

RECONHECIMENTO DOS NÚMEROS EM IMAGENS DE CARTÕES DE CRÉDITO

Fábio Takashi Hirano
fabiohirano@gmail.com

Supervisor:
Prof. Dr. Ronaldo Fumio Hashimoto

Introdução

Visando facilitar o processo de compras com cartão de crédito ou débito, pensou-se na possibilidade de obter o número do cartão apenas com uma foto do mesmo, poupando o trabalho de ter que digitar uma sequência de números.

Para executar essa tarefa, utilizaram-se técnicas de Visão Computacional, área da Ciência da Computação que estuda a forma como o computador “vê”, isto é, como o computador extrai informações de uma imagem e como ele as interpreta.

Reconhecimento de Caracteres

O Reconhecimento de Caracteres é uma área da Ciência da Computação que reúne aspectos de Visão Computacional, Reconhecimento de Padrões e Inteligência Artificial. Várias aplicações hoje em dia utilizam-se do reconhecimento de caracteres, tais como scanners e conversores de arquivos (por exemplo, de um arquivo *.pdf* para um *.odt*).

O Reconhecimento Óptico de Caracteres (*Optical Character Recognition*, ou OCR) é o nome dado a qualquer tecnologia que reconheça caracteres a partir de imagens computacionais. Através de um OCR, conseguimos, por exemplo, digitalizar um texto impresso. Exemplos de OCR disponíveis são o Tesseract, o OCRFeeder e o CuneiForm. Cada OCR se utiliza de uma sequência de algoritmos para identificar os caracteres dentro de uma imagem. O processo de reconhecimento pode ser entendido em 3 grandes etapas: Segmentação, Extração de Características e Classificação.

Extração de Características

Extração de Características é definida em [2] como o problema de “extrair de dados brutos a informação mais relevante para propósitos de classificação, com a intenção de minimizar a variação do padrão dentro de uma mesma classe e tentando maximizar a variação do padrão entre as classes”. Dentro de um OCR, a Extração de Características é a etapa que mais influencia em sua eficiência.

Um dos desafios encontrados foram as próprias imagens. Binarizar as imagens, ou seja, transforma-las em imagens digitais em “preto e branco” (isto é, cada pixel pode contar apenas um entre dois valores) pode deformar o caracter, devido, por exemplo, ao brilho da imagem. Isso dificulta a extração das características, tornando a classificação mais complexa. Para aumentar o contraste nas imagens (e, conseqüentemente, facilitar uma binarização correta), utilizamos as idéias sugeridas em [3] e [1].

Lienhart [3] sugere decompor a imagem em matrizes RGB e analisar as cores que formam o fundo da imagem. Isso pode ser feito, por exemplo, estudando os histogramas de acumulação da imagem nos 3 níveis (R, G e B). Nos picos destes histogramas estão as intensidades preponderantes na imagem. Uma vez que o *background* aparece mais nas imagens que o *foreground*, esses picos representam, em geral, as cores do fundo da imagem. Dessa forma, podemos processar a entrada de modo a incrementar o contraste entre o *background* e o *foreground*.

No exemplo abaixo, temos a imagem original e a imagem processada. O caracter 5 foi ressaltado, ficando mais nítido.

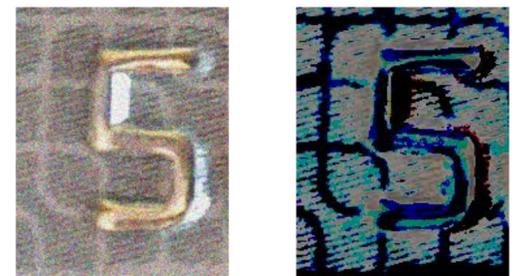


Figure 1: Imagem Original e Imagem Processada

Com as imagens processadas em mãos, os algoritmos de extração de características, como por exemplo o *X-Projection*, podem ser utilizados com mais eficiência.

Referências

- [1] Cheng, X., Yang, J., Zhang, J., and Waibel, A. Automatic detection and recognition of signs from natural scenes. *IEEE Transactions on Image Processing*, 13(1), 2004.
- [2] Devijver, P. A. and Kittler, J. *Pattern Recognition - A Statistical Approach*. Prentice-Hall, 1982.
- [3] Lienhart, R. and Wernicke, A. Localizing and segmenting texts in images and videos. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(4), 2002.