

UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

**Correção de cor para tons de pele em imagens adquiridas
por dispositivos móveis**

Relatório científico parcial de iniciação científica referente ao processo FAPESP nº 2023/15404-1

Bolsista: Rodrigo Rodrigues de Castro

Orientador: João do Espírito Santo Batista Neto

São Carlos, SP

2024

Sumário

1	Introdução	3
1.1	Projeto e cronograma originais	3
2	Atividades realizadas	5
2.1	Detecção da face e extração das regiões de interesse	5
2.2	Color grading	7
3	Resultados obtidos	11
4	Próximas atividades	13
	Referências	14

1 Introdução

Este documento tem como propósito relatar as atividades realizadas pelo bolsista Rodrigo Rodrigues de Castro no período de janeiro de 2024 a junho de 2024, referentes ao projeto de Iniciação Científica, processo FAPESP nº 2023/15404-1, intitulado "Correção de cor para tons de pele em imagens adquiridas por dispositivos móveis". O projeto visa o desenvolvimento de um método de correção de cor de pele humana (*whitebalance*) a partir de uma imagem de referência.

1.1 Projeto e cronograma originais

A representação da cor dos tons de pele em imagens frequentemente é imprecisa, uma vez que vários fatores externos, como irregularidades naturais da pele, divergências de fontes de iluminação e variações na configuração do dispositivo de captura, podem propiciar a geração de imagens com cores desproporcionais às reais. Isso representa um obstáculo para várias áreas de pesquisa onde as imagens faciais são essenciais para a análise, como fotografia, dermatologia, cosméticos e entretenimento visual, por exemplo.

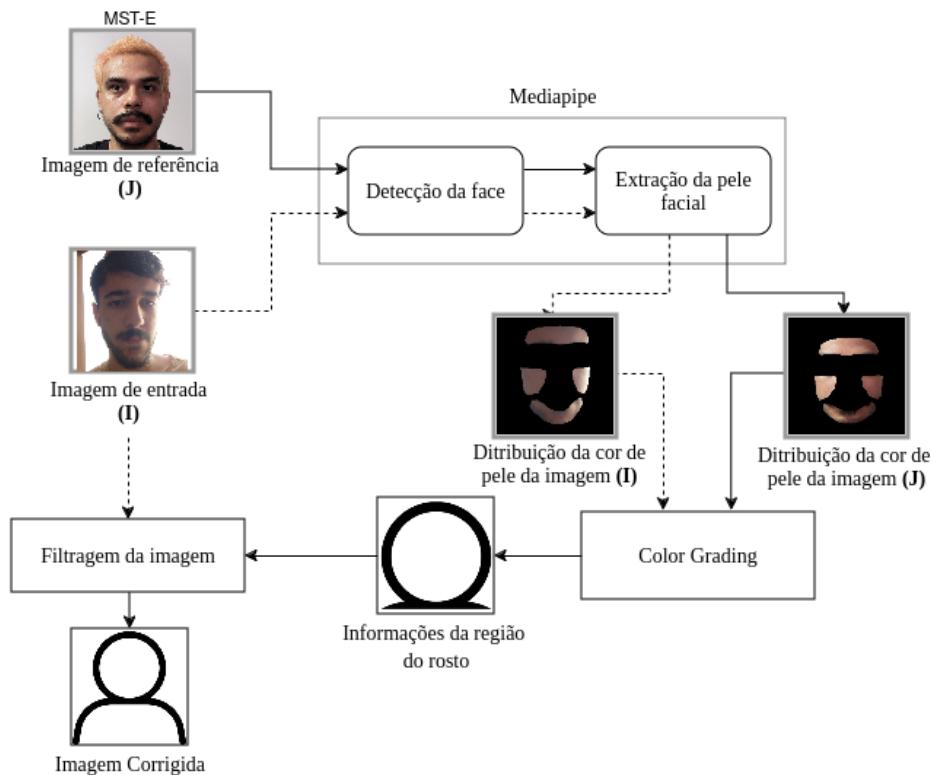
Nesse sentido, este projeto de pesquisa tem como finalidade principal o desenvolvimento de um método de correção de cor de pele humana (*whitebalance*) a partir de imagens adquiridas por dispositivos móveis, que podem apresentar uma variabilidade considerável de tonalidade entre amostras de uma mesma pessoa.

A proposta de correção de cor de uma imagem I emprega uma imagem de referência J . A imagem J é tida como idealmente precisa e possui uma cor representativa da escala de cor adotada. Este modelo de *whitebalance* é inspirado no trabalho de Shirai et al. (2021). A realização desta tarefa implica na execução das seguintes etapas:

- Detectar a face de pessoas em imagens;
- Extrair características referentes à pele facial de imagens;
- Realizar graduações de cor promovendo uma transição suave do tom de pele da imagem I segundo o tom de pele da imagem de referência J ;
- Fazer a fusão do novo tom de pele da imagem I com os elementos de fundo existentes (que não sejam pele) na imagem de entrada.

O diagrama da Figura 1 apresenta os principais módulos constituintes da metodologia proposta. O cronograma, como definido originalmente, está representado na Tabela 1.

Figura 1 – Diagrama referente a metodologia utilizada para correção de cor de tons de pele.



Fonte: Elaborada pelo autor.

Atividades	Meses de trabalho					
	1º e 2º	3º e 4º	5º e 6º	7º e 8º	9º e 10º	11º e 12º
Estudo dos mecanismos do MediaPipe	•	•				
Estudo das técnicas de Color Grading		•	•			
Extração das distribuições de cor de pele		•	•			
Redação do relatório parcial			•			
Implementação da correção de cor			•	•	•	
Avaliação e testes				•	•	
Desenvolvimento do relatório final					•	•

Tabela 1 – Cronograma de atividades original para 12 meses de trabalho.

O cronograma apresentado tem sido cumprido como originalmente proposto. As próximas seções deste relatório detalham o andamento do projeto, explicando como as atividades previstas na Tabela 1, no período decorrido, tem sido executadas. A seção 2 apresenta as principais atividades realizadas ao longo do projeto. A seção 3 expõe os resultados obtidos e as considerações levantadas sobre eles. Por fim, a seção 4 apresenta as atividades planejadas para a conclusão do processo desta iniciação científica.

2 Atividades realizadas

Nesta seção estão relatadas as atividades desenvolvidas para a implementação do projeto. Grande parte dos esforços se concentraram no estudo e na aplicação do *Color Grading*, uma técnica de fundamental importância no contexto de correção de cor. Todas as imagens de pessoas utilizadas nesta seção são parte de um conjunto de exemplos fornecidos pelo Google (SCHUMANN et al., 2023), referentes à escala de cor de pele *Monk Skin Tone* (MONK, 2023) utilizada neste projeto.

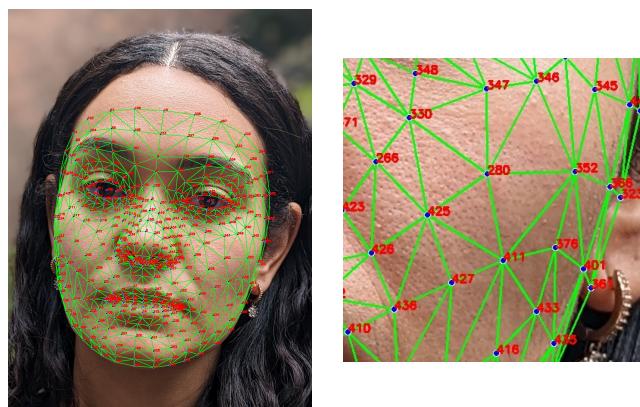
A seção 2.1 detalha o processo de detecção e extração das regiões de interesse da face humana nas imagens analisadas. Já a seção 2.2 descreve a metodologia e a aplicação dos conceitos de *color grading* baseados no trabalho de Pitié, Kokaram e Dahyot (2007).

2.1 Detecção da face e extração das regiões de interesse

A primeira etapa para o desenvolvimento do projeto é descrita no diagrama da Figura 1, como a detecção e a extração das regiões de interesse da face humana. Para isso, foi utilizado o *framework* Mediapipe (LUGARESI et al., 2019), desenvolvido pela Google.

Para detecção da face humana, o Mediapipe utiliza modelos de aprendizado de máquina treinados para identificar regiões de uma imagem ou vídeo onde rostos estão presentes. Após a detecção de uma face, o Mediapipe permite realizar a extração de características faciais, por meio de modelos de reconhecimento facial, que identificam 468 *landmarks* faciais mapeados, ou seja, pontos chave na face humana, como pontos nos olhos, nariz e boca, assim como apresentado na Figura 2.

Figura 2 – (a) Mapeamento completo dos 468 *landmarks* faciais do Mediapipe. (b) Aproximação da imagem para uma melhor visualização.

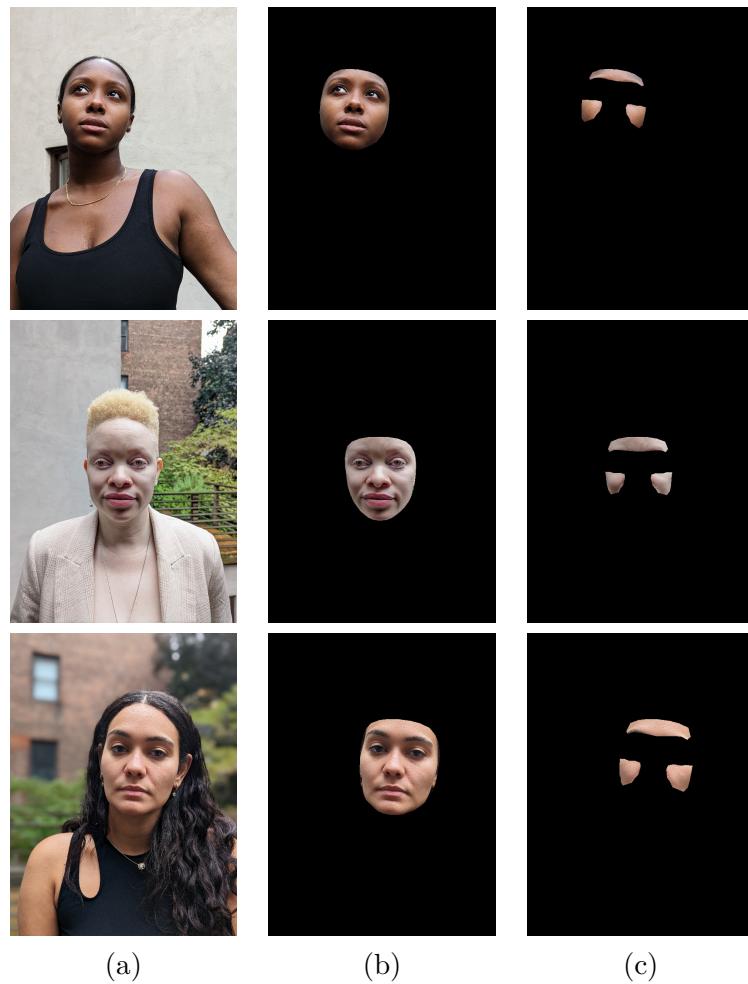


(a) (b)

Fonte: Elaborada pelo autor.

Utilizando essa ferramenta, foi possível extrair as regiões de interesse das imagens de entrada, isolando apenas as informações relevantes dos rostos dos indivíduos presentes na imagem. Este processo foi realizado por meio da segmentação do rosto, baseada na utilização de *landmarks* predefinidos, na imagem de entrada, garantindo que apenas as regiões desejadas permaneçam na imagem. A Figura 3 apresenta exemplos da aplicação do Mediapipe na extração das regiões de interesse para as etapas subsequentes do projeto.

Figura 3 – Extração das regiões de interesse. (a) Foto de entrada. (b) Segmentação do rosto da imagem de entrada. (c) Segmentação de regiões de interesse do rosto de entrada.



Fonte: Elaborada pelo autor.

Com a imagem produzida nessa etapa, é possível extrair a distribuição da cor de pele dos indivíduos nas imagens ignorando informações de fundo que não serão úteis para a correção de cor, ao considerar apenas os pixels que não representam o fundo da imagem segmentada no momento da obtenção das distribuições de cor.

2.2 Color grading

O *color grading* (gradação de cores, em português) é uma técnica que permite com que ocorra a transição suave entre tonalidades de cor entre diferentes imagens (REINHARD et al., 2001) de maneira com que a imagem resultante permaneça natural porém com a estética visual da imagem alvo.

A forma de implementar essa técnica utilizada neste projeto é descrita no trabalho de Pitié, Kokaram e Dahyot (2007). Nela, a partir de um mapeamento de cores, é possível transferir a paleta de uma imagem alvo (referência) para a imagem original (a ser modificada), conforme ilustrado na Figura 4 . Isto é realizado por meio da transformação de uma função de densidade de probabilidade (PDF) N-dimensional em outra.

Figura 4 – *Color grading* proposto.



Fonte: Pitié, Kokaram e Dahyot (2007).

Para o caso unidimensional (imagem em tons de cinza), considerando uma imagem de entrada I e uma imagem alvo G , o objetivo é encontrar uma transformação t que mapeie a função de densidade de probabilidade (PDF) de I na PDF de G , representadas por $f(u)$ e $g(v)$, respectivamente, onde u e v são variáveis aleatórias correspondentes às intensidades dos pixels nas imagens. Assim, procuramos encontrar o mapeamento das intensidades de u para v , ou seja, $t(u) = v$.

Usando notações de funções de distribuição acumulada (CDF) F e G para f e g , obtemos a expressão para o mapeamento t por meio das seguintes equações:

$$\int^u f(u) du = \int^{t(u)} g(v) dv \quad (2.1)$$

$$G(v) = F(u) \implies v = G^{-1}(F(u)) \quad (2.2)$$

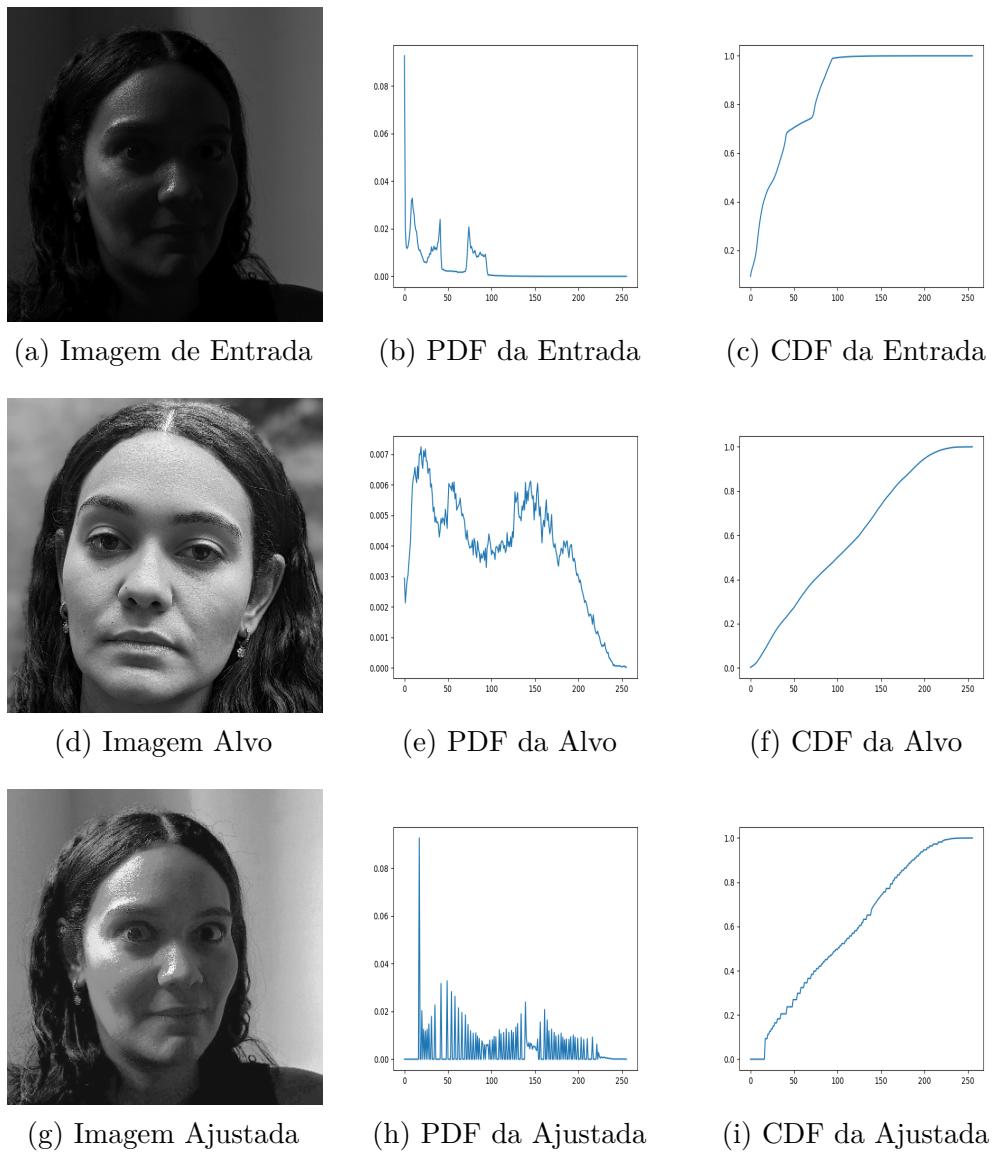
portanto,

$$\forall u \in \mathbb{R}, \quad t(u) = G^{-1}(F(u)) \quad (2.3)$$

onde $G^{-1}(a) = \inf\{u : G(u) \geq a\}$. O mapeamento pode então ser resolvido utilizando tabelas de consulta discreta (*look-up table*).

Assim, como os histogramas das imagens em tons de cinza representam a distribuição das intensidades dos pixels na imagem, a correspondência (GONZALEZ; WOODS, 2018) entre eles garante a aplicação da metodologia descrita. A Figura 5 ilustra um exemplo de aplicação de *color grading* em uma imagem tons de cinza. Nela se observa que a PDF e a CDF da imagem ajustada se assemelham às da imagem alvo.

Figura 5 – *Color grading* unidimensional de uma imagem de entrada com um alvo resultando em uma imagem ajustada, ambas imagens com suas respectivas função distribuição acumulada (CDF) e função de densidade de probabilidade (PDF).

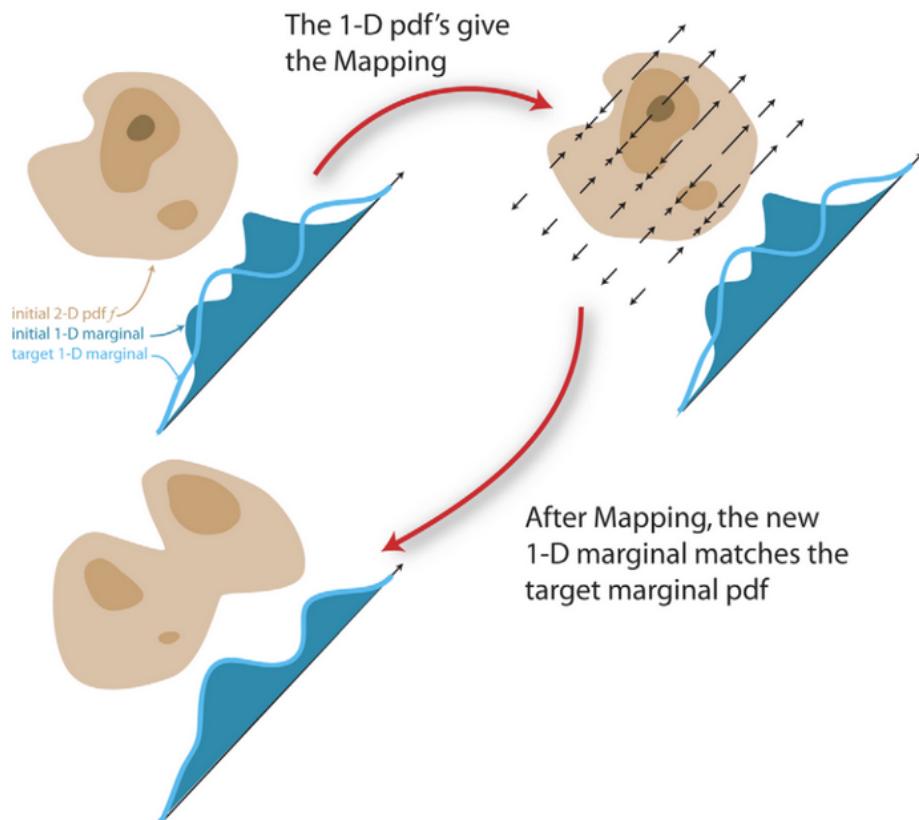


Fonte: Elaborada pelo autor.

Para o caso N-Dimensional (imagens coloridas), o método descrito por Pitié, Kokaram e Dahyot (2007) propõem a utilização da Transformada de Radon (??) em

que qualquer função contínua N-dimensional pode ser representada como uma série de projeções em eixos unidimensionais. Neste caso, a função considerada é uma PDF N-dimensional. Portanto, as projeções da Transformada de Radon resultam em uma série de PDFs marginais unidimensionais. A Figura 6 ilustra o processo de transformação de uma PDF bidimensional (2D) para ajustar suas marginais unidimensionais (1D) a uma distribuição alvo específica.

Figura 6 – Ilustração da manipulação de dados, baseada na transferência unidimensional de PDF em um eixo.



Fonte: Pitié, Kokaram e Dahyot (2007)

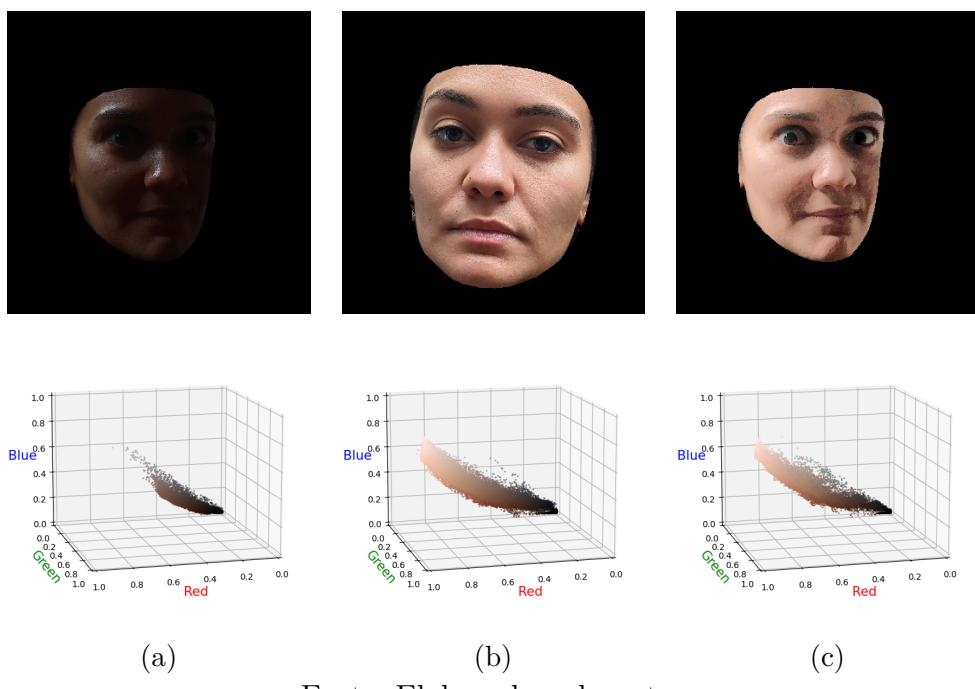
O Algoritmo 1, descrito por Pitié, Kokaram e Dahyot (2007), demonstra a aplicação dos procedimentos propostos para casos N-Dimensionais. Inicialmente, as imagens de entrada e referência são inicializadas e um loop é iniciado persistindo até a convergência de todas as marginais ou após o número definido de iterações. Com isso, uma matriz de rotação aleatória é gerada e os eixos de rotação são selecionados. Estes eixos são então aplicados sequencialmente às imagens de entrada e saída, resultando em valores variando de negativos a positivos. Os ajustes dos histogramas das imagens são então realizados com uma quantidade pré-selecionada de colunas. Em seguida, as intensidades da imagem de entrada são remapeadas de acordo com as transformações 1D encontradas, semelhante ao caso unidimensional. Por fim, as rotações são revertidas para concluir o processo.

Algorithm 1 Metodologia iterativa para o caso N-Dimensionais

- 1: Inicialização da imagem de entrada e da imagem de referência
- 2: **repeat**
- 3: Obter uma matriz de rotação aleatória $N * 3D$
- 4: **for** cada eixo de rotação 3D **do**
- 5: Obter as projeções de ambas as imagens nesse eixo
- 6: Realizar a correspondência entre os histogramas
- 7: Calcular as PDF de ambas
- 8: Encontrar as transformações 1D que criem uma correspondência entre ambas as PDFs
- 9: Remapear as amostras de cor de acordo com as transformações 1D
- 10: **end for**
- 11: Desfazer as rotações
- 12: **until** convergência em todas as marginais ou numero de interações definidas

Aplicando o Algoritmo 1 em imagens de indivíduos com suas regiões faciais de interesse segmentadas, conforme detalhado na seção 2.1, foi possível realizar o *color grading* entre uma imagem de entrada e uma imagem de referência, conforme o exemplo apresentado na Figura 7. Esse processo considerou apenas as partes da imagem onde a face humana está presente durante a correspondência entre os histogramas, garantindo que o fundo da imagem não influenciasse a aplicação dessa metodologia.

Figura 7 – Color Grading entre uma imagem de entrada (a) e uma imagem de referência (b), resultando em uma imagem (c) com a gradação de cores corretas. A linha superior apresenta a segmentações da face humana das duas imagens. Enquanto a linha inferior representa a distribuições de cores da imagem no espaço de cores RGB.



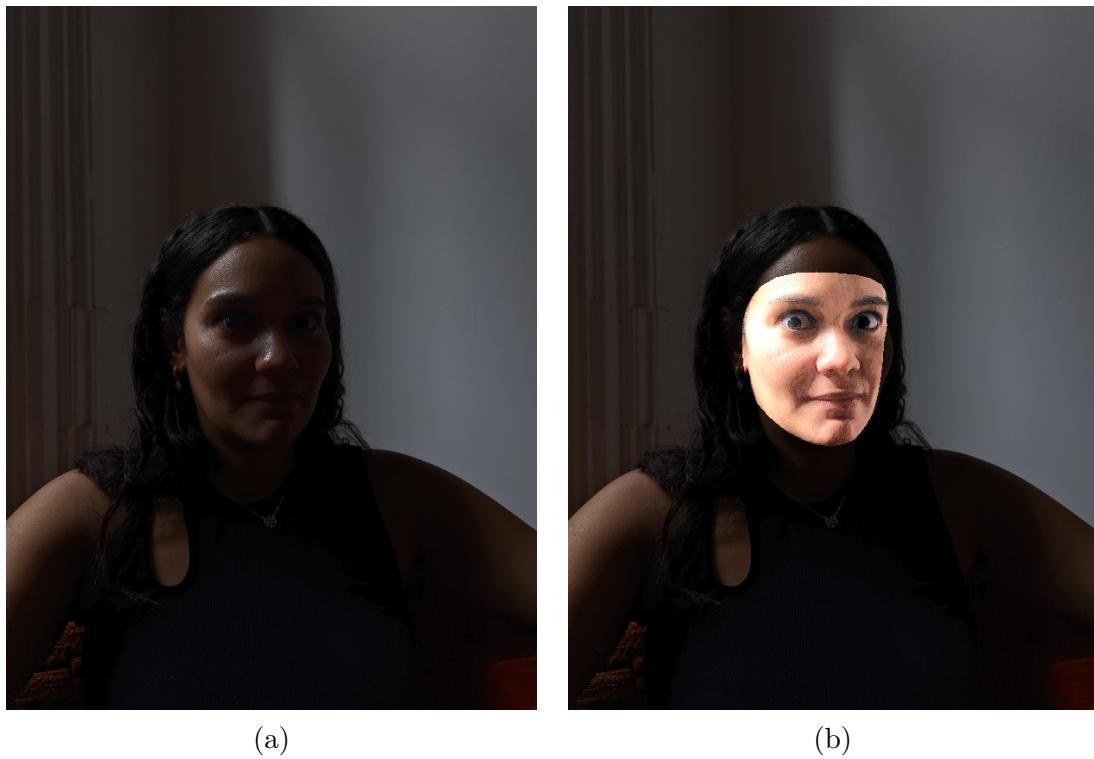
Fonte: Elaborada pelo autor

3 Resultados obtidos

O objetivo desta seção é apresentar os resultados inicialmente obtidos e as considerações levantadas sobre o desenvolvimento do projeto.

A Figura 8 apresenta um exemplo do resultado inicial obtido na aplicação da correção de cor da pele humana, utilizando a imagem central da Figura 7. Nela, embora a face da pessoa estivesse inicialmente coberta por sombras, foi possível detectar a face e aplicar o *color grading*, melhorando significativamente a representação original da cor facial da pessoa.

Figura 8 – Resultados iniciais do projeto. (a) Imagem de entrada. (b) Imagem de entrada com a cor da face corrigida.



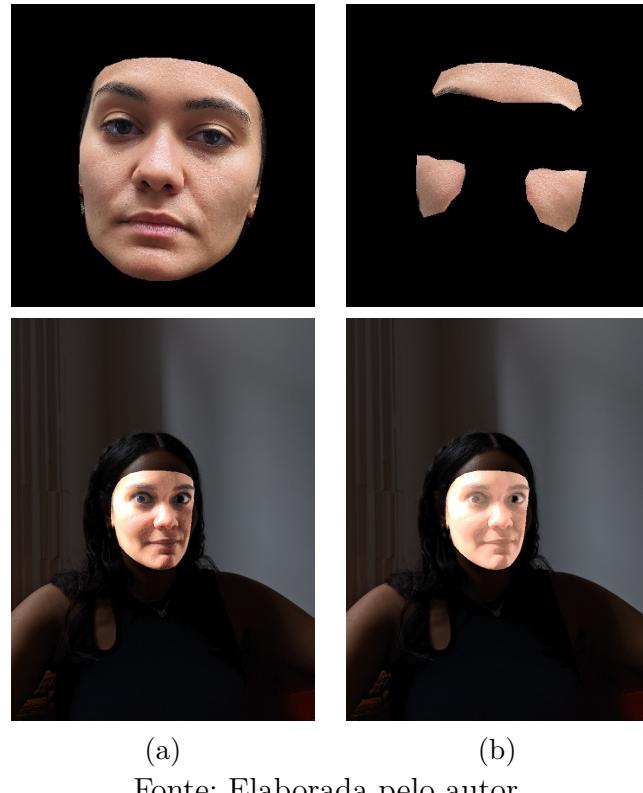
Fonte: Elaborada pelo autor.

É importante destacar que, embora a imagem obtida não apresente um visual natural devido à correção aplicada apenas ao rosto, com o restante da imagem inalterado, o projeto proposto faz parte de um esforço mais amplo que busca classificar tons de pele a partir de imagens capturadas por dispositivos móveis.

Os resultados obtidos durante a aplicação da técnica de *color grading*, descrita na seção 2.2, se mostraram mais precisos quando as imagens utilizadas representavam toda a face do indivíduo segmentada, em vez de apenas as regiões de interesse da face (testa e bochechas), como proposto no diagrama original da Figura 1. Isso fica evidente na Figura

9, onde é possível comparar os dois casos. Observa-se que, ao utilizar apenas as regiões de interesse da face, a imagem ajustada apresenta falhas em áreas como os olhos, lábios e sobrancelhas.

Figura 9 – Resultados do *color grading* utilizando (a) a face completa segmentada e (b) apenas regiões de interesse segmentadas. A linha superior apresenta as imagens de alvo utilizadas. Enquanto a linha inferior apresenta o resultado da operação.



Fonte: Elaborada pelo autor.

O método de *color grading* utilizado, entretanto, produz alguns 'ruídos', como, por exemplo, o desfoque de certas regiões da pele e a geração de artefatos em determinadas áreas do rosto, conforme ilustrado na Figura 10.

Figura 10 – Ilustração dos ruídos obtidos no método aplicado.



Fonte: Elaborada pelo autor.

Para eliminar esses ruídos, é necessário a aplicação de uma filtragem adequada. Para isso, a metodologia apresentada no trabalho de Shirai et al. (2021) oferece técnicas avançadas de filtragem guiada (HE; SUN; TANG, 2012) capazes de reduzir esses tipos de desfoques e artefatos, aprimorando significativamente a qualidade da imagem. Pretende-se, no entanto, considerar outras alternativas de filtros, em especial técnicas que preservam bordas com o Filtro Bilateral (PARIS et al., 2009). O estudo e a aplicação dessas metodologias de filtragem são, portanto, as próximas etapas fundamentais para a conclusão do projeto.

Os códigos desenvolvidos para as aplicações descritas foram escritos na linguagem de programação Python, principalmente, com a utilização da biblioteca de visão computacional OpenCV (2000) e estão disponíveis no Github¹.

4 Próximas atividades

Conforme o cronograma original e levando em conta as discussões apresentadas na seção 3, as próximas atividades a serem realizadas são as seguintes:

1. **Aplicação de uma filtragem:** estudar e implementar técnicas de filtragem que garantam a redução dos ruídos apresentados no resultados atuais.
2. **Implementação final da correção de cor:** desenvolver o programa final, garantindo a comunicação eficaz entre os códigos e as imagens de referência específicas que serão definidas, além de organizar e estruturar o programa de forma definitiva.
3. **Avaliação e testes:** testar o programa final em diversas condições de variabilidade da imagem para todos os tons de pele e inferir a qualidade das implementações feitas.
4. **Desenvolvimento do relatório final:** escrita do relatório final.

e seguirão o seguinte cronograma, visto na Tabela 2:

Atividades	Meses					
	Julho	Agosto	Setembro	Outubro	Novembro	Dezembro
Aplicação de uma filtragem	•	•	•			
Implementação final da correção de cor			•	•		
Avaliação e testes				•	•	
Desenvolvimento do relatório final					•	•

Tabela 2 – Cronograma de atividades para os próximos meses de trabalho.

¹ <https://github.com/rodrigorcz/skin-color-correction>

Referências

- GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*. 4. ed. Nova Iorque: Pearson, 2018. 761-826 p.
- HE, K.; SUN, J.; TANG, X. Guided image filtering. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 35, n. 6, p. 1397–1409, 2012.
- LUGARESI, C. et al. Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172*, 2019.
- MONK, E. The monk skin tone scale. SocArXiv, 2023.
- OPENCV. 2000. Disponível em: <<https://opencv.org>>. Acesso em: 20 feb 2021.
- PARIS, S. et al. Bilateral filtering: Theory and applications. *Foundations and Trends® in Computer Graphics and Vision*, Now Publishers, Inc., v. 4, n. 1, p. 1–73, 2009.
- PITIÉ, F.; KOKARAM, A. C.; DAHYOT, R. Automated colour grading using colour distribution transfer. *Computer Vision and Image Understanding*, v. 107, n. 1, p. 123–137, 2007. ISSN 1077-3142. Special issue on color image processing.
- REINHARD, E. et al. Color transfer between images. *IEEE Computer Graphics and Applications*, v. 21, n. 5, p. 34–41, 2001.
- SCHUMANN, C. et al. Consensus and subjectivity of skin tone annotation for ml fairness. *arXiv preprint arXiv:2305.09073*, 2023.
- SHIRAI, K. et al. Guided facial skin color correction. *Signals*, MDPI, v. 2, n. 3, p. 540–558, 2021. Disponível em: <<https://doi.org/10.3390/signals2030033>>.