

Comparação de modelos de Detecção Facial

1st Tiago de Lima

Departamento de Ciência da Computação
Universidade Federal Rural de Pernambuco

Recife, Brasil

tiago.blima@ufrpe.br

Abstract—A detecção de face é um dos problemas relacionados a visão computacional que tem ganhado bastante relevância nos últimos anos. O objetivo principal é dada uma imagem contendo um rosto, determinar a posição exata onde se encontra a face. Considerando esse problema, diferentes abordagens têm sido testadas. Dois abordagens se tornaram proeminentes, um faz uso de inferências sobre como um rosto pode ser identificado em uma imagem, e a outra utiliza algoritmos supervisionados para que possam extrair tais características automaticamente de qualquer imagem baseando-se em um dado conjunto de treino.

Index Terms—Detecção Facial, Pele, Segmentação de Imagem

I. INTRODUÇÃO

O reconhecimento facial é uma atividade bastante útil no desenvolvimento de sistemas de segurança, autenticação e entre outros [1]. Tendo despertado o desenvolvimento de várias técnicas que pudessem obter sucesso nas mais diferentes condições possíveis [1]. Sendo assim, vários modelos diferentes têm sido propostos que vão desde redes neurais a modelos estatísticos [2]. Portanto, o reconhecimento facial é de suma importância e tem sido amplamente explorado nos últimos anos [2].

Existem dois métodos principais de detecção facial que utilizam informações a priori do rosto [2]. As informações levadas em consideração são o formato do rosto, detecção de pele e cabelo e entre outros [2]. Esse tipo de abordagem foi explorada bastante no início da década de 70 se consolidando como uma abordagem baseada em características (*features based*) [2]. Por outro lado, com o desenvolvimento de técnicas de reconhecimento de padrões, abordagens baseadas na detecção de faces baseadas em característica da própria imagem passaram a surgir [2]. Tais métodos incorporam as características dos rostos por um processo de treinamento, não necessitando, portanto, de uma extração de característica baseado em uma informação a priori do rosto [2].

Na uma abordagem baseada nas características de um rosto, vários aspectos são levados em consideração, como o nível de tons de cinza e a cor [2]. Infelizmente esse tipo de segmentação, por vezes, é ambíguo devido ao baixo nível das características analisadas [2]. Portanto, características como o formato do rosto são adicionadas ao processo de detecção de uma face. Um das formas de encontrar essas características é através do reconhecimento de arestas [2]. Esse tipo de

abordagem já foi explorada em diversos trabalhos e consiste em obter características que possam determinar a posição do rosto por exemplo [2]. Outros tipos de detecção levam em consideração o padrão de cores HSI, que se mostrou bastante robusto na detecção facial, além da normalização das cores [2]. Existem ainda outras características como a informação de cinza, o formato do rosto, heurísticas, e entre outros [2].

Em uma abordagem baseada na imagem, um dos principais objetivos é fazer com que os algoritmos sejam capazes de identificar automaticamente características relacionadas aos rostos humanos [2]. Esse processo é realizado através do treinamento de modelos capazes de separar imagens contendo faces e não contendo [2]. Em relação aos métodos usados, poderia-se citar aqueles que fazem uso da análise do componente principal (PCA) e que usam redes neurais [2]. No primeiro caso, considera-se que as imagens contendo rostos se encontram em um subespaço de imagens [2]. Para detectar características como nariz, olhos e boca, são utilizados autovetores cujas características são aprendidas de imagens do conjunto de treino e podem ser reconstruídas através dessas representações [2]. Usando redes neurais, as características das faces são aprendidas por múltiplas camadas de neurônios [2]. Durante o treino do reconhecimento de imagens contendo faces a tarefa da rede neural é reconstruir a própria imagem, para imagens sem faces é feita uma aproximação com os vizinhos mais próximos de uma imagem de face [2].

Um das preocupações existentes no desenvolvimento de sistemas de reconhecimento facial é a precisão e a simplicidade do sistema [2]. Isso se dá porque os modelos precisam, em alguns casos, serem capazes de funcionar em pequenos dispositivos móveis [2]. Essas foram algumas das preocupações do trabalho desenvolvido por [1]. Nele, desenvolve-se um modelo simples e efetivo para detecção facial, levando em consideração características simples como pele e cabelo [1].

II. CONCEITOS BÁSICOS

A. Sistema de Cores HSV

O sistema de cores Hue, Saturation and Value (HSV) representa melhor as cores referentes ao tons de pele do que o sistema de cores RGB que é comumente usado para representar imagens [3]. Nele, as representações da cor da pele ficam mais evidentes porque o sistema incorpora aspectos de como os humanos veem as cores [3]. O Hue traz a informação de quão pura é a cor, enquanto a saturação diz o quanto a cor

é diluída pela luz branca e o valor informa a noção acromática da cor [3].

B. Sistema de Cores HSI

O sistema de cores Hue, Saturation and Intensity (HSI) assim como o modelo de cores HSV também faz uso da saturação e do Hue [1]. Porém ao invés de fazer uso do valor faz-se uso da intensidade existente. Também é um sistema de cores que é próximo ao olho humano [4]. A relação dele com o sistema de cores RGB é explorada no método proposto por [1].

III. METODOLOGIA

No estudo, o modelo de detecção de face é dividido em pelo menos 5 passos diferentes [1]. O primeiro passo é a detecção da pele [1]. Esse passo, é feito de maneira bastante simples, primeiro as cores vermelha e verde são normalizadas e depois disso as seguintes equações são usadas para identificar os limites superior e inferior de cores que representam a pele [1]. A equação 1 representa a normalização das cores e equação 3 especifica os limites inferiores e superiores da cor da pele. Além disso, é preciso retirar os tons de branco que aparecem na imagem 5 e garantir um refinamento da segmentação utilizando-se do sistema de cores HSI.

$$r = \frac{R}{R + G + B} \quad (1)$$

$$g = \frac{G}{R + G + B} \quad (2)$$

$$F_1(r) = -1.376r^2 + 1.0743r + 0.2 \quad (3)$$

$$F_2(r) = -0.776r^2 + 0.560r + 0.18 \quad (4)$$

$$w = (r - 0.33)^2 + (g - 0.33)^2 > 0.001 \quad (5)$$

Sendo assim, o padrão de cores da pele é dada pela equação:

$$\text{skin} = \begin{cases} 1, \text{if } (g < f_1(r) \cap g > f_2(r) \cap w > 0.001) \\ H > 20 \cup H \leq 240 \\ 0, \text{otherwise} \end{cases}$$

Além do modelo apresentado no artigo [1], existem outras formas de detectar a pele como a que se utiliza do modelo HSV de cores. O modelo proposto por [1], ainda se utiliza de outros passos para a detecção de rostos. Após de detectar a pele, o método faz a detecção do cabelo através de um método semelhante a detecção de pele, estabelecendo um intervalo onde as cores do cabelo pode ser encontrado.

$$\text{hair} = \begin{cases} 1, \text{if } (I < 80 \cap (B - G < 15 \cup B - R < 15)) \\ \cup 20 < H \leq 40 \\ 0, \text{otherwise} \end{cases}$$

$$I = \frac{R + G + B}{3} \quad (6)$$

Após um passo de quantização da cor da pele e do cabelo as imagens são combinadas para o estabelecimento da região de interesse que nesse caso é a face.

No passo de quantização da pele, é estabelecido um *threshold*, onde os pixels são divididos entre aqueles que pertencem a cor de pele e que não pertencem. Se a soma dos pixels que representam a cor preta é maior do que 12, então aquela região não pertence a pele. Isso se torna mais uma parte do processo de segmentação da imagem melhorando a detecção da pele. Algo semelhante também ocorre para a detecção do cabelo que será também ser explorado. Em último passo, será realizada a rotulação dos componentes da imagem afim de estabelecermos o componente referente ao rosto a ser encontrado. Para isso, realizarmos a binarização da imagem como está exposto de maneira semelhante na imagem 2.



Fig. 1. Uma amostra de detecção de pele usando os métodos explorados neste trabalho. Da esquerda para direita temos a imagem original, o método usando o modelo de cores HSV e pelo artigo [1], por último.



Fig. 2. Uma amostra da binarização da cor da pele ao invés da quantização.

A. Objetivo e Resultados Esperados

Apesar de todos esses passos, o objetivo principal será a comparação do método de detecção de pele proposto por [1] com o método usando o padrão de cores HSV. Portanto, o sistema de cores HSI será usado para detecção através do estabelecimento de um limite superior e inferior de cores que possam determinar a região de pele. O limite inferior em relação as cores é de [0, 58, 30] e superior o limite superior de [33, 255, 255] [5]. Sendo assim, utilizarei esses limites de cores HSV para detecção da região de pele enquanto as outras etapas permaneceram as mesmas. Isso torna o processo bem mais simples, sem a necessidade de cálculos como os apresentados na seção anterior.

Será mostrado o comportamento de ambos os métodos em diferentes situações e como isso afeta a detecção das faces. Isso poderá ajudar na definição de uma melhor abordagem para detecção de rostos. O resultado esperado é uma avaliação de ambos os métodos de detecção de cor de pele e como isso

pode afetar a detecção do rosto. Sendo assim, manteremos os passos de quantização estabelecido no artigo anterior, para melhor compararmos o modelo de detecção de pele. Após isso, iremos estabelecer os limites da detecção realizada desenhando retângulos nas regiões que não estão marcadas como preto após a quantização tomado por base aquilo mostrado na figura 1.

IV. EXPERIMENTOS

A. Estudo de Caso I

Em um primeiro experimento iremos verificar o comportamento de ambos os métodos detectando a face baseado na imagem original do artigo [1]. Essa imagem como ilustrado 1 foi utilizada na detecção em ambos os métodos considerando a detecção de pele como base para detecção do rosto. Portanto, se detectarmos a região de pele iremos por consequência detectar o rosto. Como mencionado na seção anterior, o primeiro passo será a detecção da pele. Nesse passo, usamos dois métodos. O primeiro é baseado na detecção usando o padrão de cores HSV e o segundo foi baseado no modelo de cor HSI proposto pelo artigo original. Pode-se perceber na imagem 3 que com o modelo de cores HSV, podemos obter uma maior presença da região do rosto. Enquanto isso, o modelo de cores HSI como o mencionado no artigo [1] apresenta bastante falhas e regiões em preto onde, em tese, deveria ser o rosto. Ainda sim, através de ambos modelos é possível obter de maneira satisfatória os limites da região de onde seria o rosto.



Fig. 3. Uma amostra do processo de detecção de pele usando o modelo de cor HSV (à esquerda) e HSI proposto no artigo original (à direta).

O segundo passo é a quantização da imagem que diminui o tamanho dela porém isso acelera o processo que calcula a região provável do rosto. Para quantizar a imagem foi utilizado o kmeans cluster existente na biblioteca opencv. Que é feito através de uma janela 5×5 de pixel. Se na a quantidade de pretos for maior do que 12, então aquela região é descartada, caso contrário a região é considerada pele.

Depois disso, seguimos para realizar a rotulação dos componentes da imagem a fim de encontrarmos definitivamente o componente referente ao rosto. Para isso, foi realizada a binarização da imagem para que os componentes pudessem ser identificados. A seguir a iremos mostrar o efeito gerado pela rotulação dos rótulos na imagem.



Fig. 4. Uma amostra do processo de detecção de pele usando o modelo de cor HSV (à esquerda) e HSI proposto no artigo original (à direta).

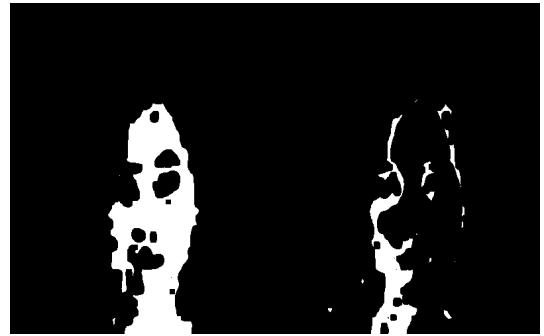


Fig. 5. Uma amostra do processo de detecção de pele usando o modelo de cor HSV (à esquerda) e HSI proposto no artigo original (à direta).

Até aqui, fica evidente que através de ambos modelo de cores é possível fazer com que os limites dos rostos sejam determinados o que pode garantir uma boa detecção facial. Porém, como será exposto na próxima subseção isso não acontece para todas as imagens.

Ainda considerando o artigo [1], nele é realizada a detecção de cabelo para que seja determinada com maior precisão onde está o rosto. A seguir vemos o processo de detecção do cabelo que é baseado nas equações que são descritas na seção III. Como pode ser visto na primeira etapa de detecção muito ruído acaba sendo confundido com a cor do cabelo em si. Porém, boa parte do ruído é removido nas etapas de quantização e rotulação dos componentes da imagem. Além disso, como as próximas figuras irão apontar, a detecção do cabelo parece bastante consistente com a área que de fato corresponde a essa região.

Por fim, foi desenhado um retângulo foi desenhado na seguindo os limites de onde as cores da pele existem no caso e do cabelo (no caso do algoritmo proposto no artigo original).

Assim como no artigo original, é possível observar que o rosto é detectado de maneira bastante centralizada tanto pela detecção de cabelo como pela detecção da pele em si. Isso acontece apesar de falhas consideráveis durante o processo de detecção que vão desde da detecção da cor da face até a rotulação dos componentes. Enquanto isso, usando o sistema de cores HSV, é possível notar que podemos obter uma qualidade maior na imagem do rosto que é detectada.



Fig. 6. Uma amostra do processo de detecção do cabelo usando o modelo de cor HSI proposto no artigo original .



Fig. 7. Uma amostra do processo de detecção do cabelo após a quantização da imagem usando o modelo de cor HSI proposto no artigo original .

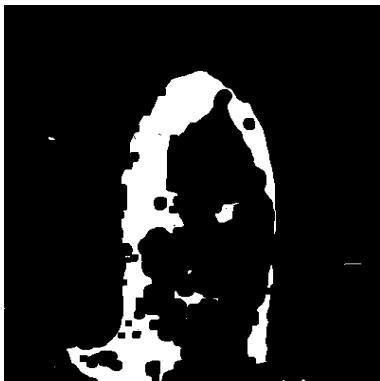


Fig. 8. Uma amostra do processo de detecção do cabelo após a quantização da imagem usando o modelo de cor HSI proposto no artigo original .

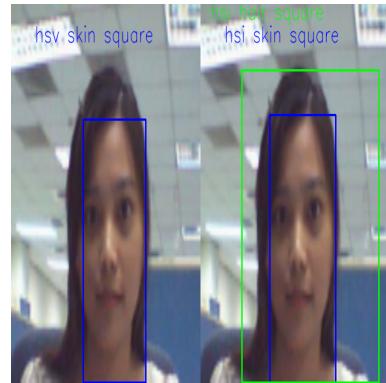


Fig. 9. Uma amostra do processo de detecção de pele usando o modelo de cor HSV (à esquerda) e HSI proposto no artigo original (à direita). O retângulo em azul representa a detecção de pele em ambos os casos, já o retângulo em verde representa a detecção da região do cabelo no caso do modelo de cor HSI.

B. Estudo de Caso II

No primeiro estudo de caso foi possível observar que apesar de todos as limitas o sistema HSI pode prover uma detecção consideravelmente boa em relação a imagem original. Para ambos os sistemas de cores, foi possível obter um resultado razoável em relação a detecção do rosto. Porém, como irei mostrar aqui, isso não acontece com todos os tipos de imagens. Na imagem a seguir, o sistema é incapaz de reconhecer o rosto e, portanto, não parece ser tão robusto. Enquanto isso, o sistema de cores HSV consegue reconhecer de maneira consistente o rosto na imagem demonstrando ser mais robusto e consistente.



Fig. 10. Uma amostra da imagem que será usada para detecção de rosto no segundo exemplo.

Assim como no primeiro estudo de caso podemos podemos verificar o uso de ambos os modelos para a detecção da região da pele. Logo aqui, pode-se notar uma diferença com o primeiro exemplo. Enquanto no primeiro exemplo pode-se verificar ao menos os limites superior e inferior da região da pele com clareza. Aqui, usando o sistema HSI, pode-se notar que a pele pode ser bem mal detectada restando poucos sinais da região exata de onde ele pode se encontrar. Enquanto isso,

o sistema HSV de cores consegue detectar de maneira bem mais clara a região da pele.



Fig. 11. Uma amostra do processo de detecção de pele usando o modelo de cor HSV (à esquerda) e HSI proposto no artigo original (à direta). Nesse exemplo é possível perceber que a cor da pele não é detectada usando o sistema de cores HSI.

Por consequência, considerando que os próximos passos se baseiam nesse mesma imagem, não será possível detectar o rosto apenas por meio da pele apenas.

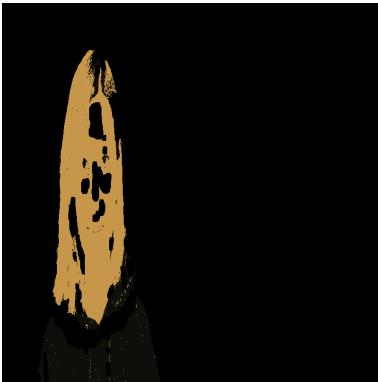


Fig. 12. Uma amostra do processo de detecção de pele usando o modelo de cor HSV (à esquerda) e HSI proposto no artigo original (à direta). Nesse exemplo é possível perceber que a cor da pele não é detectada usando o sistema de cores HSI.

A seguir realizamos a quantização de ambas imagens. É possível ver que usando o sistema de cores HSV, apesar de algumas falhas, boa parte da cor da pele é mantida. Enquanto isso usando o sistema de cores HSI os pixels de cor de pele que ainda restavam acabam por desaparecer devido ao *threshold* de 12 pixels estabelecido no processo de quantização da imagem impedindo qualquer possibilidade de detecção.

Agora, com a rotulação dos componentes é possível visualizar que o sistema de cores HSV. A rotulação permite uma maior clareza na região considerada de pele retirando boa parte dos ruídos.

Seguindo o que foi proposto no artigo [1], realizou-se uma detecção do cabelo. Nesse caso, a detecção do cabelo foi possível tal como a detecção do rosto no modelo HSV. Isso fica evidente na próxima imagem 14. Também é possível ver que a região do rosto também foi preservada. Isso deve ter

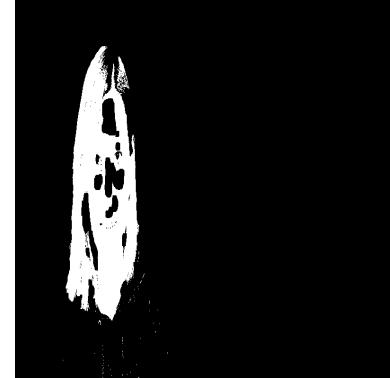


Fig. 13. Uma amostra do processo de detecção de pele usando o modelo de cor HSV (à esquerda) e HSI proposto no artigo original (à direta). Nesse exemplo é possível perceber que a cor da pele não é detectada usando o sistema de cores HSI.

acontecido devido ao fato da cor da pele estar próxima da cor do cabelo em termos de tonalidade.



Fig. 14. Uma amostra do processo de detecção do cabelo usando o modelo de cor HSI proposto no artigo original (à direta). Nesse exemplo é possível perceber que não a cor da pele não é detectada usando o sistema de cores HSI.

Os próximos passos são a quantização da imagem e a rotulação dos componentes. É possível ver que os limites do rosto são bem mantidos apesar de algumas falhas dentro da região considerada de pele.

Esses passos permitem mostrar que o modelo irá basear-se apenas na detecção do cabelo e não da cor da pele para encontrar o rosto. Isso acaba por tornar o processo de detecção da cor da pele totalmente dispensável, um vez que ele não contribui para o processo de detecção facial.

Percebe-se portanto que a detecção facial realizada por ambos os modelos é semelhantemente a mesma.

C. Considerações Finais

Considerando ambos os estudos de casos, é possível avaliar que no primeiro caso a detecção utilizando a proposta do artigo [1] foi bem sucedida quando utilizando a imagem do artigo. As regiões referente a pele e cabelo foram bem delimitadas. O resultado obtido foi bastante semelhante aquele que foi encontrado usando o sistema de cores HSV. Apesar



Fig. 15. Uma amostra do processo de detecção do cabelo após a quantização da imagem usando o modelo de cor HSI proposto no artigo original .

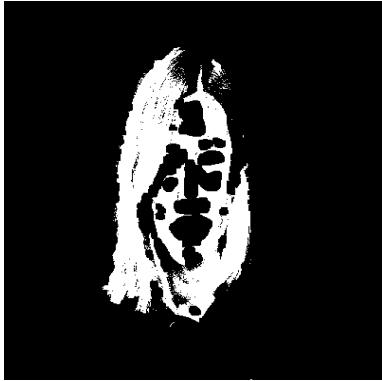


Fig. 16. Uma amostra do processo de detecção do cabelo após a quantização da imagem usando o modelo de cor HSI proposto no artigo original .

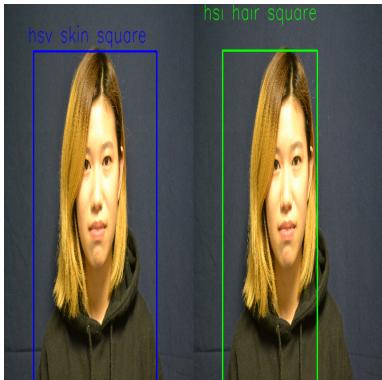


Fig. 17. Uma amostra do processo de detecção fácil usando o sistema de cores HSV (à esquerda) e HSI (à direita). O retângulo de azul representa a detecção baseada na pele e o retângulo verde representa a detecção do cabelo feita com o sistema de cores HSI.

do resultado final ser semelhante, é possível notar que a detecção de pele é muito melhor realizada usando o sistema de cores HSV do que HSI naquilo que foi proposto no artigo original [1]. Isso pode ser visto na figura 4.

Sendo assim, é evidente que não é prático o uso de detecção de dois aspectos distintos do rosto quando existem casos onde apenas um deles funciona. Enquanto, isso usando o sistema de cores HSV o mesmo resultado pode ser obtido utilizando-se apenas a detecção de pele. Portanto, em termos de praticidade é melhor usar-se o sistema de cores HSV ou a detecção da região do cabelo nesse dois casos aqui explorados. É necessário uma maior investigação sobre possíveis cenários a serem abordados usando ambos os sistemas de cores. Além disso, poderia-se citar a limitação dessa abordagem de detecção que não é capaz de se adaptar ao ambiente onde a foto foi tirado nem ao ângulo e entre outras características da imagem.

V. TRABALHOS FUTUROS

Em trabalhos futuros, além da implementação completa do artigo original, podem ser pesquisadas formas de detecção que não dependam de um *threshold* fixo, mas sejam capazes de extrair as características automaticamente da imagem. Esse tipo de abordagem é conhecida como uma abordagem baseada na imagem e em termos gerais baseia-se no fato de alguns algoritmos são capazes de aprender automaticamente as características existentes na imagem [2]. Um exemplo disso seria o uso de redes neurais convolutivas ou análise de PCA [2]. Isso permitiria a detecção facial em mais diferentes posições e formatos e ângulos. Ademais, será relevante considerar o uso de mais imagens para a comparação dos resultados.

VI. CONCLUSÃO

Nesse trabalho, foi possível verificar as diferenças significativas entre os sistemas de cores HSV e HSI no tocante a detecção de pele e reconhecimento da região do rosto. Foi possível verificar que o sistema de cores HSV apresenta-se com uma maior precisão detectando diferentes tons de pele. Isso não aconteceu com o modelo proposto no artigo [1]. Além disso, foi possível investigar que sistemas baseados em características a priori dos objetos a serem detectados podem sofrer quando esses não estão dispostos de maneira a favorecer os sistemas baseados nas premissas usadas na construção dos algoritmos de detecção. Isso mostra que é importante investigar o uso de modelos e algoritmos que sejam capazes de extrair automaticamente as características dos objetos presentes na imagem, bem como, de encontrar as regiões referentes a eles.

REFERENCES

- [1] Y. J. Chen and Y. C. Lin, "Simple face-detection algorithm based on minimum facial features," *IECON Proceedings (Industrial Electronics Conference)*, pp. 455–460, 2007.
- [2] E. Hjelmås and B. K. Low, "Face Detection: A Survey," *Computer Vision and Image Understanding*, vol. 83, pp. 236–274, 2001.
- [3] I. Aldasouqi and M. Hassan, "Smart human face detection system," *International journal of computers*, vol. 5, no. 2, pp. 210–216, 2011.

- [4] E. Welch, R. Moorhead, and J. Owens, “Image processing using the hsi color space,” in *IEEE Proceedings of the SOUTHEASTCON'91*. IEEE, 1991, pp. 722–725.
- [5] “Skin detection using opencv python.” [Online]. Available: <https://nalinc.github.io/blog/2018/skin-detection-python-opencv/>