

Uma abordagem para reconhecimento de gestos humanos Contando dedos

Iuri Bueno Drumond de Andrade
Guilherme Oliveira França
Samuel Lino de Macedo

Departamento de Ciência da Computação
Universidade Federal de Minas Gerais (UFMG)
Av. Antonio Carlos, 6627, 31270-010, Belo Horizonte, MG, Brazil
Telefone / fax: 31 3499-5589 / 31 3499-5858
{bueno, guidcc, slmacedo}@dcc.ufmg.br

Resumo

Este trabalho apresenta uma forma de detecção e subsequente contagem de dedos de uma imagem de uma mão humana capturada de uma simples webcam. Isso é feito através de segmentação de cores por tonalidade de pele no espaço YCBCR e transformação do espaço de pixels para o espaço polar. Os resultados encontrados mostram que essa abordagem é bastante eficiente, sem muita interferência do background em que encontra-se a mão.

1. Introdução

O reconhecimento de gestos humanos possui grandes aplicações no mundo digital atualmente. Aplicações como interatividade entre homem-máquina usando webcam simples vem se tornando cada vez mais populares.

O trabalho será desenvolvido abordando-se uma parte do problema de reconhecimento de gestos humanos. Mais especificamente, com a detecção da mão humana e o reconhecimento da quantidade de dedos que essa apresenta num determinado tempo usando-se uma webcam para isso.

2. Trabalhos relacionados

Em [5] o autor propõe uma maneira de reconhecer gestos como apontar para alguma coisa e saber para onde está apontando, "cliqueando" em alguma coisa ou mesmo reconhecendo a quantidade de dedos que a pessoa está mostrando para a câmera. Isso é feito sem nenhum tipo de mecanismo complicado, tais como infra-vermelho ou marcações na cena. Ele utiliza segmentação da mão através de um plano normalizado de cores e o agrupamento de pixels usando

uma gaussiana 2D. Os resultados obtidos se mostraram bastante relevantes, com grande redução de processamento, dadas as restrições descritas no artigo.

Em [3] é apresentado uma técnica para realizar o desenvolvimento de um sistema de visão computacional que reconheça gestos humanos. Sua técnica possui três partes: detecção, rastreamento e reconhecimento dos gestos. Isso é feito através do uso de algoritmos de segmentação da mão, rastreamento do centro de massa daquela superfície de pele, usando o algoritmo de Mean-Shift e um sistema de interpretação de gestos através de uma sequência de coordenadas mostrando a posição do objeto de interesse no tempo.

Em [6] é realizado uma pesquisa com os diversos métodos de cores possíveis para a segmentação eficiente da pele humana. Os autores apresentam várias formas de modelar a pele e compara tais formas, mostrando a eficiência de cada uma.

3. Metodologia

O processo adotado para contagem de dedos a partir da imagem de uma mão humana requer três passos: segmentação da mão, transformação dos pixels para a coordenada polar e contagem dos dedos.

3.1. Segmentação

Esta etapa do processo parte de uma imagem obtida através de uma câmera em ambientes não controlados com background variado. Uma das técnicas mais utilizadas é a segmentação da mão pela cor de pele, que funciona escolhendo-se um espaço de cor e um intervalo de valores que melhor representa os tons da pele naquele espaço. Nos

experimentos realizados, percebeu-se que o espaço de cor que apresentou os melhores resultados foi o YCbCr. Esse espaço de cor é muito usado em sistemas de vídeo e fotografia digitais. A componente Y corresponde à luminosidade da imagem e as componentes Cb e Cr correspondem à cromaticidade do azul e do vermelho respectivamente. O problema da segmentação pelo espaço de cor é que a luminosidade da imagem pode alterar os limiares dos pixels correspondentes aos tons da pele.

Para realizar a segmentação em diferentes luminosidades e tonalidades de pele, se utilizou um algoritmo adaptativo que inicialmente captura o tom da pele do usuário e, partir dele, define um limiar naquele espaço de cores.

Uma faixa de valores descrito em [4] e [1] apresentou resultados satisfatórios:

$$85 \leq Cb \leq 135 \quad 135 \leq Cr \leq 180 \quad Y \geq 80$$



Figura 1. Figura original

Ao aplicar o algoritmo de segmentação utilizando-se os limiares definidos anteriormente, foi obtido como resultado uma nova imagem segmentada. Essa imagem, além do objeto de foco que era a mão do usuário, continuou a mostrar objetos indesejados, como demonstrado na figura a seguir:

Para remover os objetos indesejados da imagem, foram utilizados alguns operadores morfológicos. Esses operadores revemem da imagem as partes que não correspondem ao objeto principal, como demonstrado na figura a seguir:

Com a imagem da mão totalmente segmentada, o próximo passo desta etapa é a detecção de bordas. Para se realizar a detecção utilizou-se o operador de Sobel. A imagem resultante após a aplicação do operador foi:

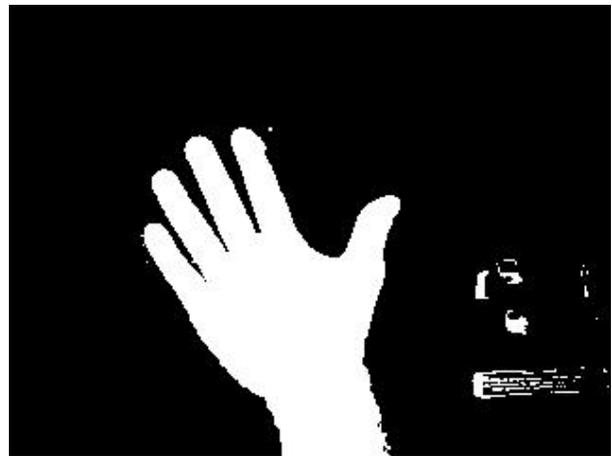


Figura 2. Resultado da aplicação do limiar entre os planos

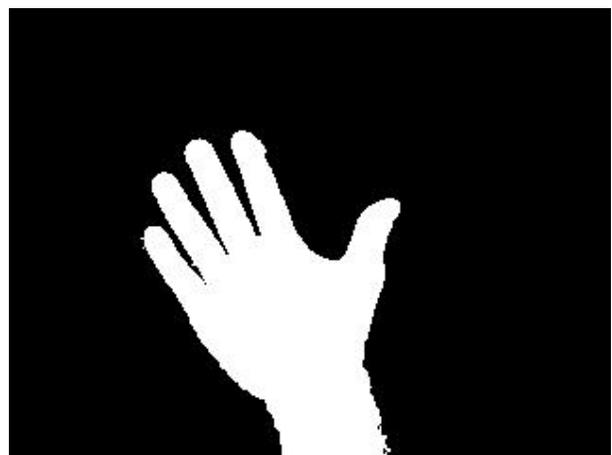


Figura 3. Resultado da aplicação dos operadores morfológicos

3.2. Transformação para coordenadas polares

A partir da imagem segmentada contendo apenas as bordas da mão, neste passo transforma-se as coordenadas dos pixels para o espaço polar. Primeiramente dado a imagem E binarizada com as bordas da mão, e sendo A o número total de pixels dessa borda, se define o centroide

$$(x, y) = \left(\frac{1}{A} \sum_{i,j} jE(i, j), \frac{1}{A} \sum_{i,j} iE(i, j) \right)$$

que corresponde ao centro de massa do contorno. De acordo com [2], esse centroide é ajustado para $(x, \zeta y)$, escolhido experimentalmente para mover o centroide para próximo ao

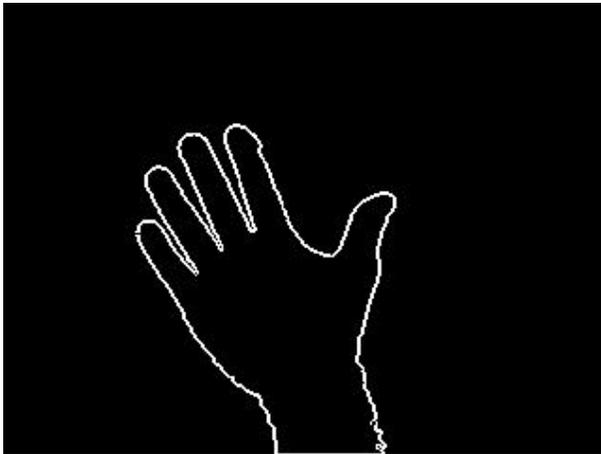


Figura 4. Resultado da detecção de bordas

centro da palma da mão, se tornando o centro de coordenada. Então se define uma região de interesse(ROI) contendo somente os pixels da borda que estão acima desse centro de coordenada. A transformação para coordenadas polares

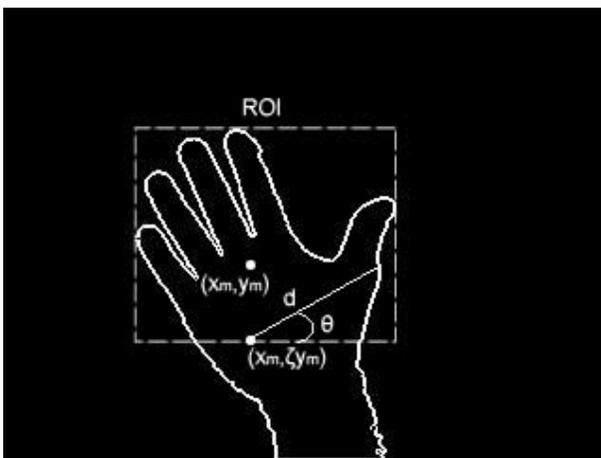


Figura 5. Transformação para coordenadas polares

res acontece através da definição do centroide como a origem do espaço de coordenadas. E assim, para cada ponto da borda da mão na imagem, calcula-se um ângulo α e uma distância d correspondentes à origem.

3.3. Contagem de dedos

Para a contagem dos dedos é definido inicialmente um limiar τ como sendo uma fração da maior distância que um pixel atingiu na região de interesse determinada: $\tau =$

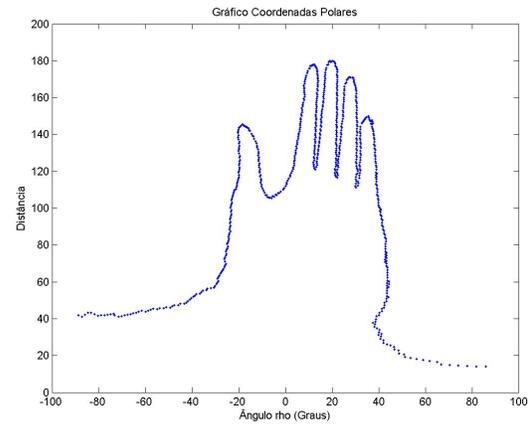


Figura 6. Gráfico resultante da transformação das coordenadas

$c \max_{\alpha} r(\alpha)$. Então percorre-se todos os pontos resultantes da transformação e compara-se suas respectivas distâncias com o limiar, realizando a contagem dos dedos. Mais especificamente, sempre que há um valor que supere o limiar e outro que fique abaixo desse, temos um possível dedo.

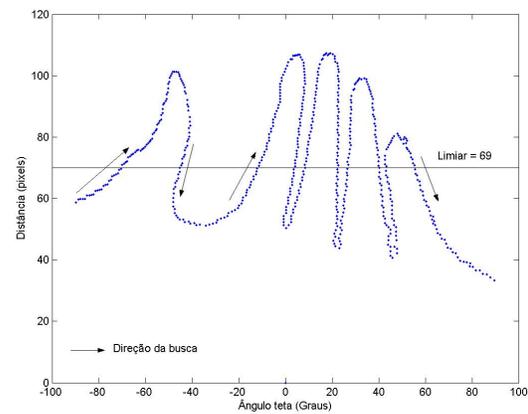


Figura 7. Limiar para contagem dos dedos

Porém, podem acontecer casos em que fatores externos como ruído e/ou iluminação atrapalhem a segmentação e, conseqüentemente dificultem a contagem de dedos. Isso faz com que regiões da imagem possam ser erroneamente detectadas como dedos. Portanto, ao invés de se comparar o limiar com apenas a distância de um ponto, no intuito de detectar se essa está acima ou abaixo do limiar, comparou-se com um conjunto de tamanho M de distâncias. Dessa forma, toda vez que um conjunto de distâncias superar o li-

miar e o próximo ficar abaixo desse, tem-se um dedo.

4. Resultados

Foram realizados vários testes para a comprovação do funcionamento do sistema de contagem de dedos. Na primeira instância testes utilizou-se um conjunto de 20 diferentes imagens, em que se variou vários parâmetros como: iluminação, fundo da imagem, tamanho da mão e forma de se mostrar os dedos. Na segunda, utilizou-se um conjunto de 7 imagens tentando-se minimizar as interferências externas ao se criar um ambiente mais controlado, variando-se apenas a quantidade de dedos mostrados. Para todos os testes definiu-se $\zeta = 1.2$, $c = 0.6$ e $M = 20$.

Teste Instâncias	Taxa	
	Acerto	Erro
Variável	60%	40%
Controlado	100%	0%

Tabela 1. Tabela de taxa de acertos e erros

Ao analisar a taxa de acerto da primeira instância de testes vemos que era de se esperar que essa não apresentasse bons resultados, uma vez que houve grande variação na captura das imagens. Os resultados mostraram que o algoritmo implementado apresentou problemas durante a segmentação da mão ao se variar a iluminação e o fundo da imagem. Isso se deve ao fato dos parâmetros para a segmentação utilizando cor não estarem sendo calibrados automaticamente e sim utilizando valores estáticos. Com isso, em diferentes ambientes os resultados encontrados não são os mesmos.

Já ao se minimizar as variações externas, houve um ganho significativo na detecção, como se pode ver na tabela. Uma vez que a contagem de dedos depende diretamente da segmentação, essa apresentando bons resultados fará com que a taxa de erro diminua.

Outro detalhe dos testes realizados foi que o algoritmo implementado apresentou bons resultados na parte de contagem de dedos, mesmo com variações das maneiras como esses dedos foram mostrados. Assim se comprovou que a técnica de transformação para coordenadas polares é viável para a solução desse tipo de problema.

4.1. Conclusão e Trabalhos futuros

O trabalho realizado atingiu os objetivos propostos, sendo que ao final se conseguiu uma boa técnica para a contagem de dedos. Também pode-se concluir que para implementar um sistema desse tipo, é necessário um bom algoritmo para segmentar a imagem, sendo que ele deve ser adaptativo para diferentes tipos de ambientes e iluminação.

O algoritmo implementado para a contagem dos dedos a partir da imagem já segmentada, se mostrou eficiente e viável para essa situação.

Uma próxima etapa desse trabalho seria a implementação de um algoritmo adaptativo para realizar a segmentação da mão. Também seria interessante a implementação de um sistema que capturasse várias imagens a partir de um vídeo, não somente uma única figuras, e realizasse algum tipo de interação com o usuário.

Referências

- [1] U. Ahlvers, U. Zolzer, and R. Rajagopalan. Model-free face detection and head tracking with morphological hole mapping. *EUSIPCO*, 2005.
- [2] S. C. Crampton and M. Betke. Counting fingers in real time: A webcam-based human-computer interface with game applications. *Proceedings of the Conference on Universal Access in Human-Computer Interaction*, pages 1357–1361, 2003.
- [3] F. Dadgostar and A. Sarrafzadeh. A component-based architecture for vision-based gesture recognition. *Image and Vision Computing New Zealand*, III, 2005.
- [4] V. NABIYEV and A. Gunay. Towards a biometric purpose image filter according to skin detection. *PCI*, 2008.
- [5] M. Storing, T. B. Moeslund, Y. Liu, and E. Granum. Computer vision-based gesture recognition for an augmented reality interface. *4th IASTED International Conference on VISUALIZATION, IMAGING, AND IMAGE PROCESSING*, III:766–771, 2004.
- [6] V. Vezhnevets, V. Sazonov, and A. Andreeva. A survey on pixel-based skin color detection techniques. *Proc. Graphicon*, pages 85–92, 2003.