

Máquinas que enxergam

Use a visão computacional com suas diversas técnicas para melhorar a acessibilidade dos sistemas.
por Alessandro de Oliveira Faria (Cabelo)



Segundo a Wikipédia, “Visão computacional é a ciência e tecnologia das máquinas que enxergam. Ela desenvolve teoria e tecnologia para a construção de sistemas artificiais que obtêm in-

formação de imagens ou quaisquer dados multidimensionais. Exemplos de aplicações incluem o controle de processos (como robôs industriais ou veículos autônomos), detecção de eventos, organização de informação,

modelagem de objetos ou ambientes e interação (atrelado à interação homem-computador).”

Projetos de biometria, realidade aumentada e robótica utilizam reconhecimento de padrões

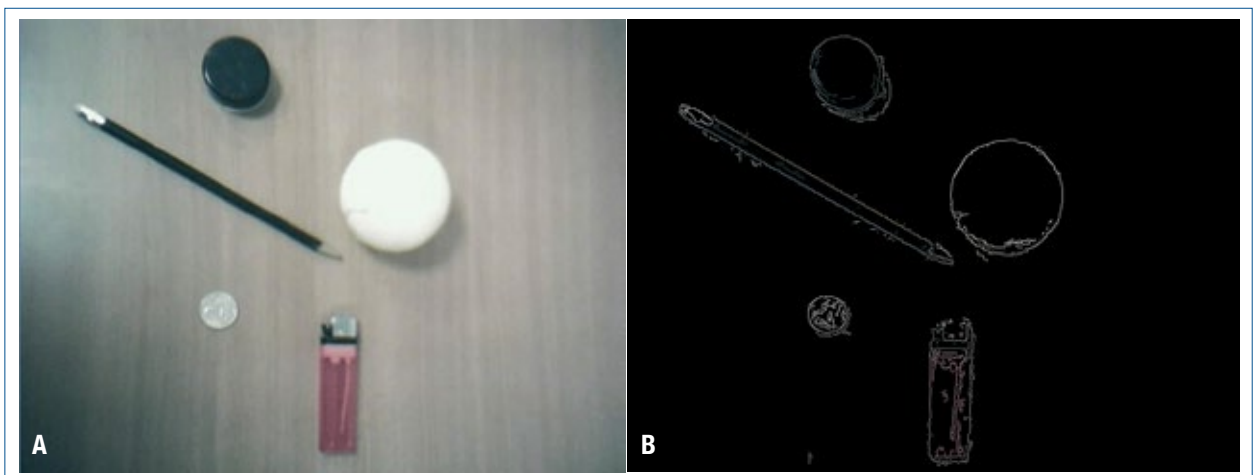


Figura 1 Esta imagem (A) pode ser tratada com o reconhecimento de arestas de forma a gerar a imagem B.



Figura 2 A técnica de detecção de cantos permite encontrar diferenças entre duas imagens.

e processamento de imagens em tempo real. Como consequência, as câmeras ultimamente vêm ganhando inteligência em função dos algoritmos de visão computacional. Essa tarefa demanda conhecimentos matemáticos, fundamentos relacionados a imagens e conhecimentos sobre manipulação de vídeo ao vivo, como também, para maior eficiência, otimização no uso da arquitetura de hardware e utilização de processadores com múltiplos núcleos.

Aplicar processos de visão computacional em fluxos de vídeo ao vivo é uma tarefa morosa devido ao grande consumo de processamento matemático. Entretanto, atualmente existem diversas bibliotecas projetadas para facilitar o aprendizado e o desenvolvimento de tarefas como essa, proporcionando assim diversas funções para o segmento do processamento de vídeo digitais.

As bibliotecas *Mimas* [1] e *OpenCV* [2], por exemplo, permitem o processamento de imagens estáticas ou ao vivo. Podem ser aplicadas em reconhecimento de face, expressão, gestos, objetos, rastreamento e detecção de movimentos e também análise de movimentos, entre outras infinitas aplicações.

As bibliotecas de visão computacional apresentam recursos de manipulação de imagem, en-

trada e saída de vídeo, cálculo de matriz, álgebra linear, morfologia matemática, análise estrutural, reconhecimento de objetos e gráficos primitivos.

Geralmente, o processamento de imagens tem como principal objetivo obter uma nova figura ou quadro, nos quais o resultado apresenta um determinado resultado comparado à imagem inicial ou anterior. Diversas técnicas para manipulação de imagens são utilizadas geralmente para um posterior processamento.

A técnica de suavização é utilizada para borrar ou embaçar a imagem, geralmente com o objetivo de eliminar ruídos ou detalhes da figura. Os processos mais utilizados são a suavização com o somatório dos pixels vizinhos, ou mediana de uma vizinhança, entre outros tipos.

Arestas

Um método de manipulação de imagens bem conhecido é a detecção de arestas, relativamente simples quando comparada a outros recursos no segmento de visão computacional. Ele consiste basicamente em processar uma imagem de entrada e localizar e exibir a descontinuidade das intensidades rastreadas (figura 1).

Cantos

O método de detecção de cantos, por sua vez, pode ser utilizado para rastrear objetos. Essa técnica consiste basicamente na extração da característica a partir de uma determinada imagem para posteriormente submetê-la a comparações de desigualdade. Nesse caso, características do quadro atual são comparadas à imagem base para ser processadas, obtendo assim as diferenças entre os dois e, por consequência, localizando os pontos com na figura 2.

Segmentação

Podemos entender como uma forma de segmentação a submissão de imagens para processamento com fins de localização de figuras ou modelos específicos. Um exemplo de uso disso é a separação entre objetos específicos do fundo e da imagem (figura 3). Portanto, essa é uma técnica muito utilizada em projetos de realidade aumentada.

Tudo junto

A junção de processos também é uma técnica muito comum. Por exemplo, é possível utilizar o realce matemático junto ao processo de separação de modelos, tendo como

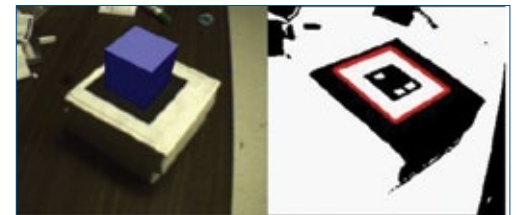


Figura 3 Ao separar o fundo do restante da imagem, é possível inserir uma figura externa à imagem original.



Figura 4 A rotulação de imagem binária permite usos interessantes em Realidade Aumentada, como a detecção de figuras pré-definidas em uma imagem apresentada à câmera.



Figura 5 Com a biblioteca Bazar, é possível inserir uma figura como base...



Figura 6 ...para usá-la como apoio para um objeto virtual.

resultado final o que se denomina *deteção de linhas*.

Em projetos de realidade aumentada, a imagem é realçada usando os processos de deteção de arestas e separação de modelos, seguidos da tarefa de traçado de contornos. A rotulação de imagem binária também é um processo importante dentro do campo da realidade aumentada, pois significa determinar um valor único de cor para cada pixel pertencente a uma determinada região (**figura 4**).

Rastreamento

No segmento de visão computacional, o rastreo de objetos (*object tracking*) é um recurso muito interessante. Com ele, é possível acompanhar o percurso de um objeto em tempo real, isto é, rastrear um objeto em movimento em frente a uma câ-

mera por meio do fluxo de cada pixel, ou com base na semelhança e na combinação de amostras.

Bazar

A biblioteca *Bazar* [3] utiliza técnicas de rastreamento e análise de pixels para produzir uma aplicação de realidade aumentada. A instalação dessa biblioteca é simples: bastam os comandos padrão `./configure && make && sudo make install` após baixar e descompactar o código-fonte. Além disso, ela é bem fácil de utilizar.

Ao executá-la, é possível apresentar um objeto à câmera (**figura 5**) e em seguida usá-lo de forma interativa com uma figura inserida na imagem pelo computador (**figura 6**).

Fluxo óptico

A técnica de fluxo óptico se baseia no rastreamento dos pixels. Ela permite estimar o deslocamento dos pixels em até três dimensões. Uma das aplicações desse recurso é o reconhecimento facial na autenticação (**figura 7**), de forma a impedir

que o simples uso de fotografias subverta o processo de verificação [4]. O projeto *ehci* [5] é um bom começo para prova de conceito.

Reconhecimento de objetos

Utilizando classificadores, é possível identificar objetos presentes num vídeo ao vivo ou em quadros estáticos. O treinamento é efetuado com a seleção de diversas imagens positivas (objeto presente na imagem), com o objeto presente em diversos ângulos e condições de iluminação. Outro grupo de imagens são agrupadas sem a presença do objeto (imagens negativas). Esse grupo de imagens é utilizado no treinamento com *cascata de classificadores* para posterior deteção.

Exemplos de aplicação para o reconhecimento de padrões treinados são os processos para contagem de pessoas em tempo real. Neles, um processo analisa em paralelo o fluxo de vídeo ao vivo, reconhecendo assim padrões de corpos humanos e rastreando seu movimento durante sua passagem pelo campo de visão da câmera (**figura 8**).

Outro exemplo de aplicação da visão computacional é o projeto *kmando* [6], um sistema de projetor



Figura 8 Usando o reconhecimento de face em três dimensões, não basta apresentar uma fotografia sua para alguém fazer login com o seu usuário.



Figura 8 Após o devido treinamento, é possível contar quantas pessoas estão presentes numa imagem e até rastrear seu movimento num vídeo.

interativo baseado em processamento em tempo real. Após o treinamento da câmera junto ao sistema, é possível mover o cursor do mouse por meio de uma caneta hidrocor ou objeto semelhante.

O *eViacam* [7] é um projeto que explora bem os recursos mencionados neste artigo. Voltado à acessibilidade, ele fornece uma interface para interpretar os movimentos da face e convertê-los em movimentos e ações para o cursor do mouse, como mostra o vídeo em [8]. A acessibilidade computacional significa prover a pessoas portadoras de necessidades especiais a capacidade de operar computadores de forma transparente. Para atingir tal objetivo, nada mais lógico e sensato que o aproveitamento da evolução do hardware e software para o desenvolvimento de mecanismos e interfaces, como faz o *eViacam* [9].

OpenCV

A biblioteca livre *OpenCV* (*Open Source Computer Vision*) reúne vários algoritmos de visão computacional para integração a qualquer aplicativo. Para instalá-la a partir do código-fonte, baixe o pacote compactado a partir

do site do projeto [2], descompacte-o e execute os tradicionais comandos:

```
./configure
make
sudo make install
```

Conclusão

A área de visão computacional é tão ampla que talvez seria necessário um livro para explicar todos os tópicos. Este artigo demonstra o potencial dos algoritmos de visão computacional agregados a qualquer dispositivo de videocaptura, pois esses conceitos matemáticos estarão presentes em todos os dispositivos (webcams, celulares, câmeras digitais, filmadoras e outros equipamentos) em curtíssimo prazo. ■

Sobre o autor

Alessandro Faria é sócio-proprietário da NETI Tecnologia, fundada em junho de 1996 e especializada em desenvolvimento de software e soluções biométricas. Ele é consultor biométrico na tecnologia de reconhecimento facial, atuando na área de tecnologia desde 1986. Leva o Linux a sério desde 1998, desenvolve soluções de código aberto, é membro colaborador da comunidade Viva O Linux, mantenedor da biblioteca de código aberto de vídeo captura, entre outros projetos dos quais faz parte.

Gostou do artigo?

Queremos ouvir sua opinião. Fale conosco em cartas@linuxmagazine.com.br

Este artigo no nosso site: <http://nm.com.br/article/3169>

Mais informações

[1] Mimas: <http://vision.eng.shu.ac.uk/mediawiki/index.php/Mimas>

[2] OpenCV: <http://sourceforge.net/projects/opencvlibrary/>

[3] Biblioteca Bazar: <http://cvlab.epfl.ch/software/bazar/index.php>

[4] Biometria facial na autenticação do usuário root: <http://tinyurl.com/y9jrbk9>

[5] Projeto ehci: <http://code.google.com/p/ehci/>

[6] Projeto kmando: http://vision.eng.shu.ac.uk/mmlwiki/index.php/Interactive_Camera-Projector_System

[7] eViacam: <http://eviacam.sourceforge.net/>

[8] Movimentos do mouse com a face: <http://tinyurl.com/o5pjb4>

[9] Uso do eViacam para acessibilidade: [http://www.vivaolinux.com.br/artigo/Acessibilidade-Movimentos-do-mouse-com-a-face-\(eViacam\)](http://www.vivaolinux.com.br/artigo/Acessibilidade-Movimentos-do-mouse-com-a-face-(eViacam))

[10] Jogo Levelhead (Realidade Aumentada): <http://tinyurl.com/cwhqrc>

[11] Zebra Barcode Reader: <http://tinyurl.com/ybwwnef>

[12] ARToolkit: Criação de aplicativos de Realidade Aumentada: <http://tinyurl.com/y8hnh4r>

[13] Acessibilidade para tetraplégicos em Li NETI: <http://www.youtube.com/watch?v=kuN1XznWRZ4>