TNM025 – Assignment 2 – 2015

Here you will experiment with histograms calculated for image content obtained from more than one image channel. In the second part of the homework, derived histograms will be used for illustrating the efficiency of using Principal Component Analysis for dimensionality reduction.

1. **2-D/3-D Histograms**

   Write programs that compute histograms for vector valued images. Do this for two different higher-dimensional pixel values.

   Some examples:
   a) Every pixel in the image consists of an HS vector (Hue and Saturation value)
   hsvim = rgb2hsv(rgbim); hsim = hsvim(:,:,1:2)
   b) Every pixel in the image consists of an RGB vector
   c) Use the matrix [1,1,1;1,-1,0;1,1,-2] to convert from RGB to a HSV-like description
   d) How critical are the values in the conversion matrix? There are many others, try one/some to see if/how the results change

   In case (a) the histogram describes the chromaticity properties of the image (i.e. properties independent of the intensity).

   Requirement: Fast code. Avoid using for loops in Matlab.

2. **Investigate how the approximation order in the PCA affects the retrieval result**

   Compute the histograms for the database images and keep them somewhere. Use the RGB histograms and use 8x8x8 = 512 bins.

   Compute the 512x512 correlation matrix of histograms and its first 20 eigenvectors

   Transform all computed histograms to their 20-D PCA form

   Use the L_2 norm to measure distances between histograms.

   Now do the following experiment:

Choose a number of random images in the database (say 100) store their indices in a vector
Use each of these images as query and find the most similar images in the database (say 20-30). Store the retrieved image numbers. Use the full histograms for retrieval.

In a loop base the retrieval on the 1, 1-2, 1-3, 1-4 etc coefficients in the eigenvector expansion. Measure how the retrieval changes. How many of the original (the list retrieved using full histogram) 20-30 images are found on average when you use k eigen coefficients?