

Trabalho 2 - Processamento Digital de Imagens

Angélica Alves Viana*

Novembro de 2019

Resumo

Aplicações de processamento digital de imagens vêm sendo cada vez mais requisitadas em diversas áreas com destaque para numerosas aplicações na área da medicina. Por esse motivo, muitas técnicas foram e vem sendo desenvolvidas com o objetivo de suprir essa demanda. Esse trabalho tem como objetivo apresentar, implementar, testar e analisar algumas dessas técnicas para compreender plenamente o seu funcionamento e verificar o comportamento das mesmas sob diferentes condições de aplicação.

Palavras-chaves: processamento de imagens. elementos estruturantes. técnicas.

Introdução

Uma imagem digital pode ser definida como uma função $f(x, y)$ em que cada ponto (x, y) de coordenadas espaciais da mesma apresenta um nível de cinza e essas coordenadas e intensidade de cores são finitas (Gonzalez, 2007).

Imagens Digitais são amplamente utilizadas em inúmeras aplicações desde o seu surgimento. Algumas delas são na obtenção de imagens médicas, meteorologia, restauração de imagens e imagens de satélite. No entanto, em muitos casos a informação de interesse na imagem a ser analisada corresponde a uma pequena parcela da mesma, por esse motivo processamentos digitais que visem reduzir a quantidade de informações na imagem deixando apenas o objeto de interesse foram implementados, o que não apenas facilita a questão do armazenamento desses dados como reduz o custo computacional de processamento das mesmas.

Por outro lado, a informação de interesse na imagem muitas vezes não pode ser obtida da imagem no domínio do tempo, utilizando técnicas sobre a imagem original. Daí surgem as transformadas do domínio da frequência, que permite visualizar aspectos da imagem que apenas são apresentados no domínio da frequência.

*angelicaalvesviana@gmail.com

Em outras aplicações, especialmente em reconhecimento de objetos, geralmente há a necessidade da obtenção de uma base de dados muito extensa e para isso muitas vezes procedimentos alternativos para a obtenção da mesma são utilizados, como o *Data Augmentation*, cujo objetivo é gerar novas imagens a partir da aplicação de algumas transformações geométricas.

Devido a essa gama de aplicações e a necessidade de em muitos casos reduzir a quantidade de informações das imagens, aplicar transformações geométricas ou analisá-las no domínio da frequência, muitas técnicas de processamento de imagens digitais vêm sendo desenvolvidas com essa finalidade, como aplicações de transformadas geométricas básicas como translação, rotação e escala, técnicas de segmentação de imagens para destacar o objeto de interesse e aplicações de transformadas do domínio da frequência.

Diante da importância dessas técnicas, esse trabalho tem como objetivo realizar um estudo sobre algumas delas, destacando aspectos teóricos relacionados ao funcionamento e essas técnicas serão comparadas e analisadas a partir de testes realizados.

1 Transformações Geométricas

As transformações geométricas modificam a relação espacial entre os pixels e consistem de duas operações básicas (SÁNCHEZ; RUIZ e MEZA, 2014):

- **Transformação Espacial:** Define a realocação dos pixels no plano da imagem.
- **Interpolação dos Níveis de Cinza:** Mapeia os níveis de intensidade dos pixels da imagem transformada.

Neste trabalho, será de interesse um caso particular das transformadas geométricas que são as transformações lineares, tais como: Translação, Rotação e Escala. Cujas representações matriciais são apresentadas abaixo.

$$\begin{bmatrix} 1 & 0 & tx \\ 0 & 1 & ty \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} \alpha & \beta & (1 - \alpha).center.x - \beta.center.y \\ -\beta & \alpha & \beta.center.x + (1 - \alpha).center.y \end{bmatrix} \quad (2)$$

$$\begin{bmatrix} w & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

No contexto de processamento de imagens, tx e ty são os valores de deslocamento no eixo x e no eixo y , respectivamente, definidos na operação de translação. $\alpha = scale.cos(\theta)$ e $\beta = scale.sin(\theta)$, em que θ é valor do ângulo de rotação, $scale$ é uma medida de escala da imagem. Além disso na matriz de transformação da escala, w e h são valores inteiros positivos se deseja-se aumentar a dimensão da imagem ou são valores fracionários e positivos, se deseja-se reduzir a dimensão da imagem.

Em geral, essas transformações são aplicadas na introdução de variações nas imagens de uma base de dados com o intuito de aumentá-la ou para obter um reconhecimento de um objeto que seja invariante a certas transformações, para estimar o movimento de um objeto ou para produzir imagens panorâmicas.

2 Morfologia Matemática

A Morfologia Matemática é descrita em Gonzales (2007) como uma ferramenta para extrair componentes das imagens que são úteis na representação e na descrição da forma de uma região. Para este trabalho, serão apresentadas as duas operações primitivas de morfologia matemática que são fundamentais para a elaboração de muitas outras operações: *erosão* e *dilatação*.

Ainda em Gonzales (2007) são dadas duas descrições em formato de conjunto para a operação de erosão.

- Com A e B como conjuntos de Z^2 , a erosão de A por B, indicada por $A \ominus B$ é definida como

$$A \ominus B = \{z | (B)_z \subseteq A\}$$

Outra definição equivalente para essa operação é dada por

$$A \ominus B = \{z | (B)_z \cap A^c = \emptyset\}$$

Em ambas as definições A representa o objeto a ser erodido na imagem binária e $(B)_z$ o elemento estruturante centrado em z.

Já para as imagens em tons de cinza, a operação de erosão é diferente. Nesse caso, o valor da imagem coincidente com o pixel central do elemento estruturante deve ser substituído pelo menor valor de intensidade contido na região da imagem delimitada pelo elemento estruturante. A representação dessa operação em forma de equação é dada por

$$[f \ominus b](x, y) = \min\{f(x + s, y + t)\} \quad (4)$$

Para a operação de dilatação, as definições novamente dadas em Gonzales (2007) são apresentadas abaixo

- Sejam A e B conjuntos de Z^2 , a dilatação de A por B, indicada por $A \oplus B$, pe definida como

$$A \oplus B = \{z | (B)_z \cap A \neq \emptyset\} \quad (5)$$

Outra definição equivalente para essa operação é dada por

$$A \oplus B = \{z | [(B)_z \cap A] \subseteq A\} \quad (6)$$

Para imagens em tons de cinza, a operação de dilatação segue a mesma ideia da erosão, no entanto, ao invés de substituir o pixel centrado no elemento estruturante pelo valor mínimo de intensidade na região delimitada pelo elemento estruturante, o mesmo é substituído pelo valor máximo de intensidade. A equação que representa essa operação está descrita abaixo

$$[f \oplus b](x, y) = \max_{(s,t) \in b} \{f(x - s, y - t)\} \quad (7)$$

Em suma, as operações de erosão e dilatação aplicadas em imagens binárias têm o efeito de "diminuir" e "aumentar" os objetos na imagem respectivamente. Já essas operações aplicadas em imagens em tons de cinza têm o efeito de tornar a imagem mais escura aumentando os baixos níveis de cinza na imagem e de torná-la mais clara aumentando os altos níveis de cinza na imagem.

3 Segmentação

Segmentação em Processamento Digital de Imagens consiste em um procedimento cuja imagem original é separada em partes menores com o intuito de analisar apenas uma pequena região de interesse. Para este fim, muitas técnicas já foram implementadas. Nessa seção serão apresentadas com detalhes algumas delas, tais como Crescimento de Regiões, *K-Means*, a Transformada *Hough* e *Watershed*.

O método de **Crescimento de Regiões** foi proposto por Adams e Bischof (1994) e nesse artigo o mesmo é descrito como um método robusto e rápido de segmentação de intensidades em imagens e tem como entrada um conjunto de sementes, que pode ser apenas um conjunto de pixels isolados ou de regiões que irão "crescer" e posteriormente gerar as regiões segmentadas na imagem. Os autores definem T como sendo um conjunto de pixels candidatos à pertencerem a uma determinada região A_i . A representação matemática de T é dada por:

$$T = \left\{ x \notin \bigcup_{i=1}^n A_i \mid N(x) \cap \bigcup_{i=1}^n A_i \neq \emptyset \right\} \quad (8)$$

Em que $N(x)$ corresponde aos vizinhos 8-conectados de x . Se há interseção entre os vizinhos de x e apenas uma região A_i , o novo pixel x será adicionado à região desde que apresente o menor valor de δ dentre os candidatos, δ consiste em uma medida de similaridade adotada para a agregação de novos pixels a região e sua definição é dada em Adams e Bischof (1996) como

$$\delta(x) = |g(x) - \text{mean}_{y \in A_i(x)}(g(y))| \quad (9)$$

Onde $g(x)$ corresponde ao nível de cinza de x . Esse procedimento é repetido até que todos os pixels tenham sido agregados.

Este método pode ser utilizado em inúmeras aplicações, como na segmentação de imagens de satélite, como pode ser visto em Bins et al (1996) em que o mesmo é implementado no sistema de processamento de informações e imagens geográficas (SPRING) desenvolvido no INPE.

Outro método de segmentação que apresenta a mesma abordagem para seleção dos pixels que serão agrupados, a similaridade, é o método ***K-Means***. O mesmo é descrito em Jyoti e Sharma (2013) como um método de agrupamento não supervisionado, não determinístico, numérico e iterativo. Ainda em Jyoti e Sharma (2013) é explicado o funcionamento desse algoritmo, cujos passos são resumidos em dois e são apresentados abaixo.

- Primeiramente, são escolhidos k elementos do conjunto de dados, que no caso de uma imagem são k pixels. Esses elementos serão os representantes dos grupos nos quais se deseja segmentar a imagem.
- Em seguida, o algoritmo entra em um *loop* até a sua convergência na qual (1)- Todas as amostras de dados são rotuladas com o mesmo rótulo do centróide do qual está mais próximo; (2) - São calculados novos centroids que serão dados pela média dos elementos pertencentes a cada grupo.

A medida de proximidade adotada nesse algoritmo é a distância euclidiana e o critério de avaliação para determinar a convergência do algoritmo estão geralmente

associados à estabilização dos grupos, ou seja, nenhum elemento migra de um grupo para outro após a convergência.

Acima foram apresentados dois métodos de segmentação por similaridade. Outra abordagem de segmentação de imagens é a segmentação por descontinuidade. Duas técnicas muito conhecidas dessa abordagem são a Transformada Hough para a detecção de linhas e círculos e o algoritmo *Watershed* utilizado para segmentar regiões.

Para a Transformada Hough existem duas abordagens distintas, uma utilizada para a detecção de linhas e a outra para a detecção de círculos. A transformada Hough pode ser entendida como uma transformação de um ponto (x, y) para um espaço de parâmetros (Duda e Hart, 1972). O algoritmo tradicional da **Transformada Hough para a detecção de linhas** é apresentado em ZhuXun, Yiren e Jianqing (2009) e seus passos são apresentados abaixo.

1. Extração das bordas da imagem usando um algoritmo de detecção de bordas.
2. Quantização do plano de parâmetros $h(m, n)$ com valores iniciais iguais a zero, onde $m = INT\left(\frac{\pi}{d_\theta}\right) + 1$ e $n = INT\left(\frac{2R}{d_\rho}\right) + 1$, d_θ e d_ρ são os intervalos de quantização, $R = \frac{\sqrt{L^2 + W^2}}{2}$. L e W são as dimensões da imagem.
3. Para cada ponto $f(x, y)$, $h(i, j) = h(i, j) + 1$, onde $\theta_i = i.d_\theta$, $\rho_i = x.\cos\theta_i + y.\sin\theta_i$ e $j = INT\left(\frac{\rho_i}{d_\rho}\right)$
4. Detectar o ponto máximo do plano de parâmetros h . As coordenadas (i, j) do ponto máximo permitem obter os parâmetros da linha detectada, $\theta_i = i.d_\theta$ e $\rho_j = j.d_\rho$

Se é desejada uma alta acurácia, d_θ deve ser um valor pequeno e consequentemente m será grande. Quanto mais preciso, maior será o tempo de execução de algoritmo. Um fator importante a se ressaltar é a questão da escolha do limiar para a obtenção do valor máximo, que em muitos casos não é uma escolha simples e pode gerar muitos falsos positivos.

Enquanto na transformada Hough para a detecção de linhas era utilizada a equação paramétrica da reta, para a detecção de círculos o procedimento é ainda mais simples, uma vez que o círculo pode ser transferido diretamente para o espaço de parâmetros. A representação paramétrica da equação do círculo $r^2 = (x - a)^2 + (y - b)^2$ é dada pelas equações:

$$x = a + r.\cos\theta \quad y = b + r.\sin\theta \quad (10)$$

Desse modo o espaço de parâmetros para o círculo se encontra no plano R^3 . Como o espaço de parâmetros aumenta com a complexidade da transformada Hough, uma alternativa para reduzir o custo computacional é utilizar um tamanho de raio fixo ou determinar um intervalo pequeno para a variação dos tamanhos dos raios. O algoritmo da **Transformada Hough para detecção de círculos** está descrito em poucos passos em Soltany, Zadeh e Pourreza (2011) e os mesmos são apresentados abaixo.

1. Encontrar as bordas na imagem utilizando um algoritmo de detecção de bordas.

2. Para cada ponto pertencente a borda, desenhar um círculo de raio r com centro no ponto da borda e incrementar no acumulador todas as coordenadas por onde o perímetro do círculo passa.
3. Encontrar um ou vários máximos no acumulador.
4. Mapear todos os parâmetros (r , a , b) encontrados no passo 3 para imagem original.

A transformada Hough para a detecção de linhas pode ser utilizada em uma vasta área de aplicação, como por exemplo em Soares (2017) a mesma é utilizada para a detecção de linhas de plantio com o intuito de estimar a produtividade de uma lavoura, realizar a contagem de plantas, dentre outras atividades. Para a detecção de círculos as aplicações são inúmeras também, uma delas é apresentada em Marroni (2002), que consiste em realizar a detecção de olhos em faces humanas sendo utilizado com um método inicial para posterior reconhecimento facial.

Já a **transformada de Watershed** é um método de segmentação baseado em morfologia matemática. Nesse método uma imagem é considerada uma paisagem topográfica com picos e vales. A elevação de água na paisagem é representada pelo nível de cinza ou pela magnitude do gradiente. Essa transformação decompõe a imagem em bacias hidrográficas (regiões). Para cada mínimo local, uma bacia hidrográfica compreende todos os pontos cuja trajetória de descida mais íngreme termina nesse mínimo. Uma das abordagens de implementação desse algoritmo foi proposta por Vincent e Soille (ANO) e o mesmo é apresentado em Roerdink e Meijster (2001).

4 Transformadas no Domínio da Frequência

Em muitas situações a análise no domínio do tempo de uma imagem não permite visualizar algumas informações que somente o domínio da frequência poderia apresentar, por esse motivo, foram desenvolvidas algumas transformadas que permitem passar uma imagem para o domínio da frequência e a partir da análise de seu espectro, realizar filtragens e obter o resultado desejado. Nesse trabalho serão apresentadas e aplicadas duas Transformadas do domínio da frequência, Wavelet e Fourier.

O processo de conversão para o domínio da frequência utilizando a transformada de Fourier provê uma forma gráfica de extrair informações da imagem, na qual no centro do espectro se encontra a componente DC de frequência zero e essa frequência aumenta conforme o pixel se aproxima mais das bordas da imagem. A representação em forma de equação da transformada de fourier é descrita em P(1997) como

$$\phi(r, s) = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \phi(i, j) e^{-i2\pi(ir + js)/K} \quad (11)$$

Em que r e s são as frequências espaciais, i e j as linhas e colunas da imagem original de tamanho $K \times K$.

O teorema da convolução resume bem a relação entre uma filtragem no domínio do tempo e no domínio da frequência e pode ser descrito da seguinte forma (Pinilla; Alcalá e Ariza, 1997)

1. A transformada de Fourier entre duas funções no espaço de domínio temporal é igual ao produto das transformadas de Fourier de cada uma delas.
2. A transformada inversa de Fourier de um produto de duas funções no domínio da frequência é igual a convolução de suas transformadas inversas.

As propriedades descritas acima são de fundamental importância para o tratamento de imagens no domínio da frequência.

Uma outra ferramenta para análise de imagens no domínio da frequência é a transformada de Wavelet, que consiste na convolução de uma imagem por um sinal conhecido, que se distingue da transformada de Fourier por apresentar semelhança entre o sinal de entrada e o sinal utilizado como base.

Dentre as funções base para a transformada de Wavelet, se destaca o operador Haar, que utiliza sinais quadrados de frequência fixa para extrair informações de interesse. A função de Wavelet é descrita matematicamente em Gonzales (2007) como

$$T = H F H \quad (12)$$

Em que H é a matriz de entrada, F é a função de Wavelet e T o resultado da transformação. A transformada de Wavelet com o operador Haar é amplamente utilizada na compressão de dados, pois a mesma reduz a imagem a aproximações e detalhes.

5 Metodologia

Neste trabalho, todas as técnicas de Processamento Digital de Imagens vistas nas seções anteriores foram implementadas na linguagem Python. Foram implementadas e aplicadas as técnicas de transformações geométricas: translação, rotação e escala; as operações de erosão e dilatação tanto para imagens binárias como para imagens em tons de cinza e utilizando essas operações fundamentais da morfologia matemática foi também implementado o gradiente morfológico; foram aplicadas todas as técnicas de segmentação de imagens descritas na seção 3: Crescimento de Regiões, K-Means, Transformada Hough e transformada de Watershed, bem como também foram aplicadas algumas técnicas de filtragem no domínio da frequência utilizando as transformadas de Fourier e Wavelet.

Foram realizados alguns testes para cada uma dessas técnicas apresentadas com o objetivo de analisar o comportamento das mesmas sob diferentes condições de aplicação. Para as transformações geométricas, foram aplicadas as 3 técnicas citadas tanto para imagens binárias como para imagens em tons de cinza e o resultado foi analisados.

Já para as operações fundamentais da morfologia matemática, erosão e dilatação, foram realizados testes com as mesmas aplicando-se 3 tipos de elementos estruturantes distintos variando seus tamanhos tanto em imagens binárias, como em imagens em tons de cinza. Os resultados foram analisados e comparados entre si.

Ainda utilizando essas operações básicas da morfologia matemática, o gradiente morfológico foi aplicado a uma imagem em tons de cinza, variando o tamanho do elemento estruturante e os resultados foram comparados.

Já com as técnicas de segmentação de imagens foram feitos alguns testes aplicando-se o método de Crescimento de Regiões com inicialização por *click* na imagem; outros teste

foi realizado utilizando a transformada de Watershed cujos marcadores são obtidos a partir da utilização de uma função do *skimage* que encontra os máximos locais. Utilizando segmentação por descontinuidade, foi aplicada a Transformada Hough para detectar linhas e círculos.

Por fim, foram feitos alguns testes utilizando transformadas Wavelet e Fourier do domínio da frequência.

6 Análise e Interpretação dos Resultados

Os resultados obtidos com a aplicação das técnicas descritas são apresentados e discutidos a seguir.

As imagens abaixo representam os resultados da aplicação das transformações geométricas de translação, rotação e escala à uma imagem binária.

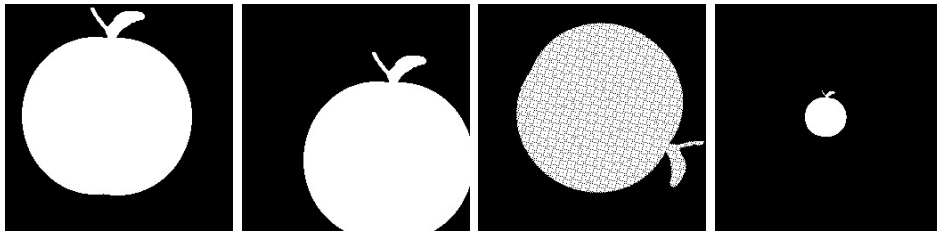


Figura 1 – Aplicações de Transformações Geométricas em uma imagem binária.

A primeira imagem acima é a imagem binária original, a segunda imagem representa o resultado da translação da imagem original por um vetor $[tx, ty] = [50, 50]$, ou seja, a imagem foi deslocada 50 pixels tanto na horizontal como na vertical. A terceira imagem representa a imagem original rotacionada por um ângulo de 90° preservando a escala original da imagem. Já a última imagem da sequência, representa a imagem original com sua escala reduzida, considerando-se o valor mínimo 0,2 e o máximo 0,3.

As mesmas técnicas foram novamente aplicadas, mas agora considerando uma imagem em tons de cinza. Os resultados obtidos são apresentados na figura 2.



Figura 2 – Aplicações de Transformações Geométricas em uma imagem em tons de cinza.

É possível observar, que as operações de translação, rotação e escala por realizarem deslocamentos da imagem, mantendo suas dimensões originais, faz com que parte da informação da imagem original seja perdida, sendo a imagem resultante preenchida com valores nulos, esse aspecto é mais fácil de perceber na aplicação em imagem em tons de cinza, cujo *background* da imagem não é preto, como pôde ser visto na Figura 2.

Os segundo teste realizado consistiu em uma série de aplicações das operações morfológicas de erosão e dilatação. Primeiramente, foi aplicada a erosão para imagens binárias com o elemento estruturante em formato de "cruz" de tamanho 3, 7 e 15. Os resultados dessas aplicações são apresentados abaixo

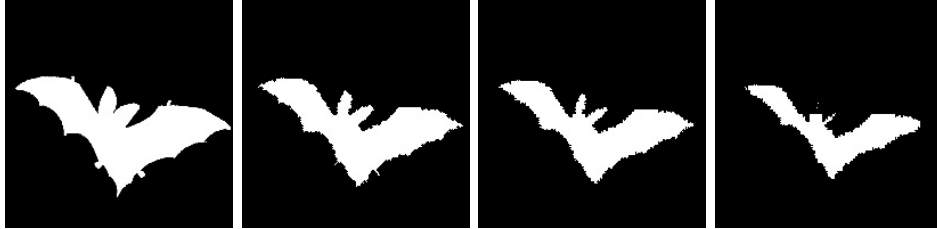


Figura 3 – Erosão em imagem binária com elementos estruturantes de tamanho 3, 7 e 15 na forma de "cruz".

A imagem mais a esquerda representa a imagem original; Na segunda imagem o resultado da aplicação de um elemento estruturante em forma de "cruz" de tamanho 3, faz com que parte das patas do morcego desapareçam e os detalhes em suas asas sumam completamente, já com um elemento estruturante de tamanho 7, as patas do morcego somem totalmente e com um de tamanho 15, a sua orelha direita some quase por completo. Como a operação de erosão tem como objetivo, reduzir o objeto na imagem, vemos que de fato é o que ocorre, no entanto esse procedimento faz com que haja perdas de partes do objetos proporcionais ao tamanho do elemento estruturante que é utilizado.

O mesmo procedimento com um elemento estruturante em formato de "quadrado" foi aplicado a mesma imagem e os resultados podem ser vistos na figura a seguir.

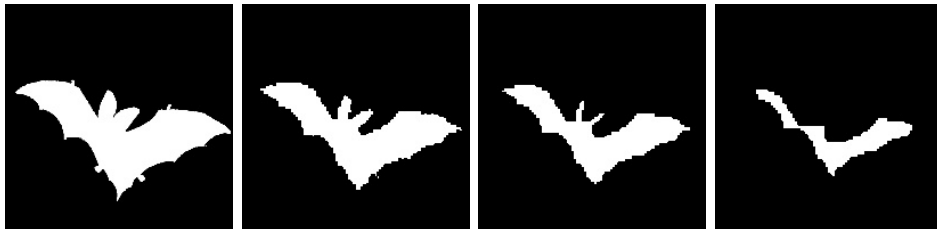


Figura 4 – Erosão em imagem binária com elementos estruturantes de tamanho 3, 7 e 15 na forma de quadrado

Podemos observar que os resultados obtidos são bem semelhantes aqueles obtidos com o elemento estruturante em formato de "cruz", no entanto, a "redução" do objeto na imagem é maior, consequentemente a perda é maior, e essa diferença se acentua ainda mais com a utilização de um elemento estruturante de tamanho maior, como pode ser visto na última imagem, a aplicação da erosão com elemento estruturante em formato de "quadrado" de tamanho 15 removeu completamente a cabeça do morcego, enquanto no teste anterior ainda restou boa parte dela.

O mesmo procedimento foi repetido mais uma vez utilizando agora um elemento estruturante em formato de "L" e os resultados são vistos na Figura 5.

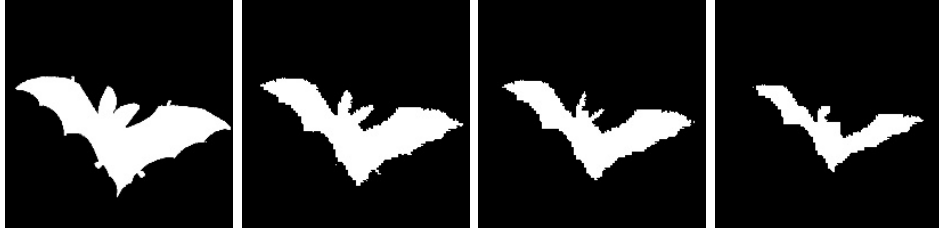


Figura 5 – Erosão em imagem binária com elementos estruturantes de tamanho 3, 7 e 15 na forma de L

Novamente é possível perceber que para os elementos estruturantes de tamanho 3 e 7, os resultados da aplicação da erosão são bem semelhantes, no entanto, quando o tamanho é aumentado consideravelmente a diferença entre as aplicações de elementos estruturantes distintos fica mais evidente, resultando em imagens com objetos de formas distintas.

Seguindo um raciocínio análogo, foram realizados também testes com a operação de dilatação nessa mesma imagem, os resultados para a aplicação de elementos estruturantes da forma de "cruz", "quadrado" e "L" são apresentados abaixo respectivamente.

Para elementos estruturantes em forma de "cruz" de tamanhos 3, 7 e 15 respectivamente.

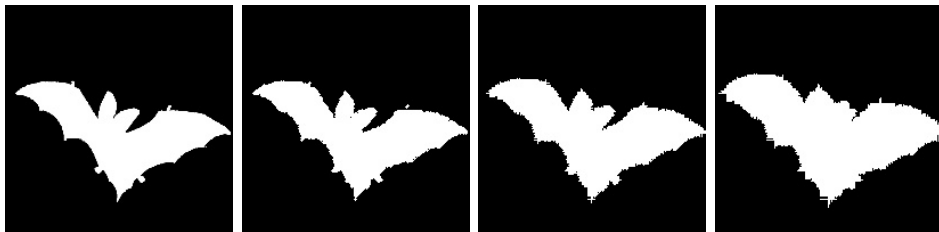


Figura 6 – Dilatação em imagem binária com elementos estruturantes de tamanho 3, 7 e 15 na forma de cruz

Para elementos estruturantes em forma de "quadrado" de tamanhos 3, 7 e 15 respectivamente.

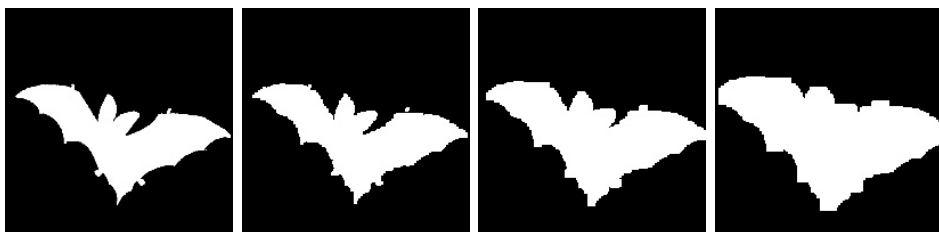


Figura 7 – Dilatação em imagem binária com elementos estruturantes de tamanho 3, 7 e 15 na forma de quadrado

Para elementos estruturantes em forma de "L" de tamanhos 3, 7 e 15 respectivamente.



Figura 8 – Dilatação em imagem binária com elementos estruturantes de tamanho 3, 7 e 15 na forma de L

A partir desses resultados obtidos com a dilatação nessa imagem binária, é possível observar que essa operação tende a "expandir" o objeto de interesse na imagem tornando-o uma estrutura convexa e além disso independente do elemento estruturante o resultado é bem semelhante.

Como visto na seção 2, as operações de erosão e dilatação para imagens em tons de cinza são diferentes das aplicadas em imagens binárias. Abaixo são apresentados alguns resultados obtidos da aplicação da erosão para imagens em tons de cinza com a mesma variação no tamanho dos elementos estruturantes aplicadas as imagens binárias.

Para elementos estruturantes no formato de "cruz"



Figura 9 – Erosão em imagem em tons de cinza com elementos estruturantes em formato de "cruz" de tamanhos 3, 7 e 15.

Para elementos estruturantes no formato de "quadrado".



Figura 10 – Erosão em imagem em tons de cinza com elementos estruturantes em formato de "quadrado" de tamanhos 3, 7 e 15.

Para elementos estruturantes no formato de "L".



Figura 11 – Erosão em imagem em tons de cinza com elementos estruturantes em formato de "L" de tamanhos 3, 7 e 15.

Observa-se que o elemento estruturante no formato de "L" preservou melhor a forma do objeto na imagem original, isso ocorre porque esse elemento estruturante é composto por uns na primeira coluna e na última linha, o que faz com que os tons de cinza centrais na imagem não se propaguem para as bordas, mantendo-as preservadas em sua maioria. Além disso, as bordas da imagem se tornaram pretas devido ao preenchimento inicial da imagem com zeros nas bordas para posterior aplicação do procedimento de convolução.

Analogamente, para a operação de dilatação tem-se os seguintes resultados

Para elementos estruturante no formato de "cruz" de tamanhos 3, 7 e 15.



Figura 12 – Dilatação em imagem em tons de cinza com elementos estruturantes em formato de "cruz" de tamanhos 3, 7 e 15.

Para elementos estruturantes no formato de "quadrado" de tamanhos 3, 7 e 15.



Figura 13 – Dilatação em imagem em tons de cinza com elementos estruturantes em formato de "quadrado" de tamanhos 3, 7 e 15.

Para elementos estruturantes no formato de "L" de tamanhos 3, 7 e 15.



Figura 14 – Dilatação em imagem em tons de cinza com elementos estruturantes em formato de "L" de tamanhos 3, 7 e 15.

É possível observar a partir dos resultados obtidos da aplicação das operações de erosão e de dilatação em tons de cinza que de fato como descrito na seção 2, a primeira aumenta a quantidade de pixels de baixos níveis de cinza, provocando um escurecimento na imagem, já a dilatação têm o efeito contrário, aumentando a quantidade de pixels de alto níveis de cinza, deixando a imagem mais clara.

Como mencionado na seção 2, as operações de erosão e dilatação são amplamente combinadas na composição de outras técnicas de morfologia matemática e uma dessas técnicas bastante conhecida e útil na detecção de bordas é o gradiente morfológico também descrito em 2. A seguir são apresentados os resultados do gradiente morfológico aplicado a uma imagem em tom de cinza. O elemento estruturante utilizado nesse teste foram "quadrados" de tamanhos 3, 5 e 7.

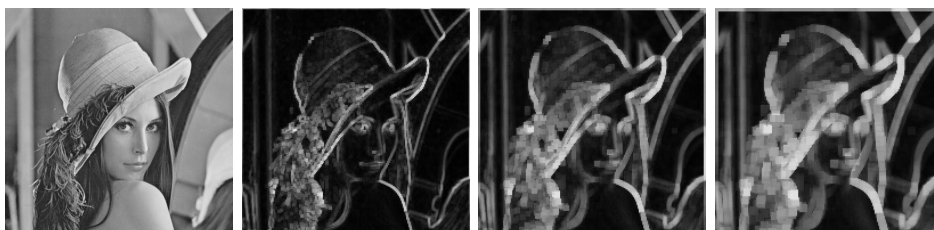


Figura 15 – Aplicação do gradiente morfológico com elementos estruturantes quadrados de tamanhos 3, 5 e 7

Na aplicação de gradiente morfológico, quanto maior o tamanho do elemento estruturante menos nítidas e mais grossas serão as bordas encontradas, no exemplo acima, o elemento estruturante de tamanho 3 foi o que gerou bordas bem mais delimitadas na imagem que os demais tamanhos.

Dos experimentos realizados com aplicações de técnicas de segmentação de imagens, os resultados são apresentados a seguir.

O método de Crescimento de Região foi aplicado na segmentação de uma imagem binária e em uma imagem em tom de cinza. As imagens abaixo representam algumas iterações do algoritmo na segmentação da imagem binária.

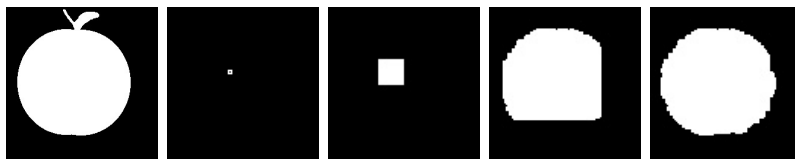


Figura 16 – Etapas de execução do método de Crescimento de Regiões em uma imagem binária.

As imagens acima são resultados da segmentação pelo algoritmo de segmentação Crescimento de Regiões com inicialização com click na imagem.

O mesmo método foi aplicado em uma imagem em tom de cinza, com o mesmo procedimento de inicialização com click na imagem, o resultado obtido pode ser visto na Figura 17.

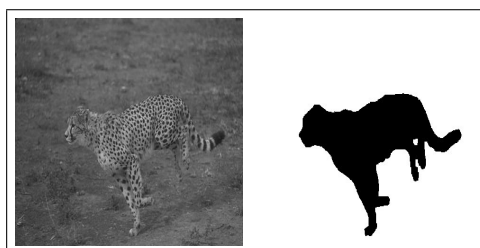


Figura 17 – Imagem em tom de cinza segmentada pelo algoritmo de Crescimento de Regiões

Outro método de segmentação de imagens que foi aplicado a uma imagem em tom de cinza, foi o método *K-Means*. O resultado de sua aplicação com o parâmetro k , variando de 1 a 3 pode ser visto a partir da figura abaixo.

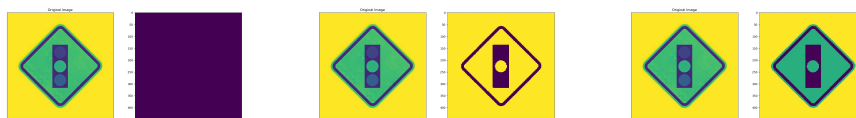


Figura 18 – Placa segmentada pelo método *K-Means* com k igual a 1, 2 e 3 respectivamente

Os resultados obtidos permitem concluir que o valor de k determina em quantas regiões a imagem será segmentada. Um aspecto importante a se ressaltar é que se esse valor de k exceder a quantidade de regiões na imagem, a mesma será segmentada no máximo na quantidade de regiões que ela contém.

A esquerda está a imagem original da placa e a direita a segmentação dada pelo método *K-means* para $k = 1, 2$ e 3 respectivamente.

Passando para os métodos de segmentação baseados em descontinuidade, foram realizados testes com a Transformada Hough em imagens tons de cinza para detecção de

linhas e círculos.

A figura abaixo representa uma imagem em tom de cinza da bandeira do Brasil e o resultado da aplicação da transformada Hough para a detecção de linhas nessa imagem.

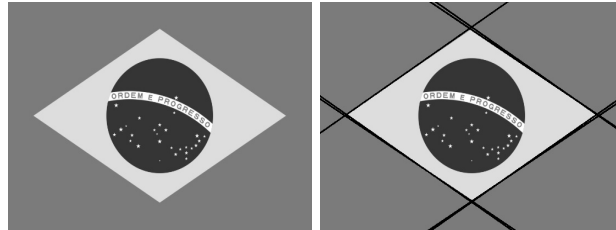


Figura 19 – Linhas detectadas na bandeira do Brasil pela Transformada Hough

O resultado da aplicação da Transformada Hough para detecção de círculos é apresentado abaixo

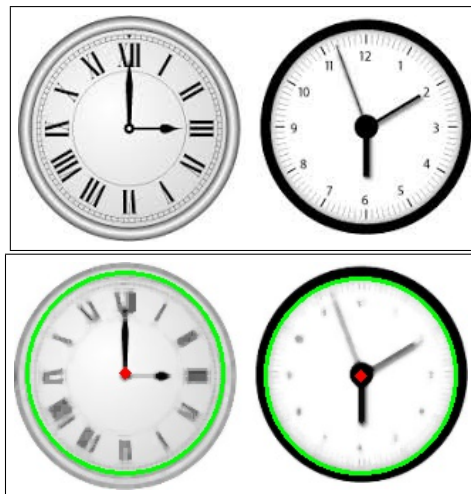


Figura 20 – Círculos detectados em imagens de relógios pela Transformada Hough

Ainda sobre segmentação baseada em descontinuidade, a Transformada de Watershed foi aplicada a uma imagem em tons de cinza para a segmentar objetos nela contidos. As imagens na sequência representam, respectivamente, a imagem original, a imagem segmentada para posterior aplicação da transformada e o resultado final da aplicação.

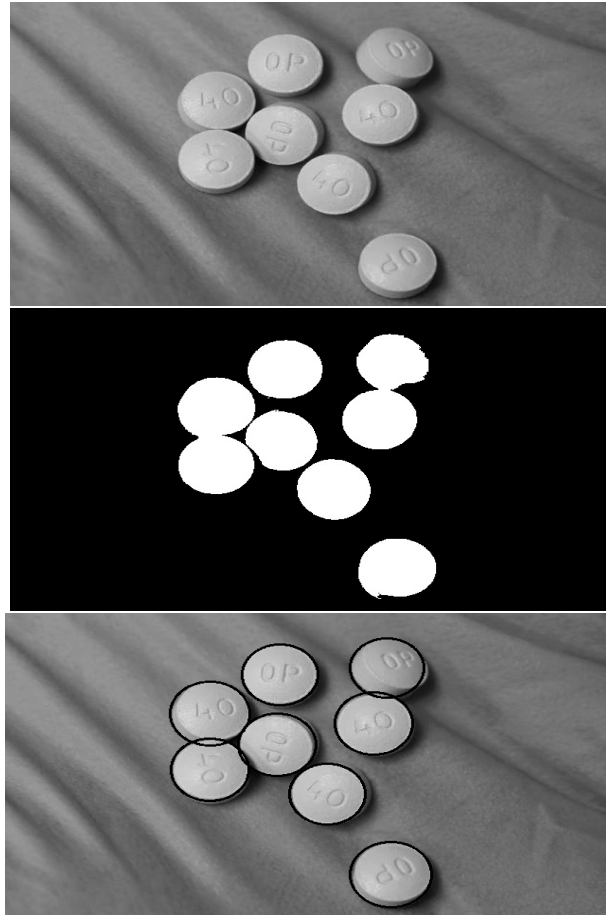


Figura 21 – Aplicação da Transformada *Watershed*

Terminados os experimentos com os métodos de segmentação de imagens, foram feitos alguns testes com as transformadas no domínio da frequência, Wavelet e Fourier. Primeiramente, a Transformada Haar foi aplicada a uma imagem em tons de cinza e os detalhes obtidos nas 3 primeiras decomposições foram analisados. A Figura 22 apresenta a imagem original e esses detalhes obtidos com essa transformada são apresentadas.

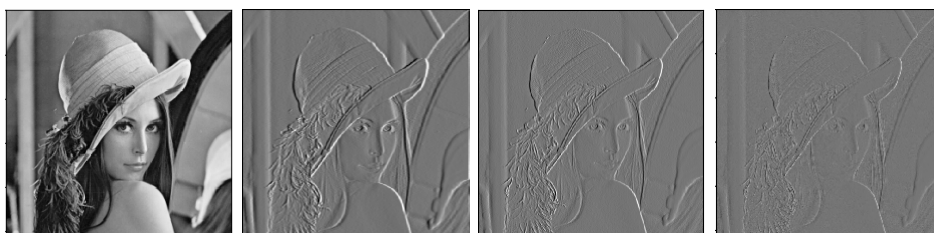


Figura 22 – Detalhes obtidos nas 3 primeiras decomposições da imagem.

É possível observar a partir desses resultados que os detalhes obtidos da transformada de Wavelet a cada nível vão ficando cada vez mais suaves, o que é natural já que a imagem está sendo aproximada recursivamente, resultando na perda de detalhes a cada aproximação.

Esse resultado, além disso, representa 3 subprodutos da imagem original obtidos a partir da filtragem passa-alta dessa imagem, que representam os coeficientes com informações a respeito das linhas horizontais, verticais e diagonais. Após esse processo, a mesma passa por uma filtragem passa-baixas com o intuito de preservar as características da imagem original e por fim, passam por um processo chamado *downscaling*.

Por fim, o último teste consistiu em aplicar as transformadas Wavelet e Fourier e analisar a primeira e a segunda componente.

A aplicação da transformada de Wavelet em uma imagem em tons de cinza resulta em uma tupla, sendo o primeiro elemento da tupla uma aproximação para a imagem e o segundo elemento da tupla é uma outra tupla contendo 3 detalhes obtidos da imagem, tanto na horizontal como na vertical e na diagonal.

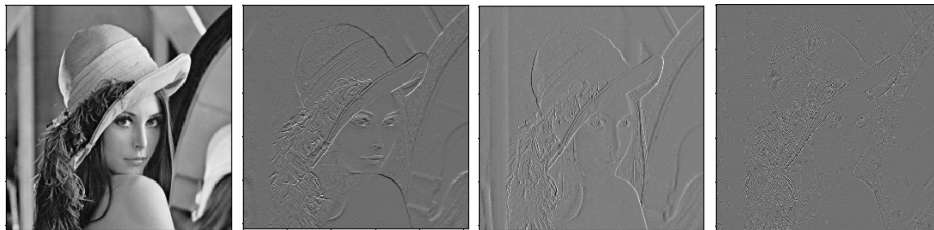


Figura 23 – Detalhes na horizontal, vertical e diagonal obtidos pela aplicação da Wavelet

A aplicação da transformada de Fourier em uma imagem em tons de cinza resultando na magnitude calculada a partir das componentes real e imaginária da mesma é apresentada da Figura 24.



Figura 24 – Espectro de frequência da imagem.

No espectro da frequência, operações de filtragem são bem mais simples de serem realizadas pelo motivo explicado no Teorema da Convolução, convoluções são substituídas por simples multiplicações ponto a ponto. Os filtros passa-baixas e passa-altas são imagens do tamanho da imagem original mais claras no centro e escuras nas bordas e mais escuras no centro e claras nas bordas respectivamente. As imagens abaixo representam a aplicação de um filtro passa-baixas gaussiano e um filtro passa-altas laplaciano no domínio da frequência.



Figura 25 – Aplicações dos filtros gaussiano e laplaciano na imagem no domínio da frequência.

Percebe-se que o resultado é bem semelhante a filtragem no domínio do tempo, mas como mencionado a operação não apenas tem um custo computacional mais baixo como o resultado da aplicação de um filtro passa-altas na frequência já é uma imagem binária, ainda há o ganho de não necessitar utilizar uma outra operação para limiarizar a imagem resultante.

7 Conclusão

A partir da análise dos resultados obtidos foi possível concluir que embora haja muitas técnicas elaboradas para um mesmo fim, devido ao princípio de funcionamento de cada uma, é possível fazer uma escolha mais apropriada para aplicação dependendo de fatores inerentes a própria imagem com a qual se deseja trabalhar, como luminosidade e intensidade dos pixels por exemplo. Para se atingir os resultados esperados com a aplicação de técnicas dependentes diretamente da escolha de parâmetros, como a Transformada Hough para a detecção de linhas e círculos, foi necessário a compreensão plena dos mesmos e essa escolha varia de acordo com o objetivo do usuário. Outro fator interessante a se ressaltar é que embora seja possível realizar as mesmas transformações do domínio do tempo no domínio da frequência, esse procedimento nem sempre é útil devido ao custo de se realizar a transformada inversa, mas existem casos em que a única maneira de conseguir visualizar a informação desejada é aplicando filtragens no domínio da frequência e que compensa a redução do custo computacional na operação de filtragem, já que se trabalha com multiplicações ponto a ponto ao invés de convoluções.

Quanto as técnicas de segmentação baseadas em descontinuidade, de acordo com a escolha da semente o resultado final obtido pode ser diferente, convergindo em um tempo maior ou menor. E pra finalizar, a operação de gradiente morfológico é preferida em muitos casos devido a sua simplicidade e por a sua saída já ser uma imagem binária, não necessitando da aplicação adicional de técnicas de limiarização.

8 Referências

- ADAMS, R.; BISCHOF, L.. Seeded region growing. IEEE Trans. Pattern Analysis and Machine Intelligence 16, 641-647, 1994.
- DUDA, R. O.; HART, P. E.. Use of the Hough transformation to detect lines and curves in pictures, Communications of the Association for Computing Machinery 15 :11-15, 1972.
- GONZALEZ, R. C.; WOODS, R. E. 3rd. Edition. Digital Image Processing, Prentice Hall, 2007. 976p.

- JYOTI, Y.; SHARMA, M. A Review of K-Mean Algorithm. *Int. J. Eng. Trends Technol.* 4, 2972–2975, 2013.
- MARRONI, L. S.. Aplicação da Transformada de Hough para localização dos olhos em faces humanas. Dissertação (Mestrado): Escola de Engenharia de São Carlos da Universidade de São Paulo: USP, 2002.
- PINILLA, C.; ALCALÁ, A.; ARIZA, A. Filtrado de Imágenes en el Domínio de La Frecuencia. Departamento de Ingeniería Cartográfica, Geodésica y Fotogrametría, Universidad de Jaén. 1997.
- ROERDINK, J.B.T.M.; MEIJSTER, A. The Watershed Transform: Definitions, Algorithms and Parallelization Strategies. *Fundamenta Informaticae*, 41, 187-228, 2001.
- SÁNCHEZ, S. A. G.; RUIZ, C. A.; MEZA, C. E. A. GEOMETRIC TRANSFORMATIONS AND ITS APPLICATION IN DIGITAL IMAGES. *Jornal of Theoretical and Applied Information Technology*, v. 60, n. 1, 2014.
- SHARMA, P.; DIWAKAR, M.; CHOUDHARY, S.. Application of Edge Detection for Brain Tumor Detection, *International Journal of Computer Applications*, vol.58, no.16, pp 21-25, 2012.
- SOARES, G. A. Identificação de linhas de plantio por meio da Transformada de Hough. Monografia (Graduação): Faculdade de Computação da Universidade Federal de Uberlândia, Minas Gerais, 2017.
- SOLTANY, M.; ZADEH, S. T.; POURREZA, H. R. Fast and Accurate Pupil Positioning Algorithm Using Circular Hough Transform and Gray Projection, in *International Conference on Computer Communication and Management (CSIT)*, Vol. 5, pp. 556- 561, 2011.
- VICENT, L.; SOILLE, P. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):583–598, 1991.
- ZHUXUN, Z.; YIREN, M.; JIANQIQ, Z. Image Segmentation Based on Hough Transformation. Commission III of ISPRS. Wuhan Technical Unlversty of surveying and mapping P.R.China, 2009.