

# Projeto Final Processamento de Imagens

Ana Beatriz Santos Pontes  
Departamento de Ciência da Computação  
Universidade de Brasília  
Brasília, Brasil  
180012428@aluno.unb.br

Hyago Gabriel Oliveira Figueiredo  
Departamento de Ciência da Computação  
Universidade de Brasília  
Brasília, Brasil  
170105067@aluno.unb.br

**Abstract**—O reconhecimento de gestos manuais tem se mostrado uma área de grande interesse em várias aplicações, como interações homem-máquina, realidade virtual, jogos, reabilitação e assistência a pessoas com deficiências motoras, linguagem de sinais e comunicação não verbal como um todo. Com o avanço da tecnologia de sensores e técnicas de processamento de imagem, tornou-se possível desenvolver sistemas capazes de interpretar e reconhecer gestos realizados pelo ser humano.

Este trabalho tem como objetivo replicar o que foi abordado no artigo “Hand Gesture Recognition based on Shape Parameters” de 2011, adequando e melhorando a execução proposta, explorando o reconhecimento de gestos manuais com base na forma de parâmetros. A forma de parâmetros é uma abordagem que descreve o movimento de um gesto através de características geométricas, como ângulos, distâncias e proporções. Ao analisar essas características, é possível extrair informações relevantes e identificar padrões nos gestos realizados.

**Index Terms**—processamento de imagem, reconhecimento de gestos, distância, classificação, segmentação e binarização.

## I. INTRODUÇÃO

Neste trabalho, propomos um sistema de reconhecimento de gestos manuais baseado na forma de parâmetros. Inicialmente, capturamos os gestos através de uma imagem. Em seguida, utilizamos técnicas de pré-processamento para melhor adequar a imagem a ser trabalhada. Utilizamos técnicas de pré-processamento para binarizar as imagens e preencher eventuais buracos, garantindo uma representação mais clara dos gestos. Após isso, extraímos características da forma de parâmetros dos gestos para poder classificar a imagem de acordo com picos (dedos levantados ou abaixados).

Além disso, é necessário ajustar a orientação das imagens para que todas estejam na posição vertical, tornando o processo de classificação mais consistente. A identificação das bordas dos gestos é fundamental para segmentar a imagem em seções e obter informações relevantes. O dedão desempenha um papel crucial nesse processo, sendo localizado e utilizado como referência para a classificação.

Ao encontrar o centro da imagem, é possível determinar o número de dedos levantados e abaixados, permitindo a classificação final da imagem por meio de bits. Essa abordagem oferece uma maneira eficiente e precisa de reconhecer

gestos manuais com base em sua forma e parâmetros específicos.

## II. METODOLOGIA

### A. Artigo original

No artigo usado como base o autor segue o seguinte fluxo de implementação e metodologia:

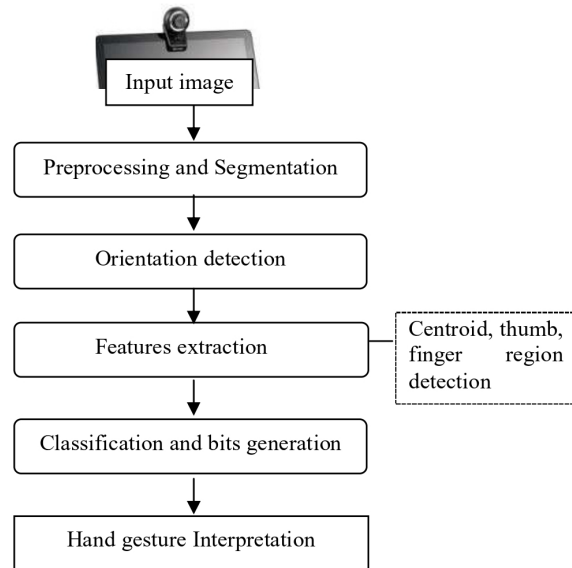


Fig 1. Diagrama do artigo original

A imagem de entrada passa por um pré processamento, sendo convertida de RGB para YCrCb. Após isso, uma segmentação K-means de dois clusters é feita na imagem pré processada para que a parte do gesto seja totalmente separada da parte do fundo. O algoritmo K-means também calcula o centróide de cada cluster.

Após isso a imagem é binarizada e é passado um preenchimento de buracos para que esteja pronta para a classificação. Para que a classificação seja feita de maneira mais fácil as “bordas” do gesto são identificadas, usando um scanearamento da esquerda da direita do topo ao fim para encontrar a “borda”

esquerda, da direita para a esquerda do topo ao fim para encontrar a “borda” direita e do topo ao fim para encontrar a “borda” do topo. A “borda” fim não precisa ser encontrada pois já é delimitada pelo final da imagem. Depois a orientação da imagem é detectada e todas as imagens são colocadas na vertical(as que já não estão). Isso é necessário para o passo de extração de características e classificação do gesto.

Na fase de extração de características a imagem é avaliada em vários aspectos

- Possui ou não dedão levantado - isso é feito comparando a quantidade de pixels brancos existente em áreas da imagem com 7% da quantidade total de pixels brancos da imagem
- Área onde o dedão se encontra - a imagem é dividida em duas áreas, a área esquerda e a área direita, cada uma com 30 pixels de largura e usando o cálculo da quantidade de pixels brancos contidos na área é possível identificar se o dedão está nessa área ou não
- Parte onde os dedos estão levantados e parte onde os dedos estão abaixados - nessa extração de características é usando a distância entre os picos da imagem e o centróide calculado no K-means. Caso a distância seja maior que 75% da distância entre o centróide e o ponto mais elevado(borda topo) esse ponto é declarado como pico da imagem e nele possui um dedo levantado, caso não o dedo está abaixado.
- Classificação e geração de bits - a partir de todas as características retiradas a classificação é feita da seguinte forma: começando a partir do dedão caso o dedo esteja levantado ele recebe o bit 1(e é declarado como pico) e caso esteja abaixado ele recebe o bit 0(e é declarado como não pico).
- Na fase da interpretação do gesto os bits são comparados aos das imagens que o programa recebeu como treino e a classificação desse gesto sai dessa comparação de bits.

### B. Replicação e adaptação do que foi abordado no artigo original

Com base nas mudanças e adaptações o seguinte fluxo de implementação e metodologia foi seguido:

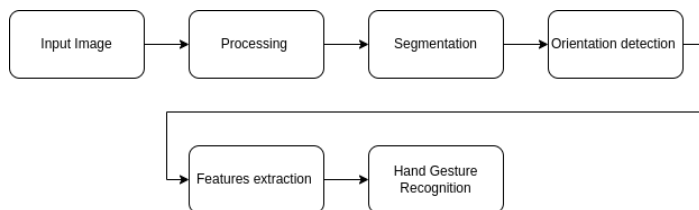


Fig 2. Diagrama da adaptação

A imagem de entrada passa por um pré processamento, sendo convertida de RGB para YCrCb. Após isso a imagem é binarizada e é passado um preenchimento de buracos para que esteja pronta para a classificação.

Para que a classificação seja feita de maneira mais fácil as “bordas” do gesto são identificadas, usando um scanearno

da esquerda pra direita do topo ao fim para encontrar a “borda” esquerda, da direita para a esquerda do topo ao fim para encontrar a “borda” direita, do topo ao fim para encontrar a “borda” do topo e a “borda” fim do final da imagem ao topo da “borda” esquerda a “borda” direita já calculadas.

Na fase de extração de características a imagem é avaliada em diversos aspectos:

- Orientação - para os fins deste estudo, avaliar a orientação tem papel importante, uma vez que os métodos empregados estão direcionados para imagens na vertical. Os cálculos seguem exatamente os moldes descritos nas referências utilizadas, preocupando-se com as dimensões do objeto na imagem binarizada;
- Possui ou não dedão levantado - isso é feito comparando a quantidade de pixels brancos existente em áreas da imagem com 7% da quantidade total de pixels brancos da imagem
- Área onde o dedão se encontra - a imagem é dividida em duas porções, esquerda e direita, com metade da quantidade de pixels usada no artigo original(15 pixels) e, usando o cálculo da quantidade de pixels brancos contidos na área, é possível identificar se o dedão está nessa área ou não;
- Identificação dos demais dedos: o processo de extração desta segmentação específica se dá num processo bem simples. Primeiro, todos os primeiros pixels não nulos da esquerda para a direita, de cima para baixo, são cotados como pontas dos dedos. Após, com técnicas meramente matemáticas, foi possível extrair deste grupo um subconjunto de máximos locais. Cada um destes novos pontos é, novamente, cotado para ser o topo do dedo na imagem vertical.
- Reconhecimento dos gestos - para o propósito deste relato, todos os gestos identificados foram classificados como tentativas de expressar números com as mãos. Assim sendo, o resultado do estudo, e da aplicação do método, resultam numa imagem segmentada, com todos os possíveis pontos cotados como ponta dos dedos, tanto os confirmados, quanto os assumidos, e um pequeno texto que acompanha cada uma das imagens. As curtas linhas possuem informações importantíssimas, juntamente com um apanhado dos tópicos acima.

### III. RESULTADOS

Breve explicação das linhas que acompanham as imagens: Orientation; Total de pixels brancos na imagem; Pixels brancos identificados nas áreas esquerda e direita, respectivamente; O limite para considerarmos a presença do dedão na área; Presença ou não do dedão na imagem; Quantidade de dedos encontrados durante a aplicação do método; Centros de massa do objeto segmentado;

Todas as linhas seguem o mesmo modelo, no intuito de facilitar o acompanhamento das informações.

Legenda para os pontos: pontos vermelhos são todos os assumidos como topo num primeiro momento. Os amarelos apresentam incerteza, por estarem próximos o suficiente de

outros pontos cotados também como máximos locais. Os verdes apresentam os pontos certos como máximos locais e coincidem, em sua maioria, à quantidade de dedos encontrada ao final da aplicação dos métodos



Fig 3. Entrada 1

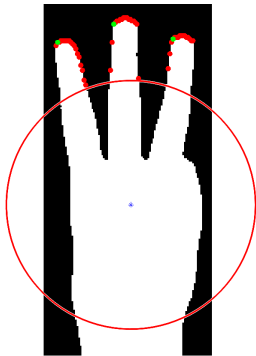


Fig 4. Saida 1

Orientation: Vertical;  
totalWhitePixels: 12818;  
leftWhitePixels: 1504;  
rightWhitePixels: 1975;  
whitePixelLimit: 897.26;  
No thumb;  
Dedos: 3;  
Weighted Centroid: (53, 122)



Fig 5. Entrada 2

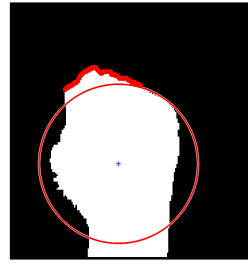


Fig 6. Saida 2

Orientation: Vertical  
totalWhitePixels = 9857  
leftWhitePixels = 705  
rightWhitePixels = 1339  
whitePixelLimit = 689.99  
No thumb  
Dedos = 0  
Weighted Centroid: (80, 118)



Fig 7. Entrada 3

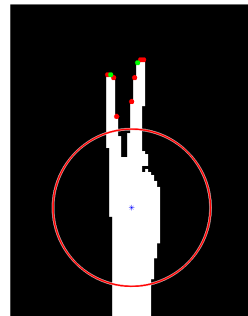


Fig 8. Saida 3

Orientation: Vertical  
totalWhitePixels = 1033  
leftWhitePixels = 988  
rightWhitePixels = 960  
whitePixelLimit = 72.310  
No thumb  
Dedos = 2  
Weighted Centroid: (41, 68)



Fig 9. Entrada 4

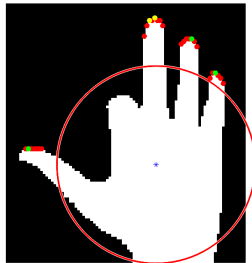


Fig 10. Saída 4

Orientation: Vertical  
totalWhitePixels = 3407  
leftWhitePixels = 127  
rightWhitePixels = 1084  
whitePixelLimit = 238.49  
Thumb: To the left  
Dedos = 3(?)  
Weighted Centroid: (57, 62)



Fig 11. Entrada 5

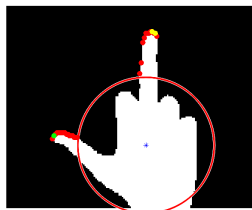


Fig 12. Saída 5

Orientation: Vertical  
totalWhitePixels = 4472  
leftWhitePixels = 133  
rightWhitePixels = 949  
whitePixelLimit = 313.04  
Thumb: To the left

Dedos = 1  
Weighted Centroid: (90, 90)



Fig 13. Entrada 6

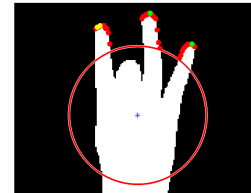


Fig 14. Saída 6

Orientation: Vertical  
totalWhitePixels = 5980  
leftWhitePixels = 1577  
rightWhitePixels = 496  
whitePixelLimit = 418.60  
No thumb  
Dedos = 2  
Weighted Centroid: (89, 80)



Fig 15. Entrada 7

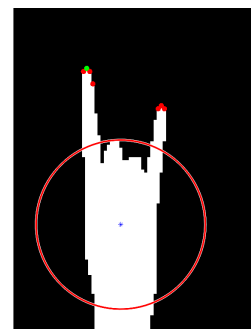


Fig 16. Saída 7

Orientation: Vertical  
totalWhitePixels = 1458  
leftWhitePixels = 949

rightWhitePixels = 800  
whitePixelLimit = 102.06  
No thumb  
Dedos = 1  
Weighted Centroid: (34, 70)

#### IV. CONCLUSÃO

O reconhecimento de gestos manuais com base na forma de parâmetros é uma área de pesquisa importante, com aplicações diversas em interação homem-máquina e sistemas de realidade virtual. Neste trabalho, apresentamos um método que utiliza técnicas de pré-processamento para binarizar as imagens de gestos e preencher buracos, garantindo uma representação adequada para a classificação.

A orientação correta das imagens é essencial para obter resultados consistentes, e, portanto, implementamos um procedimento para verificar a orientação das imagens. A identificação das bordas dos gestos e a localização do dedão são etapas cruciais, que permitem dividir a imagem em seções e obter informações relevantes para a classificação.

Ao encontrar o centro da imagem e analisar o número de dedos levantados e abaixados, conseguimos gerar um número de informações sobre características da imagem do gesto, mostrando a posição da imagem inicial, a quantidade de pixels brancos contidos na imagem e a quantidade de dedos levantados na mesma.

Os resultados obtidos demonstram a eficácia do método proposto e a sua viabilidade para a aplicação em sistemas de reconhecimento de gestos manuais. A abordagem baseada na forma de parâmetros oferece uma alternativa promissora para a detecção de características, podendo ser melhor trabalhada e melhorada ao nível de identificação e classificação de gestos contribuindo para o avanço no campo da interação humano-computador.