

# Identificação de Logos em Imagens

Jorge Lúcio

DCC-UFMG

jlucio@dcc.ufmg.br

Vitor Villani

DCC-UFMG

vitorvcf@dcc.ufmg.br

20 de março de 2010

## Resumo

O nosso trabalho consistirá em criar um método de busca de logos em imagens. Um problema com o qual muitas empresas se deparam é a verificação da quantidade de vezes que o seu logo aparece em determinado evento. Normalmente, as companhias que patrocinam certos eventos desejam saber se a propaganda da sua logo foi realmente realizada e quantas vezes a mesma apareceu nas fotos divulgadas, por exemplo. Para isso, espera-se criar um método "inteligente" de análise de imagens em busca de determinada logo levando em conta que a mesma pode estar distorcida na imagem devido ao ângulo, contraste, claridade, entre vários outros motivos.

## 1 Trabalhos Relacionados

Antes de especificar os métodos e trabalhos diretamente relacionados à identificação de padrões diversos em imagens, deve-se detalhar os modelos nos quais se baseiam esses artigos para desenvolver seus algoritmos específicos. Para reconhecer características locais nos modelos que devem ser identificados, o algoritmo mais popular entre os pesquisados é o SIFT (Scale-Invariant Features Transform) [1]. Ele permite o reconhecimento independentemente da escala, rotação, translação e luminosidade que a imagem possa sofrer. Isso se dá através da extração da imagem de vários vetores de pontos de interesse, e da posterior filtragem desses pontos apenas para os mais significativos. O modelo SURF (Speeded Up Robust Features) [2] é o que mais se distancia

do SIFT ao se basear em somas de respostas da wavelet de Haar 2D aproximadas. Isso faz dele um modelo mais rápido do que o SIFT e com uma aplicação tão boa quanto no reconhecimento de objetos.

Outros algoritmos pesquisados se preocupam mais em aperfeiçoar o SIFT do que propor um modelo novo para se identificar os pontos de interesse. Uma variação do SIFT é o PCA-SIFT (Principal Components Analysis-SIFT) é detalhado e comparado com o próprio SIFT em [4]. Tal algoritmo é utilizado para extrair descritores locais das imagens. O mesmo é utilizado em [5], cujo objetivo é determinar um processo eficaz de detecção de imagens parecidas e modificadas. Tal método é útil para encontrar imagens forjadas e violações de direitos autorais. O artigo descreve uma combinação de três algoritmos para a criação de várias tabelas que descrevem todas as imagens e seus pontos de interesse de um banco de dados escolhido. Com esses dados, uma imagem qualquer é escolhida e seus pontos de interesse são comparados com os dados da tabela. Quando alguma outra imagem semelhante é encontrada, através de outro algoritmo chamado RANSAC (Random Sample Consensus), algumas transformações de pontos são realizadas e analisadas para verificar se a imagem resultante pode realmente ser uma transformação real. Esse método é eficaz devido ao fato de que as transformações dos pontos de interesse em imagens forjadas muitas vezes não batem com os pontos da imagem original. Por se tratar de um método eficiente de se comparar imagens e encontrar pequenos detalhes modificados, tal artigo é selecionado em vários outros trabalhos da área.

Conhecidos os algoritmos que executam a primeira parte do processo de identificação de objetos em imagens, serão focadas agora as partes posteriores se baseando nos trabalhos que desenvolveram assuntos mais semelhantes ao em questão. São inúmeros os artigos que tratam desse tema [3], cada um focado num objeto específico e em suas características particulares, porém as etapas são quase sempre as mesmas: 1) identificação e descrição de pontos de interesse; 2) construção do vocabulário visual; 3) treinamento do classificador; 4) classificação. A primeira etapa já foi discutida, já a segunda consiste no clustering dos pontos de interesse detectados num subconjunto de imagens, e cada subconjunto é uma palavra do vocabulário. A terceira recebe os dados classificados da segunda e, com uma fração específica do conjunto de imagens, é testado para treinar o classificador, tendo conhecimento prévio se as imagens sendo analisadas atendem ou não ao objeto em questão. A parte final é a execução do algoritmo desenvolvido e treinado para a fração restante das imagens.

## Referências

- [1] David G. Lowe, “Distinctive image features from scale-invariant keypoints”, *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [2] H. Bay, T. Tuytelaars, and L.V. Gool, “Surf: Speeded up robust features”, *European Conf. Computer Vision, LNCS 3951, Springer, 2006*, pp. 404–417, 2006.
- [3] Natália C. Batista, Ana Paula B. Lopes, and Arnaldo de A. Araújo, “Detecting buildings in historical photographs using bag-of-keypoints”, *2009 XXII Brazilian Symposium on Computer Graphics and Image Processing*, pp. 276–283, 2009.
- [4] Yan Ke and Rahul Sukthankar, “Pca-sift: A more distinctive representation for local image descriptors”, *2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’04)*, vol. 2, pp. 506–513, 2004.
- [5] Yan Ke and Rahul Sukthankar, “Efficient near-duplicate detection and sub-image retrieval”, *Proceedings of the 12th annual ACM international conference on Multimedia (MULTIMEDIA ’04)*, pp. 869–876, 2004.