Beer label recognition and classification
EE 368 Project Proposal

Andrew Weitz, aweitz@stanford.edu
Akshay Chaudhari, akshaysc@stanford.edu

Project description:

The goal of this project is to develop an image-processing algorithm that can recognize and classify beer bottle labels in near real-time. Mobile applications that implement such an algorithm could be used to provide information not available on the bottle (e.g. consumer ratings and reviews) in different on-the-go situations like grocery shopping. This project is analogous to the final project offered in spring 2007-2008, in which algorithms were developed to identify CD covers.

Goals and Implementation:

Since this project revolves around machine learning, we will first need to collect images for the training and testing datasets. For the training set, we will collect "clean" beer labels using images available online (Figure 1, left). For the testing set, we will use photographs of bottles with corresponding labels, similar to those that would be acquired with a mobile phone (Figure 1, right). We plan to use a training set of at least 100 images/beer labels, including some that come from the same brewery to make the classification more difficult.

Figure 1: Example training image (left) and matching test image (right).

For the test images, we will perform the following pre-processing steps before feature extraction:

1. Color balancing to account for differences in acquisition conditions [1]
2. Conversion from RGB to grayscale (for some features)
3. 8:1 or 4:1 downsampling with low-pass filtering to improve runtime speed
4. Segmentation to extract the label from the rest of the photograph
5. De-warping to make the label flat using an approximate cylindrical projection [2]

Three sets of features will be investigated:

1. *Scale-invariant features*, identified by finding local extrema in difference-of-gaussian image pyramids using the SIFT algorithm [3].
2. *RGB color histograms* of the extracted label. Although this method is sensitive to subtle changes in noise and lighting, it can be useful for labels that are largely one color (for example, the label shown in Figure 1). Feature comparisons can be computed using the histogram intersection kernel for support vector machines [4].

3. *Prevalence of text*. Since labels can include many different fonts and sizes of letters, any attempt to identify specific characters via morphological image processing would probably not succeed. Therefore, instead of character recognition, we will perform text detection – that is, detecting regions where there is text versus regions where there is no text [5]. How well a test label matches with each image in the training set can then be quantified by the fraction of the label that is taken up by text.

We will perform the final label classification taking each of these features into account and weighting them according to their individual success.

DROID statement:
We do not plan to implement this system on a DROID phone.

References:

[1] F. Gasparini and R. Schettini, "Color balancing of digital photos using simple image statistics," *Pattern Recognition,* vol. 37, pp. 1201-1217, 2004.

[2] N. Stamatopoulos*, et al.*, "A two-step dewarping of camera document images," in *Document Analysis Systems, 2008. DAS'08. The Eighth IAPR International Workshop on*, 2008, pp. 209-216.

[3] D. G. Lowe, "Object recognition from local scale-invariant features," in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, 1999, pp. 1150-1157.

[4] A. Barla*, et al.*, "Histogram intersection kernel for image classification," in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, 2003, pp. III-513-16 vol. 2.

[5] C. Liu*, et al.*, "Text detection in images based on unsupervised classification of edge-based features," in *Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on*, 2005, pp. 610-614.