



UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
DEPARTAMENTO DE COMPUTAÇÃO

COMP0432 - PROCESSAMENTO DE IMAGENS

GRUPO 09

Willian Mota Oliveira - 201800017520

Yohan Alexander Dantas de França - 201800017208

Relatório do Projeto Reconhecedor de Caracteres em PBM

São Cristóvão - SE

Fevereiro de 2020

SUMÁRIO

- 1. INTRODUÇÃO**
- 2. OBJETIVOS**
- 3. METODOLOGIA**
- 4. RESULTADOS E DISCUSSÕES**
- 5. CONCLUSÕES**

1. INTRODUÇÃO

Imagens digitais são frequentemente armazenadas de forma computacional como matrizes bidimensionais de valores que indicam intensidade luminosa, em uma escala de cinza, por exemplo, podemos usar um tipo char de 1 byte para termos 256 tons diferentes.

O processamento de imagens digitais se dá por meio da manipulação dos pixels, ou seja, cada célula que compõe a matriz dessa imagem. Existem diversos tipos de operações possíveis para as transformações dessa matriz, desde operações ponto a ponto, como também o uso de filtros ou máscaras n-dimensionais que são aplicadas por meio de uma convolução com a imagem.

Neste trabalho iremos usar algumas dessas operações para destacar caracteres em imagens de texto, que são imagens essencialmente monocromáticas, ou seja, tem apenas dois tons diferentes, preto ou branco.

As imagens utilizadas como forma de validar o processamento serão do formato PBM. O formato PPM é usado para armazenamento de imagens coloridas, que podem ser salvas tanto em arquivos com o padrão ASCII quanto arquivos com o padrão binário, o padrão binário é chamado PBM. O diferencial do formato é que ele é muito fácil na questão de acesso e manipulação, sendo assim excelente para entender o funcionamento da aplicação de filtros em imagens, ou até mesmo a criação e teste de filtros ou efeitos próprios.

O padrão de um arquivo PPM é o seguinte (cada número representa uma linha):

1. Dois caracteres de identificação do tipo de arquivo.
 - P1 no caso do PBM
2. Largura e altura (necessariamente nesta ordem!), sendo ambas um número inteiro, representando pixels e separadas por espaço.
 - 512 512 (por exemplo)
3. O valor máximo de cada cor, representado por um número inteiro. O mesmo deverá ser maior que 0 até 255 (inclusive).
 - Ignorado no formato PBM pois sempre é binário
4. O valor para cada cor do plano RGB, sendo o menor valor possível 0 e o maior valor o número definido na linha anterior.
 - 0 ou 1 no caso do formato PBM

2. OBJETIVOS

Este trabalho tem como objetivo a construção de um programa que recebe como entrada um arquivo que contém uma imagem PBM, a qual contém um texto com uma ou mais colunas, e como saída o programa retorna quantas linhas e palavras o texto contém, além de gerar uma imagem PBM com as palavras do texto circunscritas por um retângulo.

Será assumido durante a execução deste trabalho que o texto da imagem de entrada está todo escrito na fonte Arial e que o mesmo tamanho de fonte é usado em todo o texto, além disso a imagem pode conter espaços vazios e ruído sal e pimenta, que serão removidos pela aplicação de filtros.

Este programa é o primeiro passo no reconhecimento de caracteres (OCR), uma vez definidas as áreas da imagem que contém palavras, o passo seguinte seria isolar e reconhecer as letras.

3. METODOLOGIA

Para a implementação do programa foi escolhido o uso da linguagem C++, pois a mesma une a velocidade do código em C, necessária para as operações em matrizes, com a abstração do paradigma de orientação a objetos, que facilita a visualização das operações nas imagens.

Inicialmente, a estrutura de uma imagem PBM foi abstraída em uma classe Image, que contém os atributos inteiros de largura e altura em pixels, e uma matriz de inteiros com os valores binários de cada pixel. Dessa forma, todas as imagens que são processadas pelo programa são instâncias dessa classe.

Com o objeto no qual seriam realizadas as operações construído, foi necessário implementar as funções de escrita e leitura em arquivos PBM. As funções cumprem o papel de mapear a imagem do disco rígido para a memória RAM, onde serão feitas todas as operações sobre a imagem.

O primeiro passo estabelecido para o pré-processamento da imagem, para tornar a circunscrição das palavras mais simples, foi a remoção de eventuais ruídos do tipo sal e pimenta. Para isso, utilizamos os conceitos de operadores morfológicos e implementamos os filtros de dilatação e erosão como métodos do objeto Image.

Como o ruído sal é basicamente pontos brancos em regiões pretas, e o ruído pimenta são pontos pretos em regiões brancas, a aplicação de uma erosão seguida por uma dilatação, traz bons resultados na remoção do ruído. Pois, a erosão filtra os ruídos pimenta menores que o tamanho da máscara, e a dilatação fecha os buracos causados pelo ruído do tipo sal.

Para os casos de testes as máscaras foram implementadas no formato 3 x 3, e consistem no uso de operadores lógicos (bit-wise) disponível na linguagem C++. Porém, para casos arbitrários foi implementada a abstração de uma máscara como uma classe Mask, que contém os atributos inteiros de largura e altura da máscara em pixels. Dessa forma, uma instância da mesma com formato n x m é então aplicada como uma convolução com a instância de uma imagem i x j.

Na segunda etapa do processamento da imagem foi necessário escolher uma estratégia para a seleção das palavras. Para isso, imaginamos cada linha de texto como um bloco único, onde seriam marcadas duas retas, uma que indica o início de uma linha de texto, e outra que indica o final. Dessa forma, o texto é varrido de cima para baixo delimitando as linhas contidas no texto.

Com as linhas de texto delimitadas, bastava apenas imaginar a passagem de um cursor (linha vertical) entre os delimitadores. É importante lembrar que para circunscrever as palavras bastam apenas dois pontos, estes que obrigatoriamente seriam simétricos um ao outro, para a delimitação do retângulo ao redor da palavra.

Com isso, a implementação seguiu deste conceito simples, regiões com espaçamento pequeno (palavras) eram identificadas pela passagem do cursor pelo texto, e quando o cursor identificava uma região com um espaçamento maior que um limiar, aquela região era identificada como um espaço.

Dessa forma, ficava mais fácil marcar os pontos dos quais seriam derivados os retângulos que cumprem o papel de destacar as palavras. Pois, para a marcação do ponto final, que é simétrico em relação ao começo da palavra, bastava voltar ao último ponto onde havia uma região de texto, depois de ser identificada uma região de espaçamento.

É importante lembrar que a marcação das palavras com um retângulo será feita na imagem original. Qualquer tipo de mecanismo de pré-processamento da imagem para facilitar o destaque das palavras, tem os resultados intermediários descartados.

Assim, é garantido que com as coordenadas corretas de todos os pontos simétricos, podemos delimitar todas as regiões que representam palavras. Além disso, seguindo a lógica do programa podemos facilmente contar quantas palavras existem no texto, além da quantidade de linhas.

4. RESULTADOS E DISCUSSÕES

Foi possível notar no processamento dos casos de teste, que apesar de assumidas as condições iniciais para a imagem, a mesma fonte e tamanho durante todo o texto, a robustez do programa foi afetada pela diferença nas resoluções de pontos (dpi) utilizadas na construção da imagem de texto.

Para resoluções menores, a aplicação de máscaras 3 x 3 para a remoção de ruído sal e pimenta, acabava destruindo os caracteres presentes na imagem, prejudicando assim o reconhecimento e circunscrição dos mesmos nas próximas etapas do programa.

Outro problema encontrado no pré-processamento, foi que ao utilizar a abstração de uma máscara arbitrária, a otimização do programa diminuía drasticamente, resultante em maiores tempos de processamento.

Além disso, em textos de mais de uma coluna a execução do programa pode ser afetada, pois seguindo a lógica de marcação de linhas horizontais, as linhas de texto das colunas podem estar desalinhadas em relação uma a outra, o que acaba resultando em uma falha na circunscrição das palavras.

5. CONCLUSÕES

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt. Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim ad minima veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur? Quis autem vel eum iure reprehenderit qui in ea voluptate velit esse quam nihil molestiae consequatur, vel illum qui dolorem eum fugiat quo voluptas nulla pariatur?

At vero eos et accusamus et iusto odio dignissimos ducimus qui blanditiis praesentium voluptatum deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non provident, similique sunt in culpa qui officia deserunt mollitia animi, id est laborum et dolorum fuga. Et harum quidem rerum facilis est et expedita distinctio. Nam libero tempore, cum soluta nobis est eligendi optio cumque nihil impedit quo minus id quod maxime placeat facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum necessitatibus saepe eveniet ut et voluptates repudiandae sint et molestiae non recusandae. Itaque earum rerum hic tenetur a sapiente delectus, ut aut reiciendis voluptatibus maiores alias consequatur aut perferendis doloribus asperiores repellat.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt. Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim ad minima veniam,