

Interfaces baseadas em técnicas de visão computacional

Prof. Dr. Carlos Hitoshi Morimoto

Departamento de Ciência da Computação - IME/USP
<http://www.ime.usp.br/~hitoshi>

Tópicos

- Interação Homem Computador (IHC)
 - ▶ Definição
 - ▶ Aspectos principais
 - ▶ Princípios de usabilidade
- Visão Computacional
 - ▶ Definição
 - ▶ Análise de movimento e forma
- IHC e Visão

Projetos em andamento

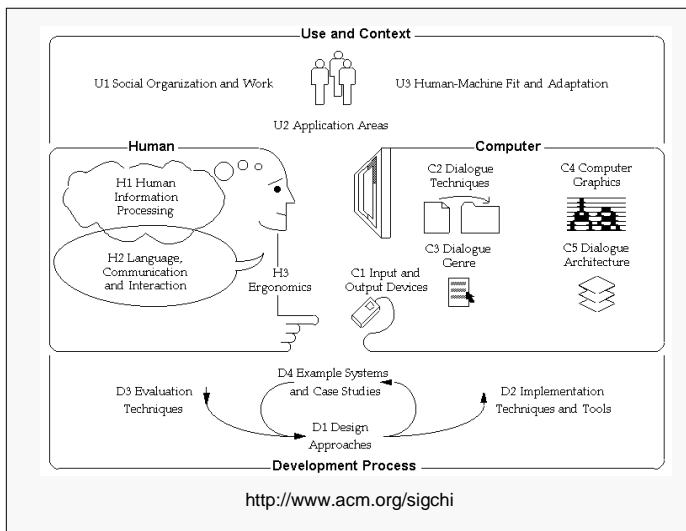
- Rastreamento
 - ▶ de objetos
 - ▶ de olhar
 - ▶ de faces
- MAGIC - Manual And Gaze Input Cascaded
- Sutor
- MAGIC Mirror (Reconhecimento de expressões faciais)

IHC - Motivação

- Computadores cada vez mais permeiam o nosso ambiente.
- Apesar disso, os aspectos principais de projeto de sistemas era fundamentados na funcionalidade e eficiência apenas (eng. de software).
- Como explicar também que alguns aplicativos são odiados por usuários mesmo atendendo a todas as especificações de seus projetistas?
- O problema é que uma ferramenta tão complexa quanto um computador deve levar em conta mais do que funcionalidade, especialmente quando pessoas sem experiência técnica estão envolvidas.

IHC - Componentes

- **Homem (mulher, humano):**
 - ▶ O usuário final de um programa aplicativo (todos nós).
- **Computador:**
 - ▶ A máquina onde o programa roda
- **Interação:**
 - ▶ O usuário "diz" ao computador o que ele quer
 - ▶ O computador "comunica" os resultados



Diversos aspectos da área de IHC

- Aspectos sobre o uso e contexto
 - ▶ U1 - Organização social e do trabalho
 - ▶ U2 - Áreas de aplicação
 - ▶ U3 - Ajuste/adaptação entre homem e máquina
- Aspectos humanos
 - ▶ Processamento de informações
 - ▶ Linguagem, comunicação e interação
 - ▶ ergonomia

Aspectos de IHC

- Aspectos computacionais (sistemas e interfaces)
 - ▶ Dispositivos de entrada e saída
 - ▶ Técnicas de diálogo
 - ▶ Formas de diálogo
 - ▶ Computação gráfica
 - ▶ Arquiteturas para diálogo
- Processo de desenvolvimento
 - ▶ Técnicas de projeto
 - ▶ Técnicas e ferramentas de implementação
 - ▶ Técnicas de avaliação

Para que serve IHC?

- Histórico:
 - ▶ Mainframes
 - ▶ Desktops (shells)
 - ▶ WIMP
 - ▶ Pervasive
- Objetivo de IHC: melhorar a usabilidade
 - ▶ produzir aplicações que são seguras, úteis, e funcionais (de fácil utilização).

O que é usabilidade?

- Fácil de aprender
 - ▶ requer pouco treinamento formal
- Fácil de memorizar
 - ▶ lembrar como usar depois de algum tempo.
- Maximizar a Produtividade
 - ▶ Fazer a tarefa de forma rápida e eficiente
- Minimizar a taxa de erros
 - ▶ se eles acontecerem, avisar o usuário e permitir uma fácil correção
- Maximizar a satisfação do usuário
 - ▶ dar-lhe confiança e segurança

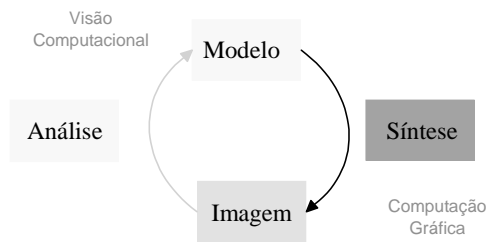
In December 1995, American Airlines Flight 965 departed from Miami on a regularly scheduled trip to Cali, Colombia. On the landing approach, the pilot of the 757 needed to select the next radio navigation fix, named "ROZO." He entered an "R" into his navigation computer. The computer returned a list of nearby navigation fixes starting with "R" and the pilot selected the first of these, whose latitude and longitude appeared to be correct. Unfortunately, instead of "ROZO," the pilot selected "ROMEO," 132 miles to the northeast. The jet was southbound, descending into a valley that runs north-south, and any lateral deviation was dangerous. Following indications on the flight computer, the pilots began an easterly turn and slammed into a granite peak at 10,000 feet. One hundred and fifty two passengers and all eight crewmembers aboard perished. Four passengers survived with serious injuries. The National Transportation Safety Board investigated, and-- as usual-- declared the problem human error. The navigational aid the pilots were following was valid but not for the landing procedure at Cali. In the literal definition of the phrase, this was indeed human error, because the pilot selected the wrong fix. However, in the larger picture, it wasn't the pilot's fault at all.

-- The Inmates are Running the Asylum (Alan Cooper)

Resumo IHC

- **Objetivo:** estudar formas de interação, para desenvolver canais de comunicação mais:
 - ▶ **eficientes:** rapidez de interação (maximizem a taxa de transmissão de dados entre usuários e computadores)
 - ▶ **eficazes:** otimizem os recursos utilizados
 - ▶ **satisfatórios:**
 - de fácil aprendizagem
 - flexíveis (vários modos de operação)
 - tolerantes a falha

Processamento de Imagens



Processamento de Imagens

- Análise e Síntese de imagens e vídeo
 - ▶ sensoriamento e percepção: análise
 - ▶ interação (feedback): síntese
 - ▶ transformações: ferramentas
- Representação e codificação de imagens e vídeo
 - ▶ armazenamento
 - ▶ busca
 - ▶ transmissão

Visão Computacional: histórico

- Origem
 - ▶ Inteligência Artificial: é o estudo de como fazer com que os computadores façam coisas que, até o momento, as pessoas fazem melhor [Rich & Knight - 91]
- Primeiros trabalhos:
 - ▶ Roberts
 - ▶ Block's world
 - wire frames
 - origami world
 - ▶ Marr

Visão Computacional: o problema

- Definição:
 - ▶ Dada um imagem, determinar os objetos que a compõem, bem como os atributos desses objetos
- Aplicações:
 - ▶ navegação / manipulação / automação
 - ▶ análise de imagens de satélite
 - ▶ processamento de imagens médicas
- Dificuldades
 - ▶ Sensor até o modelamento (1D, 2D, 3D, 4D,)
 - ▶ Análise (baixo e alto nível) com custo computacional elevado

Visão computacional: áreas

- Segmentação (processamento de imagens)
- Reconhecimento de objetos
- Análise de Profundidade/distância
 - ▶ stereo images
 - ▶ range images (radar, ultrasom, etc)
- Imagens médicas
- Modelos computacionais de visão biológica
- Análise de movimentos

Análise de movimentos

- Procura estimar o movimento de regiões de uma seqüência, e agrupá-las segundo algum modelo.
- Aplicações:
 - ▶ visualização
 - estabilização de imagens
 - imagens panorâmicas (mosaicos)
 - ▶ representação
 - armazenamento
 - transmissão
 - ▶ interação homem máquina
 - reconhecimento de gestos, expressões, etc

Processamento de Vídeo

- Processamento em *batch*:
 - ▶ a seqüência de imagens (frames) é processada e o resultado armazenado para ser analisado quando disponível
- Tempo real
 - ▶ disponibilidade "imediate" dos resultados
 - ▶ NTSC: 30 quadros(imagens) por segundo
 - ▶ Imagem colorida: 640 x 480 x 3 ~ 1MBytes
 - ou seja, 30 MBytes por segundo são recebidos pelo computador, provenientes câmara.

Construção de Panoramas

- Ampliar o campo visual de câmeras

Estabilização de Imagens

- Remoção do movimento global da câmara

Detecção e Rastreamento de Características

- Detecção de características: para cada frame da seqüência, o sistema seleciona algumas características
- Rastreamento: tais características são "localizadas" no frame seguinte
- O deslocamento dessas características são utilizadas para estimar o "movimento" da câmera, segundo algum modelo, em geral paramétrico.

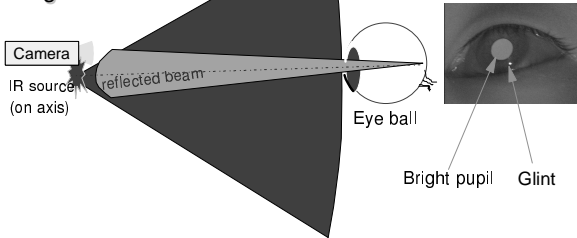


Detecção de Pupilas

- As pupilas são detectadas explorando um fenômeno similar ao que causa as "pupilas vermelhas" em fotos tiradas com um flash intenso.
- Esse fenômeno ocorre quando a fonte de luz brilhante se encontra próximo ao centro óptico da câmera.

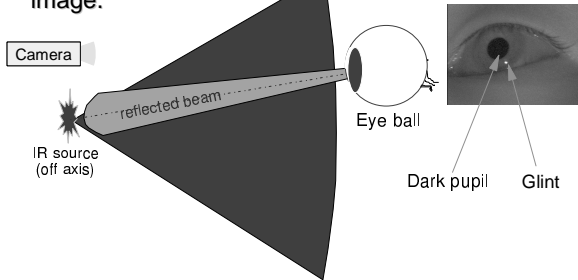
The bright pupil effect

- The on-axis IR produces a bright pupil image.



The dark pupil effect

- The off-axis IR produces a dark pupil image.



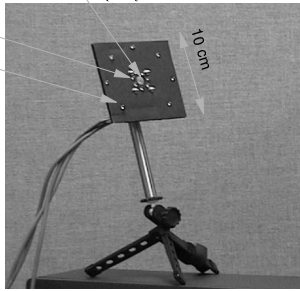
Sabemos portanto que:

- IR no eixo óptico => pupila brilhante
- IR longe do eixo óptico => pupila escura

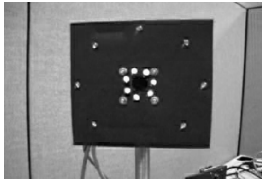
Isso é verdade independentemente da localização do olho ou da direção do olhar (desde que as pupilas sejam visíveis à câmera)

Vista do sistema

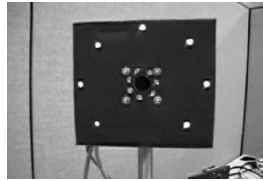
- Câmera monocromática (sensitiva ao IR).
- Duas fontes de luz infra vermelho (IR)
 - ▶ perto do eixo óptico
 - ▶ longe do eixo óptico



Fontes de luz



Sobre o eixo

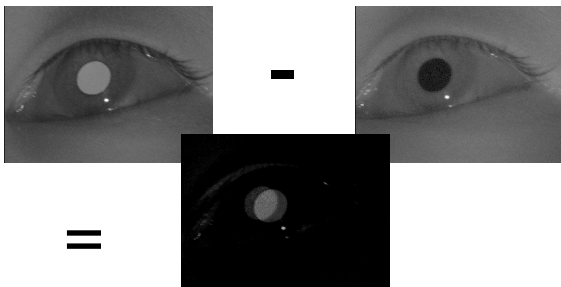


Fora do eixo

- As fontes de luz são sincronizadas com os campos par e ímpar do sinal de vídeo (vídeo entrelaçado).

Detecção de pupilas

- As pupilas são detectadas da subtração dos campos par e ímpar (pupilas escuras das pupilas brilhantes).



Detecção e rastreamento

- Componentes conexos são considerados como candidatos à pupilas
 - ▶ restrições: tamanho e forma
- As pupilas podem então serem rastreadas utilizando as imagens brilhantes ou escuras.
- Pares de pupilas podem ser agrupadas em faces, e então rastreadas em conjunto.

Implementação em tempo real

- Pentium II 400Mhz
- Digitalizador de imagens padrão PCI
- Câmera monocromática (com um filtro de luz visível)
- Dois anéis de 7 LED's IR, cada
- Um sistema bem simples para sincronização dos anéis com o sinal de vídeo, sem uso de nenhum outro hardware especial.

Desempenho

- Frame rate: 30fps, entrelaçado (60 campos por segundo)
- Imagens: 640 x 480 (8 bits).
 - ▶ câmeras baratas
- Distância de detecção: até cerca de 5m
- Bastante robusto a ruídos

Aplicações

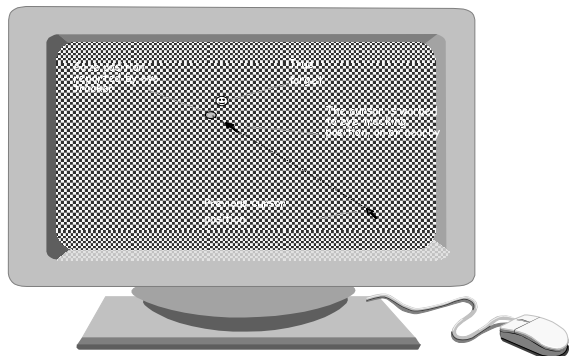
- MAGIC Pointing
- Sutor: Simple User Interest Tracker
- Contato visual

Utilizando o olhar como "mouse"

- "What you look at is what you get!" (J.L. Lavine 1981, C. Ware 1987, R.J.K Jacob 1990)
- Por que usar o olhar?
 - ▶ Mãos estarem indisponíveis
 - deficiência
 - fadiga das mãos
 - ▶ O olho é mais rápido que outros órgãos
- Limitações do método tradicional
 - ▶ O apontamento com o olhar não é preciso (0,5 grau)
 - ▶ Modelo incorreto: sobrecarregamento de funções no olho. O olho é um dispositivo de exploração e não foi desenvolvido para "manipular" objetos

MAGIC Pointing

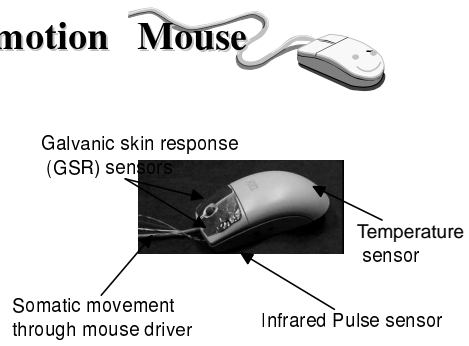
- Manual And Gaze Input Cascaded pointing
- Manual acquisition with gaze initiated cursor



Pong

Comportamento afetivo

Emotion Mouse



Emotion Mouse

- Medições: Pulso, Galvanic skin response (GSR), Temperatura.
- Entendendo o usuário
 - ▶ Obter informação fisiológica sem intervenção
 - ▶ Construção de um modelo emocional
- Objetivos
 - ▶ Identificar situações de stress
 - ▶ Melhorar condições de trabalho
 - ▶ Entender comportamentos do usuário

Trabalhos em andamento

- ✓ Detecção e rastreamento de multiplas faces
- ✓ Rastreamento do olhar (calibração)
- Detecção e rastreamento de características faciais
- Interfaces multi-modais

No animal was harmed in the making of this research.
