

Detecção de componentes conectados e classificação de imagens

Lucas Nascimento Santos Souza, 14/0151010

I. RESUMO

O Trabalho tem como objetivo realizar a detecção da pele em imagens com uma única pessoa, e com base nisto, classificá-la entre duas classes: Classe 1, a pessoa na imagem está completamente vestida; Classe 2, a pessoa na imagem está com o corpo descoberto. Para a realizar esta diferenciação entre classes, o algoritmo deve:

- Realizar a binarização da imagem, onde a cor de pele, definida pela paleta de cores adotada se equivale a 1, e o fundo da imagem a 0;
- Detectar a quantidade de elementos conectados na imagem binária;
- Determinar a área do maior elemento conectado e a partir desta área, classificar a imagem entre as duas classes existentes.

Após a realização de todos estes passos, a *F-measure*, uma análise estatística de classificação binária, deverá ser calculada e utilizada como critério para definir o desempenho do projeto.

II. INTRODUÇÃO

Para o entendimento do projeto, alguns relacionamentos básicos entre *pixels* devem ser explicitados, o que será feito no próximo parágrafo, sendo estes o conceito de vizinhança, conectividade, região e contorno.

Um *pixel* p com coordenadas (x, y) possui quatro vizinhos horizontais e verticais, cujas coordenadas são $(x + 1, y)$, $(x - 1, y)$, $(x, y + 1)$, $(x, y - 1)$. Este conjunto de *pixels* são denominados *vizinhança-de-4* de p , sendo representado pela notação $N_4(p)$. Cada *pixel* está a uma unidade de distância de p . Os pares ordenados dos quatro vizinhos diagonais de p são $(x + 1, y + 1)$, $(x + 1, y - 1)$, $(x - 1, y + 1)$, $(x - 1, y - 1)$ e são denotados por $N_D(p)$. A junção dos pontos das vizinhanças $N_4(p)$ e $N_D(p)$ é chamada de *vizinhança-de-8* de p . Nem todos os vizinhos permanecem dentro da imagem quando p se encontra na borda da mesma.

A conectividade entre *pixels* é um conceito relevante, utilizado no estabelecimento dos componentes conectados e também da borda de um objeto. Para determinar se dois *pixels* estão conectados, deve-se determinar alguma forma de adjacência entre eles, como estes serem vizinhos-de-4 e seus níveis de cinza iguais.

Seja V , o conjunto de valores de níveis de cinza utilizados para definir a conectividade, consideramos três tipos de conectividade:

a) conectividade-de-4: Dois *pixels* p e q , assumindo níveis de cinza em V , são conectados-de-4 se q está na *vizinhança-de-4* de p ;

b) conectividade-de-8: Dois *pixels* p e q , assumindo níveis de cinza em V , são conectados-de-8 se q está na *vizinhança-de-8* de p ;

c) conectividade-de-m: Dois *pixels* p e q , assumindo níveis de cinza em V , são conectados-de- m se:

- q está na *vizinhança-de-4* de p , ou
- q está em $N_4(p)$ e o conjunto $N_4(p) \cap N_4(q)$ for vazio.

Se p e q pertencerem a um subconjunto S de uma imagem, p estará conectado a q se existir um caminho entre eles constituído apenas de *pixels* pertencentes a S . Para qualquer *pixel* em S , o conjunto de *pixels* em S que estão conectados a este *pixel* é denominado *componente conectado* de S . Sendo assim, quaisquer par de *pixels* de um componente conectado estão conectados entre si, sendo que os componentes conectados distintos são disjuntos.

Uma região R é um *componente conectado*, e finalmente, um contorno C de uma região R é composto por todos os *pixels* que possuem vizinhos não pertencentes a esta região R .

III. METODOLOGIA

O Projeto foi desenvolvido na plataforma *MATLAB* no sistema operacional *Elementary OS (Linux)* e possui 6 arquivos, sendo 1 arquivo corrente *Principal.m* e outras 5 funções: *Paleta.m*, *Binarizacao.m*, *Rotulacao.m*, *Calibragem.m*, *Classificacao.m* e *Fmeasure.m*. Vários métodos foram utilizados para ser possível chegar a um resultado final. Os seguintes métodos, serão listados a seguir:

A. Definição a cor de pele

O primeiro método utilizado neste projeto foi a definição do tom de pele a ser detectado pelo algoritmo. Uma paleta de cores selecionada a partir de 10 imagens (5 homens e 5 mulheres) de cada classe (20 imagens no total), construindo uma única imagem a partir de recortes de regiões de cor de pele. Esta paleta de cores é passada da escala de cores *RGB* para a escala *YCbCr*, e após isso calcula-se a média e o desvio padrão das camadas *Cb* e *Cr* para definir o intervalo a ser considerado como cor de pele nas imagens a serem analisadas.

B. Binarização

O processo de binarizar as imagens consiste em mudar a escala de cores da imagem do *RGB* para a escala *YCbCr*, e verificar se os *pixels* das camadas *Cb* e *Cr* estão no intervalo estabelecido como cor de pele, se estes estiverem, o *pixel* correspondente na imagem binarizada recebe o valor 1, caso

contrário, este *pixel* recebe o valor 0. Portanto a cor de pele na imagem binarizada é representada pelo 1, e o que não é caracterizado como cor de pele recebe 0.

C. Rotulação

Neste método, uma nova matriz que é composta pelos rótulos é criada a partir da matriz binária obtida no processo de binarização da imagem original. O processo de rotulação consiste em:

- (i) Analisar todos os *pixels* $p = (x, y)$ da matriz binária;
- (ii) Se o valor de p é igual a 0, este pixel possui valor 1 na matriz de rótulos;
- (iii) Se o valor de p é igual a 1, analisam-se os vizinhos $b = p(x - 1, y)$ e $e = p(x, y - 1)$:
 - Se ambos os vizinhos são iguais a 0, um novo rótulo é atribuído a p ;
 - Se apenas um dos vizinhos é igual a 1, o rótulo corrente é adicionado a p ;
 - Se ambos os vizinhos são iguais a 1 e possuem o mesmo rótulo, este rótulo é adicionado a p ;
 - Se ambos os vizinhos são iguais a 1 e possuem rótulos diferentes, um rótulo é equivalente ao outro, e o menor rótulo é adicionado a p .
- (iv) Os rótulos equivalentes são substituídos por apenas um rótulo;
- (v) Os rótulos são organizados de 1 até $N+1$, sendo N o valor total de *componentes conectados*.

D. Calibragem e análise

A calibragem é realizada para obter a média das áreas dos maiores *componentes conectados* de cada classe, por meio da execução do algoritmo em 10 imagens (5 homens e 5 mulheres) de cada classe (20 imagens no total), calculando a área do maior componente conectado de cada imagem e definindo a média para cada classe.

Se a área do maior elemento conectado da imagem é menor ou igual a média da área das fotos Classe 1, esta imagem é classificada como Classe 1. Se a área do maior elemento conectado da imagem é maior ou igual que a média da área das fotos Classe 2, esta imagem é classificada como Classe 2.

E. Cálculo da F-measure

O cálculo é baseado no número de verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos. Onde os verdadeiros positivos são o número de fotos Classe 2 que são realmente da Classe 2, sendo assim os verdadeiros negativos são o número de fotos classificadas erroneamente como Classe 1. Os falsos positivos são o número de fotos Classe 1 que são realmente da Classe 1 e os falsos negativos são o número de fotos classificadas erroneamente como Classe 2. A fórmula utilizada para o cálculo é:

$$F_{measure} = \left(\frac{2V_P}{2V_P + F_P + F_N} \right)$$

Sendo V_P o número de verdadeiros positivos, F_P o número de falsos positivos e F_N o número de falsos negativos.

IV. RESULTADOS

Exemplificando a metodologia utilizada, uma imagem exemplo de cada classe será escolhida e os processos realizados com estas imagens serão explicados:



Figura 1. Imagem pertencente a Classe 2 binarizada.



Figura 2. Imagem pertencente a Classe 2 original e ao lado, a mesma sobreposta com a imagem binarizada com vermelho.



Figura 3. Imagem pertencente a Classe 1 binarizada.



Figura 4. Imagem pertencente a Classe 1 original e ao lado, a mesma sobreposta com a imagem binarizada com vermelho.

Os resultados encontrados para o intervalo considerados como pele das camadas C_b e C_r a partir dos cálculos descritos

anteriormente com a paleta de cores mostrada a seguir, são:

- (i) C_b : [98.5743; 120.1590];
- (ii) C_r : [140.1341; 162.0751].



Figura 5. Paleta de cores utilizada.

As figuras 1 e 2 representam uma imagem da Classe 1, o resultado da binarização obtido está de acordo com o esperado, pois as imagens desta classe devem possuir menos *componentes conectados* do que uma imagem da Classe 2 e vice-versa. O resultado da área do maior componente conectado desta imagem é: 2048 *pixels*.

As figuras 3 e 4 representam uma imagem da Classe 2, o resultado obtido também está de acordo com o esperado, a área do maior componente conectado desta imagem é: 21010 *pixels*.

Para se obter o resultado individual da avaliação de cada imagem do banco de dados, um vetor de resultados é criado para cada classe. No vetor da *Classe 1* o valor 0 representa a classificação efetuada de maneira correta, e o 1 a classificação errônea. O mesmo se aplica para o vetor da *Classe 2*, onde o valor esperado é igual a 2, e o valor que significa o erro é 3. Os valores dos vetores estão representados pelas figuras abaixo:

```
resultado_nao_nude =
Columns 1 through 24
3 2 2 2 3 2 2 2 3 2 2 2 2 2 3 2 3 2 2 2 2 2 2
Columns 25 through 48
2 2 3 2 2 3 3 2 2 3 3 2 3 2 2 3 2 2 2 2 2 3 2
Columns 49 through 72
2 2 3 2 3 3 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 3
Columns 73 through 93
2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 3 2 2
```

Figura 6. Vetor de resultados para as imagens da *Classe 2*.

```
resultado_nude =
Columns 1 through 24
0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 1 0 0 0 0 1 0
Columns 25 through 48
0 0 0 0 1 0 1 0 0 0 0 0 1 1 1 0 1 0 0 1 0 0 1 1
Columns 49 through 72
0 1 1 0 0 0 1 0 1 1 0 1 0 0 0 0 0 0 0 0 0 0 0
Columns 73 through 95
1 1 1 1 1 0 0 1 0 0 0 0 0 0 1 0 1 1 0 0 0
```

Figura 7. Vetor de resultados para as imagens da *Classe 1*.

A média das classes é calculada por meio da média das áreas do maior componente conectado das 20 imagens utilizadas na calibragem. O valor das médias para as duas classes são:

- (i) Média da *Classe 1* = 7497 *pixels*;

- (ii) Média da *Classe 2* = 5046 *pixels*.

Os valores dos verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos baseando se no total de 186 imagens, são:

- (i) V_P = 64 imagens;
- (ii) V_N = 70 imagens;
- (iii) F_P = 29 imagens;
- (iv) F_N = 21 imagens.

Utilizando estes valores como parâmetros, o valor da $F_{measure}$ é igual a 72%.

V. CONCLUSÕES

Com base nos dados obtidos e comparando-os com os esperados, o algoritmo teve um resultado satisfatório, pois na maioria das imagens avaliadas, estas são classificadas corretamente. A F-measure obtida foi média, considerando que esta medida vai de 0 até 1, o resultado de 0.7191 foi um resultado razoável.

REFERÊNCIAS

- [1] Zaghetto, Alexandre, **Introdução ao processamento de imagens - Notas de aula e transparências**. Brasília, Brasil: Universidade de Brasília, 2015.
- [2] Gonzalez, Rafael C. e Woods, Richard E., **Processamento de imagens digitais**, 1a ed. São Paulo, Brasil: Editora Edgard Blücher Ltda., 2000, ISBN 85-212-0264-4.