

# Trabalho 1

Levy Gurgel Chaves – RA: 264958

MO433/MC920 - Introdução ao Processamento de Imagem Digital  
Universidade Estadual de Campinas  
Prof. Hélio Pedrini

1 de Abril de 2020

## 1 Introdução

O objetivo deste trabalho é implementar técnicas fundamentais de processamento de imagens, onde conceitos importantes básicos como luminosidade, resolução, quantização e transformações são abordados de maneira prática.

Juntamente com o relatório, está sendo enviado o arquivo LEVY\_MO443T1\_264958.zip, contendo todo o conteúdo apresentado neste relatório. A seção 2 comenta detalhadamente sobre a entrada e as saídas do programa, enquanto a seção 3 apresenta a explicação da solução dos problemas descritos no primeiro trabalho, assim como o resultado dos experimentos realizados.

## 2 O programa

O programa foi desenvolvido em Python em sua versão 3.6.9 com o auxílio das seguintes bibliotecas nas suas respectivas versões: Numpy 1.18.2; Matplotlib 3.2.1; Scikit-image 0.16.2 e Scipy 1.4.1. O mesmo pode ser conferido com o auxílio do Jupyter notebook executando o arquivo *notebook-t1.ipynb*.

### 2.1 Entrada

O programa realiza a leitura de uma imagem, em escala de cinza, a partir da pasta */images*. Por convenção, a Figura 1 representa a imagem de referência adotada para os experimentos.



Figura 1: Imagem de referência para experimentos

### 2.2 Saída

As imagens relativas aos resultados dos experimentos possuem formato *PNG* e são armazenadas na pasta */output*.

## 3 Soluções e decisões tomadas

### 3.1 Leitura das imagens

A leitura das imagens foram feitas através da função **skimage.io.read** onde a mesma retorna um array Numpy possuindo as dimensões  $M \times N \times C$ , onde  $M \times N$  representam as dimensões da imagem lida e  $C$  indica a quantidade de diferentes canais.

### 3.2 Resolução de imagens

A resolução espacial de uma imagem está inteiramente ligada à densidade de pixels. Uma imagem ao ser digitalizada assume tamanho adimensional, em pixels. Após o processo de amostragem, maior será a resolução se possuir uma maior densidade de pixels, geralmente calculado através do DPI.

Com o intuito de diminuir pela metade a resolução de uma imagem, sem alterar as dimensões, uma abordagem utilizando janela deslizante foi adotada. Dada uma imagem com dimensões  $R \times C$  e uma determinada resolução desejada  $R_d \times C_d$ , uma janela possuindo dimensões  $\lceil \frac{R}{R_d} \rceil \times \lceil \frac{C}{C_d} \rceil$  foi criada com a finalidade de realizar um processo de varredura na imagem. A janela realiza a varredura na imagem nos eixos horizontais e verticais, de maneira em que não ocorra nenhuma interseção entre elementos em janelas distintas. A cada passo durante a varredura, todos os elementos atualmente avaliados serão substituídos pela média da janela. Com média da janela, entende-se como a média entre todos os elementos presentes na janela atual.

O resultado final do procedimento descrito pode ser visto na Figura 2. O efeito da redução da resolução afeta diretamente o conteúdo visual da imagem, onde implica em perda gradual de detalhes. Isso acontece pois cada pixel da imagem original agora está sendo representado por  $R_d \cdot C_d$  blocos e cada bloco comprime informações de intensidades de diversos pixels em apenas um valor.

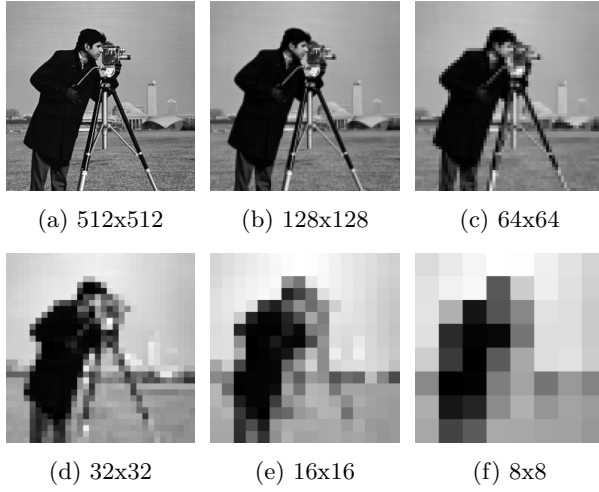


Figura 2: Imagem em diferentes resoluções

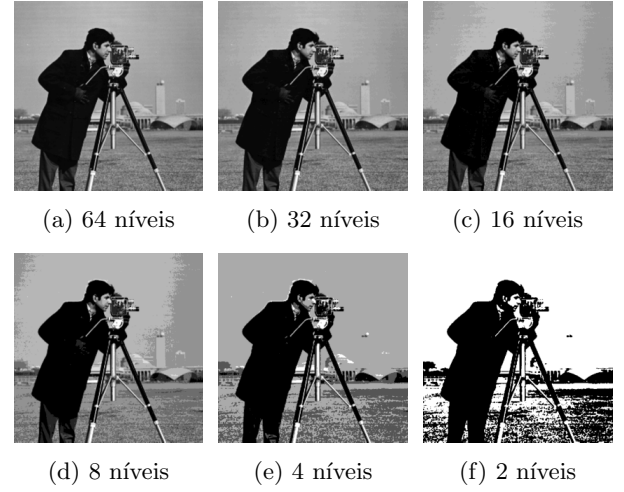


Figura 3: Imagem em diferentes níveis de quantização

Como limitação o método assume que a janela será capaz de percorrer completamente a imagem, ou seja, não há qualquer tipo de preenchimento lateral ou vertical.

### 3.3 Quantização de imagens

Quantização refere-se ao processo de converter um intervalo discreto ou contínuo em um outro de valores discretos. Por questão de convenção, adota-se que os pixels da imagem de entrada estejam entre intervalo discreto  $[0 - 255]$ . Reduzir o número de intensidades diferentes em uma imagem pode ser útil principalmente para a visualização em dispositivos que possuem limitação no número de cores a serem armazenados (KLINGER, 2003).

Para realizar este processo, é empregado a operação de *shift right* no valor da representação binária de cada intensidade. A ideia principal é zerar os planos de bits mais significativos de cada representação binária. A quantidade de bits deslocados à direita será de  $8 - k$ , onde  $k$  é o número de bits necessário para representar a quantidade de níveis subtraída de um.

A Figura 3 ilustra o resultado da técnica de quantização em diferentes níveis. Podemos perceber que quanto menor o nível escolhido, menos há uma distinção entre intensidades bem similares. Nota-se que, por exemplo, o detalhe em algumas construções presentes na Figura 3d desapareceram na Figura 3e devido à redução do intervalo de intensidade, tornando alguns objetos e contornos indistinguíveis.

Uma limitação do método desenvolvido é assumir que o número de níveis diferentes será sempre da forma  $2^n$ , onde  $n \leq 8$ .

### 3.4 Escalas de cinza

#### 3.4.1 Transformação logarítmica

Na transformação logarítmica toma-se o logaritmo do valor de cada pixel adicionado de uma unidade. Matematicamente, temos:

$$g = c \log(f + 1), \quad (1)$$

onde  $g$  é imagem resultante,  $f$  é a função de entrada e  $c$  uma constante. Nota-se que a operação é efetuada no valor de cada pixel adicionado de um, uma vez que a função  $f$  assume valores discretos entre  $[0 - 255]$  e o valor da função logarítmica é indefinida para o valor 0.

Para um melhor entendimento, a Figura 4 representa a função supracitada de maneira gráfica. Podemos perceber que a função apresenta rápido crescimento entre os valores de intensidade  $0 \leq x \leq 50$  e menos acentuado para valores maiores que 50.

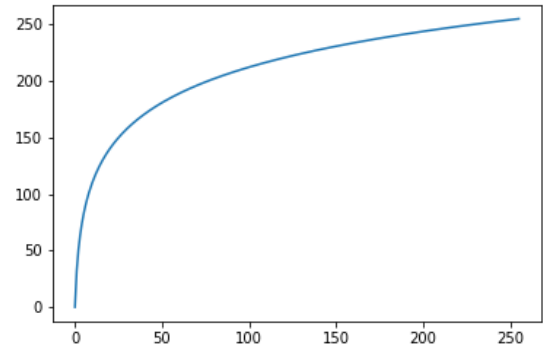


Figura 4: Comportamento da função  $c \log(x + 1)$

O comportamento apresentado pela função indica que esta transformação podem ser utilizadas para aumentar a iluminação de uma imagem. Mais frequentemente, é usada para aumentar o detalhe (ou contraste) de valores de intensidade mais baixos. Sendo assim, a Figura 5 ilustra o resultado da transformação. A imagem apresentou uma maior iluminação em comparação com a imagem de referência, ou seja, possuindo intensidades mais voltadas à cor branca.

A escolha da constante  $c$  é de fundamental importância para a geração do resultado final. Esta serve como um fator multiplicativo para ajustar o valor de saída. Como utilizamos valores entre  $[0 - 255]$ , o valor adotado  $c = \frac{255}{\log(255+1)}$ , pois o valor máximo da escala é 255, servindo como fator de normalização para o intervalo discreto  $[0 - 255]$ . Uma constante com valor



Figura 5: Imagem de referência em escala logarítmica

diferente resultará em imagens mais luminosas ou mais escuras.

### 3.4.2 Transformação exponencial

Essencialmente, transformações exponenciais representam funções inversas às funções logarítmicas, indicando que são transformações complementares. Representada matematicamente como:

$$g = c(e^x - 1) \quad (2)$$

O gráfico da transformação (ver Figura 6) indica uma maior acentuação na intensidade apenas próximo ao fim do intervalo de valores de entrada. Portanto, a imagem resultante apresentará um contraste bem menor do que a operação anterior.

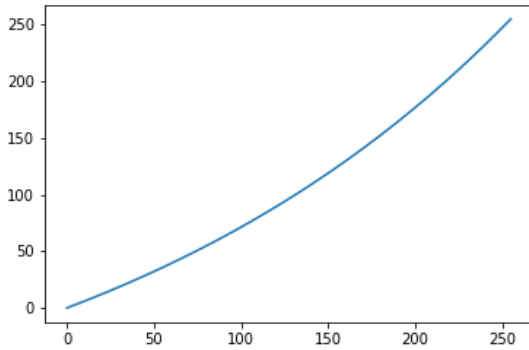


Figura 6: Comportamento da função  $c(e^x - 1)$

Se observarmos a distribuição de cores entre a escala exponencial e a logarítmica, a exponencial gerará maior distribuição de intensidades escuras, enquanto que logarítmica produzirá uma escala em que intensidades mais claras sejam maioria. Para ilustrar o resultado, a Figura 7 mostra a transformação exponencial aplicada à imagem de referência.

Para o cálculo da transformação, é necessário que a matriz esteja em sua forma normalizada, ou seja, entre o intervalo contínuo  $[0 - 1]$ , pois o cálculo de  $e^x$  resultará em *overflow* caso o intervalo discreto  $[0 - 255]$  seja adotado. Neste caso, a constante  $c$  empenha o mesmo papel do que na transformação anterior, porém, com valor  $c = \frac{255}{e-1}$ .



Figura 7: Imagem de referência em escala exponencial

### 3.4.3 Transformação quadrática

A nova imagem é resultado da operação de elevar a intensidade ao quadrado. Matematicamente:

$$g = cf^2 \quad (3)$$

Possui um comportamento muito parecido com a transformação exponencial, porém nesta a curva da função é um pouco menos suave na transição entre intensidades. A partir da figura 8 pode-se esperar que a imagem após a transformação quadrática apresente uma menor luminosidade em comparação com a transformação exponencial (KLINGER, 2003).

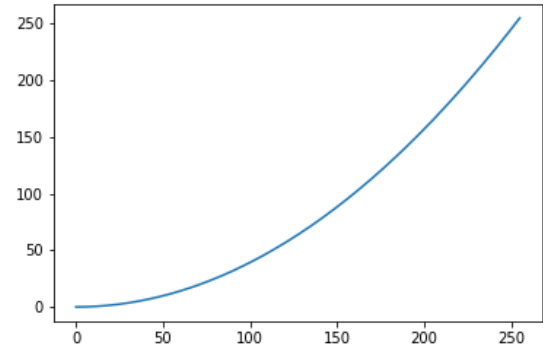


Figura 8: Comportamento da função  $cf^2$

Em escala de cinza, estas transformações diminuem a luminosidade geral de uma imagem e aumentam o contraste em regiões que possuem maior brilho ao custo do contraste em áreas mais escuras, assim como mostrado na Figura 9.



Figura 9: Imagem de referência em escala quadrática

Nesta transformação pode-se adotar tanto imagens

possuindo valores discretos entre  $[0 - 255]$  quanto no intervalo contínuo entre  $[0 - 1]$ , sendo assim, o valor da constante  $c$  é extremamente relacionada ao intervalo dos valores de entrada. Durante a implementação, adotou-se o intervalo entre  $[0 - 1]$  e  $c = 255$ . A mesma implementação poderia ser utilizada para  $[0 - 255]$ , todavia, o novo valor da constante seria de  $c = \frac{1}{\sqrt{255}}$ .

### 3.4.4 Transformação raiz quadrada

A imagem resultante é obtida após a realização da operação de raiz quadrada de todos os valores de intensidade. Pela formula e graficamente (Figura 10), temos:

$$g = c\sqrt{f} \quad (4)$$

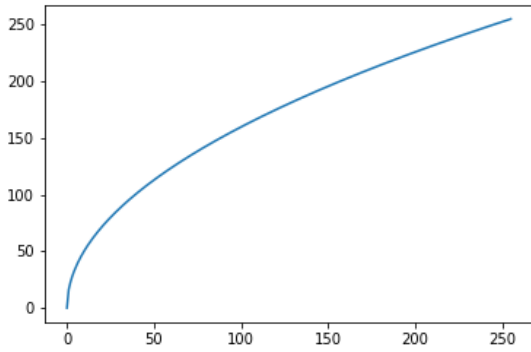


Figura 10: Comportamento da função  $c\sqrt{f}$

Em comparação com a transformação logarítmica, ambas aumentam o contraste dos pixels possuindo baixa intensidade ao custo do contraste dos pixels possuindo alta intensidade. Portanto, ambas são adequados para melhorar os detalhes contidos em valores de baixa intensidade. No entanto, produzem melhorias ligeiramente diferentes, uma vez que as formas das suas curvas não são idênticas.



Figura 11: Imagem de referência em escala raiz quadrada

Nesta transformação também assume-se que a entrada seja uma imagem normalizada entre  $[0 - 1]$  e a constante adotada foi  $c = 255$ . Todavia, o mesmo resultado seria obtido caso a imagem estivesse no intervalo discreto  $[0-255]$ , mas com o ajuste da constante para  $c = \sqrt{255}$ .

### 3.4.5 Alargamento de contraste

Pode ser analisada como uma função linear por partes, onde, a partir de um certo intervalo, a função apresenta comportamento diferente.

$$g = \begin{cases} \alpha f & \text{se } 0 \leq f \leq a \\ \beta(f - a) + \alpha a & \text{se } a < f \leq b \\ \gamma(f - b) + \beta(b - a) + \alpha a & \text{se } b < f \leq L \end{cases}$$

Graficamente, a função pode ser visualizada na Figura 12. Esta transformação pode ser pensada com o objetivo de expandir(ou diminuir) dinamicamente intervalo de intensidades. A função possui parâmetros  $\alpha, \beta, \gamma, a$  e  $b$ , onde estes parâmetros definem, essencialmente, em quais intervalos ocorrerá um aumento ou diminuição de contraste.

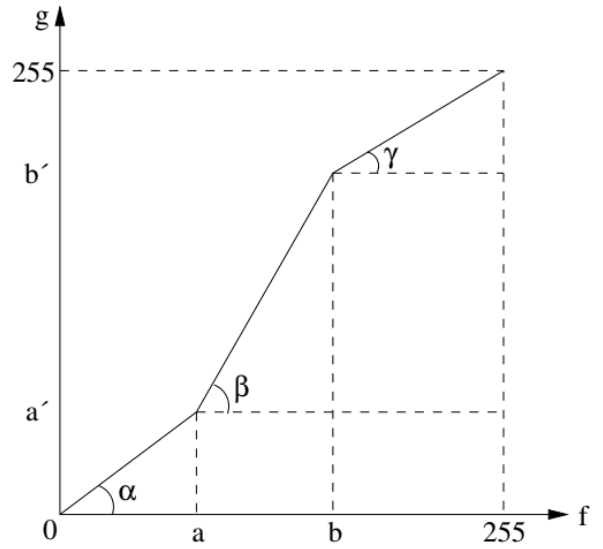


Figura 12: Alargamento linear graficamente (PEDRINI, 2020)

O conjunto de parâmetros para testes presentes na Tabela 1 foram selecionados de maneira empírica; escolhendo-se aqueles que mostravam diversas configurações da função mostrada na Figura 12. Os parâmetros  $\alpha, \beta, \gamma$  foram escolhidos de maneira que possibilitasse um aumento ou diminuição abrupta no contraste.

| Parâmetros              | Valores             |
|-------------------------|---------------------|
| $\alpha, \beta, \gamma$ | $\{0.5, 1.0, 2.0\}$ |
| $a$                     | $\{50, 100, 130\}$  |
| $b$                     | $\{150, 170, 200\}$ |

Tabela 1: Valores utilizados para experimentos

Alguns resultados obtidos a partir do parâmetros da tabela anterior podem ser conferidos na figura 13. O experimento consistiu em combinar todos os parâmetros indicados da tabela anterior com a finalidade de analisar diferentes configurações de contraste.

Por exemplo, para  $\alpha$ ,  $\beta$  ou  $\gamma$  possuindo valor  $< 1$  indicam uma diminuição da intensidade original. Assim, como valores  $> 1$  aumentam a intensidade original. Nota-se que para  $\alpha$ ,  $\beta$ ,  $\gamma$  possuindo valor maior que 1, há aumento substancial na intensidade da imagem, principalmente quando mais de um parâmetros assumam esse valor.

Nota-se que na Figura 13b e 13c, devido aos intervalos selecionados e o valor do  $\gamma$ , a diferença básica entre as imagens está nos valores de alta intensidade (observe os detalhes do tripé da câmera). Assim como as imagens 13d e 13f apresentam parâmetros semelhantes, no entanto, há realces em regiões totalmente distintas.

Como decisões tomadas, assumiu-se que nenhum dos parâmetros pode ser um valor negativo e que o valor de  $a$  é sempre menor ou igual a  $b$ . Para o caso de multiplicações em *double*, a função *around* do Numpy foi utilizada. Esta arredonda o número para o seu teto se a parte decimal é maior que 0.5, caso contrário o número é substituído pelo seu respectivo piso. Na implementação, tomou-se cuidado em nunca exceder o valor de 255 nas multiplicações efetuadas.

## Referências

KLINGER, T. **Image Processing with LabVIEW and IMAQ Vision**. Prentice Hall PTR, 2003.

(National Instruments virtual instrumentation series). ISBN 9780130474155. Disponível em: <https://books.google.com.br/books?id=8tqeVwCyZJ4C>.

PEDRINI, H. Introdução ao processamento de imagem digital – fundamentos. 2020. Disponível em: <https://www.ic.unicamp.br/~helio/>.



Figura 13: Resultados para alargamento de contraste