

Operadores Conexos Aplicados a Localização de Texto em Imagens

Aluno: Alexandre Morimitsu
Supervisor: Ronaldo Fumio Hashimoto

5 de junho de 2011

1 Tema da monografia

Estudo e implementação de algoritmos que visam melhorar o desempenho de um sistema que realiza a localização de regiões de texto em imagens.

2 Resumo da monografia a ser desenvolvida

Localização de textos em imagens é uma ferramenta bastante poderosa: ter conhecimento da posição das regiões de texto nos permite obter uma importante informação no que se refere a indexação de imagens, por exemplo, ou nos permite aplicar algoritmos de reconhecimento de texto (OCR) apenas em regiões específicas, obtendo resultados mais precisos. Por outro lado, trata-se também de um problema difícil: em uma imagem de cena, fatores como ângulo de visão ou luminosidade podem tornar difícil a detecção de tais regiões.

Pelas razões acima citadas, trata-se de um problema bastante atual. Em particular, este foi o tema escolhido pelo aluno Wonder Alexandre Luz Alves, orientando do professor Ronaldo Fumio Hashimoto, para o seu mestrado, no qual foi desenvolvido um sistema para tal utilidade.

Entretanto, por se tratar de um tema atual, constantemente surgem novas publicações sobre o assunto, relatando como modificar certas etapas do algoritmo de localização para torná-lo mais robusto ou eficiente.

O propósito desta monografia é focar principalmente no estudo de duas das etapas do sistema desenvolvido: extração e geração de regiões candidatas. Tal estudo será feito através da compreensão e implementação de

algoritmos apresentados em novas publicações da área. Os algoritmos serão integrados ao sistema já existente para analisar seus impactos, referentes tanto ao tempo de execução quanto à taxa de reconhecimento das regiões de texto.

Explicando resumidamente, a fase de extração de regiões candidatas consiste em filtrar a imagem de entrada para se obter regiões candidatas de texto. A abordagem escolhida para fazer a filtragem é baseada em um operador morfológico, o *ultimate attribute opening*, que se baseia no contraste entre os caracteres de texto e o fundo. Um dos focos desta monografia é o estudo e implementação deste operador usando a estrutura de dados *maxtree*. Esta implementação baseia-se nas referências [1] e [2].

Por sua vez, a fase de geração de regiões é uma operação pré-classificatória. Nela, já temos as regiões candidatas de texto, e estas estão prontas para serem classificadas em regiões de texto ou não-texto. Porém, o sistema atual que foi desenvolvido não é robusto a transformações afins, sendo que textos inclinados ou em perspectiva podem não ser classificados corretamente. Assim sendo, outro foco desta monografia é estudar algoritmos que permitam realizar estas transformações em regiões candidatas a fim de obtermos uma região horizontal, vista frontalmente, permitindo uma melhor classificação dos textos. O estudo será baseado na referência [3].

Havendo tempo hábil, pretende-se também desenvolver algum tipo de aplicação com o sistema a ser aprimorado.

3 Objetivos do trabalho

O objetivo principal do trabalho é, ao final de sua execução, ter implementado um algoritmo eficiente que detecte, com um alto grau de sucesso, regiões de textos em imagens.

4 Atividades realizadas

Até o presente momento, estudos e implementações referentes à primeira etapa do sistema estão praticamente concluídas, com o operador *ultimate attribute opening* implementado e integrado ao sistema.

Um dos passos do operador consiste em binarizar uma dada imagem de entrada. Aproveitando-se disto, foi também desenvolvido um algoritmo para tal finalidade, e este submetido a uma competição de binarização de imagens, o DIBCO [4].

5 Cronograma de atividades

Atividade	Mar-Mai	Jun	Jul	Ago	Set	Out	Nov
1	x	x	x	x	x		
2		x	x	x	x	x	
3				x	x	x	x
4			x	x	x	x	x
5							x

1. Estudo dos algoritmos
2. Implementação
3. Realização de Testes e Análise de Resultados
4. Escrita da monografia
5. Preparação do Pôster e da Apresentação do Trabalho

6 Estrutura esperada da monografia

A estrutura da monografia seguirá o roteiro de preparação de monografias, ou seja, será dividida em duas seções, parte técnica e parte subjetiva, compostas pelos seguintes itens:

6.1 Parte técnica

- Introdução
- Conceitos estudados
- Atividades realizadas
- Conclusões
- Bibliografia

6.2 Parte subjetiva

- Desafios e frustrações encontrados
- Disciplinas mais relevantes para a realização do trabalho
- Observações sobre a aplicação de conceitos estudados nas disciplinas
- Futuros passos na área

Referências

- [1] MARCOTEGUI, B.; HERNÁNDEZ, J.; RETORNAZ, T. Ultimate opening and gradual transitions. 2011.
- [2] FABRIZIO, J.; MARCOTEGUI, B. Fast implementation of the ultimate opening. In: International Symposium on Mathematical Morphology. c2009. v. 5720. p. 272–281.
- [3] MYERS, G. K.; BOLLES, R. C.; LUONG, Q.-T.; HERSON, J. A.; ARADHYE, H. B. Rectification and recognition of text in 3-d scenes. *IJDAR*, v. 7, p. 147–158, 2004.
- [4] Document image binarization contest. <http://utopia.duth.gr/ipratika/DIBCO2011/> (Acessado em 5 de junho de 2011).