

Uma implementação do jogo Pedra, Papel e Tesoura utilizando Visão Computacional

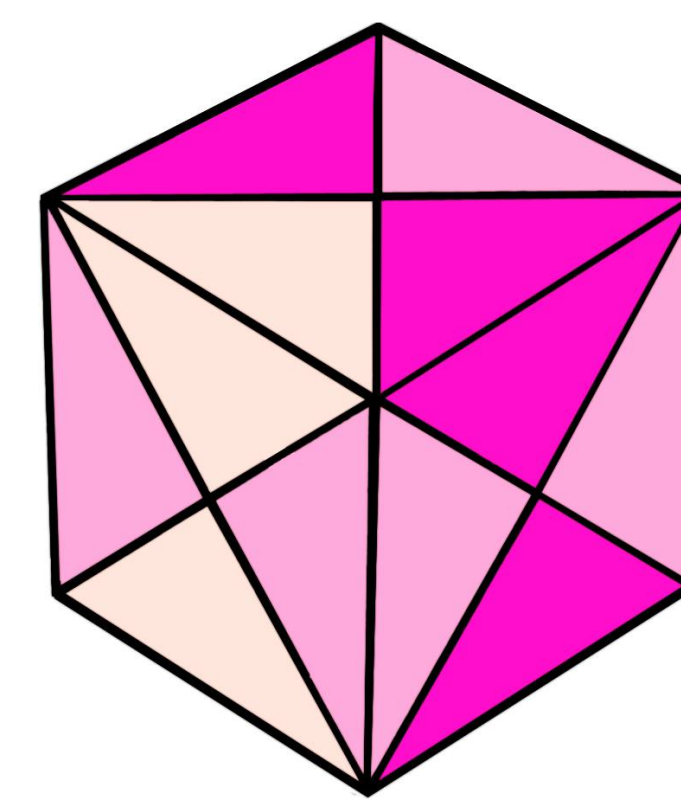
Ezequiel França dos Santos, Gabriel Fontenelle Senno Silva

Bacharelado em Ciência da Computação

Centro Universitário SENAC - Campus Santo Amaro (SENAC-SP)

Av. Engenheiro Eusébio Stevaux, 823 – Santo Amaro, São Paulo – CEP 04696-000 – SP – Brasil

ezefranca.br,coleccionador.gabriel, (@gmail.com)



Resumo

Este trabalho apresenta um jogo, controlado através de visão computacional, na identificação de gestos da mão (hand-tracking). O trabalho proposto baseia-se na segmentação da imagem e construção de um fecho convexo com algoritmo de Jarvis e determinação do padrão com base na extração de características de sua área.

1. Introdução

A busca por meios que tornam os jogos mais interativos tem sido muito explorada. Muitos destes novos meios envolvem a área de visão computacional. Este trabalho apresenta um estudo sobre a viabilidade de utilização de uma webcam como dispositivo de interação baseado em gestos da mão, especificamente para o jogo Pedra, Papel e Tesoura.

2. Objetivos

Este trabalho apresenta um estudo sobre a viabilidade de utilização de uma webcam como dispositivo de interação baseado em gestos da mão, especificamente para o jogo Pedra, Papel e Tesoura.

Neste trabalho, para o reconhecimento de gestos da mão utilizamos a combinação de algumas técnicas para um reconhecimento satisfatório, inicialmente fizemos um pré-processamento da imagem, passando a mesma para escala de cinza e em seguida binarizando, e por fim aplicamos um filtro para detecção de bordas. Em posse da imagem esquelizada pelas bordas da mão, trabalhamos o fecho convexo do conjunto de pontos e extraímos algumas características de sua área.

3. Metodologia

Aquisição da imagem

A biblioteca utilizada neste trabalho[7] permite o acesso a imagens de câmeras através da plataforma OpenCV. Ela permite trabalhar com a imagem como se a mesma fosse uma matriz tridimensional. A primeira dimensão da matriz representa a altura, a segunda dimensão representa a largura e a terceira representa os canais vermelho, verde, azul (RGB) e alfa (α) de cada pixel. O fator alfa de cada pixel é utilizado para determinar como as cores serão unidas quando imagens de cores diferentes estiverem sobrepostas.

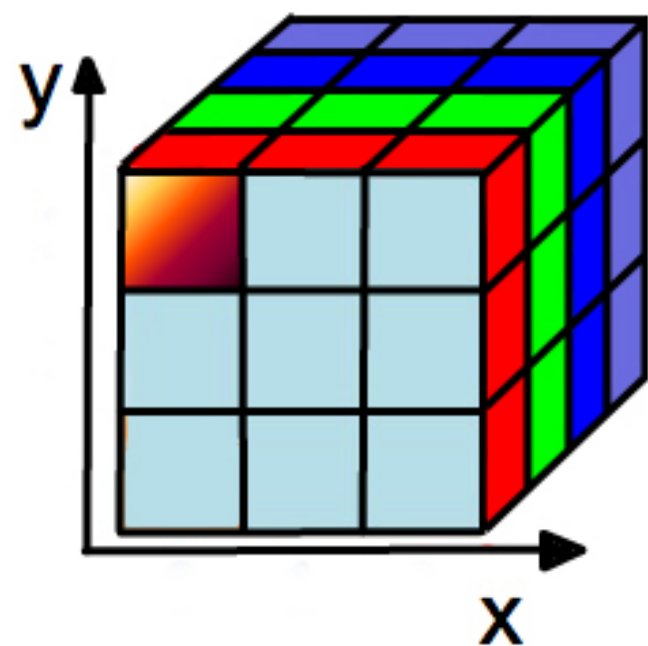


Figura 1: Representação da matriz tridimensional da imagem de XY

Normalização para escala de cinza

A primeira fase de pré-processamento trata-se da normalização da imagem colorida para tons de cinza. A normalização foi feita com base no valor médio dos canais de cores da imagem (RGB).

A Figura 2 apresenta o resultado deste processo.

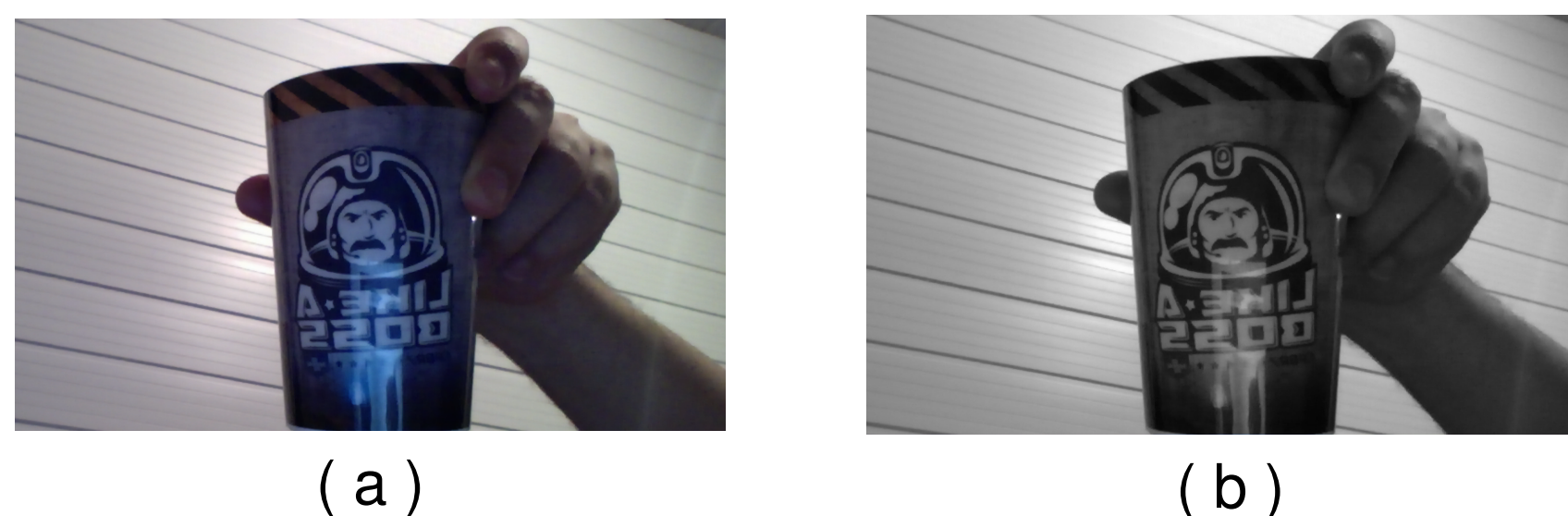


Figura 2: Imagem normal (a) e imagem normalizada em cinza (b).

Binarização da imagem

Existem diversos algoritmos para binarização de imagens, dentre a lista de soluções o adotado foi o algoritmo de Otsu, por ser de fácil implementação e apresentar resultados satisfatórios nos experimentos realizados. O método de Otsu é um método de *thresholding* global, isto é, o valor obtido é uma constante, para escolha do melhor limiar. O resultado da binarização com limiar ajustado segundo o método de Otsu pode ser observado na Figura 3.

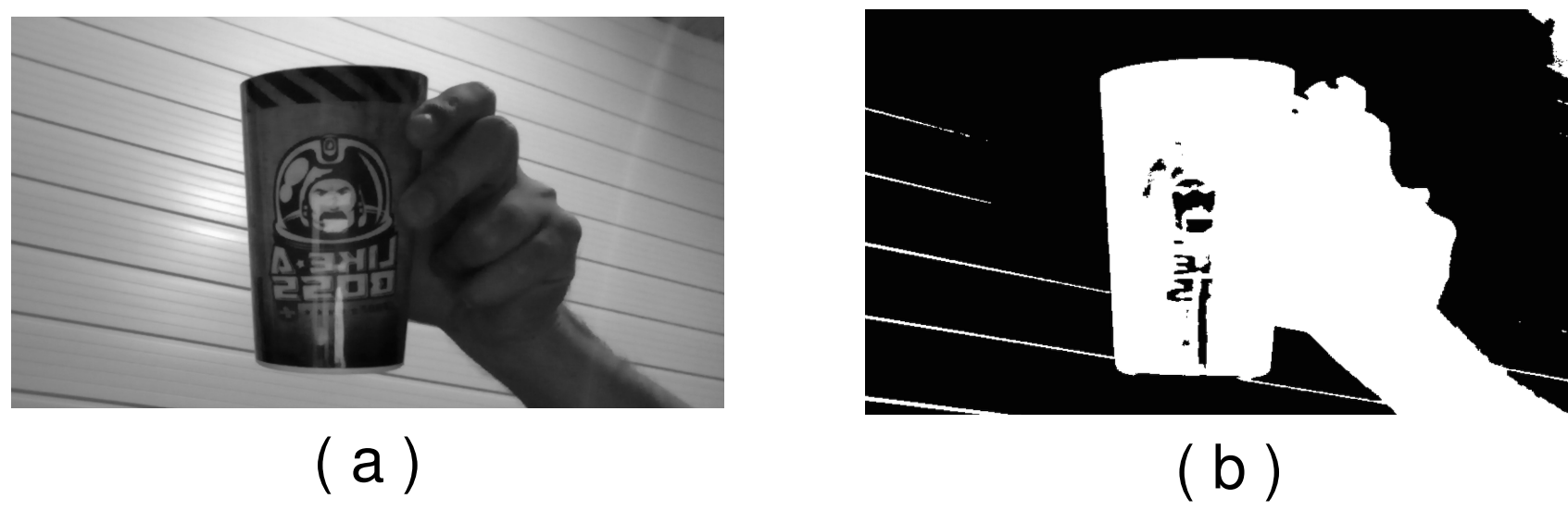


Figura 3: Imagem binarizada (b) com limiar definido pelo método de Otsu.

Detecção de bordas com filtro Sobel

O filtro Sobel calcula o gradiente da intensidade da imagem em cada ponto, dando a direção da maior variação de claro para escuro e a quantidade de variação nessa direção, através de duas matrizes 3x3, que são convoluídas com a imagem original para calcular aproximações das derivadas - uma para as variações horizontais G_x e uma para as verticais G_y . O resultado da utilização do filtro Sobel pode ser observado na Figura 4.



Figura 4: Imagem binarizada (a) e imagem com borda detectada com o filtro Sobel (b).

Reconhecimento dos padrões

Determinação do fecho convexo

Um conjunto $S \subset \mathcal{R}^d$ é *convexo* se $\lambda z_1 + (1 - \lambda)z_2 \in S$ sempre $z_1, z_2 \in S$, e $0 \leq \lambda \leq 1$. Resumidamente, S contém todos os segmentos de linha que conectam pares de pontos em S .

O algoritmo da *marcha de Jarvis*, popularmente conhecido como *gift wrapping algorithm* / *algoritmo do embrulho de presente*, visita os pontos do fecho convexo de maneira ordenada.

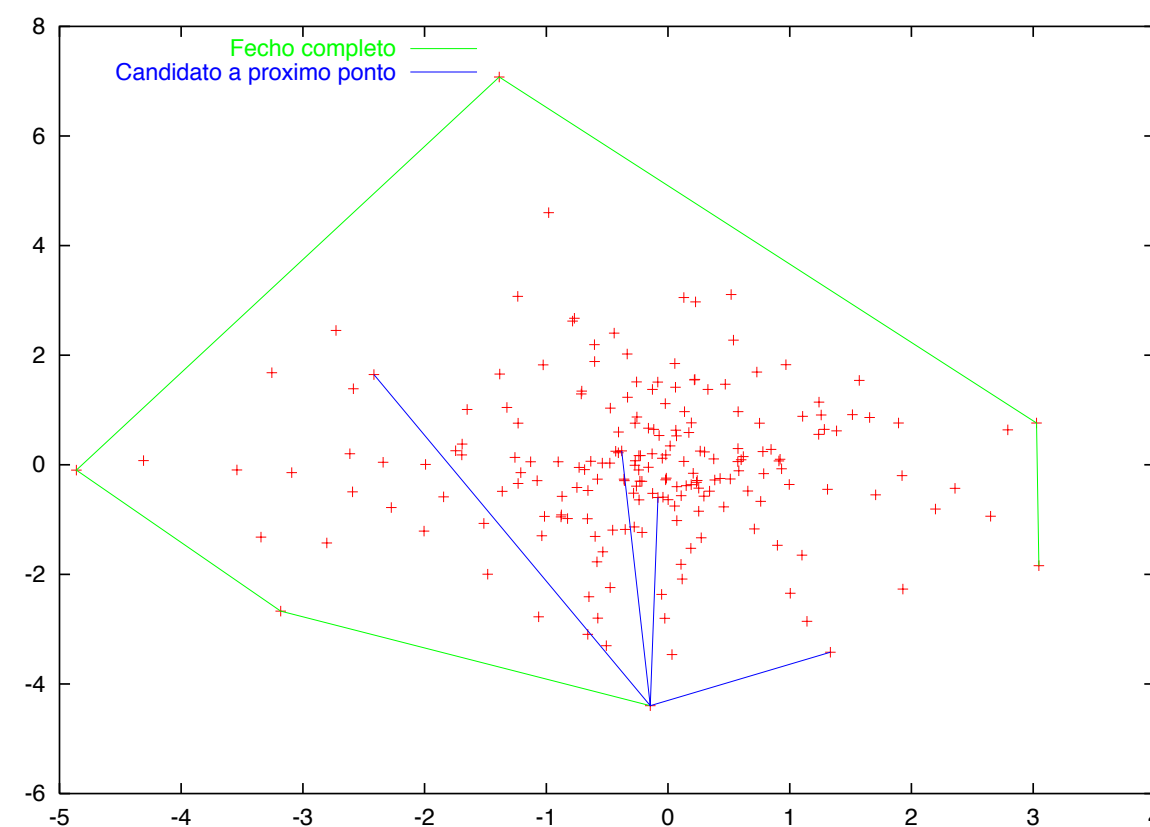


Figura 5: Representação do fecho convexo.

O algoritmo de marcha Jarvis tem no pior caso complexidade $O(n^2)$, o que ocorre se todos os pontos estão no fecho. Em geral, se h pontos estão no fecho, a complexidade é $O(nh)$.

Cálculo da área total do fecho

Para o cálculo da área do fecho utilizamos o *shoelace*, ou algoritmo cadarço de sapato. Este algoritmo matemático é utilizado para determinar a área de um polígono simples cujos vértices são descritos por pares ordenados no plano. Neste algoritmo, multiplica-se de forma cruzada as coordenadas correspondentes para encontrar a área que abrange o polígono e subtrai-se do polígono circundante para encontrar a área do polígono interno.

Cálculo da área "branca"

Para o cálculo da área "branca" foi utilizado uma cópia da matriz binarizada, sem aplicação do filtro de Sobel, e cada pixel dos vértices do fecho é gravado nessa cópia, e em seguida são verificados se os pixels dentro do fecho são "pretos" ou "brancos", sendo os pixels "brancos" contabilizados. A Figura 6 demonstra as áreas em questão.

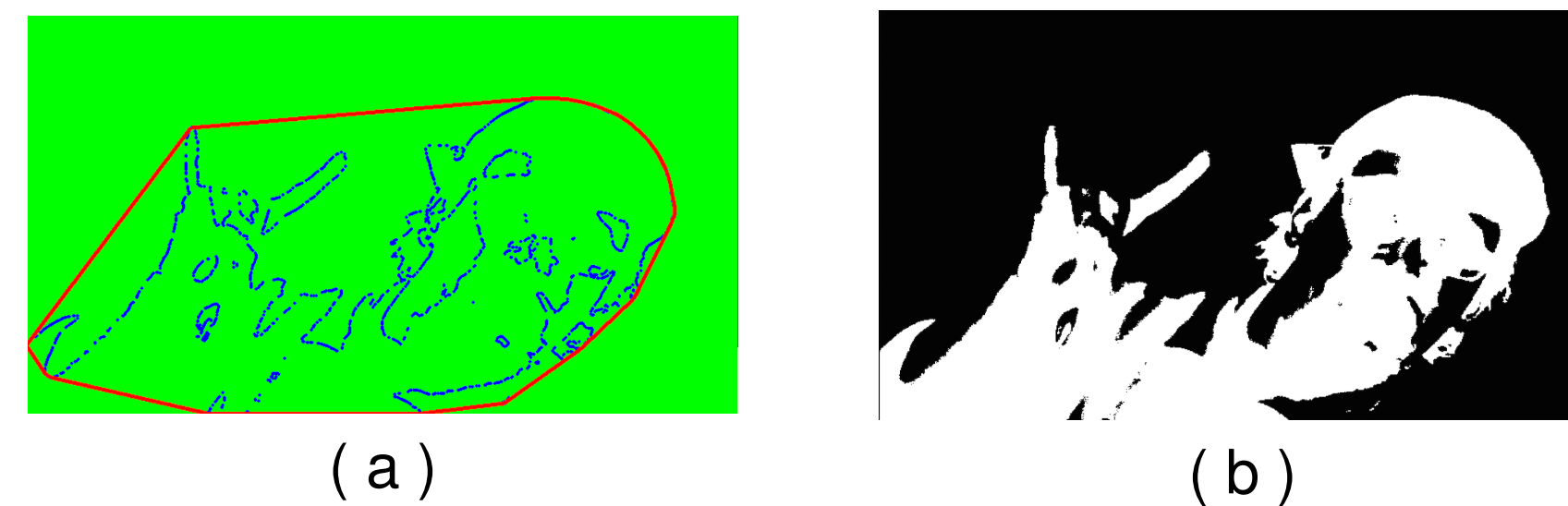


Figura 6: Fecho convexo (a) e imagem binarizada (b) utilizada no cálculo.

Determinação do padrão

A determinação do padrão foi feita com base na proporção entre área total x pixels "brancos".

4. Resultados e Discussão

Os resultados obtidos possibilitaram o desenvolvimento de um algoritmo para detecção de padrões pré-estabelecidos de gestos, através da simples análise do comportamento dos pontos da imagem segmentada. Entretanto, este tipo de análise, por mais rudimentar que seja apresentou-se eficaz para este problema.

Ao longo desse trabalho, foi possível obtermos uma visão de alguns dos problemas no campo da visão computacional e da geometria computacional, casos em que demandavam uma maior atenção no desenvolvimento de seus algoritmos, além de cuidado na performance em que algoritmos destes ramos demandam.

5. Conclusão

O trabalho apresenta como principal contribuição, uma possibilidade de reconhecimento de padrões pré-determinados para controle de jogos utilizando métodos simples, porém com resultados, dentro de seus limites, precisos.

Os algoritmos propostos são de fácil implementação e não requerem uma abordagem matemática profunda para sua compreensão e aplicação. O trabalho mostra ainda, que estes métodos, com pouca modificação poderiam ser utilizados em qualquer outro tipo de interface por visão computacional, uma vez que seus algoritmos possuem complexidades relativamente médias.

Referências

- [1] MARQUES FILHO, Ogê; VIEIRA NETO, Hugo. Processamento Digital de Imagens, Rio de Janeiro, Brasport, 1999.
- [2] MARENGONI, Maurício; STRINGHINI, Denise. Tutorial: Introdução à Visão. Computacional usando OpenCV, Revista Rita, Porto Alegre, v. XVI, n 1, 2009
- [3] BRAZIL, Emilio Ashton Vital. Algoritmos de Fecho Convexo, Geometria Computacional, IMPA - Instituto de Matemática Pura e Aplicada, 02 de Maio de 2005.
- [4] MATTEUCCI, Matteo. Lecture 4 (2000), disponível em: http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MORSE/threshold.pdf.
- [5] HAMBERG, Charles L; VAVRINEK Ronald. Shoelace Algorithm.. *Illinois Mathematics and Science Academy disponível em:* <http://choosgs3math.wiki.hci.edu.sg/file/view/Shoelace.pdf>.
- [6] SOBEL, Irwin A 3x3 isotropic gradient operator for image processing. Never published but presented at a talk at the Stanford Artificial Project, 1968.
- [7] HASHIMOTO, Marcelo - Biblioteca fornecida para este projeto. disponível em <https://github.com/senacbcc/Hashimoto-Camera-lib>.