



APLICAÇÃO DAS TÉCNICAS DE BINARIZAÇÃO E SEGMENTAÇÃO DE IMAGENS PARA O RECONHECIMENTO DE TEXTO A PARTIR DE UM DOCUMENTO - IMAGEM

WEBER, Fernando Arthur¹
ANTONIAZZI, Rodrigo Luiz²

Palavras-chave: Processamento Digital de Imagens. *Watershed. Threshold.*

Com a popularização da comunicação por meios eletrônicos houve grande incremento no compartilhamento de informações digitais, dentre elas imagens. A partir disso, surge o interesse em aperfeiçoar ou identificar padrões por meio de métodos de processamento digital de imagens. Processamento digital de imagens é definido por qualquer técnica de processamento, manipulação ou análise onde a entrada e o resultado obtido são imagens. Essas técnicas tiveram as primeiras aplicações no início da década de 1920 onde imagens eram enviadas entre Londres e Nova York por meio de cabos subaquáticos. Os cabos subaquáticos eram ligados a aparelhos que codificavam as imagens e realizavam a transmissão para poder ser impressa do outro lado do oceano Atlântico. Atualmente, não há praticamente nenhuma área que não possua ao menos relação com o processamento digital de imagens. O seu objetivo é melhorar a informação visual para a interpretação humana e para a percepção automática, ou seja, a utilização por máquinas. O trabalho segue a premissa da dificuldade de transcrever de modo automatizado uma determinada imagem digital contendo texto de forma que esse possa ser editado posteriormente, utilizando o método de segmentação de imagens digitais *Watershed* alinhado com a técnica de limiarização *Threshold*. O método *Watershed* está contido na categoria dos métodos de processamento de imagem baseados em região, onde o processamento é iniciado em um *pixel* pré-definido e a partir desse os demais *pixels* ao seu redor são analisados de acordo com suas características semelhantes que podem ser nível de cinza, cor, textura ou demais particularidades do *pixel*. O método utilizado na limiarização, *Threshold*, é uma técnica que transforma uma imagem em duas classes predominantes, fundo e objeto. A divisão entre essas classes ocorre durante o processamento, onde a partir de uma imagem binária ele agrupa os pontos do primeiro plano (objeto) com valor "1" e os demais recebem o valor "0" definidos então como fundo. Até o presente momento, o ambiente possibilita que o usuário faça o upload da imagem a ser identificada nos principais formatos de extensão, são eles: gif, jpeg, png e bmp, pois o próprio ambiente converte todas as imagens para jpg e as redimensiona para um tamanho reduzido o que proporciona melhora no desempenho. Depois de selecionadas as imagens são submetidas ao filtro da escala de cinza e logo em seguida são limiarizadas com o método *Threshold*, o resultado é uma matriz binária que fica armazenada em um banco de dados. Pode-se destacar ainda o desempenho da limiarização nos testes realizados, pois aliada a etapa da escala de cinza ela permite que obtenhamos uma matriz binária fiel à imagem original, isso porque o processo avalia pixel à pixel individualmente.

¹ Acadêmico do curso de ciências da computação da Universidade de Cruz Alta. E-mail: fernando_aweber@hotmail.com.

² Docente do curso de ciências da computação da Universidade de Cruz Alta. E-mail: rantoniazzi@unicruz.com.br