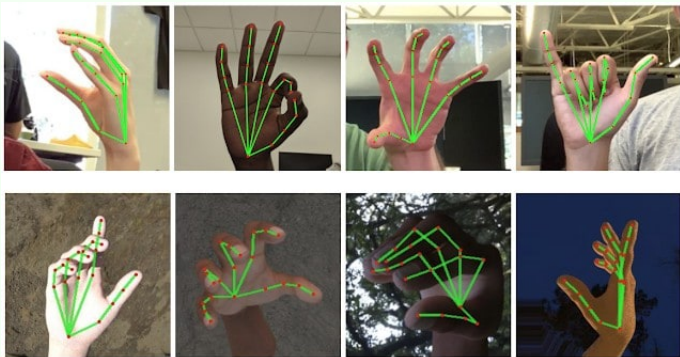
Exemplo



Exemplo

- ▶ Em bases de dados de imagem/vídeo:
 - ▶ Recuperação de imagens
- ▶ Em HCI:
 - ▶ Reconhecimento de expressão.
 - ▶ Estimação de olhar.
 - ▶ Reconhecimento de gestos.

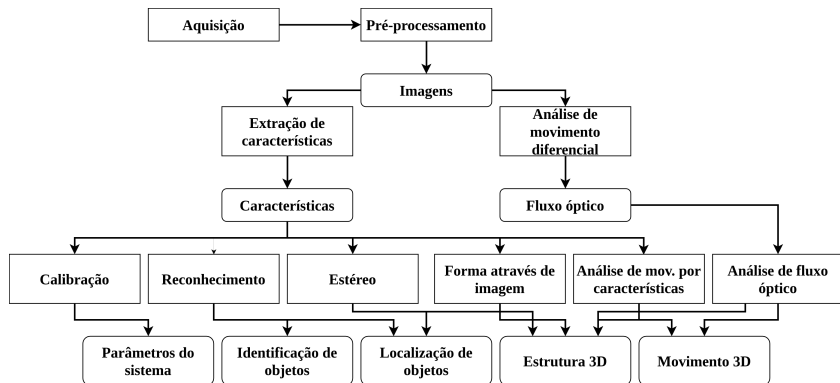
Exemplo



Níveis de visão computacional

- ▶ Visão de baixo nível: processamento para extração de características (bordas, cantos, fluxo óptico, etc.).
- ▶ Visão de nível médio: Reconhecimento de objetos, análise de movimento, reconstrução 3D a partir de características de baixo nível.
- ▶ Visão de alto nível: interpretação da informação fornecida pelo nível médio ao longo do tempo, coordenação das tarefas de nível baixo e médio.

Níveis de visão computacional



Visão computacional como um problema

- ▶ Em termos matemáticos, um problema é considerado **bem colocado (well-posed)** se ele atende aos seguintes pré-requisitos:
 - ▶ O problema possui uma solução.
 - ▶ A solução é única.
 - ▶ A solução muda de maneira contínua com as condições iniciais do problema.
- ▶ Qualquer problema que não atenda a pelo menos uma destas condições é considerado **mal colocado (ill-posed)**
- ▶ Evidentemente, problemas bem colocados possuem resolução mais fácil do que problemas mal colocados.

Visão computacional como um problema

- ▶ É possível argumentar que os problemas de visão computacional podem apresentar mais de uma solução correta, ao se considerar diferentes graus de especificidade.
- ▶ No entanto, certamente eles estão sujeitos a variações súbitas na solução quando há pequenas variações nas condições iniciais.

Exemplo

- ▶ Um sistema de diagnóstico médico pode mudar sua avaliação drasticamente a partir de pequenas mudanças na imagem analisada.
-
- ▶ Isto é extremamente problemático quando se considera a presença de perturbações e ruído no sistema.

Visão computacional como um problema

- ▶ Ainda, em geral o processamento de imagens em geral constitui um problema inverso.
- ▶ Isto é, a imagem é o resultado de algum evento cuja causa se deseja conhecer.

Exemplo

- ▶ Restauração de imagens.
 - ▶ Verificar a presença de um objeto.
 - ▶ Detectar movimento.
-
- ▶ Sendo assim, a visão computacional engloba problemas inversos mal colocados, o que a torna um tópico de estudo desafiador.

- ▶ A visão computacional enquanto área de estudo é formada pela junção de áreas do conhecimento de aplicação mais geral.
- ▶ Dentre estas, duas são de particular importância:
 - ▶ Processamento de imagens digitais.
 - ▶ Reconhecimento de padrões.

- ▶ O reconhecimento de padrões utiliza técnicas matemáticas diversas, como métodos estatísticos, redes neurais e máquinas de vetor de suporte, para classificar diferentes padrões de informação.
- ▶ *A priori*, não há restrições para o tipo de dado utilizado como entrada do processo.
- ▶ Diversos dos problemas de visão podem ser formulados como problemas de classificação, o que significa que este ferramental pode ser aplicado

- ▶ O processamento de imagem neste contexto se refere aos processos que transformam uma imagem em outra.
- ▶ Algumas operações típicas são:
 - ▶ Compressão.
 - ▶ Restauração (tratamento de ruído).
 - ▶ Melhoria de qualidade.
- ▶ Algoritmos de visão computacional assumem que foram realizados processos para melhorar a qualidade da imagem de entrada.

Objetivos

- ▶ Entender os conceitos fundamentais da visão computacional.
- ▶ Projetar e implementar um subconjunto dos algoritmos de visão computacional.
- ▶ Aplicar técnicas de visão computacional à aplicações de interesse.