



Universidade Federal de Pernambuco - UFPE

Centro de Informática – CIn

Graduação em Engenharia da Computação

DESCOLORIZAÇÃO DE IMAGENS BASEADA EM PIRÂMIDES LAPLACIANAS

Matheus de Farias Cavalcanti Santos

TRABALHO DE GRADUAÇÃO

Recife, Dezembro de 2017



Universidade Federal de Pernambuco - UFPE

Centro de Informática - CIn

Matheus de Farias Cavalcanti Santos

DESCOLORIZAÇÃO DE IMAGENS BASEADA EM PIRÂMIDES LAPLACIANAS

Monografia apresentada ao Centro de Informática (CIn) da Universidade Federal de Pernambuco (UFPE), como requisito parcial para conclusão do Curso de Engenharia da Computação, orientada pelo professor **Carlos Alexandre Barros de Mello**.

Recife, Dezembro de 2017
Universidade Federal de Pernambuco - UFPE
Centro de Informática - CIn
Curso de Bacharelado em Engenharia da Computação

Matheus de Farias Cavalcanti Santos

DESCOLORIZAÇÃO DE IMAGENS BASEADA EM PIRÂMIDES LAPLACIANAS

Monografia submetida ao corpo docente da Universidade Federal de Pernambuco (UFPE), defendida e aprovada em 05 de Dezembro de 2017.

Banca Examinadora:

Carlos Alexandre Barros de Mello

Doutor

Orientador

Tsang Ing Ren

Doutor

Examinador

Dedico este trabalho à minha família, que muito me apoiou e me incentivou à realizá-lo, ao meu orientador por toda dedicação, suporte e paciência e aos meus familiares e amigos pela torcida.

AGRADECIMENTOS

Agradeço primeiramente à Deus por me iluminar e me prover forças para superar as adversidades da vida e continuar seguindo rumo aos meus objetivos.

Agradeço à minha mãe Sandra de Farias Cavalcanti e à minha irmã Thaisa de Farias Cavalcanti Santos por sempre estarem ao meu lado me apoiando, me estimulando positivamente e torcendo por mim e ao meu pai Geraldo Santos pelos seus ensinamentos e que, com toda certeza, sempre me guia de onde estiver.

Agradeço ao meu orientador, Professor Carlos Alexandre Barros de Mello, pela sua dedicação, paciência e orientação comigo, além de todo o conhecimento compartilhado.

Agradeço aos meus familiares pela torcida eterna, aos amigos que sempre me apoiaram e me ajudaram e à todos que de alguma forma, contribuíram para que este sonho se tornasse realidade.

“A visão é o mais avançado dos nossos sentidos, de forma que não é de surpreender que as imagens exerçam o papel mais importante na percepção humana.”

Rafael C. Gonzalez

RESUMO

Neste trabalho de graduação é apresentada uma estratégia de descolorização de imagens que se baseia no resultado da utilização do operador Laplaciano. Descolorização é um processo que converte uma imagem representada por três matrizes (canais) de cor (uma matriz para cada componente, no caso do sistema RGB) para uma imagem representada por apenas uma matriz, ou seja, uma imagem em escala de tons de cinza. Esse processo é muito importante, pois, é amplamente utilizado em aplicações de processamento de imagens monocromáticas. Contudo, a descolorização pode resultar em perda de detalhamento em regiões de fronteira (regiões de alto contraste entre as cores) e pode resultar também em perda da discriminabilidade entre as cores da imagem original na imagem em escala de tons de cinza após a descolorização.

Além da distribuição espacial de cores, a informação local desempenha um papel importante na manutenção da discriminabilidade das cores da imagem descolorizada. A estratégia utilizada neste trabalho utiliza os três canais de cores Vermelho, Verde e Azul (*Red, Green, Blue – R,G,B*) como entradas, guiados por dois mapas de pesos que filtram as transições locais e medem os valores dominantes das regiões utilizando o resultado da aplicação do operador Laplaciano. Com o objetivo de reduzir os ruídos na imagem introduzidos pelos mapas de pesos, o princípio de fusão de imagens é utilizado no final do processo e a imagem descolorizada é obtida como saída.

Palavras-chave: Processamento digital de imagens, Conversão para tons de cinza, Operador Laplaciano, Espaço de cor RGB, Fusão Multiescala.

ABSTRACT

In this undergraduate work is presented a strategy for image decolorization that is based on the result from the Laplacian operator. Decolorization is a process that converts an image represented by three color matrices (channels), a single matrix for each component, in the case of the RGB system, to an image represented by only one matrix, i.e. a grayscale image. This process is very important, due to its use in monochrome image processing applications. However, decolorization can result in loss of detail in the edge regions (regions of high contrast between colors) and may also result in loss of discrimination between the colors of the original image in grayscale image after decolorization process.

In addition to the spatial color distribution, the local information plays an important role in maintaining discrimination of the decolorized image. The strategy used in this work use the three color channels Red, Green and Blue (R, G, B) as inputs, guided by two weight maps that filter the local transitions and measure the dominant values of the regions using the results from the Laplacian operator. In order to reduce image noise introduced by weight maps, the principle of image fusion is used at the end of the process and a decolorized image is obtained as output.

Keywords: Digital image processing, Color to grayscale conversion, Laplacian Operator, RGB Color space, Multi-scale fusion.

Sumário

1	Introdução.....	12
1.1	Motivação.....	12
1.2	Objetivos	15
1.3	Estrutura do Trabalho	16
2	Métodos para representação de uma imagem colorida em níveis de brilho	17
2.1	Representação da imagem em tons de cinza por um dos Canais de Cor	18
2.2	Representação da imagem em tons de cinza pelo Máximo	18
2.3	Representação da imagem em tons de cinza pela Média.....	19
2.4	Representação de imagem em tons de cinza pela Luminância	19
2.5	Representação de imagem em tons de cinza via Contraste Percebido	19
3	Algoritmo em Estudo.....	22
3.1	Visão Geral	22
3.2	1ª Etapa: Aplicação do Mapa de Pesos Laplaciano e de Contraste Global	24
3.3	2ª Etapa: Normalização dos Mapas de Pesos	25
3.4	3ª Etapa: Construção da Pirâmide Gaussiana e da Pirâmide Laplaciana	26
3.5	4ª Etapa: Fusão Multiescala das Pirâmides Gaussiana e Laplaciana	29
4	Experimentos e Análises	31
5	Conclusões e Trabalhos Futuros	34
5.1	Contribuições	34
5.2	Limitações	34
5.3	Trabalhos Futuros.....	35
	Referências Bibliográficas.....	36

Lista de Figuras

Figura 1 - Descolorização de uma imagem. A imagem em (a) é representada por três canais de cores e a imagem em (b) é representada por apenas um canal de cor, resultante da descolorização. Fonte: [2].	12
Figura 2 - Em (a) é mostrada a imagem original, em (b) é mostrada a imagem resultante da descolorização manipulando-se o canal de luminância e em (c) é mostrada a imagem corretamente descolorizada. Fonte: [5].	13
Figura 3 - Na primeira imagem observa-se uma imagem em escala de tons de cinza sem preservar corretamente as arestas. Na outras três imagens tem-se os canais de cor vermelho, verde e azul respectivamente em que pelo menos em um dos três não há omissão das arestas. Fonte: [1].	14
Figura 4 - A figura mostra a aplicação dos mapas de peso Laplaciano e dos mapas de peso de contraste global e seus resultados. Fonte: [4].	15
Figura 5 - Imagem digital e os três canais de cor que a constitui.	17
Figura 6 - Imagem digital em tons de cinza (imagem monocromática).	17
Figura 7 - Imagem digital colorida utilizada como exemplo. Fonte:[8].	18
Figura 8 - Imagens resultantes dos métodos de representação de imagens em tons de cinza descritos.	21
Figura 9 - Imagem de exemplo. Fonte: [4].	22
Figura 10	23
Figura 11 - Resultados obtidos na 1ª Etapa do algoritmo de descolorização.	25
Figura 12 - Resultados obtidos na 2ª Etapa do algoritmo de descolorização.	26
Figura 13 - Representação Multiescala de uma imagem (Pirâmide). Fonte: [10].	27
Figura 14 - Construção da Pirâmide Laplaciana. Fonte: [11].	28
Figura 15 - Pirâmides Gaussianas para cada canal de cor.	29
Figura 16 - Pirâmides Laplacianas para cada canal de cor.	29
Figura 17 - Pirâmide Final (a) e a saída do algoritmo (b).	30
Figura 18 – Comparação das imagens em tons de cinza obtidas pelos métodos de descolorização abordados neste trabalho.	33

Lista de Tabelas

Tabela 1 - Média e desvio padrão do CCPR, avaliando as imagens geradas por métodos clássicos e geradas pelo algoritmo [4] tendo como base as imagens geradas pela média dos três canais de cor.....	32
Tabela 2 - Média e desvio padrão do SSIM, comparando as imagens geradas por métodos clássicos com as imagens geradas pelo algoritmo [4].....	32
Tabela 3 - Tabela 3. Análise de variância aplicada aos resultados de comparação gerados pelo SSIM.	33

1 Introdução

Nesta seção é apresentada a motivação para o desenvolvimento deste trabalho de graduação, os objetivos propostos e por fim é apresentada a estrutura do trabalho.

1.1 Motivação

O processo de descolorização, exemplificado na Figura 1, converte uma imagem representada por três matrizes de cores (uma matriz para cada canal de cor, no caso do sistema RGB) em uma imagem em escala de tons de cinza (uma matriz de cor apenas) [1]. Muitas aplicações de processamento de imagens utilizam imagens com um único canal de cor como base para realizar, por exemplo, tarefas de reconhecimento/detecção de objetos. Com isso a perda de informação que pode ocorrer devido ao processo de descolorização tem que ser minimizada.

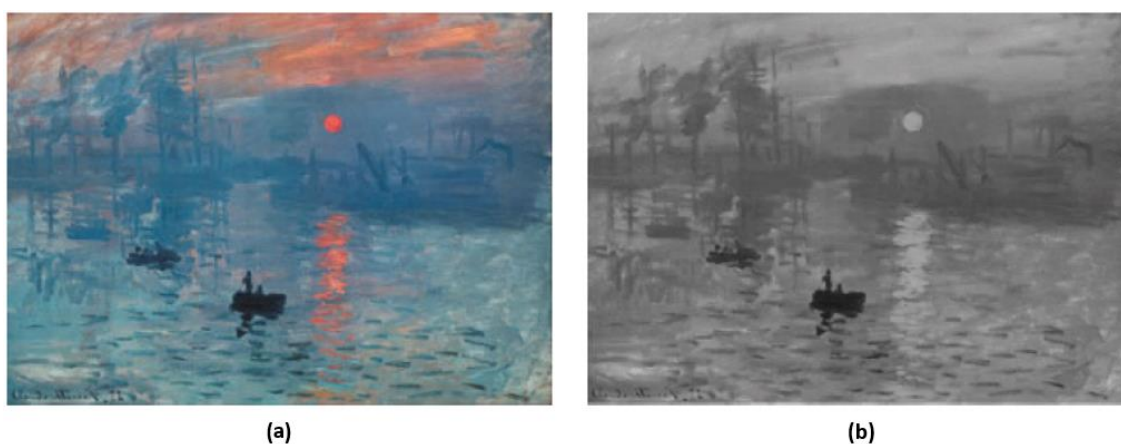


Figura 1 - Descolorização de uma imagem. A imagem em (a) é representada por três canais de cores e a imagem em (b) é representada por apenas um canal de cor, resultante da descolorização. Fonte: [2].

A forma mais comum de se realizar o processo de descolorização é manipulando o canal de luminância da imagem para os diferentes espaços de cores ($CIEL^*a^*b^*$, $YCbCr$, HSL/HSV) [3], obtendo-se como resultado a imagem em escala de tons de cinza (representada por um canal de cor). Contudo, esse mapeamento global realizado no canal de luminância não leva em consideração importantes informações cromáticas, podendo produzir imagens em escala de tons de cinza, cujos contrastes existentes na imagem original sejam perdidos. Além disso, importantes detalhes e transições podem acabar sendo omitidos [4][5]. Essas limitações existentes no processo de descolorização manipulando-se o canal de luminância podem ser observadas na Figura 2.

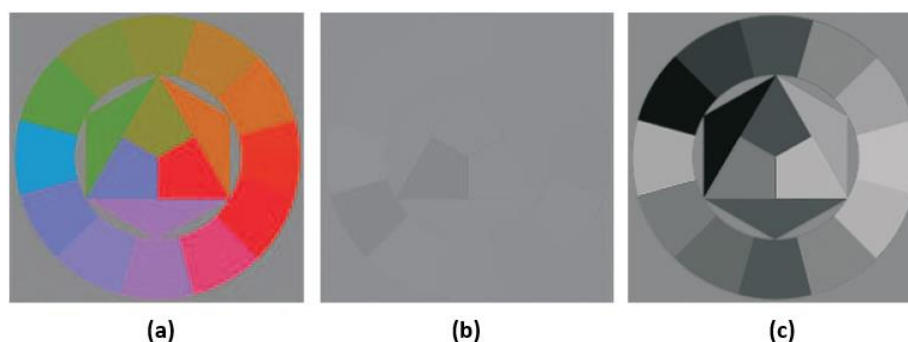


Figura 2 - Em (a) é mostrada a imagem original, em (b) é mostrada a imagem resultante da descolorização manipulando-se o canal de luminância e em (c) é mostrada a imagem corretamente descolorizada. *Fonte: [5].*

Pode-se observar que os detalhes e transições presentes na Figura 2a, foram omitidos na Figura 2b. Além disso, a maior parte das cores na imagem original (Figura 2a) foi mapeada com o mesmo tom de cinza na Figura 2b após a descolorização. Esses são os problemas que podem ocorrer no processo de descolorização e que devem ser contornados.

Portanto, objetivando contornar as limitações citadas acima que podem ocorrer no processo de descolorização de imagens, uma estratégia de mapeamento local em escala de tons de cinza, o qual filtra as características e transições tendo como base a informação derivada do operador Laplaciano como filtro local é proposta em [4]. O operador Laplaciano [1] foi utilizado como filtro local, pois, ele analisa a vizinhança do pixel central ao qual será atribuído um valor de intensidade de tom de cinza. Para pixels vizinhos com intensidades constantes ou com baixo contraste o valor da intensidade atribuído é o correspondente ao preto. Caso os pixels vizinhos possuam alto contraste entre cada um deles e o pixel central, um valor de intensidade de tom de cinza será atribuído ao pixel central e esse mapeamento local irá destacar na imagem como um todo as regiões de fronteiras ou arestas [1].

O processo de descolorização utilizado neste trabalho foi projetado no espaço de cor RGB e utiliza como entradas individuais os três canais de cor (RGB), cada um sendo uma entrada. Como pode ser observado na Figura 3, apesar das arestas não estarem bem definidas no canal de luminância L^* (devido à estratégia de mapeamento linear utilizado), elas estão bem definidas em pelo menos um dos três canais de cor RGB [1].



Figura 3 - Na primeira imagem observa-se uma imagem em escala de tons de cinza sem preservar corretamente as arestas. Na outras três imagens tem-se os canais de cor vermelho, verde e azul respectivamente em que pelo menos em um dos três não há omissão das arestas. *Fonte: [1].*

Como o objetivo é preservar a maior parte das arestas e dos detalhes da imagem original na imagem em escala de tons de cinza, o algoritmo utilizado neste trabalho é guiado por dois mapas de pesos. O primeiro mapa de pesos, o *Laplacian Weight*, transfere para a imagem de saída as informações mais significativas (contrastes locais e transições locais) de cada canal de cor que são utilizados como entradas individuais e isso é possível devido à informação resultante do operador Laplaciano. O segundo mapa de pesos, o *Global Contrast Weight*, foi projetado para enfatizar não só os valores mais significativos de cada canal de cor, mas também as regiões caracterizadas por altos contrastes e com isso mapeando duas cores contrastantes da imagem original com seus respectivos e corretos valores de cor em escala de tons de cinza na imagem descolorizada de saída [2][4].

Como pode-se observar na Figura 4, a aplicação dos mapas de peso Laplaciano, resulta em imagens dos canais de cor RGB, cujas transições e detalhes são preservados. Já a aplicação dos mapas de peso de contraste global, resulta em imagens dos canais de cor RGB, os quais preservam os contrastes das cores da imagem original na imagem em escala de tons de cinza, porém, as transições e detalhes são, de certa forma, omitidos. A ideia da técnica utilizada neste trabalho é unir o melhor dos dois filtros para obter uma imagem em escala de tons de cinza corretamente representada como na imagem da Figura 4, a qual mostra o resultado do método proposto em [4].

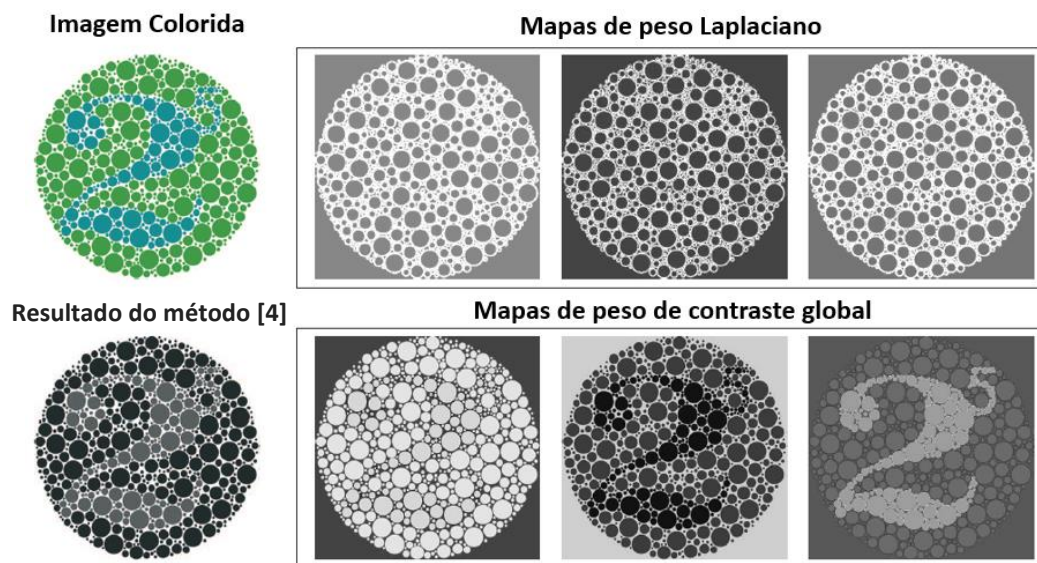


Figura 4 - A figura mostra a aplicação dos mapas de peso Laplaciano e dos mapas de peso de contraste global e seus resultados. *Fonte: [4].*

Visando minimizar a inclusão de ruídos devido à utilização dos mapas de pesos local e global, a abordagem utilizada neste trabalho é projetada na utilização da Pirâmide Laplaciana [6] das entradas combinada com a Pirâmide Gaussiana dos pesos normalizados [7]. Após a combinação da Pirâmide Laplaciana com a Pirâmide Gaussiana, combina-se as camadas da pirâmide final obtendo-se a imagem final descolorizada, preservando as arestas e detalhes da imagem original, e com as cores contrastantes da imagem original mapeadas corretamente em cores na escala de tons de cinza.

1.2 Objetivos

O presente trabalho possui como objetivo geral explicar o funcionamento das técnicas de descolorização de imagens já existentes e as mais utilizadas, e avaliar a pertinência da utilização da técnica de descolorização de imagens abordada neste trabalho.

Como objetivos específicos, tem-se:

- Implementar o algoritmo proposto em [4] baseando-se nos resultados positivos que o mapeamento local e global produzem. Após isso executar a fusão desses mapeamentos;
- Como saída do algoritmo proposto em [4], produzir imagens em tons de cinza com detalhamento próximo ou igual da imagem colorida e tons de cinza coerentes com as cores da imagem colorida;
- Comparação com outras técnicas;
- Avaliação da eficácia e desempenho da técnica proposta.

1.3 Estrutura do Trabalho

Este trabalho está estruturado em cinco capítulos dos quais, este corresponde ao capítulo da introdução. Os demais capítulos estão descritos a seguir.

Capítulo 2 - Métodos para representação de uma imagem colorida em níveis de brilho – apresenta os métodos mais conhecidos e utilizados para converter uma imagem para tons de cinza.

Capítulo 3 – Algoritmo em Estudo - descreve o funcionamento da técnica de descolorização de imagens proposta em [4] implementada neste trabalho.

Capítulo 4 - Experimentos e Análises – apresenta os experimentos feitos e as análises realizadas nos resultados obtidos.

Capítulo 5 – Conclusões e Trabalhos Futuros – apresenta as conclusões obtidas a partir das análises dos resultados após a realização dos experimentos e sugere trabalhos futuros que podem ser realizados tendo como base este trabalho.

2 Métodos para representação de uma imagem colorida em níveis de brilho

As imagens digitais, no modelo de cor RGB, são constituídas por três canais de cor: Vermelho, Verde, Azul (*Red, Green, Blue - R,G,B*), como ilustrado na Figura 5. Estas imagens são chamadas de cromáticas, pois são representadas por diversas cores e cada uma destas cores possuem diversos tons.

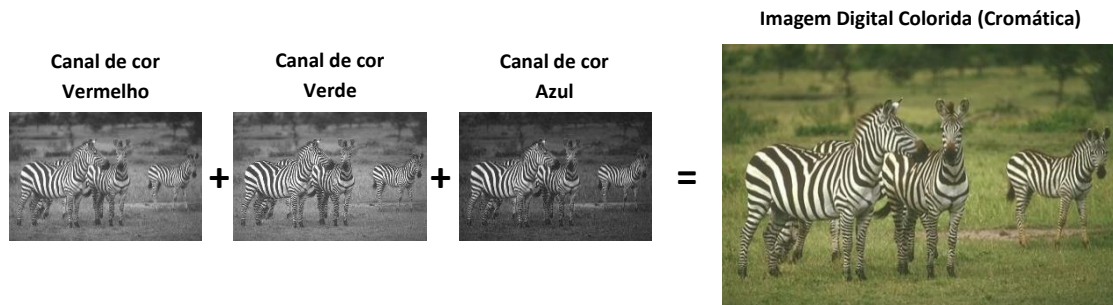


Figura 5 - Imagem digital e os três canais de cor que a constitui.

Em contrapartida, as imagens monocromáticas são representadas por uma única cor, a qual possui diversos tons. Por exemplo, as imagens em tons de cinza (ou escalas de cinza), como mostrado na Figura 6, são imagens representadas pela cor cinza e seus diversos tons. Estes tons de cinza variam do tom mais claro, o branco, até o tom mais escuro, o preto. Em imagens, cuja cor dos pixels é representada por 8 bits, totalizando 256 tons de cinza possíveis para cada pixel na representação da imagem, a cor preta equivale à intensidade 0, a cor branca equivale à intensidade 255 e as intensidades entre 0 e 255 equivalem aos tons de cinza intermediários.



Figura 6 - Imagem digital em tons de cinza (imagem monocromática).

Existem diversos métodos utilizados para representar imagens coloridas (três matrizes de cor - RGB) em imagens em tons de cinza (uma única matriz de cor), alguns dos quais são

descritos, a seguir. A imagem da Figura 7 é utilizada como exemplo para ilustrar cada um desses métodos.



Figura 7 - Imagem digital colorida utilizada como exemplo. *Fonte:[8].*

2.1 Representação da imagem em tons de cinza por um dos Canais de Cor

Neste método de representação, a extração dos canais de cor no modelo RGB da imagem colorida é realizada e apenas um desses canais é utilizado como a representação da imagem colorida em tons de cinza. Esta representação está definida na expressão:

$$R, G, B = \text{extração}(\text{ImagemColorida})$$

$$C = R \text{ ou } C = G \text{ ou } C = B \quad (1)$$

onde C é a imagem em tons de cinza.

2.2 Representação da imagem em tons de cinza pelo Máximo

Neste método de representação, é realizada a extração dos canais de cor no modelo RGB da imagem colorida. Em seguida, o máximo valor do pixel na mesma posição, entre os três canais de cor, é atribuído ao pixel de mesma posição na imagem em tons de cinza. Esta representação está definida na expressão:

$$R, G, B = \text{extração}(\text{ImagemColorida})$$

$$C = \text{máximo}(R, G, B) \quad (2)$$

onde C é a imagem em tons de cinza.

2.3 Representação da imagem em tons de cinza pela Média

Neste método de representação, é realizada a extração dos canais de cor no modelo RGB da imagem colorida e em seguida, a média do valor do pixel na mesma posição, entre os três canais de cor, é atribuído ao pixel de mesma posição na imagem em tons de cinza. Esta representação está definida na expressão:

$$R, G, B = \text{extração}(\text{ImagemColorida})$$
$$C = \text{média}(R, G, B) \quad (3)$$

onde C é a imagem em tons de cinza.

2.4 Representação de imagem em tons de cinza pela Luminância

Neste método de representação, a imagem em tons de cinza é obtida através da soma das multiplicações das intensidades dos pixels de mesma posição por pesos. Esta representação, para uma imagem no modelo RGB, está definida na expressão:

$$C = (0,3 * R + 0,59 * G + 0,11 * B) \quad (4)$$

onde C é a imagem em tons de cinza.

2.5 Representação de imagem em tons de cinza via Contraste Percebido

Neste método de representação proposto em [9], a imagem em tons de cinza é obtida via contraste percebido, o qual, utiliza a saliência de canal (valor de distinção que mede a força percebida de cada canal) para selecionar o canal de cor da imagem que será utilizado no processo de descolorização. O espaço de cor utilizado é o IHSL (*Improved Hue, Saturation, Luminance*). A saliência de canal está definida na expressão:

$$CS(ch) = CS_D(ch) * CS_B(ch) \quad (5)$$

em que ch representa um dos três canais de cor do espaço de cor IHLS, o primeiro termo, CS_D , é o termo de distinção, que mede a distinção entre duas entradas, o segundo termo, CS_B , é o termo de equilíbrio, que mede a uniformidade quantitativa entre duas entradas.

A informação saliente é recebe duas entradas, e os pixels podem ser agrupados em dois clusters. Para isso, um Modelo de Mistura Gaussiana (*GMM - Gaussian Mixture Module*) de

dois centros é aplicado resultando em um modelo de probabilidade de distribuição de pixels do canal.

A conversão da imagem em tons de cinza se dá através de ajustes globais do canal de luminância L da imagem. É construído o mapa de contrastes M para realizar os ajustes em L e com isso obter a imagem em tons de cinza. Esta representação está definida na expressão:

$$C = L + \alpha M \quad (6)$$

onde C é a imagem em tons de cinza e α é o grau de aprimoramento da imagem. O contraste de cor saliente, o qual é obtido a partir do contraste percebido ΔM , é armazenado no mapa de contraste M .

Há dois tipos de contraste percebido: o contraste percebido sem ajuste espacial e o contraste percebido com ajuste espacial.

No contraste percebido sem ajuste espacial, é calculada a saliência de canal para os três canais de cor do espaço de cor IHSL, e após este cálculo é escolhido o canal de cor com maior valor, ou seja, o canal saliente ch^s . Este canal saliente possui o contraste mais importante que é percebido da imagem original.

No contraste percebido com ajuste espacial, primeiramente é realizada a elaboração de tom, na qual, é obtida uma amostragem não-uniforme, para selecionar o mesmo número de pontos de cada cluster do GMM de acordo com a distribuição Gaussiana do canal saliente ch^s encontrado. Após a realização da elaboração de tom, é feita a propagação do tom elaborado para todos os pixels levando em consideração a informação da cor como a informação espacial. Por fim, é realizada a aproximação da solução para o contraste percebido ΔM .

Utilizando a Figura 7 como exemplo para ilustrar os métodos de representação de imagens em tons de cinza descritos acima, os resultados desses métodos estão ilustrados na Figura 8.



Imagem colorida



**Representação por
um dos Canais de Cor**



**Representação pelo
Máximo**



**Representação pela
Média**



**Representação pela
Luminância**



**Representação pelo
Contraste Percebido**

Figura 8 - Imagens resultantes dos métodos de representação de imagens em tons de cinza descritos.

3 Algoritmo em Estudo

3.1 Visão Geral

O algoritmo proposto em [4] recebe como entrada os canais de cor Vermelho, Verde e Azul individuais, considerando uma imagem no modelo de cores RGB. Além disso, o algoritmo é guiado por dois mapas de pesos no processo de descolorização de uma imagem: o primeiro mapa de pesos é denominado Mapa de Pesos Laplaciano e o segundo é denominado Mapa de Pesos de Contraste Global. Por fim, é realizada a Fusão Multiescala com o objetivo de reduzir os ruídos inseridos na imagem após a aplicação dos Mapas de pesos.

O processo de descolorização executado pelo algoritmo está dividido em quatro etapas: aplicação do Mapa de Pesos Laplaciano e do Mapa de Pesos de Contraste Global, normalização dos mapas de pesos, construção da Pirâmide Gaussiana e da Pirâmide Laplaciana, execução da Fusão Multiescala das pirâmides obtendo a imagem final descolorizada. A imagem de exemplo da Figura 9 é utilizada para ilustrar o que cada etapa do algoritmo produz como resultado. Todas essas etapas estão ilustradas no fluxograma da Figura 10 e são descritas com mais detalhes a seguir.

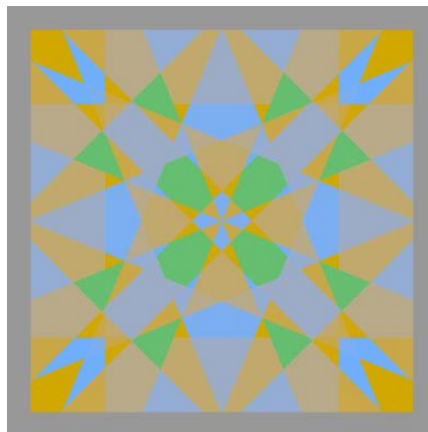


Figura 9 - Imagem de exemplo. *Fonte: [4].*

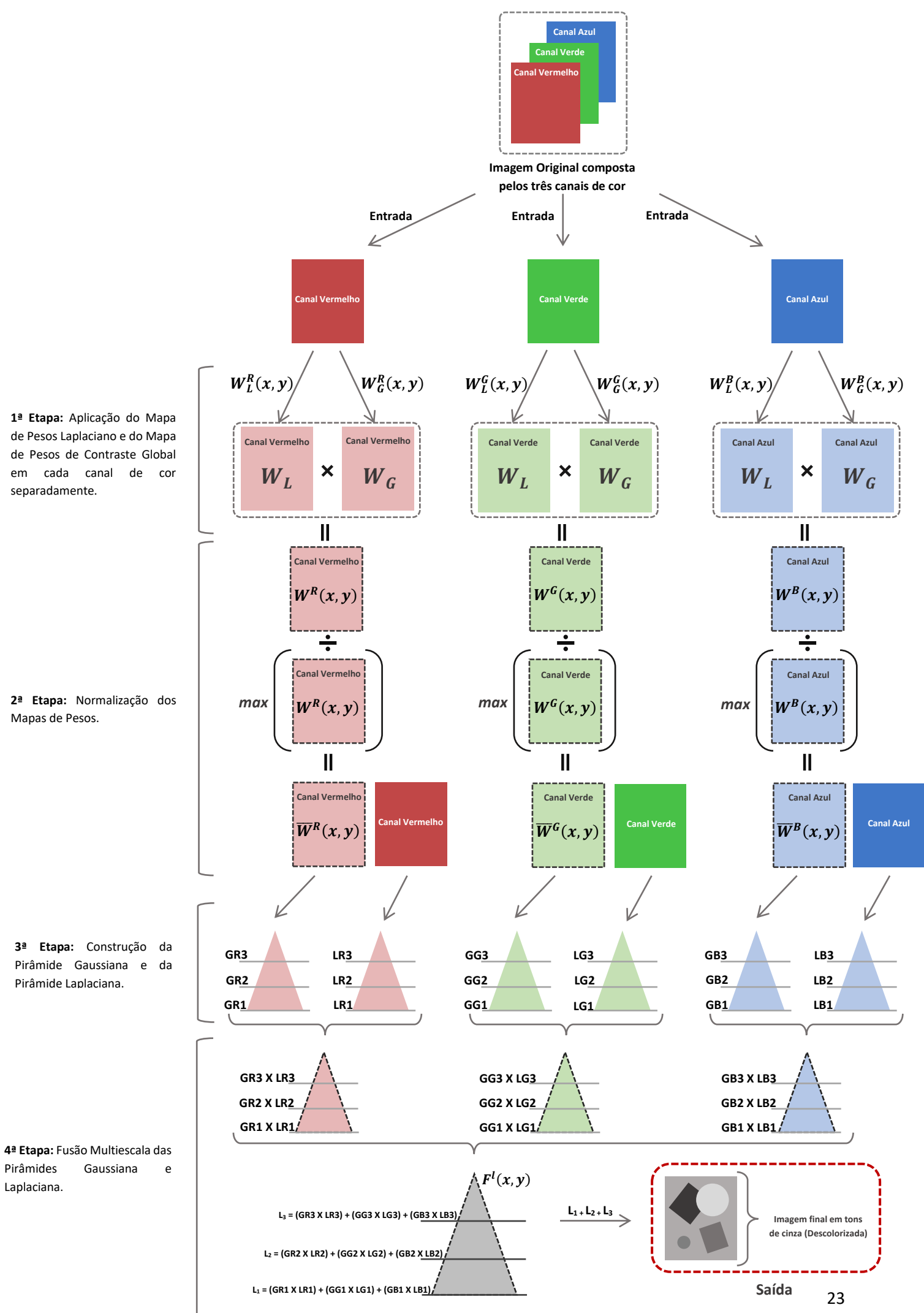


Figura 10 - Fluxograma das etapas do algoritmo de descolorização.

3.2 1ª Etapa: Aplicação do Mapa de Pesos Laplaciano e de Contraste Global

O **Mapa de Pesos Laplaciano** utiliza como base o Operador Laplaciano [1], $L \{I(x, y)\} = \nabla^2 I(x, y)$, o qual sinaliza contrastes locais em uma imagem. Utilizando a Informação extraída da aplicação do operador Laplaciano, o Mapa de Pesos Laplaciano transfere para a imagem filtrada resultante as informações mais significativas (contrastos locais e transições locais) de cada canal de cor que são utilizados como entradas individuais. O Mapa de Pesos Laplaciano é definido pela equação:

$$W_L^k(x, y) = \overline{L \{I^k(x, y)\}} + |L \{I^k(x, y)\}| \quad (7)$$

em que, $W_L^k(x, y)$ representa o Mapa de Pesos Laplaciano, k é o índice que representa os canais de cor individuais que são as entradas do algoritmo variando seu valor de 1 a 3, dado que são três canais de cor: R, G, B. O primeiro termo calcula a média aritmética do Laplaciano de cada canal de cor e esse cálculo representa a quantidade de transições existentes na imagem. O segundo termo calcula o valor absoluto do Laplaciano de cada canal de cor com o objetivo de informar os locais na imagem em que há transições e evitar perda de informação, pois, ao invés de normalizar um possível valor negativo para zero, é utilizado o seu valor positivo, o qual possui informação.

O **Mapa de Pesos de Contraste Global** foi projetado para enfatizar não só os valores mais significativos de cada canal de cor, mas também as regiões caracterizadas por altos contrastes, com isso, mapeando duas cores contrastantes da imagem original com seus respectivos valores de cor em escala de tons de cinza na imagem filtrada de saída. O Mapa de pesos de Contraste Global é definido pela equação:

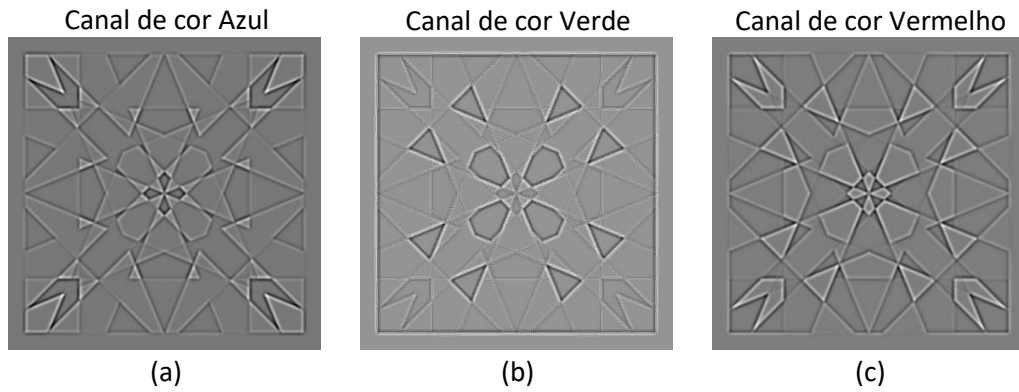
$$W_G^k(x, y) = \left[I^k(x, y) - \overline{L \{I^k(x, y)\}} \right]^\delta \quad (8)$$

em que, o Mapa de Pesos de Contraste Global de cada canal, $W_G^k(x, y)$, é a diferença entre o canal de cor e a média aritmética do Laplaciano deste canal de cor. O índice δ possui valor 2 como padrão, de acordo com o algoritmo original.

Após realizar a extração dos três canais de cor da imagem original, os quais são as entradas do processo de descolorização, são aplicados os Mapas de Pesos Laplaciano e de Contraste Global em cada canal de cor separadamente. Os Mapas de Pesos Laplaciano e de Contraste Global são denotados por $W_L^k(x, y)$ e $W_G^k(x, y)$ respectivamente, em que k representa o canal de cor RGB ($k = [1, 3]$) ao qual o Mapa de Pesos Laplaciano e de Contraste Global está sendo aplicado. A filtragem utilizando o Mapa de Pesos Laplaciano resulta em uma nova imagem, denotada por W_L , para cada canal de cor assim como a filtragem utilizando o Mapa de Pesos de Contraste Global resulta também em uma nova imagem, denotada por W_G , para cada canal de cor. Portanto, para cada canal de cor, duas imagens são geradas, uma resultante do Mapa de Pesos Laplaciano e uma resultante do Mapa de Pesos de Contraste Global, totalizando em seis imagens, como ilustrado na Figura 10, 1ª Etapa. Como resultado,

esta etapa produz as imagens ilustradas na Figura 11 para cada canal de cor. As Figuras 11a, 11b e 11c ilustram os resultados obtidos após a aplicação do Mapa de Pesos Laplaciano em cada canal de cor. As Figuras 11d, 11e e 11f ilustram os resultados obtidos após a aplicação do Mapa de Pesos de Contraste Global em cada canal de cor.

Mapa de Pesos Laplaciano



Mapa de Pesos de Contraste Global

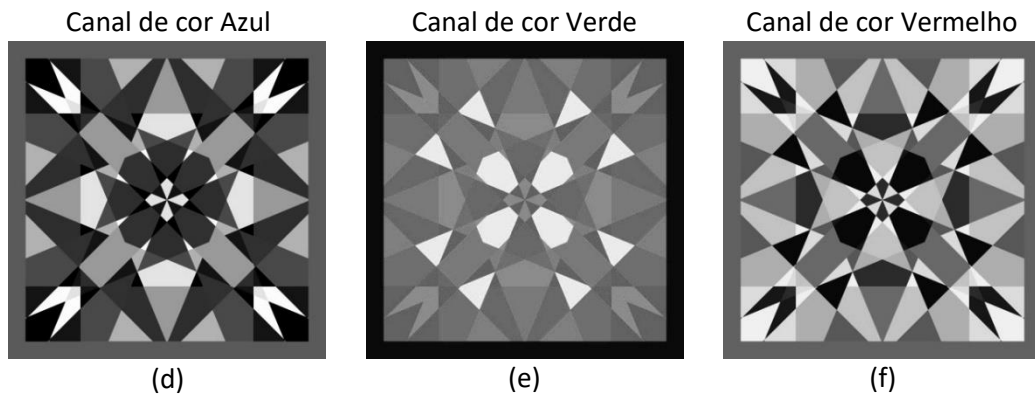


Figura 11 - Resultados obtidos na 1ª Etapa do algoritmo de descolorização.

3.3 2ª Etapa: Normalização dos Mapas de Pesos

Após a execução da 1ª Etapa, são geradas, para cada canal de cor, as matrizes W_L e W_G , as quais são resultados dos processos de filtragem utilizando os Mapas de Pesos Laplaciano e de Contraste Global, respectivamente. Neste momento, para cada canal de cor, as matrizes W_L e W_G são multiplicadas, isto é, um elemento com coordenada (x, y) da matriz W_L é multiplicado por um elemento, de mesma coordenada (x, y) , na matriz W_G . Essa multiplicação resulta em um Mapa de Pesos representado por uma matriz e denotado por $W^k(x, y)$, em que k representa o canal de cor RGB ($k = [1, 3]$). Para realizar a normalização do Mapa de Pesos,

divide-se a matriz $W^k(x, y)$ pelo maior valor de um dado elemento desta mesma matriz. Esta normalização é definida pela expressão a seguir:

$$\bar{W}^k(x, y) = W^k(x, y) / \max(W^k(x, y)) \quad (9)$$

em que, $\bar{W}^k(x, y)$ é o Mapa de Pesos normalizado para cada canal de cor e k representa o canal de cor RGB ($k = [1, 3]$). Esta etapa está ilustrada na Figura 10, 2ª Etapa. Como resultado, esta etapa produz as imagens ilustradas nas Figuras 12a, 12b e 12c, as quais representam o Mapa de Pesos Normalizado para cada canal de cor.

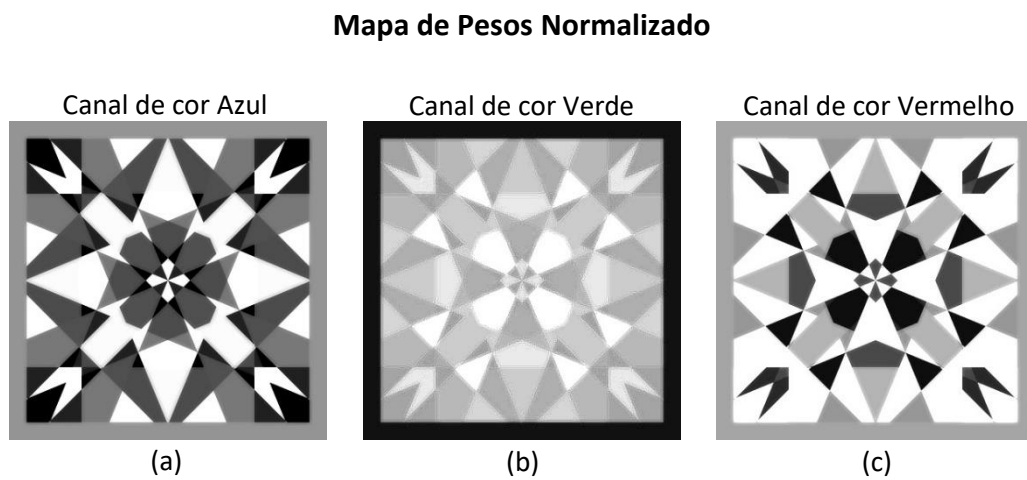


Figura 12 - Resultados obtidos na 2ª Etapa do algoritmo de descolorização.

3.4 3ª Etapa: Construção da Pirâmide Gaussiana e da Pirâmide Laplaciana

Nesta etapa, são construídas as pirâmides Gaussiana e Laplaciana. O número de níveis, que é igual para as duas pirâmides, é determinado pela seguinte expressão:

$$N = \log_{10}(H) \quad (10)$$

em que, N representa o número de níveis das pirâmides e H representa a altura da imagem original.

Definido o número de níveis das pirâmides, estas são representadas, cada uma, por uma lista de imagens, em que cada imagem representa um nível da pirâmide e este número de imagens é determinado pelo número de níveis da pirâmide (N) calculado anteriormente. A primeira imagem dessas listas são as imagens sem modificação em sua largura e altura. As imagens seguintes possuem suas larguras e alturas reduzidas pela metade, referente à imagem na posição anterior, cada vez que são armazenadas na posição seguinte da pirâmide (lista) como mostrado na Figura 13. Desta forma, cada imagem que sofre redução em suas dimensões se torna um pouco embaçada caracterizando a Pirâmide Gaussiana.

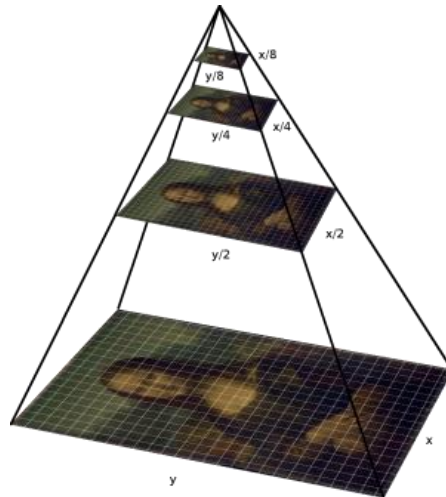


Figura 13 - Representação Multiescala de uma imagem (Pirâmide). *Fonte: [10].*

Neste trabalho, a matriz utilizada como base na construção da Pirâmide Gaussiana é a matriz do Mapa de Pesos Normalizados. Para cada canal de cor, é construída uma Pirâmide Gaussiana tendo como base na construção, a matriz do Mapa de Pesos Normalizados de cada canal de cor. Este procedimento de construção da Pirâmide Gaussiana está ilustrado na Figura 10, 3ª Etapa, na qual para cada canal de cor, as pirâmides da esquerda são as Pirâmides Gaussianas e as pirâmides da direita são as Pirâmides Laplacianas. Ainda na Figura 10, 3ª Etapa, por exemplo para a Pirâmide Gaussiana, **GR1** significa: **G**(Pirâmide Gaussiana), **R**(Canal de cor Vermelho - *Red*) e **1**(Nível da pirâmide); **GG3** significa: **G**(Pirâmide Gaussiana), **G**(Canal de cor Verde - *Green*) e **3**(Nível da pirâmide); **GB2** significa: **G**(Pirâmide Gaussiana), **B**(Canal de cor Azul - *Blue*) e **2**(Nível da pirâmide) e assim por diante.

A Pirâmide Laplaciana [6] é construída a partir da Pirâmide Gaussiana e devido à isso, em todo processo, no qual utiliza Pirâmide Laplaciana, se faz necessário primeiramente a construção da Pirâmide Gaussiana.

A Pirâmide Laplaciana é formada por imagens obtidas pela diferença de duas imagens da Pirâmide Gaussiana. O processo de construção da Pirâmide Laplaciana está ilustrada na Figura 14. De acordo com [6], esse processo cria imagens semelhantes ao resultado alcançado pela aplicação de um filtro Laplaciano.

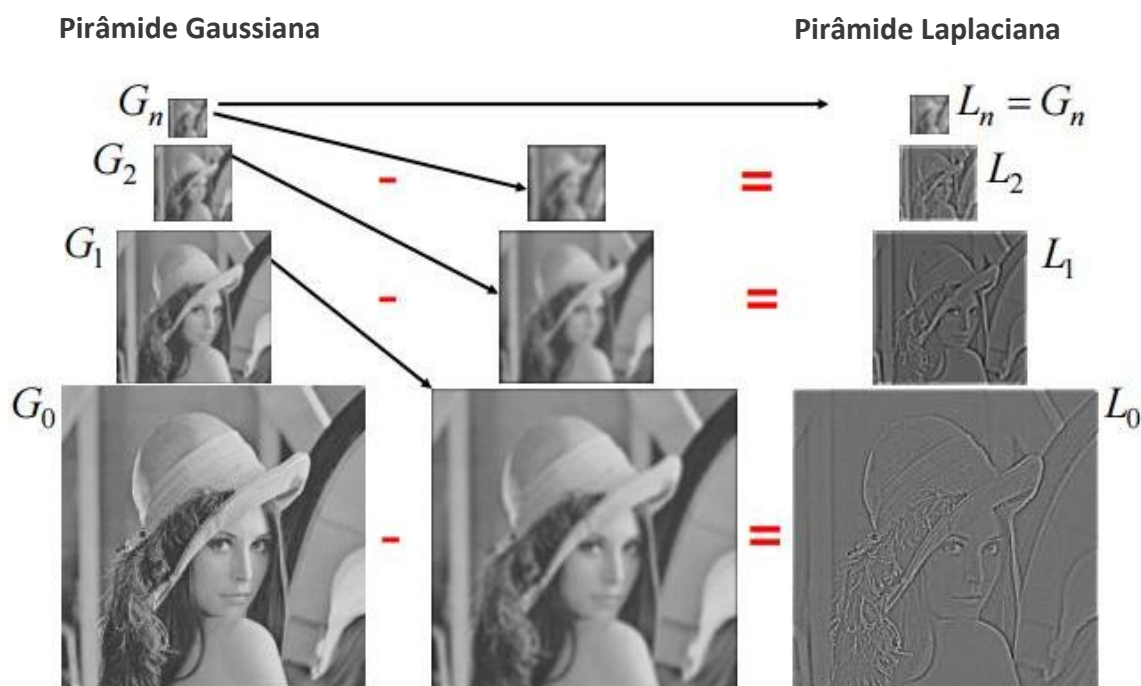


Figura 14 - Construção da Pirâmide Laplaciana. Fonte: [11].

Na Figura 14, a imagem G_0 é a imagem do primeiro nível da Pirâmide Gaussiana e a imagem G_n é a imagem do último nível da Pirâmide Gaussiana. Assim, a imagem L_0 , a qual representa o primeiro nível da Pirâmide Laplaciana, é obtida pela diferença entre a imagem G_0 e G_1 (redimensionada com as dimensões de G_0). O próximo nível da Pirâmide Laplaciana, L_1 , é obtido pela diferença entre a imagem G_1 e G_2 (redimensionada com as dimensões de G_1) e assim por diante. Para obter a imagem do último nível da Pirâmide Laplaciana, L_n , a estratégia adotada foi a de usar a última imagem, G_n , da Pirâmide Gaussiana como a imagem do último nível, L_n , da Pirâmide Laplaciana pelo fato de não haver uma imagem no nível superior para ser redimensionada e aplicada a diferença [6].

De forma semelhante à Pirâmide Gaussiana, na Figura 10, 3ª Etapa, por exemplo para a Pirâmide Laplaciana, **LR1** significa: **L**(Pirâmide Laplaciana), **R**(Canal de cor Vermelho - **Red**) e **1**(Nível da pirâmide); **LG3** significa: **L**(Pirâmide Laplaciana), **G**(Canal de cor Verde - **Green**) e **3**(Nível da pirâmide); **LB2** significa: **L**(Pirâmide Laplaciana), **B**(Canal de cor Azul - **Blue**) e **2**(Nível da pirâmide) e assim por diante.

Neste trabalho, os três canais de cor utilizados como entrada no processo de descolorização são utilizados como base para a construção da Pirâmide Laplaciana, uma para cada canal de cor. Primeiro, são construídas três Pirâmides Gaussianas tendo como base os três canais de cor, um canal de cor para cada Pirâmide Gaussiana e, posteriormente, tendo como base essas Pirâmides Gaussianas, são construídas três Pirâmides Laplacianas, uma pirâmide para cada canal de cor. Desta forma, as Pirâmides Laplacianas são criadas e este processo está representado na Figura 10, 3ª Etapa. As Pirâmides Gaussianas e Laplacianas para cada canal de cor estão ilustradas nas Figuras 15 e 16 respectivamente.

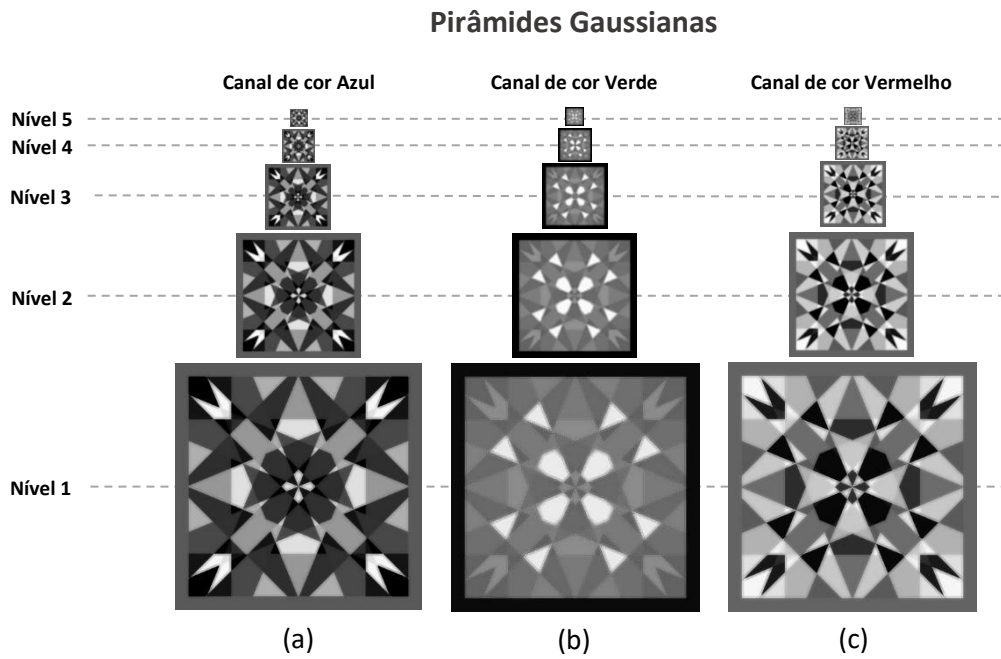


Figura 15 - Pirâmides Gaussianas para cada canal de cor.

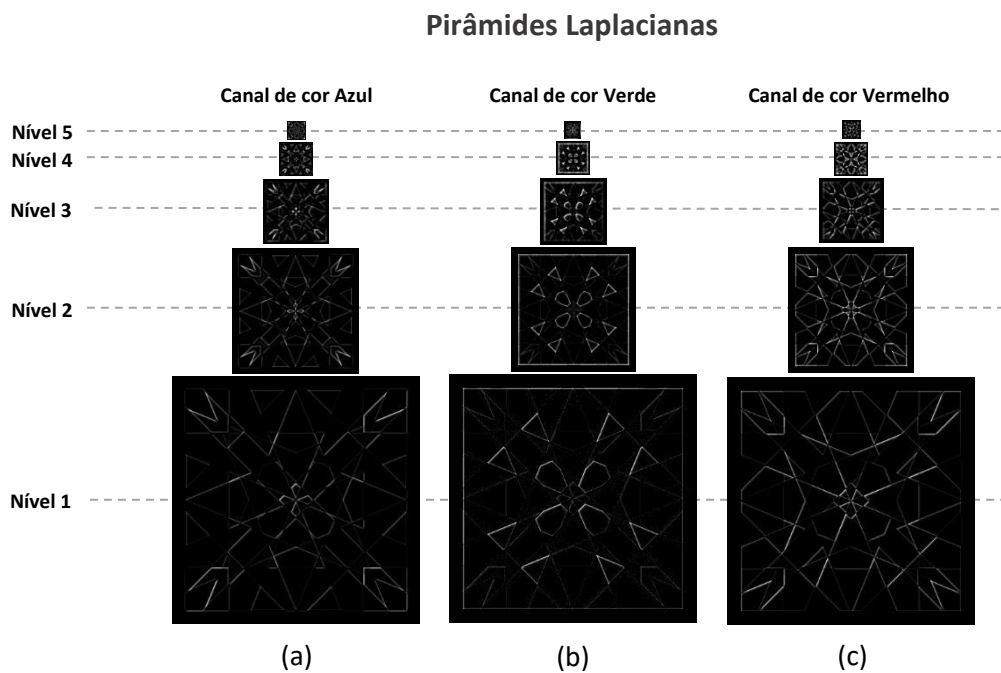


Figura 16 - Pirâmides Laplacianas para cada canal de cor.

3.5 4ª Etapa: Fusão Multiescala das Pirâmides Gaussiana e Laplaciana

A Fusão Multiescala é responsável por gerar a imagem final descolorizada e reduzir ao máximo os ruídos que foram introduzidos pelos Mapas de Pesos Laplaciano e de Contraste Global. A Fusão Multiescala é definida pela equação:

$$F^l(x, y) = \sum_{k=1}^3 G^l\{\bar{W}^k(x, y)\} L^l\{I^k(x, y)\} \quad (11)$$

em que, l representa o número de níveis das pirâmides calculado através da equação (10) na 3ª Etapa do algoritmo, k representa o canal de cor RGB ($k = [1, 3]$), $G^l\bar{W}^k(x, y)$ representa o Mapa de Pesos normalizado sob efeito semelhante ao do Filtro Gaussiano [1] para um determinado nível l da pirâmide e para um determinado canal de cor k , $L^l I^k(x, y)$ representa o canal de cor sob efeito semelhante ao do Operador Laplaciano [1] no mesmo nível l da pirâmide e para o mesmo canal de cor k e $F^l(x, y)$ é a pirâmide final, na qual ocorre a fusão das imagens, isto é, as imagens de todos os níveis são somadas e como resultado tem-se a imagem final descolorizada (saída do algoritmo destacado em vermelho pela linha tracejada na Figura 10, 4ª Etapa). A pirâmide final, a qual possui as imagens que sofrem fusão para gerar a imagem descolorizada e esta, estão ilustradas na Figura 17a e 17b respectivamente.

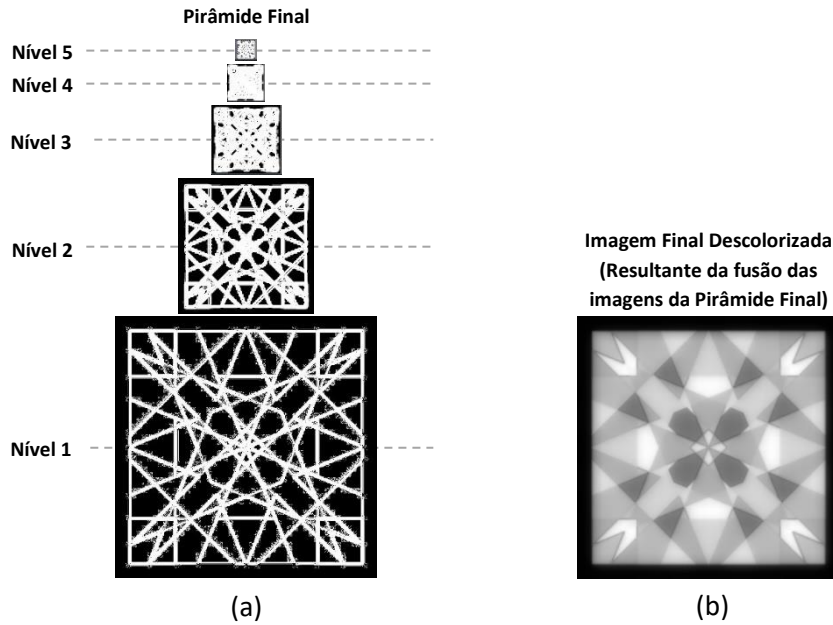


Figura 17 - Pirâmide Final (a) e a saída do algoritmo (b).

O grande objetivo de se utilizar a Fusão Multiescala das Pirâmides Gaussiana e Laplaciana é de reduzir os possíveis ruídos introduzidos pelos Mapas de Pesos Laplaciano e de Contraste Global, cujos objetivos são de manter os detalhes, fronteiras, transições de intensidade, contrastes da imagem original na imagem descolorizada [4].

4 Experimentos e Análises

Experimentos foram realizados com dois objetivos: (i) verificar se as imagens em tons de cinza preservam os contrastes das cores presentes na imagem original constituída de três canais de cor e (ii) verificar se as imagens em tons de cinza geradas pelos algoritmos são estatisticamente diferentes. As medidas de avaliação adotadas em [12] não foram utilizadas neste trabalho e os experimentos foram realizados utilizando o MatLab.

Em uma primeira avaliação, para verificar se os contrastes das cores existentes na imagem RGB foram preservados na imagem em tons de cinza, utilizou-se a Taxa de Preservação de Contraste (CCPR – *Color Contrast Preserving Ratio*), a qual tem como base a diferença de cor δ ser menor que um limiar τ . O processo de descolorização preservando o contraste significa manter as mudanças de cor capazes de serem percebidas pelos humanos [12]. A CCPR é definida pela expressão a seguir.

$$CCPR = \frac{\#\{(x, y) | (x, y) \in \Omega, |g_x - g_y| \geq \tau\}}{\|\Omega\|} \quad (12)$$

em que Ω é o conjunto da imagem fonte que contém todos os pares de pixels que são vizinhos, na qual sua diferença de cor $\delta_{x,y} \geq \tau$. $\|\Omega\|$ é a quantidade de pares de pixels em Ω , onde (x, y) é um par de pixels. $\#\{(x, y) | (x, y) \in \Omega, |g_x - g_y| \geq \tau\}$ é a quantidade de pares de pixels em Ω que continuam distintivos depois da conversão para tons de cinza.

Devido ao fato de a CCPR se tratar de uma taxa, o seu valor varia de 0 a 1, em que 0 significa que não houve nenhuma preservação de contraste na imagem em tons de cinza, e 1 significa que houve total preservação de contraste na imagem em tons de cinza em relação à imagem RGB original.

Na avaliação, foi calculado o CCPR para cada uma das cem imagens em tons de cinza geradas pelos métodos apresentados no capítulo 2 e pelo algoritmo de Ancuti e Ancuti tendo como referência as mesmas imagens RGB originais convertidas para tons de cinza pelo método da média dos três canais de cor. Assim, foi possível verificar se os contrastes nas imagens em tons de cinza foram preservados ou não. A tabela 1 apresenta a média do CCPR e o desvio padrão referente à cada método de conversão de imagem em tons de cinza.

Diante dos dados presentes na Tabela 1, pode-se observar que a média do CCPR para todos os métodos de conversão de imagem em tons de cinza com exceção do algoritmo Ancuti e Ancuti apresentou valor próximo de 1. Isto significa que esses métodos preservam os contrastes presentes na imagem RGB original na imagem em tons de cinza. Já no algoritmo Ancuti e Ancuti, observa-se o contrário.

Tabela 1 - Média e desvio padrão do CCPR, avaliando as imagens geradas por métodos clássicos e geradas pelo algoritmo [4] tendo como base as imagens geradas pela média dos três canais de cor.

<i>Método de Conversão</i>	<i>Média</i>	<i>Desvio Padrão</i>
<i>Canal de Cor Verde</i>	0,94	0,06
<i>Canal de Luminância</i>	0,92	0,06
<i>Máximo entre os canais RGB</i>	0,95	0,05
<i>Média dos canais RGB</i>	0,97	0,03
<i>Contraste Percebido</i>	0,85	0,20
<i>Algoritmo Ancuti e Ancuti</i>	0,32	0,15

Em uma segunda avaliação, comparamos as imagens geradas pelos métodos apresentados, tomando como base a medida de similaridade de imagens SSIM (*Structural Similarity Index*). O objetivo aqui foi definir se as imagens geradas pelos algoritmos são estatisticamente diferentes ou não. Para tanto, comparamos os resultados gerados pelo algoritmo de Ancuti e Ancuti com os resultados gerados pelos outros métodos apresentados no capítulo 2. Em relação ao uso de apenas um dos canais da imagem (R, G ou B), optamos por usar apenas o canal verde (G) nesta comparação. O SSIM foi calculado para as cem imagens geradas por cada método, comparando com o método de Ancuti e Ancuti. A Tabela 2 apresenta a média e o desvio padrão desses resultados.

O valor do SSIM varia de -1 a 1, em que 0 significa que as imagens comparadas são diferentes, e 1 significa que as imagens comparadas são bastante semelhantes.

Tabela 2 - Média e desvio padrão do SSIM, comparando as imagens geradas por métodos clássicos com as imagens geradas pelo algoritmo [4].

<i>Método de Conversão</i>	<i>Média</i>	<i>Desvio Padrão</i>
<i>Canal de Cor Verde</i>	0,36	0,18
<i>Canal de Luminância</i>	0,35	0,17
<i>Máximo entre os canais RGB</i>	0,34	0,17
<i>Média dos canais RGB</i>	0,35	0,17
<i>Contraste Percebido</i>	0,38	0,19

Em seguida, executamos uma Análise de Variância (ANOVA) para analisar se os valores de similaridade encontrados pelo SSIM indicam diferenças estatisticamente significativas entre as imagens ou não. O *p-value* encontrado foi igual a 0,5484, indicando que as diferenças entre os valores de SSIM encontrados não são estatisticamente significativas. Isso indica que, sob o ponto de vista de comparação via SSIM, não há diferença na escolha dentre os métodos. A Tabela 3 apresenta os resultados gerados pela ANOVA.

Tabela 3 - Tabela 3. Análise de variância aplicada aos resultados de comparação gerados pelo SSIM.

Fonte	SS	df	MS	F	Prob>F
Colunas	0,0984	4	0,02461	0,76	0,5484
Erro	15,9267	495	0,03218		
Total	16,0252	499			

A Tabela 3 mostra a variação entre grupos (em "Colunas") e intra grupos ("Erro"). SS é a soma dos quadrados (*Square Sum*) e df apresenta os graus de liberdade (*degrees of freedom*). MS é o erro quadrático médio (*Mean Squared Error*), calculado como SS/df. A estatística F é a razão entre os erros quadráticos médios. Por último, o *p-value* (representado por Prob>F) é a probabilidade que o teste estatístico assuma um valor maior que F. O alto valor (maior que 0,05) indica que as diferenças não são estatisticamente significativas como reportado anteriormente.

Na Figura 18, estão ilustrados os resultados comparativos das imagens em tons de cinza resultantes do algoritmo proposto em [4] implementado neste trabalho com as imagens em tons de cinza resultantes dos métodos de descolorização abordados no capítulo 2.

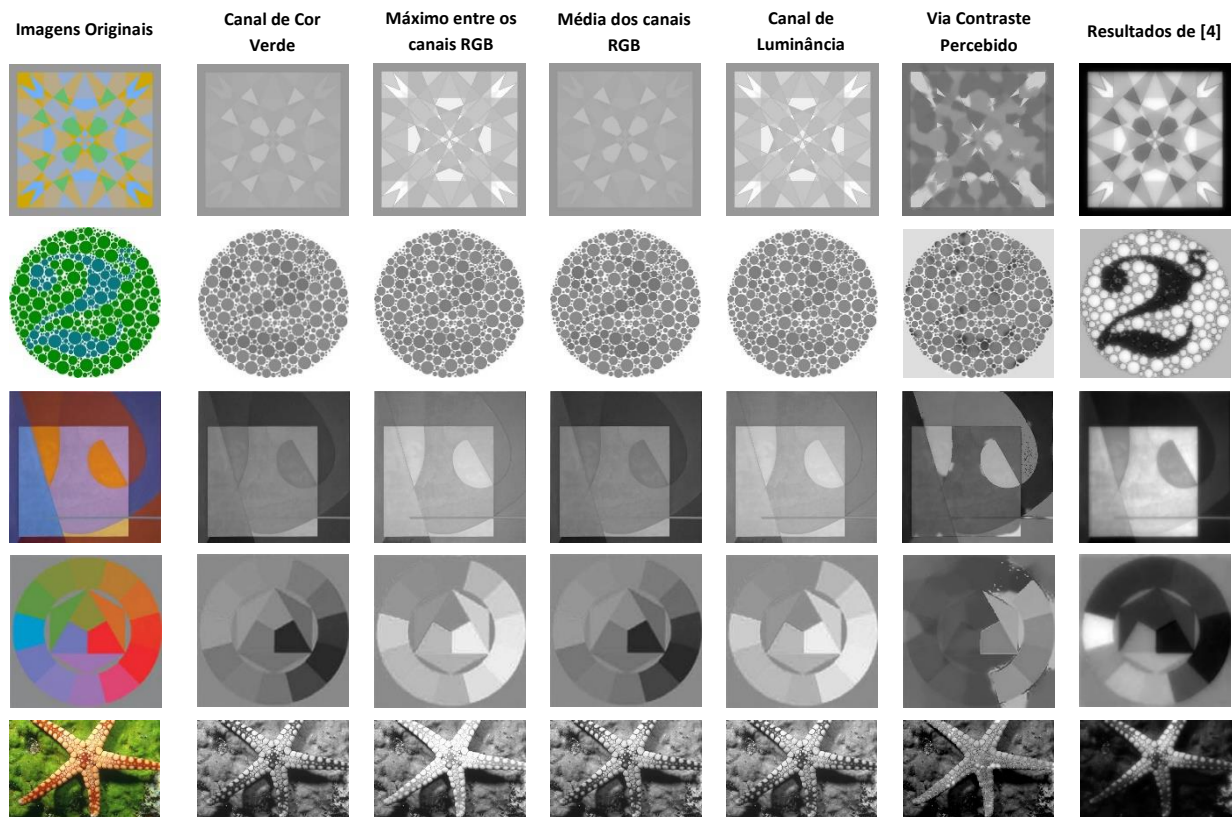


Figura 18 – Comparação das imagens em tons de cinza obtidas pelos métodos de descolorização abordados neste trabalho.

5 Conclusões e Trabalhos Futuros

Neste trabalho, foi realizado o estudo da técnica de conversão de imagens coloridas em tons de cinza proposta em [4] e das técnicas clássicas que foram descritas no capítulo 2. Para verificar a pertinência da utilização do algoritmo proposto por Ancuti e Ancuti, foram realizados experimentos para avaliar o quanto os contrastes das imagens RGB originais foram preservados na imagem em tons de cinza utilizando o CCPR para cada um dos métodos de conversão. Além disso, comparou-se as imagens geradas pelos métodos apresentados, tomando como base a medida de similaridade de imagens SSIM (*Structural Similarity Index*) com o objetivo verificar se as imagens geradas pelos algoritmos são estatisticamente diferentes ou não. Com base nos dados obtidos, observou-se que o método proposto em [4] não preserva de forma adequada os contrastes da imagem RGB original na imagem em tons de cinza e que as imagens geradas não são estatisticamente diferentes. Isto significa que, sob o ponto de vista de comparação via SSIM, não há diferença na escolha dentre os métodos mostrados neste trabalho.

5.1 Contribuições

Este Trabalho de Graduação deixa como contribuição um estudo do algoritmo definido em [4], o qual foi completamente implementado. Nessa transcrição da descrição do artigo para o algoritmo, observamos que diversos ajustes não estavam descritos de forma completamente correta e tiveram que ser modificados para alcançar os resultados apresentados. O código implementado será disponibilizado para uso futuro.

5.2 Limitações

Ausência de detalhes na explicação do algoritmo proposto em [4] e ausência de uma base de imagens para verificação da corretude da implementação do algoritmo neste trabalho.

Além disso, no capítulo 4 foi apresentada uma medida de avaliar o quanto o contraste da imagem original é preservado na imagem em tons de cinza após a conversão, esta medida é o CCPR. Um dos cálculos realizados pelo CCPR é mensurar a quantidade de pixels vizinhos da imagem original possui diferença maior que um limiar. Dado que a imagem original possui três canais de cor (RGB) e por não haver documentação suficiente desta medida de avaliação, foi necessário, para este trabalho, realizar o cálculo da média dos três canais de cor da imagem original para obter a quantidade de pixels vizinhos. Esta estratégia pode não ser a melhor forma de utilizar o CCPR.

5.3 Trabalhos Futuros

Estender a técnica de conversão de imagens para tons de cinza apresentada em [4] utilizando Deep Learning.

A extensão sugerida se dará da seguinte forma: uma imagem em tons de cinza utilizada na detecção/reconhecimento de objetos de uma dada aplicação será a entrada de uma Rede Neural Profunda (*Deep Learning*), a qual será treinada para gerar a imagem em tons de cinza de entrada preservando o máximo dos detalhes e contrastes da imagem original.

Uma forma de avaliar se a Rede Neural Profunda produzirá resultados satisfatórios, será avaliando a corretude na detecção/reconhecimento de objetos desta mesma imagem, a qual é uma imagem de referência que preserva os detalhes e contrastes de forma adequada. Caso os resultados sejam satisfatórios, outras imagens serão utilizadas como entrada para serem ajustadas e os contrastes e detalhes sejam preservados, caso contrário, a Rede Neural Profunda teria os pesos das camadas de convolução ajustados e o treinamento continuaria até resultados satisfatórios serem obtidos.

Referências Bibliográficas

- [1] Rafael C. Gonzalez e Richard E. Woods, Processamento Digital de Imagens, Pearson Education, 3a Edição, 2008.
- [2] Amy A. Gooch, Sven C. Olsen, Jack Tumblin e Bruce Gooch, "Color2Gray: Saliency-Preserving Color Removal", SIGGRAPH, ACM Transactions on Graphics, vol. 24, no. 3, pp.634–639, 2005.
- [3] G. Wyszecki e W.S. Stiles, Color science: Concepts and methods, quantitative data and formulae, Wiley-Interscience, 2a Edição, 2000.
- [4] Cosmin Ancuti e Codruta O. Ancuti, "Laplacian-Guided Image Decolorization", IEEE International Conference on Image Processing, Phoenix, EUA, pp.4107-4111, 2016.
- [5] Codruta Orniana Ancuti, Cosmin Ancuti e Philippe Bekaert, "Decolorizing Images for Robust Matching", IEEE 17th International Conference on Image Processing, Hong Kong, China, pp.149-152, 2010.
- [6] Peter J. BURT e Edward H. Adelson, "The Laplacian Pyramid as a Compact Image Code", IEEE Transactions on Communications, Vol. Com-31, no. 4, pp.532-540, 1983.
- [7] Codruta O. Ancuti, Cosmin Ancuti, Chris Hermans e Philippe Bekaert, "Image and Video Decolorization by Fusion", Asian Conference on Computer Vision, pp.79-92, 2010.
- [8] University of California, Berkeley, Computer Vision Group, 2011. Contour Detection and Hierarchical Image Segmentation.
<https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html>.
- [9] ZHU, W., HU, R., LIU, L.: "Grey conversion via perceived-contrast", The Visual Computer, vol. 30, Issue 3, pp. 209-309, Mar. 2014.
- [10] Ruven Pillay, IIPImage » Images, 2015. Acessado em 22 de Outubro de 2017.
<http://iipimage.sourceforge.net/documentation/images/>
- [11] *HYPJUDY, [CVPR] Panorama Image Stitching. Acessado em 22 de Outubro de 2017.*
<https://hypjudy.github.io/2017/05/10/panorama-image-stitching/>
- [12] LU, C., XU, L., JIA, J.: Contrast preserving decolorization. In: 2012 IEEE International Conference on Computational Photography (ICCP), pp. 1–7 (2012).