

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Comparação de Métodos de Detecção de Texto em Imagens de Sensoriamento Remoto

Thalita Botuem Honorio¹Marilaine Colnago²Wallace Casaca³

Departamento de Engenharia de Energia, UNESP, Rosana

1 Introdução

As imagens digitais estão cada vez mais presentes em nosso cotidiano pois constituem uma maneira eficiente de transmitir informações e enquadrar elementos reais através de representações gráficas; por esse motivo, o Processamento Digital de Imagens (PDI) tem se tornado uma ferramenta essencial com inúmeras aplicações em áreas diversas. Assim, é natural que alguns tipos de imagens, tais como fotos de placas de carros, cenários urbanos, imagens capturadas por satélite com especificações técnicas, etc, retenham textos. Isso significa que, mesmo havendo diferentes algoritmos para detecção textual, há uma alta complexidade na identificação do texto nas imagens devido à particularidades como: diferentes fontes de texto, presença de ruído, etc, que dificultam o processo de detecção.

O presente trabalho apresenta uma análise comparativa entre dois modelos de detecção de textos em imagens: Modelo Baseado em Bordas e o *MSEER* (*Maximally Stable Extremal Regions*), sob a perspectiva de detecção textual para imagens de sensoriamento remoto.

2 Metodologia

O modelo de detecção baseado em bordas foi proposto por Tahim [2], onde os autores utilizam as bordas das letras para a detecção de candidatos a texto. As intensidades dos *pixels* da imagem que estão presentes no texto possuem valores superiores às do resto dos pixels da imagem. Desta forma, as bordas do texto destacam-se dos demais objetos da imagem, o que leva à identificação do texto [2]. Nesse método, a identificação é realizada em quatro etapas: na primeira etapa, converte-se a imagem original numa imagem em tons de cinza; na segunda etapa, calcula-se o vetor gradiente em cada *pixel* da imagem; na terceira etapa, há um realce das bordas da imagem através de um processo de limiarização; por fim, na quarta etapa, a imagem binária final é obtida via componentes conexos [2].

¹thalita.botuem@unesp.br²marilaine.colnago@unesp.br³wallace.casaca@unesp.br

Já o modelo MSER opera a partir de mudanças de perspectiva, escala e iluminação. Este localiza pontos extremos na imagem (pontos de mínimo/máximo definidas em vizinhanças da imagem) a fim de identifica as regiões conexas da mesma [1]. Para facilitar a identificação do texto, o método inicialmente aplica um filtro para destacar as bordas, de forma a aumentar o contraste, o que permite maior precisão na identificação do texto [1].

3 Resultados

A Figura 1 ilustra os resultados obtidos pelos modelos: baseado em bordas (Centro), e MSER (Direita). Ambos os modelos permitiram a detecção do texto, no entanto, o MSER identificou alguns falsos positivos como o “galpão” branco (destacado em vermelho).



Figura 1: (Esquerda) Imagem original, (Centro) resultado obtido pelo modelo baseado em bordas e (Direita) resultado obtido pelo modelo MSER.

4 Conclusão

Concluiu-se que ambos os modelos apresentaram resultados satisfatórios na tarefa de detecção de textos em imagens remotamente sensoriadas, embora haja diferenças significativas entre as abordagens quando aplicadas à imagens de grandes dimensões.

Agradecimentos

Os autores gostariam de agradecer ao CNPq (PIBIC) pelo financiamento à pesquisa.

Referências

- [1] H. Chen et al. Robust Text Detection in Natural Images with Edge-Enhanced Maximally Stable Extremal Regions. *IEEE ICIP*, p. 2609-2612, 2011.
- [2] A. P. N. Tahim, Localização e Extração Automática de Textos em Imagens Complexas, Dissertação de Mestrado, UFSC, 2010.