

# MC920 - Trabalho 3: Operadores Morfológicos

Naomi Takemoto  
RA 184849  
n184849@dac.unicamp.br

## I. INTRODUÇÃO

Morfologia matemática é um ramo do processamento de imagens baseada na teoria de conjuntos. Inicialmente essas técnicas eram aplicadas apenas em imagens binárias, com diversos objetivos, tais como encontrar regiões conexas, busca de padrões, delimitação de fechos convexos, etc.

## II. OBJETIVOS

Neste trabalho, utilizaremos operadores morfológicos, como fechamento, abertura, dilatação, erosão, etc para processar uma imagem binária, que consiste em uma fotocópia de um artigo científico. O objetivo é estimar algumas estatísticas básicas como número de linhas, palavras, figuras, etc.

## III. TEORIA

## IV. EXPERIMENTOS

Os passos do trabalho, assim foram definidos na especificação do trabalho [1].

- 1) Dilatação da imagem original com um elemento estruturante de 1 pixel de altura e 100 pixels de largura;
- 2) Erosão da imagem resultante com o mesmo elemento estruturante do passo (1);
- 3) dilatação da imagem original com um elemento estruturante de 200 pixels de altura e 1 pixel de largura;
- 4) Erosão da imagem resultante com o mesmo elemento estruturante do passo (3);
- 5) Aplicação da intersecção (AND) dos resultados dos passos (2) e (4);
- 6) Fechamento do resultado obtido no passo (5) com um elemento estruturante de 1 pixel de altura e 30 pixels de largura;
- 7) Aplicação de algoritmo para identificação de componentes conexos sobre o resultado do passo (6);
- 8) para cada retângulo envolvendo um objeto, calcule
  - a) razão entre o número de pixels pretos e o número total de pixels (altura  $\times$  largura);
  - b) razão entre o número de transições verticais e horizontais branco para preto e o número total de pixels pretos;
- 9) criação de uma regra para classificar cada componente conexo, de acordo com as medidas obtidas no passo (8), como texto e não texto.
- 10) aplicação de operadores morfológicos apropriados para segmentar cada linha do texto em blocos de palavras. Coloque um retângulo envolvendo cada palavra na imagem original. Calcule o número total de linhas de texto e de blocos de palavras na imagem.

### A. Passo 1

310

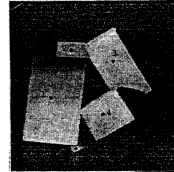


Fig. 4. Range image of an AMBIGUOUS scene and the corresponding graph.

IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION, VOL. 7, NO. 3, JUNE 1991

sensory feedback is carried out in a local reflexive mode rather than in a planned mode with one exception, that is, when a pathological state is detected.

3) States: This is a finite set of states describing the environment of the Turing machine as perceived by the sensors. If new sensors are added, the set of states is partitioned to describe the scene as perceived by the additional sensors. For example, if a sensor capable of determining the "touch" relations of objects in the scene is added, then the set of five states, can be partitioned (a finer partition) to describe both the "touch" and "on-top-of" relations. The states of the machine are:

Empty	If there are no vertices in the diagram, i.e., an empty diagram.
Dispersed	If there no edges in the diagram, i.e., a null diagram (Fig. 2).
Overlapped	If there are at least two vertices connected with an edge (Fig. 3).
Ambiguous	If there is one or more directed cycles in the diagram (Fig. 4).
Unstable	This category is not tested by the analysis of the graph but through analysis of the contact point/line of the object with the support plane. If this contact is a point or a line, it is classified as unstable. See Fig. 5.

Figura 1: Fotocópia da imagem a ser processada

Após execução do passo 1, descrito anteriormente obteve-se o resultado da figura:



Figura 2: Resultado do passo 1.

O resultado do uso de um elemento estruturante de altura igual a 1 pixel de largura 100 é ressaltar o a dimensão horizontal dos elementos da imagem. Como pode ser visto acima, alguns elementos relativamente pequenos, como as notações sobre as arestas to grafo presente na imagem original, acabaram sendo "esticados".

### B. Passo 2

Para evitar essa distorção, é então aplicada uma erosão com o mesmo elemento estruturante do passo anterior. O efeito

dessa transformação é excluir lacunas, que são por exemplo espaços entre palavras. Assim elementos Após a execução do passo 2 percebe-se que o efeito desejado foi de fato atingido:

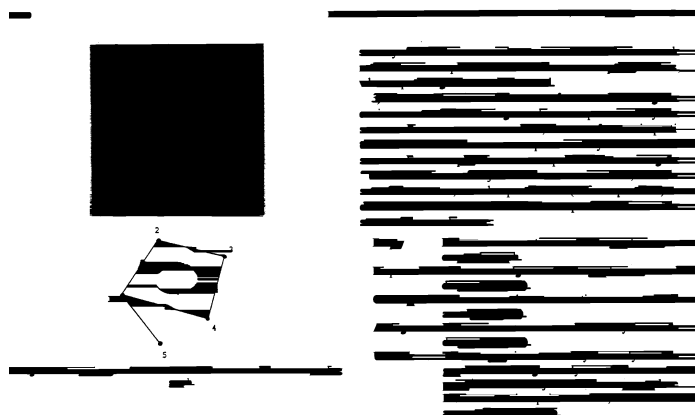


Figura 3: Resultado do passo 2.

### C. Passo 3

Com a dilatação da imagem com elemento estruturante de 200 pixels de altura e largura 1, o efeito esperado é uma expansão das dimensões verticais dos componentes presentes na imagem. Assim como pode ser visto na figura se seguir:



Figura 4: Resultado do passo 3.

### D. Passo 4

Usando a erosão, assim como foi descrito no passo 2, corrige-se um pouco da distorção de dimensões. No entanto, nota-se que vários componentes, principalmente texto foram fundidos.

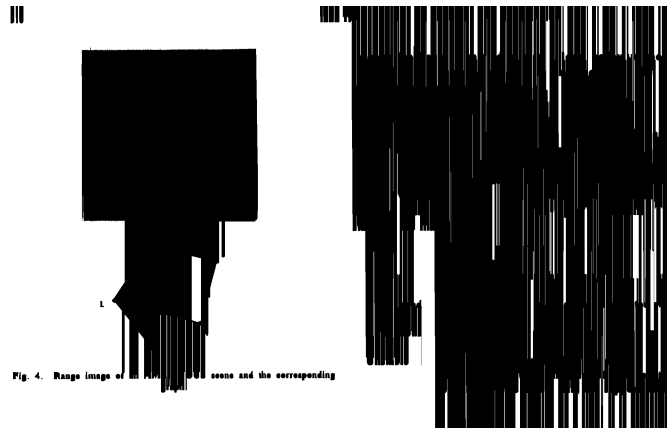


Figura 5: Resultado do passo 4.

### E. Passo 5

Neste passo combinamos os resultados obtidos nos passos 2 e 4 através do operador lógico AND. Analisando o a figura 6 percebe-se que foram aproximadamente segmentadas linhas de frases (texto) e imagens (uma a fotografia e a outra o grafo).

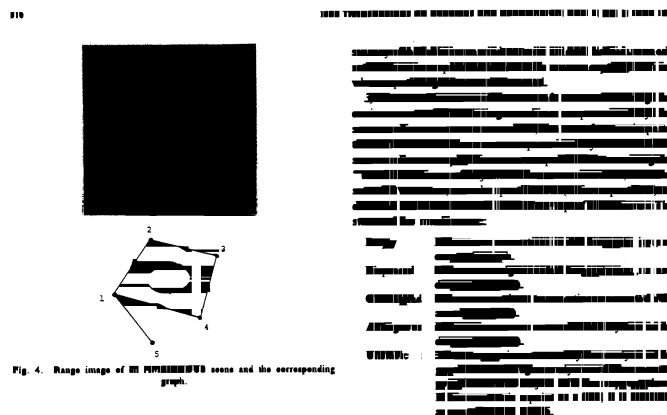


Figura 6: Resultado do passo 5.

### F. Passo 6

Efetuamos o fechamento da imagem resultante do passo anterior, um elemento estruturante de 1 pixel de altura e 30 pixels de largura.

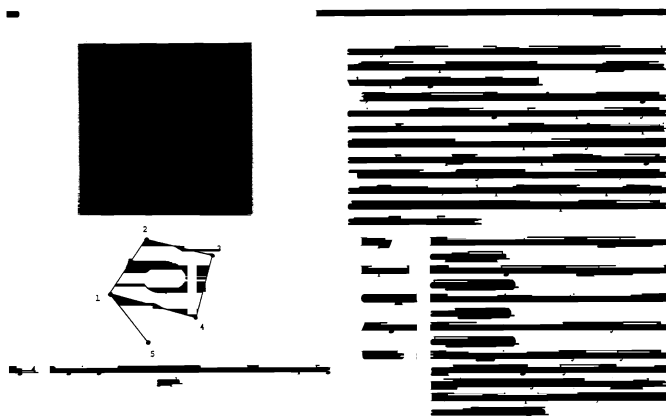


Figura 7: Resultado do passo 6.

Após a realização desse procedimento, lacunas que existiam entre algumas palavras e que ainda era visíveis na figura 6 foram apagadas.

#### G. Passo 7

Nesta etapa, aplicamos um algoritmo de detecção de componentes conexas, que traçou retângulos em torno de cada componente encontrada. Em uma primeira abordagem, utilizando o código oferecido junto a especificação deste trabalho, foram obtivemos 56 componentes.

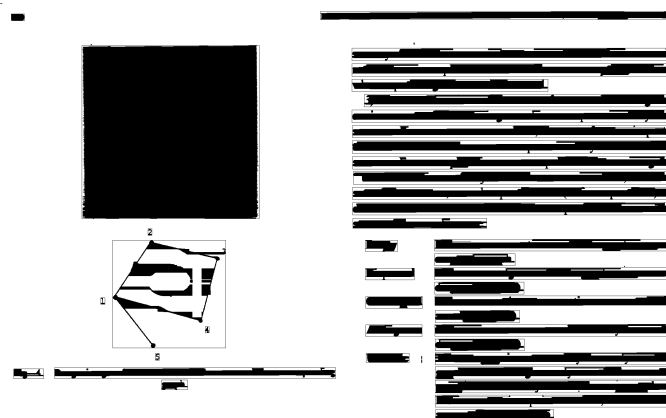


Figura 8: Resultado do passo 7.

#### H. Passo 8

No passo 8, com o auxílio de funcionalidades prontas do *framework OpenCV*, refizemos o passo 7 e calculamos algumas estatísticas. Primeiramente o número de componentes conexas encontradas com o algoritmo utilizado nessa ferramenta foi igual a 57.

Calculamos a razão entre o número de pixels pretos e o total, esses valores se encontram no intervalo (0, 1), para melhor visualização criamos a figura 9. Pode-se perceber que a maior parte dos componentes conexas são mais pretos do que brancos, indicando que tem poucas lacunas, ou que são de formatos retangulares.

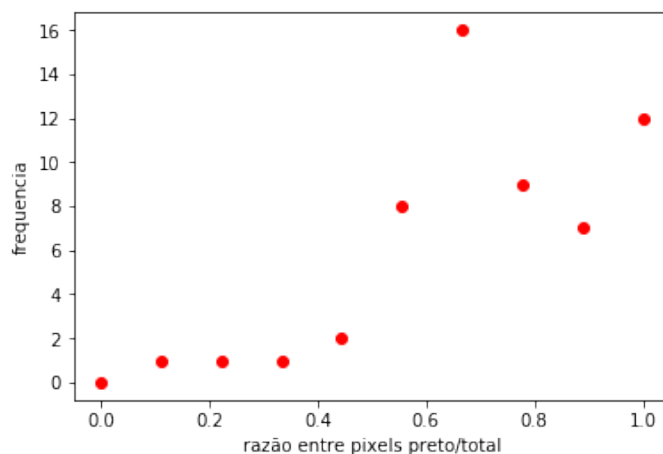


Figura 9: Resultado do passo 8. No x, os valores de razão entre número de pixels pretos sobre o total, no eixo y o número de componentes conexas.

A segunda estatística calculada foi a razão entre transições verticais e transições horizontais. Os resultados são ilustrados no gráfico da figura 10

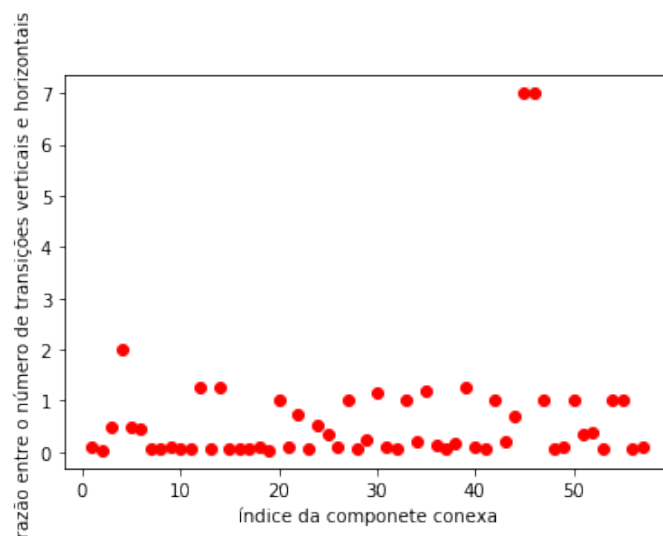


Figura 10: Resultado do passo 8. Razão entre transições verticais e horizontais para cada uma das componentes conexas encontradas.

#### I. Passo 9

Examinando os resultados do item anterior, em especial a razão entre transições verticais e horizontais, figura 10, podemos propor um limiar para separar texto e não texto.

Usando um valor de limiar  $th = 0.3$  e selecionando os componentes que obedecem a essa regra:

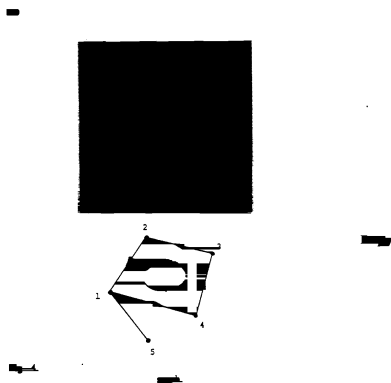


Figura 11: Resultado do passo 9. Em preto as componentes conexas identificadas como figuras pelo critério estabelecido no passo 9.

#### *J. Passo 10*

No passo 10, para encontrar o número de palavras, repetiu-se os passos 1 e 2, agora usando um elemento estruturante de menor dimensão, no caso de 1 pixel de altura e 11 de largura. Essa escolha foi tomada estimando com base na ideia de que um texto normalmente é escrito com tamanhos de fonte que variam entre 10 a 12 pixels. Assim caracteres vizinhos seriam unidos na dilatação.

### V. CONCLUSÕES

Operadores morfológicos são fortes ferramentas para o processamento de imagens, mas a obtenção de bons resultados depende do conhecimento de características básicas da imagem que se está processando. Como por exemplo no passo 10, conhecer a fonte pode ajudar na escolha das dimensões dos elementos estruturantes.

### REFERÊNCIAS

- [1] PEDRINI Hélio. Trabalho 3 MC920 - 1ºSemestre de 2019. Disponível em: <http://www.ic.unicamp.br/~helio/disciplinas/MC920/trabalho3.pdf>. Acesso em 22 de maio de 2019.