

# Introdução ao Processamento de Imagens

## Projeto Final

1<sup>st</sup> Filipe Maia Soares

15/0051409

2<sup>st</sup> Tiago de Souza Fernandes

18/0131818

3<sup>st</sup> Victor Ferreira Gonçalves

18/0047345

Universidade de Brasília

filipemaia-soares@gmail.com

**Abstract**—The purpose of this project was about inserting special elements into digital images of people's faces. For this, eyes, mouths and hats of cartoon characters were used.

### I. INTRODUÇÃO

A facilidade com que as fotos são tiradas a partir das câmeras de celulares constituiu fator determinante para o crescimento de redes sociais como o *Instagram*, *Pinterest* etc. Essas e outras mídias contêm uma alta variedade de filtros e efeitos para que o usuário possa customizar suas fotos da forma que preferir.

Grande parte dos efeitos tem como premissa inserir elementos inesperados na face do usuário, tais como orelhas de cão, feições felinas, corações nos olhos etc. Esse trabalho tem como objetivo modificar a face das pessoas com base em personagens de anime, tentando substituir os olhos e bocas das faces das imagens por partes dos personagens de forma proporcional ao rosto da imagem.

#### A. Anime

O termo anime se refere à animação que é produzida por estúdios japoneses. A palavra é a pronúncia abreviada de "animação" em japonês, onde esse termo se refere a qualquer animação. Para os ocidentais, a palavra se refere às animações oriundas do Japão.

As proporções corporais dos personagens de anime humanos tendem a refletir com precisão as proporções do corpo humano na realidade, porém uma das características mais marcantes dos animes é o tamanho exagerado dos olhos, o artista adiciona sombreamento de cor variável aos olhos e, particularmente, à córnea, para proporcionar maior profundidade. Geralmente, é usada uma mistura de sombra leve, cor de tom e uma sombra escura.

#### B. Canal Alpha

O canal alpha é um elemento que define a opacidade de determinado píxel em uma imagem. Caso sejam utilizados valores inteiros, seus valores variam em um intervalo de 0 à 255, onde o valor 0 representa a transparência máxima, 255 a mínima, e os demais representam valores intermediários. Esta

feature está presente apenas em alguns formatos de imagem, sendo o mais comum deles o PNG.

Uma das aplicações do canal alpha seria juntar duas ou mais imagens em uma única só evitando que o fundo de uma delas aparecesse em cima da outra. Isto poderia ser feito definindo regiões mais opacas como os objetos que se deseja sobrepor na imagem original e regiões transparentes como aquelas indesejadas na imagem final. Esta técnica pode ser utilizada para inserção de diversos elementos ilustrativos em uma imagem, como olhos, boca, chapéus e outros.

#### C. Identificação de Elementos da Face

Um dos primeiros passos para adicionar elementos em uma face é identificar as regiões nas quais se deseja inseri-los. Um dos métodos que pode ser utilizado para tal identificação seria o algoritmo de Viola-Jones.

Este algoritmo foi implementado para identificar faces, mas também pode fazer o mesmo para alguns de seus elementos, como olhos, boca, nariz etc. Sua efetividade está totalmente relacionada com o ângulo que a pessoa está olhando pra câmera, isto é, quanto mais direto pra câmera o olhar, melhores os resultados e, à medida que este ângulo vai aumentando, sua taxa de acerto diminui.

Apesar de esses requisitos parecerem extremamente limitantes, o algoritmo apresenta bons resultados para um grande grupo de imagens. Ele possui uma baixa taxa de resultados falso-negativos e pode ser utilizado em algumas aplicações em tempo real para identificação de faces.

O algoritmo específico para identificação de olhos tem como entrada uma imagem qualquer que contenha rostos. Sua saída consiste em uma matriz com o número de linhas correspondente à quantidade de olhos encontrados na imagem e 4 colunas. As duas primeiras colunas representam as coordenadas do ponto inicial de um retângulo que delimita a região dos olhos. Já as duas últimas representam, respectivamente, seu comprimento e sua largura.

#### D. Proporções do rosto humano

(a distância entre dois olhos é aproximadamente a distância entre a boca e a altura dos olhos)

## II. METODOLOGIA

### A. Inserção dos Olhos

O algoritmo se inicia recebendo a imagem original e outras 3 (chapéu, olhos e boca) juntamente com seus respectivos canais alpha. Caso alguma delas já não tenha um canal alpha, cria-se um preenchendo todos os seus elementos com o valor 255.

Em seguida, é realizada a identificação dos olhos utilizando o algoritmo de Viola Jones, que retorna em cada linha a quantidade de olhos detectados e nas colunas os valores referentes à posição inicial e dimensões do retângulo que delimita a região dos olhos.

O algoritmo implementado, por padrão, seleciona apenas o primeiro par de olhos identificado na imagem original e, em seguida, redimensiona a imagem dos olhos desenhados, de maneira a adaptar seu tamanho a rostos diferentes. O redimensionamento é feito segundo a seguinte equação.

$$Eye = imresize(Eyesize, propeye * scale); \quad (1)$$

$$scale = bbox(k, 3)/y1; \quad (2)$$

Onde  $scale$  é definido como a quantidade de vezes que o tamanho dos olhos da pessoa é maior do que o tamanho dos olhos desenhados, no eixo x. Esta medida de comprimento dos olhos da pessoa é estimada a partir do comprimento do retângulo delimitador. O redimensionamento é feito em seguida, tanto na imagem dos olhos quanto em sua matriz alpha, utilizando essa proporção de  $scale$  nos dois eixos.

Para identificar a posição dos olhos, calculou-se metade da diferença de comprimento ( $dy$ ) e largura ( $dx$ ) entre os olhos detectados e os redimensionados.

Por fim, realizou-se um loop que insere os novos olhos na imagem original. Para isso, o algoritmo se inicia da posição inicial do retângulo deslocada da distâncias calculadas previamente ( $dx$  e  $dy$ ) e incrementa esses valores para cada linha ou coluna avançada. Para cada posição a imagem de saída recebe uma combinação linear das matrizes original e a de olhos desenhados, segundo a equação abaixo.

$$Img(m, n, :) = (1 - \alpha) * Img(m, n, :) + \alpha * Eye(i, j, :); \quad (3)$$

Onde  $\alpha$  representa o valor, de 0 a 1, de opacidade em cada ponto. Dessa maneira, caso a opacidade valha 0 (fundo), a imagem de saída recebe apenas o valor da imagem original. Porém, caso a opacidade valha 1, ela recebe apenas o valor da imagem do olho.

### B. Substituição do Chapéu e Boca

A face humana possui proporções conhecidas entre suas dimensões, e essas proporções são frequentemente confirmadas na maioria dos rostos. Uma delas é a distância entre a altura dos olhos e a boca, que é aproximadamente igual a distância entre os dois olhos, e ela é a referência do algoritmo para estimar a posição da boca e analogamente a do chapéu.

Para inserir a boca, primeiramente ela passa pelo mesmo processo de redimensionamento da imagem e do canal alpha que foi implementado para inserir os olhos, contudo utilizando de uma nova constante de proporcionalidade encontrada empiricamente. Logo após, a boca é inserida utilizando o centro do retângulo, onde foi identificado os olhos, transladado para baixo utilizando uma constante de altura vezes o comprimento do retângulo.

Para a inserção do chapéu, um procedimento similar ao processo de inserção da boca, diferindo apenas das constantes de proporção e de altura, essa última que agora é negativa pois o chapéu é localizado acima da altura dos olhos. As constantes foram encontradas empiricamente analisando os resultados obtidos.

## III. RESULTADOS

### A. Acertos

Uma imagem de saída é considerada um acerto quando os olhos, chapéu e boca são posicionados no local exato ou muito próximo de onde deveriam. Este critério é subjetivo, mas foi o adotado para a realização dos testes. A figura 2 abaixo representa um exemplo deste caso. Nota-se da figura 1 que os retângulos e circunferências indicando os olhos, boca e chapéu foram muito bem posicionados.

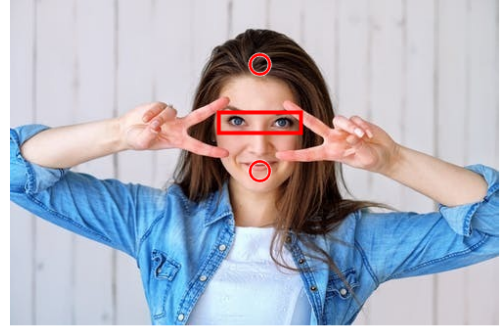


Figure 1. Exemplo 1 onde os olhos foram corretamente identificados.

### B. Detecção de Olhos Deslocada

O segundo caso representa as vezes em que o algoritmo não detecta com precisão a posição dos olhos e, consequentemente, o de posicionamento de bocas e chapéus não é capaz de posicioná-los corretamente. Nota-se que a face inclinada da mulher na figura 3 não permite uma identificação correta, uma vez que o algoritmo só detecta olhos horizontalmente, mesmo sentido do retângulo.

### C. Detecção de Chapéu e Boca Deslocada

O terceiro caso ocorre quando o rosto da pessoa desrespeita as proporções utilizadas para as estimativas dos locais da boca e chapéu. Dessa maneira, mesmo que os olhos sejam corretamente identificados, os outros elementos são posicionados erroneamente. Isto pode ser visto nas figuras 5 e 6, abaixo.



Figure 2. Exemplo 1 com os enfeites inseridos.



Figure 3. Exemplo 2 com a identificação deslocada dos olhos.

#### D. Detecção Incorreta de Olhos

Esta opção é caracterizada pela detecção de olhos que não existem na imagem. Dessa maneira, algumas regiões totalmente diferentes de faces acabam sendo erroneamente preenchidas com olhos, boca e chapéu. Isto pode ser visto nas figuras 7 e 8.

#### E. Não Detecção

O último caso apenas ocorreu uma vez a partir do banco de dados testado e consiste na falta de detecção de faces. Este fator pode estar relacionado com elementos que influenciam



Figure 4. Exemplo 2 com a enfeites inseridos.

diretamente nos olhos da pessoa, como por exemplo o reflexo nos óculos, mostrado na figura 9.

#### F. Percentuais

Os erros falso-positivos ocorrem quando há um par de olhos na imagem, mas o algoritmo identifica outros elementos da imagem, que não são olhos, como sendo o elemento desejado. Já os erros falso-negativos representam as situações em que não são identificados olhos na imagem, mesmo que esses elementos de fato existam.

Utilizando-se de um banco de dados com 30 imagens, foram obtidos os dados apresentados no gráfico 10 abaixo. Percebe-se uma alta taxa de acerto, indicando a robustez do algoritmo utilizado.

Os erros falso-positivos podem se tornar menos frequentes utilizando o algoritmo de Viola-Jones para uma identificação prévia de rostos humanos, para então uma posterior utilização do mesmo algoritmo para a identificação do par de olhos apenas em uma região pré-selecionada. A abordagem acima não foi implementada no projeto mas é uma melhoria digna de ser mencionada para a uma melhor performance do algoritmo.

## IV. CONCLUSÕES

O algoritmo implementado pelo grupo teve uma alta taxa de acerto, ou seja, os olhos, boca e chapéu foram devidamente posicionados na maior parte dos casos, entretanto, foi possível observar uma alta dependência do desempenho do algoritmo

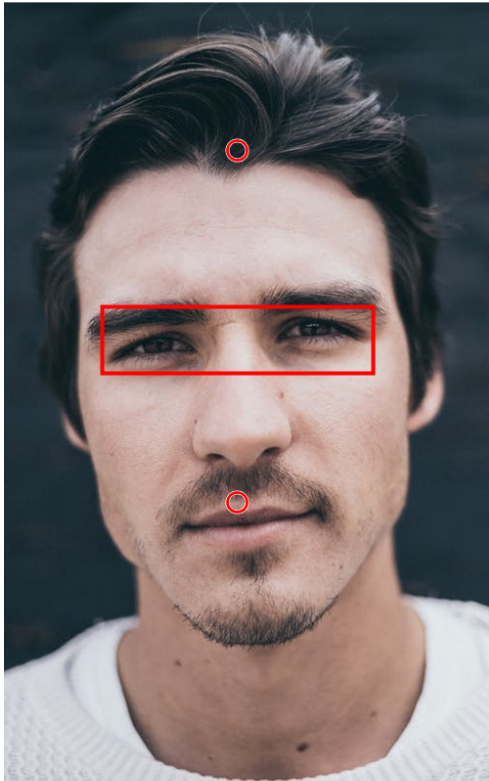


Figure 5. Exemplo 3 com a identificação deslocada do chapéu e boca.



Figure 6. Exemplo 3 com enfeites inseridos.

em relação às imagem de entrada. Este fato pode ser verificado em imagens inclinadas, com muitos detalhes ou em que a região dos olhos sofre interferência de muitos elementos, como o de reflexos ou iluminações muito diferentes.

Um possível fator de melhora do código seria a aplicação do algoritmo de identificação dos olhos em um espaço reduzido, previamente delimitado pela identificação de faces. Apesar dessas observações, os resultados obtidos foram muito bons utilizando-se uma solução com baixo grau de complexidade.

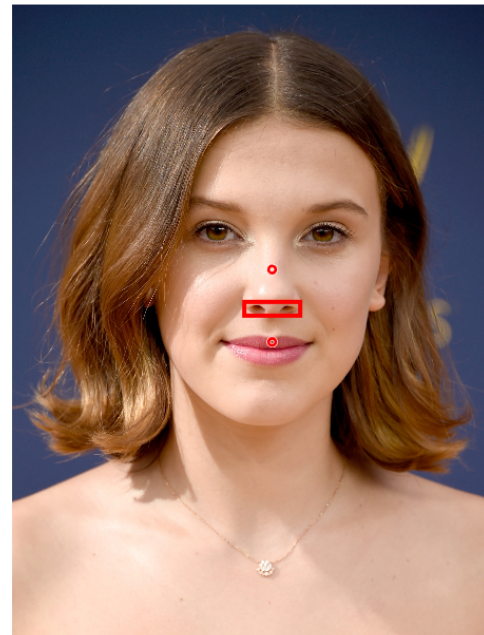


Figure 7. Exemplo 5 onde os olhos foram identificados erroneamente.





Figure 8. Exemplo 5 com enfeites inseridos.



Figure 9. Exemplo 6 onde nada foi encontrado.

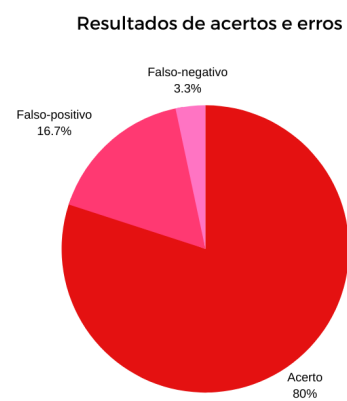


Figure 10. Gráfico de porcentagem dos resultados obtidos.