

# Operadores Conexos Aplicados à Localização de Texto em Imagens

Aluno: Alexandre Morimitsu - Orientador: Prof. Dr. Ronaldo Fumio Hashimoto

Trabalho de Formatura Supervisionado - Instituto de Matemática e Estatística, Universidade de São Paulo

## Introdução e objetivos

Este trabalho está relacionado ao problema de detecção de regiões de texto em imagens, e um possível algoritmo para tratar deste problema é descrito brevemente abaixo:

- ▶ Dada uma imagem  $I$ , converta-a em uma imagem em níveis de cinza  $I_2$ ;
- ▶ Compute o *Ultimate Opening* de  $I_2$  para obter uma imagem de contrastes  $I_3$ ;
- ▶ Binarize  $I_3$ , obtendo uma imagem binária  $I_4$ ;
- ▶ Agrupe objetos de  $I_4$ , gerando possíveis candidatos a texto;
- ▶ Classifique cada candidato gerado em texto ou não-texto.

Visa-se desenvolver duas das partes do algoritmo descrito. A primeira trata do uso da estrutura de dado *Maxtree* para a aceleração do processo *Ultimate Opening*. A segunda é a introdução de uma etapa pré-classificatória: a retificação das regiões candidatas, a fim de se obter imagens com maior chance de serem classificadas corretamente.

## Maxtree

Uma imagem em níveis de cinza pode ser vista como um fundo preto, sobreposto por camadas de *pixels* cada vez mais claras. Conjuntos conexos de *pixels* de uma mesma camada formam o que definimos com uma componente conexa (CC) da imagem.

A *Maxtree* é uma estrutura de dados que representa uma imagem através de uma árvore, baseada nas CCs da imagem: cada nó da árvore está associada a uma CC. Quando uma CC  $C_2$  se sobrepõe à uma CC  $C_1$  de uma camada anterior,  $C_1 \neq C_2$ , um nó é criado na *Maxtree*, e o nó de  $C_2$  é definido nó como filho de  $C_1$ .

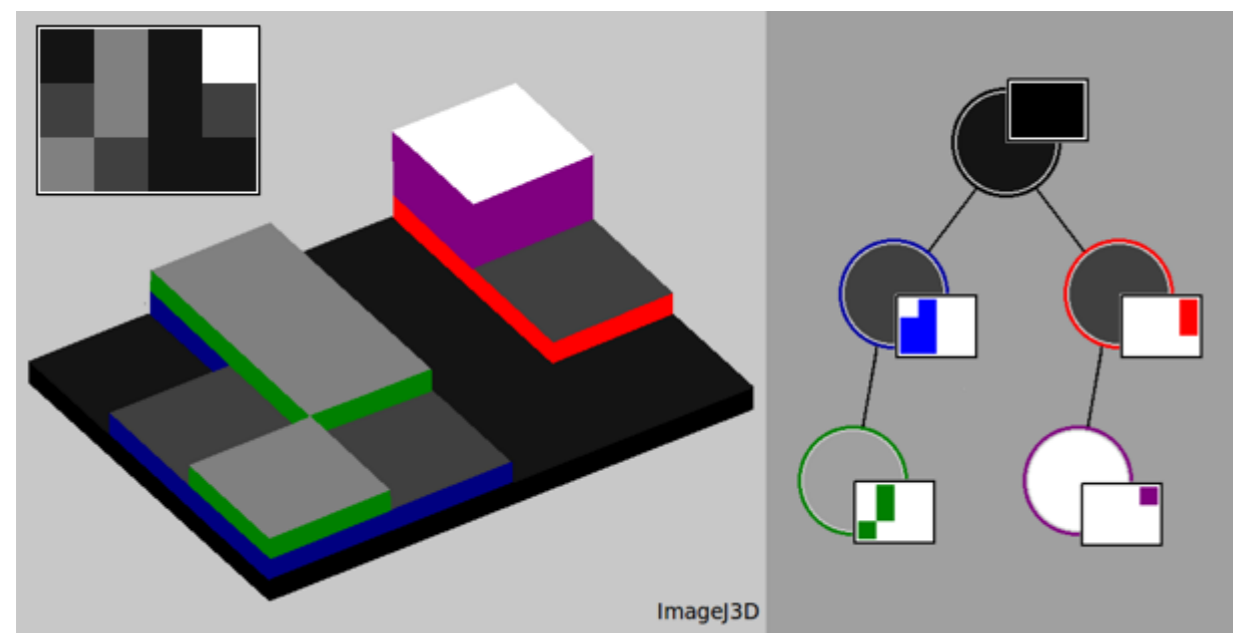


Figura: Exemplo de uma imagem, sua representação em camadas e sua *Maxtree*. Cada cor representa uma CC diferente, e em cada nó é mostrado a sua respectiva CC, além de seu nível de cinza associado. A componente verde é uma só pois foi considerada uma 8-vizinhança



## Ultimate Opening

*Ultimate Opening* (UO) é um operador morfológico que destaca regiões de alto contraste em uma imagem. Como o próprio nome sugere, o operador é baseado em aberturas (*openings*). São computadas aberturas sucessivas e é computada a maior diferença residual entre duas aberturas consecutivas. Formalmente, a UO é definida como:

$$v = \sup(r_i) = \sup(\gamma_i - \gamma_{i-1})$$
$$q = \max(i) + 1; r_i = v(\neq 0),$$

onde  $\gamma_i$  é a abertura de tamanho  $i$ .

Como saída, o operador devolve duas imagens: a imagem  $v$  do contraste e a imagem  $q$  do resíduo. Neste trabalho, as aberturas são feitas com relação ao atributo altura de uma CC, e apenas a imagem  $v$  é utilizada.

A *Maxtree* é uma estrutura adequada para a computação do UO porque, como visto, a cada CC da imagem é associado um nível de cinza, e o contraste pode ser facilmente computado através da diferença entre o nível de cinza entre um nó pai e um nó filho. Computa-se a imagem de contraste executando uma busca em profundidade na árvore.



Figura: Imagem original e sua respectiva imagem de contraste gerado pelo UO. Quanto mais claro o nível de cinza do pixel, maior o seu contraste

## Referências e detalhes na página

<http://www.linux.ime.usp.br/~alem/MAC499/>

## Ultimate Opening com transições graduais

Transição gradual foi um conceito introduzido para aumentar o contraste de regiões "embaçadas", um efeito que pode ocorrer devido à, por exemplo, alteração de escala de uma imagem. Usando transições graduais, pequenas variações de atributos de duas CCs são desconsideradas, gerando uma imagem com melhor contraste.



Figura: imagem original, UO sem e com transição gradual

## Retificação de texto

Na geração de regiões candidatas a texto, pode ocorrer de a região a ser classificada não estar no plano da imagem, o que pode afetar o resultado de sua classificação. Assim sendo, uma etapa pré-classificatória é introduzida, projetando tal região para o plano da imagem.

Para a retificação é calculada uma matrix que mapeia cada pixel da região candidata a um pixel da imagem de saída. Após o mapeamento, é aplicada uma interpolação para o preencher *pixels* que não forem mapeados.

Peacocks Peacocks Peacocks

Figura: Região de texto candidata, retificação da região de texto e resultado final

## Resultados

Exemplo de execução de um algoritmo de localização de texto, aplicados os conceitos apresentados:

