

# Controle Remoto Visual

## *Controle de transparências através de gestos*

Bernardo Cunha Vieira

`bcvieira@dcc.ufmg.br`

Departamento de Ciência da Computação

Instituto de Ciências Exatas

Universidade Federal de Minas Gerais



# Objetivo

O objetivo deste trabalho é implementar uma Interface Humano Computador que controle um programa de apresentação de transparências utilizando uma WebCam.



# Referências

- G. Rigoll, A. Kosmala, and S. Eickeler. High performance real-time gesture recognition using hidden markov models. In Proceedings of the International Gesture Workshop on Gesture and Sign Language in Human-Computer Interaction, pages 69-80, London, UK, 1998. Springer-Verlag.
- R. M. Resende. Desenvolvimento de uma interface Humano-Robô utilizando visão computacional e Sistemas a Eventos Discretos. Master's thesis, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Agosto 2005.



# Metodologia

- Utilizar segmentação por cores, no padrão HSV.
- Identificar o centróide do objeto de interesse.
- Identificar a direção do movimento entre frames espaçados por um  $\Delta t = 30 \text{ frames}$ .
- Discretizar o movimento.
  - Treinar a HMM; ou
  - Identificar o movimento a partir de uma HMM treinada.



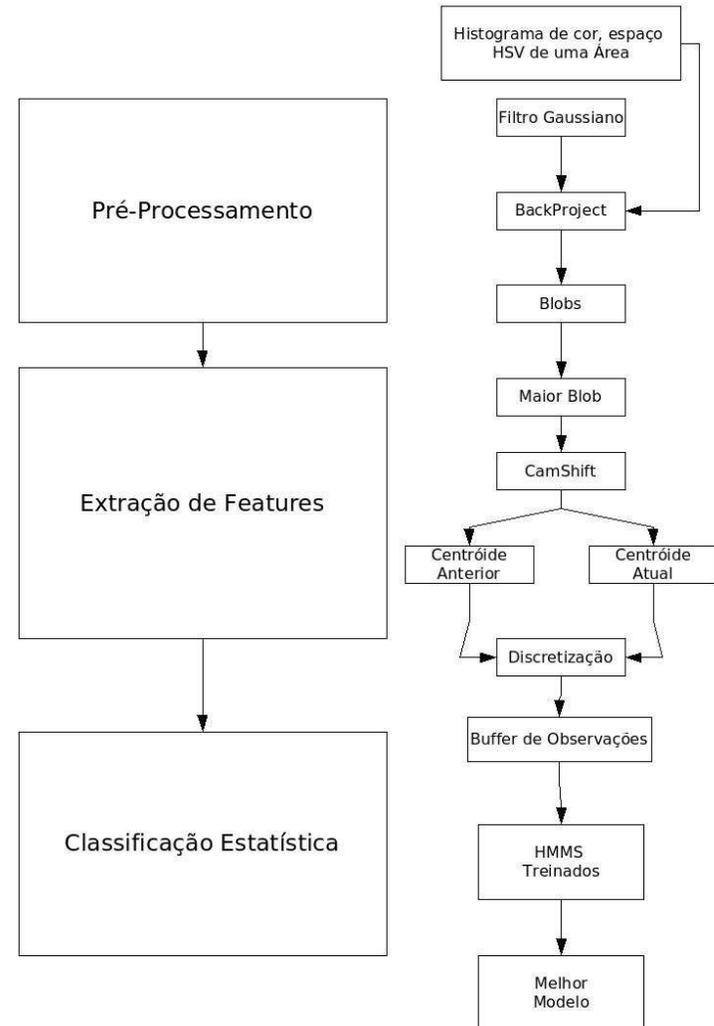
# Aparato

- Webcam  $320 \times 240 \text{ px}$ ,  $30 \frac{\text{frames}}{\text{s}}$ .
- 10 Vídeos de cada movimento.
- O usuário deve utilizar uma luva amarela para facilitar a segmentação.



# Divisões do Trabalho

- Pré-processamento
- Extração de Features
- Classificação Estatística



# Pré-processamento

Para limpar os ruídos, a primeira fase do pré processamento utiliza um filtro gaussiano de kernel tamanho 5.



# Pré-processamento

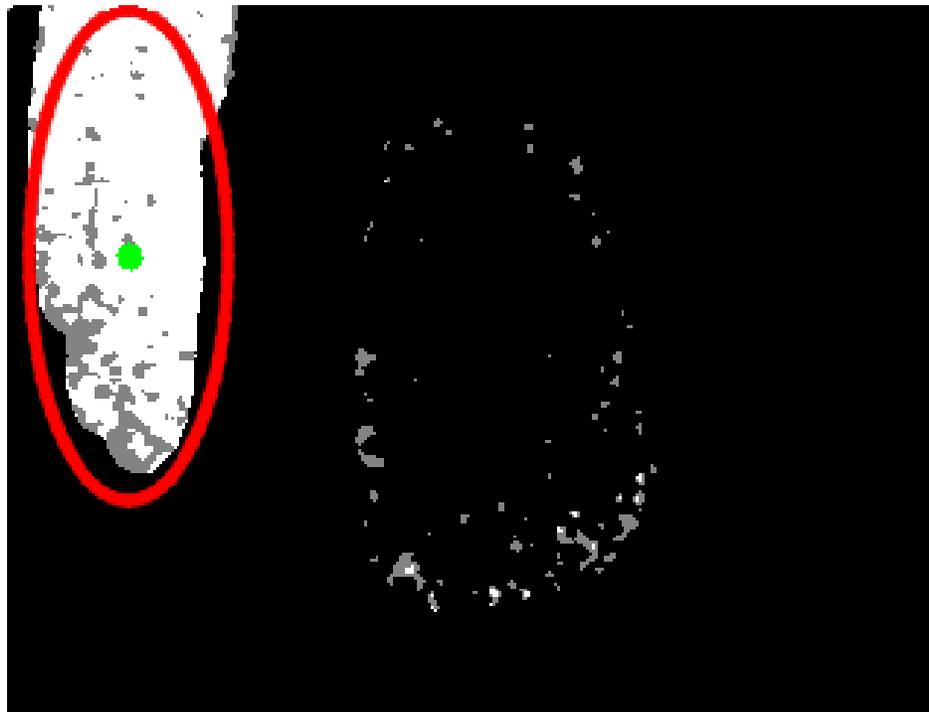
Para segmentar a imagem utiliza-se um histograma da parte H do espaço de cores HSV. Este histograma é calculado utilizando uma área demarcada no início da execução do programa, com 16 bins.

O histograma é então submetido ao algoritmo do OpenCV Backproject, que segmenta a imagem probabilisticamente. A imagem final sofre um threshold (valor 200).



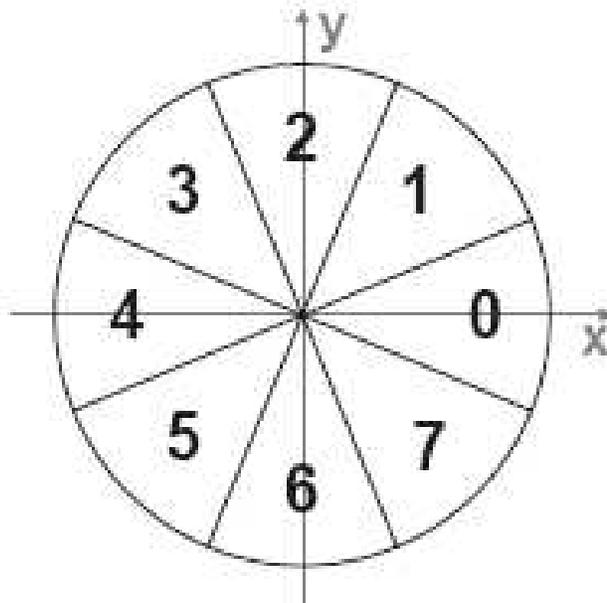
# Extração de Features

Para identificação do objeto de interesse utilizou-se a identificação de blobs por vizinhos 8-conectados das regiões brancas da imagem binarizada no passo anterior. A maior área representa o objeto de interesse. Para identificar o centróide e a orientação da mão utilizou-se o Camshift do OpenCV.



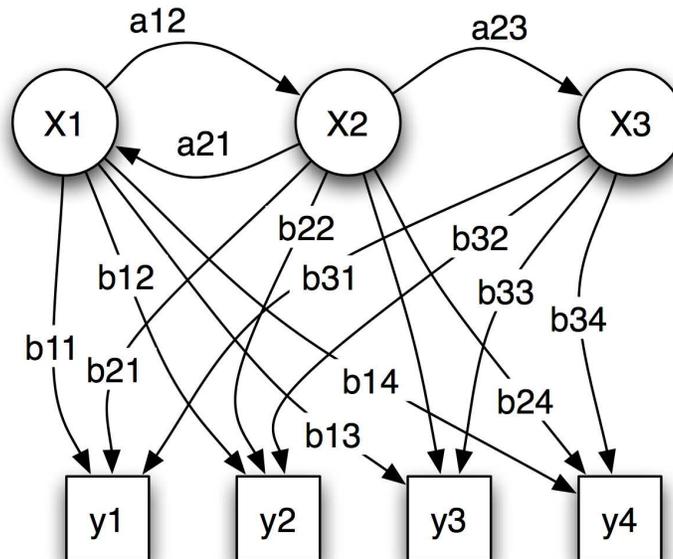
# Extração de Features

Após a identificação do centróide, no passo anterior, pegamos a diferença entre os centróides de dois frames consecutivos, separados de um  $\Delta t$  (30 frames). A diferença é utilizada para cálculo do ângulo, que é então discretizado através do círculo abaixo.



# Classificação Estatística

- Cada gesto é um modelo escondido de markov, treinado com 10 vídeos, com probabilidade de emissão e transição uniformemente distribuídas.
- Basta verificar qual modelo melhor se encaixa nas observações obtidas da discretização.



# Trabalhos Futuros

- Verificar a aplicação e desempenho do algoritmo Condensation para tracking e classificação temporal dos movimentos.
- Aumentar a funcionalidade identificando poses das mãos.

