

Fundamentos e Tecnologias para Percepção – Trabalho 1

Profa. Dra. Silvia da Costa Botelho

Felipe Albring Guth - 99781

Rio Grande, Setembro de 2012.

Sistema de reconhecimento de placas (ANPR – auto number plate recognition). http://en.wikipedia.org/wiki/Automatic_number_plate_recognition#Algorithms



É um sistema de vigilância que utiliza um método de reconhecimento ótico de caracteres em imagens para leitura de placas de veículos obtidas, por exemplo, através de câmeras de tráfego. Sistemas deste tipo geralmente usam infravermelho para possibilitar a fotografia em qualquer período do dia. A tecnologia tende a ser específica para cada região dado os diferentes formatos de placas.

ANPR usa a reconhecimento ótico de caracteres (OCR – optical character recognition), obtidas através de câmeras.

Existem seis algoritmos primários para identificar uma placa de licença:

- 1: Localização da placa - responsável por descobrir e isolar a placa na imagem.
- 2: Orientação e tamanho da placa - compensa a inclinação da placa e ajusta as dimensões para o tamanho desejado.
- 3: Normalização - ajusta o brilho e o contraste da imagem.
- 4: Segmentação de caracteres - encontra os caracteres individuais nas placas.
- 5: Reconhecimento ótico de caracteres.
- 6: Análise sintática / geométrica - verifica caracteres e posições em relação a regras específicas regionais.

A complexidade de cada uma destas subsecções do programa determina a precisão do sistema. Durante a terceira fase (normalização), alguns sistemas utilizam técnicas de detecção de borda para aumentar a diferença de imagem entre as cartas e o suporte de placa. Um filtro de mediana pode também ser utilizado para reduzir o ruído visual sobre a imagem.

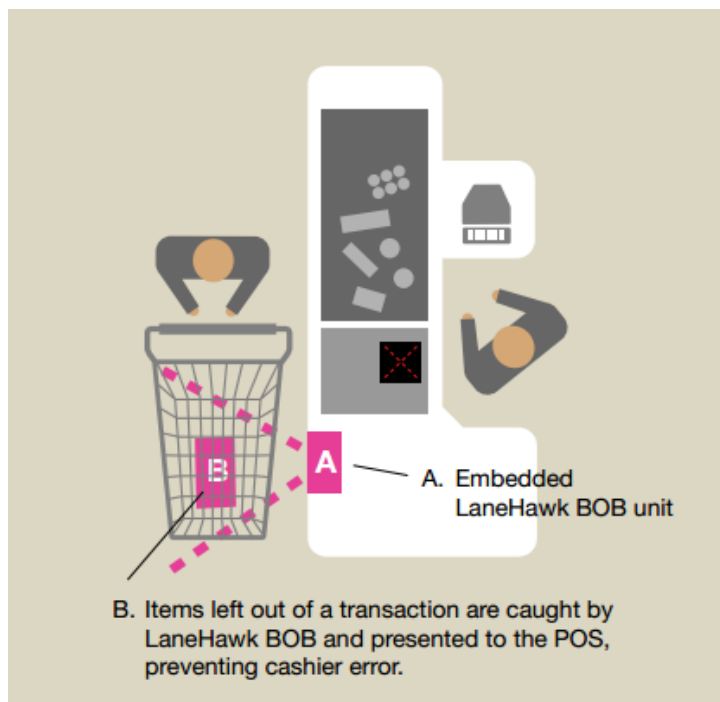
Smile Detection <http://www.sony.co.uk/hub/learnandenjoy/2/2>



Quando Obturador de Sorriso está ativado a câmera irá tirar fotos quando as pessoas sorrir. É um recurso para a captura de sorrisos espontâneos e naturais.

Reconhecimento de objetos em supermercados

<http://www.evoretail.com/lanehawk/>



LaneHawk é uma solução de prevenção de perdas que evita a perdas de produtos em caixas eletrônicos propiciando um maior lucro a empresa.

A solução inicia com um scanner visual que reconhece os itens sem ter que ler o código de barras. A solução integra-se facilmente com qualquer sistema POS, incluindo self-checkout. Uma câmera inteligente está embutida na esteira de verificação geral, observando continuamente por itens. Quando um item é detectado e reconhecido usando a tecnologia patenteada, suas informações UPC são enviadas diretamente através de uma conexão Ethernet para o PDV. O caixa verifica os itens que foram encontrados embaixo da cesta e continua a fechar a transação. O item pode permanecer sob a cesta, com o sistema LaneHawk BOB, tem-se a certeza que o item será cobrado.

Vision-based biometrics – How the Afghan Girl was Identified by Her Iris Patterns

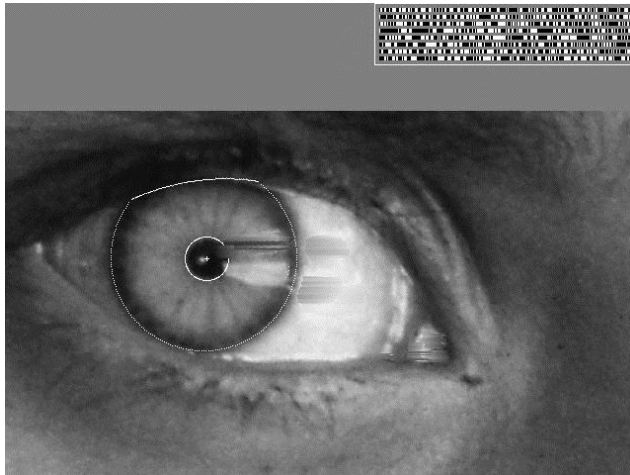
<http://www.cl.cam.ac.uk/~jgd1000/afghan.html>



O link apresenta a história da garota afegana que foi reconhecida através de seu padrão de íris depois de um período de 18 anos da sua primeira fotografia.

Primeiramente foi rodado o algoritmo de reconhecimento de íris sobre ambos os olhos da foto tirada em 1984. Gráficos sobrepostos mostram a localização da íris e os seus limites, conforme imagem abaixo. Após isso, foi rodado o

algoritmo de reconhecimento de íris na foto de 2002 para obter-se os códigos de identificação da mesma.



O terceiro passo foi executar o algoritmo de correspondência dados as codificações das íris. O sistema apresentou então uma chance de 6 milhões contra uma sobre a correspondência do olho esquerdo entre as duas imagens e uma correspondência de 10^{15} para o olho direito.

A National Geographic aceitou e publicou esta conclusão em uma edição de capa com Sharbat Gula, 18 anos depois da primeira foto, e foi lançada o “Fundo menina afegã” para ajudar a educação de meninas muçulmanas em culturas que desencorajam ou proíbem a educação feminina.

Sistemas de reconhecimento de face para login sem senha

<http://www.sensiblevision.com/>

Sistemas de reconhecimento de face são usados de forma abrangente para controle de acesso em computadores comerciais e pessoais. Propiciando maior segurança e agilidade o reconhecimento de face tem sido empregado em diversos setores da indústria a fim de proteger dados importantes como senhas de contas bancárias, informações confidenciais de trabalho, dados pessoais.



Lembrar de todas as senhas é uma tarefa complexa e difícil. Sistemas de reconhecimento de face reconhecem e rastreiam continuamente a face do usuário para a entrada segura e automática em web sites com acesso por senha, reduzir o tempo de login com o uso de senhas, providenciar tempo integral de segurança para os dados quando o usuário se afasta do computador.

Este tipo de tecnologia é confiável a ponto de ser usada por bancos e hospitais e mesmo assim está disponível para usuários domésticos e pequenas empresas. Fácil de usar, muito preciso e confiável.

Reconhecimento de objetos em dispositivos móveis

Atualmente os smartphones são presentes em todo o mundo, potencializando a execução das mais diversas tarefas, o que tempos atrás só poderia ser feito com computadores. Um dos itens mais importantes de um smartphone é a sua câmera, através desta pode-se fazer o uso de algoritmos de visão computacional para a realização de várias tarefas de forma rápida e eficaz propiciando utilidades importantes ao usuário, como mostrado com os exemplos a seguir.

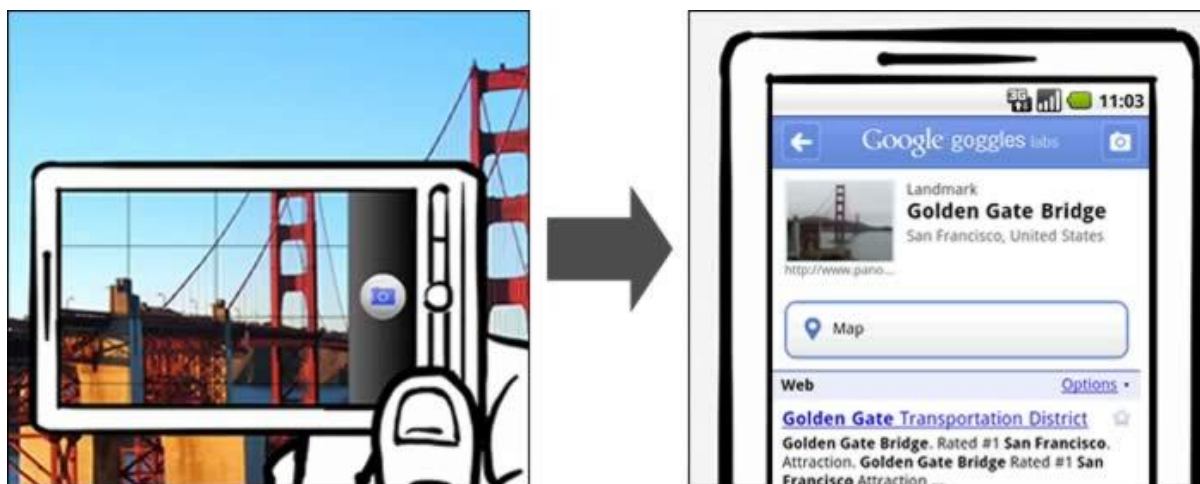
Google googles - <http://www.google.com/mobile/goggles/>

Nokia - <http://www.infoworld.com/d/networking/nokia-taps-silicon-valley-761>

Reconhecimento de texto e tradução automática:

Tirando uma foto de um texto o sistema automaticamente reconhece os caracteres e traduz para a língua padrão do usuário.

Reconhecimento de pontos de referência: Ao tirar foto de um ponto de referência o algoritmo faz uma busca na internet para encontrar uma imagem correspondente e apresentar ao usuário informações sobre a mesma.



Esse mesmo sistema é usado para buscar informações de livros, obras de arte, vinhos, logos de empresa, etc. Usa-se um algoritmo que busca na internet imagens semelhantes e apresenta ao usuário informações do mesmo.

How the First-Down Line Works

<http://entertainment.howstuffworks.com/first-down-line.htm>

Um dos principais objetivos do futebol Americano é ganhar um “first down”. Para isto, a ofensiva deve ganhar 10 metros dentro de uma série de quatro jogadas, ou quedas. Para lidar com o problema de medir exatamente os 10 metros requeridos nas jogadas, empresas de TV fazem o uso de um vasto aparato tecnológico e técnicas de visão computacional para demonstrar uma linha no televisor dos espectadores (realidade virtual) demarcando a marca dos 10 metros.



A fim de determinar onde a linha deve ser desenhada, um computador central utiliza diferentes informações:

O campo virtual modelado a partir de medições do campo real (medida antes do jogo), e os dados a partir da montagem das câmeras mostra a escala de cada vista de câmara.

A alimentação de vídeo da câmera que está atualmente no ar (determinada por um computador separado no caminhão da produção Sportvision).

Duas paletas de cores distintas, uma que representa as cores no campo que deve ser alterado para amarelo para representar a primeira linha para baixo, e outra representando cores que não devem ser alterados (como aquelas dos uniformes dos jogadores e funcionários - isto permite que um jogador apareça para “ocultar” a linha, fazendo a linha aparecer como se estivesse realmente “pintada no campo”).

Uma vez que o computador determina exatamente que pixels devem ser de cor amarela, esta informação, juntamente com o sinal bruto de vídeo (no ar) da câmara, é enviada para um computador que desenha a linha amarela, 60 vezes por segundo. A linha é então enviada para um manipulador linear para sobrepor a linha amarela para o vídeo do programa.

Smart cars – Mobileye

<http://www.mobileye.com/>

Sistemas de visão computacional atualmente estão presentes em vários modelos de carros, principalmente no mercado asiático, europeu e americano. Estes sistemas proporcionam diversas ferramentas de auxílio ao motorista, como por exemplo:

Aviso de colisão com pedestre, aviso de colisão frontal, aviso de troca de pista, monitoramento e aviso referente ao limite de velocidade encontrado em placas, monitoramento de distância frontal em referencia a outros carros e aviso de perigo ou segurança referente a distância.

The screenshot displays the Mobileye website's main content area. At the top, there are navigation tabs for 'manufacturer products' and 'consumer products'. The central banner features the slogan 'Our Vision. Your Safety.' above a top-down view of a car with yellow beams representing the fields of view of its 'rear looking camera', 'forward looking camera', and 'side looking camera'. Below the banner, there are three main product sections: 'EyeQ Vision on a Chip' with an image of a chip, 'Vision Applications' showing a pedestrian on a crosswalk, and 'AWS Advance Warning System' with a circular radar-like graphic. To the right of these sections is a sidebar with 'News' and 'Events' categories, each containing links to recent press releases and events.

Google cars

http://www.nytimes.com/2010/10/10/science/10google.html?_r=2&ref=artificialintelligence

Automóvel autônomo, também conhecido como "robótico", "inteligente" ou também "sem motorista", é um veículo automático capaz de preencher as necessidades humanas de transporte atendidas por um automóvel tradicional. Como um veículo autônomo, ele é capaz de perceber todo ambiente à sua volta e dirigir sozinho. Um humano pode escolher o destino, mas não é necessário para realizar nenhuma operação mecânica do veículo.



O coração do sistema é um localizador a laser instalado no teto do carro. O dispositivo, um feixe de laser Velodyne, gera um mapa detalhado em 3D do entorno. O carro então combina as medições do laser com os mapas globais em alta resolução, produzindo diferentes tipos de modelos de dados para que o veículo se auto-dirija enquanto desvia de obstáculos e respeita as leis de trânsito.

O carro também conta com outros sensores: quatro radares instalados nos para-choques dianteiro e traseiro, para que o carro “enxergue” mais longe e reconheça o tráfego rápido das vias expressas; uma câmera posicionada no espelho retrovisor, que detecta as luzes dos semáforos; e um GPS, unidade de medida inercial e codificador de posicionamento, que determinam a localização do veículo e rastreiam seus movimentos.

Duas coisas parecem particularmente interessantes na visão do Google. Primeiro, o protótipo usa mapas muito detalhados de estradas, vias e do tipo de terreno, o que é essencial para determinar localização exata do veículo. Se usassem apenas técnicas baseadas em GPS, a localização poderia falhar por vários metros, explicou.

A segunda coisa é que, antes de enviar o carro autônomo para um teste nas ruas, os engenheiros do Google dirigem pelo trajeto uma ou mais vezes para

obter dados do entorno. Quando o veículo passa a ser auto-controlado, ele compara os dados que está captando com os registrados previamente, um recurso útil para diferenciar pedestres de objetos imóveis, como postes e caixas de correio.

Vision-based interaction (and games)

O uso de câmeras para interações em jogos também representa um vasto número de aplicações. Algoritmos são usados para capturar expressões e rastrear pessoas e seus movimentos através de imagens de câmeras especiais que utilizam muitas vezes tecnologias de infravermelho. Em jogos compatíveis, os movimentos realizados pelo usuário são rastreados e executados no ambiente virtual do jogo.



O surgimento de tecnologias como o Microsoft Kinect permitiu o desenvolvimento de um série de aplicações que fazem uso dos dados sensoriais em aplicações de reconhecimento de objetos, rastreamento de movimentos, controle e demais aplicabilidades em cenários virtualmente construídos.



Vision in space and Mobile robots

<http://marsrovers.jpl.nasa.gov/gallery/images.html>

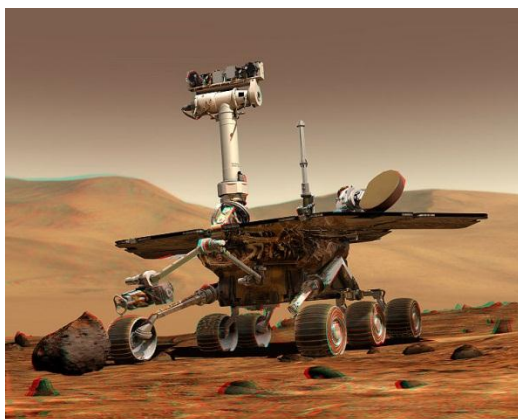
http://www.ri.cmu.edu/publication_view.html?pub_id=5719

A Visão computacional tem e continuará a desempenhar um papel importante no aumento da autonomia de sondas e veículos robóticos. Algoritmos de visão computacional usados em Marte na Exploração de Marte da NASA / JPL Rover missão (MER), que foi um grande passo em frente no uso de visão computacional no espaço. Esses algoritmos de visão estéreo faz odometria visual para a navegação e rastreamento para a estimativa da velocidade horizontal para o Landers.



Outros sistemas de visão computacional são utilizados para diversas tarefas como:

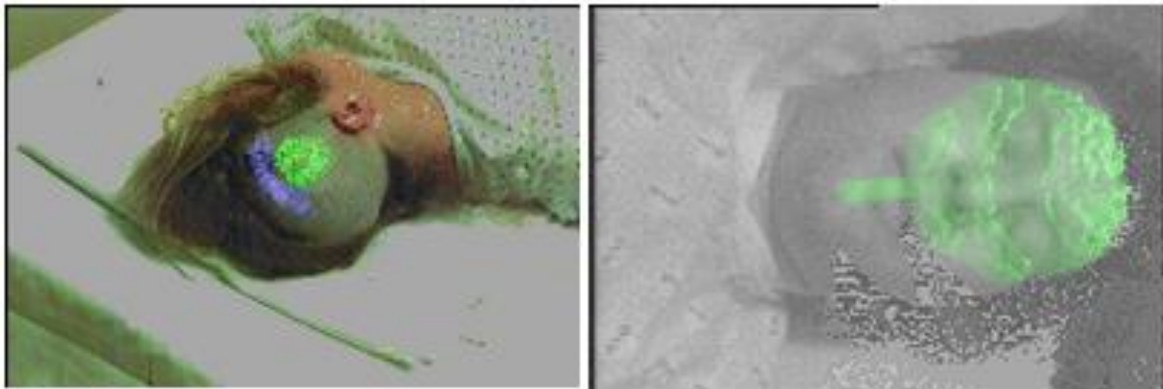
Moldagem de panorama, modelagem do terreno em 3D. Detecção de obstáculos e rastreamento de posição.



Medical Imaging

http://groups.csail.mit.edu/vision/medical-vision/surgery/surgical_navigation.html

O Grupo de Visão Computacional do Laboratório de Inteligência Artificial do MIT tem vindo a colaborar de perto por vários anos com o Laboratório de Planejamento Cirúrgico da Brigham e do Hospital da Mulher. Como parte da colaboração, as ferramentas estão sendo desenvolvidas para apoiar a cirurgia guiada por imagem. Tais ferramentas permitem cirurgiões visualizar estruturas internas através de uma sobreposição automática de reconstruções em 3D da anatomia interna em cima de pontos de vista de vídeo ao vivo de um paciente.



Usa-se técnicas de renderização de superfície para exibir as estruturas segmentadas de ressonância magnética. Este procedimento consiste em primeiro extrair superfícies limitantes do volume segmentado do MRI utilizando o algoritmo de marcha cubos. Este algoritmo gera um conjunto de triângulos ligados para representar a superfície 3D para cada estrutura segmentada. Estas superfícies são exibidas ao selecionar um local virtual de visualização de câmera e orientação no quadro de coordenadas da ressonância magnética e usando técnicas de computação gráfica para projetar a superfície sobre a visão da câmera. Este processo remove porções escondidas da superfície, as máscaras da superfície de acordo com o seu local normal e, opcionalmente, varia a opacidade da superfície para permitir vislumbres de estruturas internas.

Pode-se retirar a “skin” da ressonância magnética e ver onde as estruturas internas são localizadas em relação ao ponto de vista da câmera. Assim, o cirurgião tem visão de raio-x, uma capacidade que será cada vez mais necessária à medida que continuamos caminhando para cirurgias minimamente invasivas.

