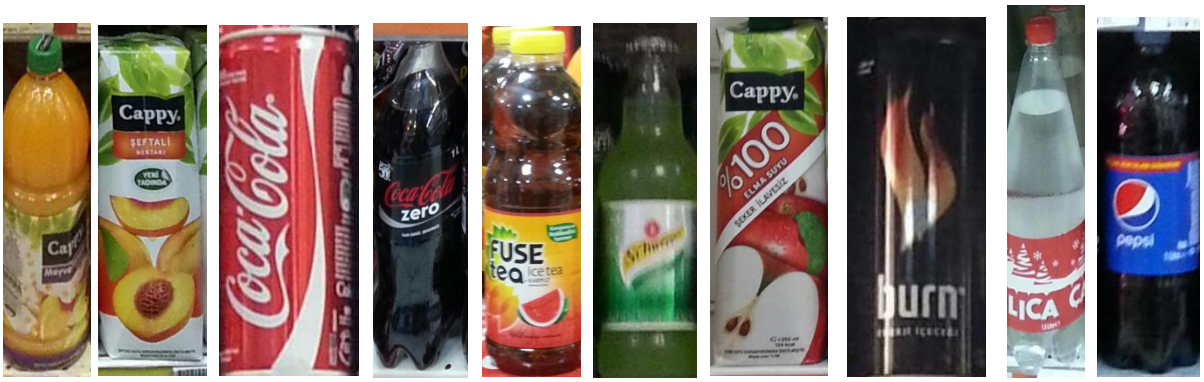
ECE 527 Computer Vision 2015
Assignment 04

Color Histogram Descriptors

In this homework, you will experiment with color descriptors and the K-nn classifier on a real dataset, provided here: https://www.dropbox.com/s/n8zh37udjmrydad/20150320_DM4VM_dataset.zip?dl=0

The dataset consists of 10 classes of tightly cropped color images of packaged supermarket products (see figure below). Each class contains approximately 100 instances, 30 of which are allocated as *test* instances, and the remaining as *training* instances. In your experiments, first off, you should normalize the size of each image into 256x128 pixels (height x width). You will use this size-normalized version of the images in all cases (Why?).



In each case below, you will calculate color histograms (with various definitions of color information) on (A) non-overlapping and (B) sliding blocks.

(A) **Non-overlapping blocks.** Divide the image into non-overlapping blocks of size 32x16. There will be 64 such blocks. For each such block, you will calculate a color histogram, then you will normalize the histogram so that the sum of its entries is one. For this option, the image will then be described by 64 L1-normalized color histograms.

(B) **Overlapping blocks.** Apply a stride parameter of 16 pixels in the vertical and 8 pixels in the horizontal direction. You will have 240 such blocks some of which are overlapping. For each such block, you will calculate a color histogram, then you will normalize the histogram so that the sum of its entries is one. For this option, the image will then be described by 240 L1-normalized color histograms.

Experiments

(1) Plain RGB Histograms. Each block should be described by a concatenation of three 1-D histograms based on the R, G and B channels of the RGB color space. Quantize each of the channels into 8 bins and construct a normalized histogram for each channel. A block is thus described by 3 histograms of length 8. Then, for (A), you will obtain 64x3 = 192 histograms, each of size 8, that you should concatenate into a global descriptor of size 1536 of the image. Then, for (B), you will obtain 240x3 = 720 histograms, each of size 8, that you should concatenate into a global descriptor of size 5760 of the image. In both cases separately, run the K-nn classification algorithm using L1-distance and K=1 on test instances. That is, you will

match each test instance against the training set instances, the class label of the best matching instance will be the estimate of the class label of the test instance (there are 300 instances to test). Report per class accuracies and overall accuracy as well as a confusion matrix (showing which classes are confused the most with which).

(2) rg Histograms. Chromaticity values *r*, *g*, and *b* are obtained by the following formulae

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B} \text{ and } b = \frac{B}{R+G+B}.$$

You perform this for each pixel in the image, then calculate *r* and *g*-histograms as described in part (1). Note that you don't need the *b*-histogram, why? Repeat the whole procedure described in (1).

(3) Transformed RGB Histograms. This time you will apply the following transformation at each pixel:

$$\begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = \begin{pmatrix} \frac{R-\mu_R}{\sigma_R} \\ \frac{G-\mu_G}{\sigma_G} \\ \frac{B-\mu_B}{\sigma_B} \end{pmatrix}$$

where μ_i and σ_i are the mean and standard deviation of the respective color channel $i \in \{R, G, B\}$ calculated from the entire image. After this, what you should do is entirely the same as in (1).

(4) Compare and discuss your results obtained above. How could you improve the performance? Which are parameters would be more effective (stride parameter, histogram bins, distance metric, K)? If possible, experiment with different parameters and report your results (this would provide you with some BONUS if successful).

(5) What are the limitations of the approach you experimented with in this assignment?