



UNIVERSIDAD NACIONAL DEL LITORAL
Facultad de Ingeniería y Ciencias Hídricas

Processamento Digital de Imagens

Entrega IV

Aluna: Carla de Oliveira Barden

Santa Fe, Mayo de 2020

Introdução

A análise de imagens através do domínio frequencial auxilia, computacionalmente, em uma série de inferências sobre suas características (no caso deste trabalho, a orientação de um texto). O domínio frequencial também facilita a criação e a aplicação de *filtros* à imagem, sendo esses úteis para a detecção de bordas ("passa-alta", permite a passagem de altas frequências), diminuição de ruído ("passa-baixa", permite a passagem de frequências baixas), "passa-banda" ou "rejeita-banda", entre outros.

A imagem é transformada do domínio do espaço para o domínio da frequência através da **Transformada Discreta de Fourier (TDF)**, que decompõe a imagem de entrada nas frequências que a geram. O resultado da TDF é uma matriz no domínio complexo, que pode ser interpretada como o resultado do produto interno entre as "amplitudes" e "fases" da imagem original com as *imagens base* da transformada. Dentre as diversas propriedades da TDF, é importante ressaltar que todas as alterações na transformada (convolução, correlação, rotação, etc) refletem-se na imagem quando esta é convertida novamente ao domínio do espaço, através da *Transformada Discreta de Fourier Inversa*.

Na prática, em Processamento de Imagens, é utilizada a magnitude (Espectro de Fourier - domínio real) da imagem, com a frequência (0,0) centralizada e aplicando-se a função de *logaritmo* para comprimir a "energia" da frequência (0,0) (que representa o brilho médio da imagem) e expandir outras baixas frequências importantes, enquanto a fase detém toda a informação de contornos e bordas da imagem.

Problema Proposto

Desenvolva uma aplicação de pré-processamento para OCR (*Optical Character Recognition*). A imagem de entrada é um texto escaneado (pode-se usar a imagem 1(a) e a imagem 1(b)). Deve-se identificar se o texto está rotacionado e, se o estiver, corrigir sua orientação.

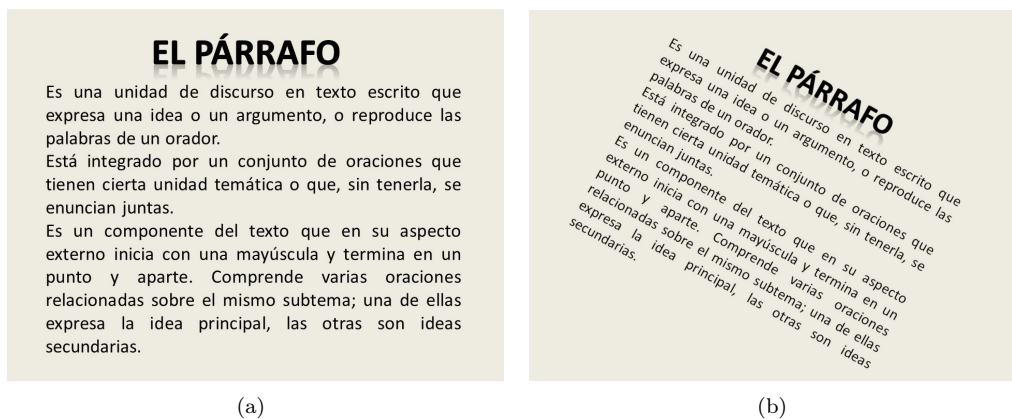


Figura 1: Imagens Fornecidas

Solução Proposta

A solução proposta encontra-se em <https://tinyurl.com/y74ztj7c>. Para executá-la, é necessário passar como parâmetro o caminho da imagem a ser analisada.

Metodologia

A primeira parte da solução adotada consistiu em gerar a TDF da imagem de entrada, obtendo-se sua fase e magnitude (figura 2) - já centralizada na frequência(0,0), com a função de *logaritmo* aplicada e normalizada. Observando-se a magnitude(2) de cada imagem, é possível observar as direções principais, onde há mais *energia*, e a rotação na transformada causada pela rotação na imagem.

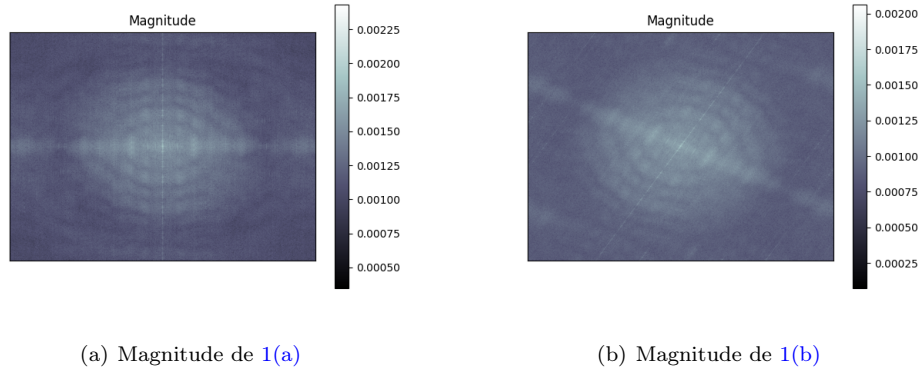


Figura 2: Magnitudes

A seguir, utilizou-se uma Transformação de Limiar em cada magnitude. Outra abordagem cogitada foi o cálculo do vetor gradiente da TDF da imagem, porém, mesmo após a aplicação de um filtro, a frequência $(0,0)$ permaneceu com maior energia. Assim, o vetor gradiente não foi gerado na direção necessária. Então, aplicando-se a Transformação de Limiar, foi possível perceber que há algumas "linhas" principais no espectro, como demonstrado em 3.

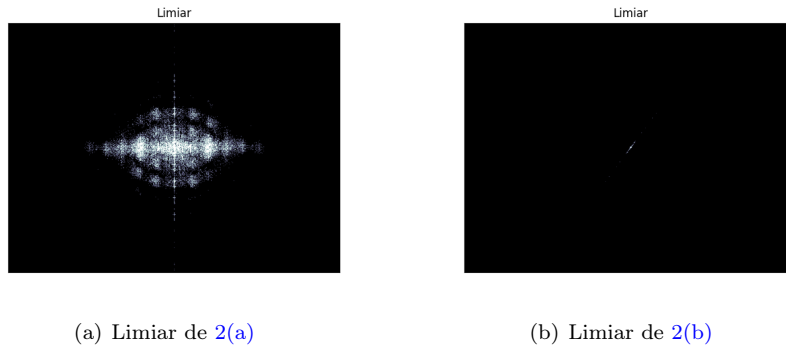


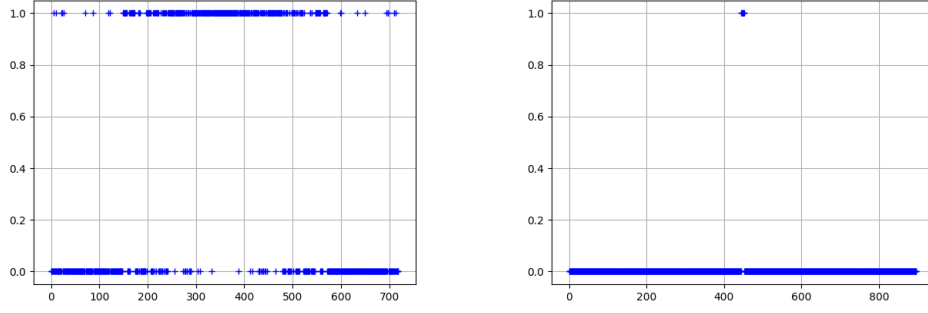
Figura 3: Limiares

Observando-se a figura 3(a), é possível notar uma "nuvem" de alta energia. Isto ocorre porque é difícil encontrar um valor de limiar que se aplique às duas magnitudes, o que explica o porquê de a solução apresentada aqui não ser totalmente genérica. Ainda assim, é possível perceber que o espectro limiarizado da imagem 1(a) possui uma linha divisória vertical ao centro (indicando que há alta variação horizontal de frequências). Já o espectro limiarizado de 1(b), apresenta inclinação nesta linha, indicando a rotação do texto.

Após a limiarização, foram gerados os perfis de intensidade da coluna central de cada imagem. Observando-se os perfis de intensidade apresentados em 4, é possível perceber que o perfil de intensidade da imagem que está na orientação correta possui um grande número de pontos "1", enquanto no vetor de perfil de intensidade da imagem rotacionada, acontece o oposto.

Assim, a estratégia usada foi:

1. Caso o perfil de intensidade da coluna central apresentar muitos elementos iguais à 1, significa que a imagem está na orientação correta.
2. Caso o perfil de intensidade da coluna central apresentar poucos (menos que 10) elementos iguais à 1, significa que a imagem está rotacionada.



(a) Perfil de Intensidade da coluna central de 3(a) (b) Perfil de Intensidade da coluna central de 3(b)

Figura 4: Perfis de Intensidade

Para obter-se o ângulo de rotação no caso 2, considerando que o Perfil de Intensidade foi gerado a partir da coluna central da imagem, o procedimento adotado foi:

- A partir das coordenadas centrais (x, y) da imagem (ponto médio), geraram-se dois outros Perfis de Intensidade verticais, da forma: $vp1 = x - 5$ e $vp2 = x + 5$;
- Nos dois perfis, procuraram-se as coordenadas onde os elementos eram maiores que zero;
- A partir de um par de coordenadas de cada um dos perfis, foi gerado um vetor ($V = B - A$, sendo V o vetor, A, B pontos no plano, e vx, vy as componentes do vetor - 5);
- Após normalizar as componentes do vetor, para torná-lo unitário, foi realizado o cálculo do ângulo $alpha = arctan(vy/vx) * 180/PI$, obtendo-se o ângulo em graus.
- Por fim, a imagem foi rotacionada, usando-se $alpha * -0.5$.

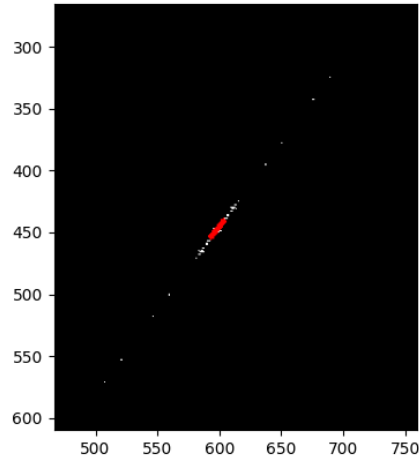


Figura 5: Vetor para se obter ângulo em vermelho, ampliado.

Resultados

A aplicação foi capaz de identificar a orientação (e corrigi-la, quando necessário - 6) nos dois textos fornecidos . O ângulo de rotação, em graus, é exibido no *shell* onde a aplicação é invocada. Exemplo de execução: `python3 000-ocr.py -i imgs/parrafo1.jpg`



Figura 6: Resultados apresentados pela aplicação.

Referências

Videoaulas da Disciplina;
Processamento Digital de Imagens - Rafael C. Gonzalez e Richard E. Woods - 3a edição.