CSE252C – Object Recognition – Assignment #1

Instructor: Prof. Serge Belongie.

`http://www-cse.ucsd.edu/classes/fa07/cse252c`

Target Due Date: Thursday Oct. 11, 2007.

1. Handwritten Digits.

   (a) Download the MNIST training and testing data from http://yann.lecun.com/exdb/mnist.

   (b) Write a utility to extract the images (of size 28×28) and labels $(0, \ldots, 9)$. Use it to import the first $M = 2000$ training digits and the first $N = 1000$ testing digits.

   (c) Display the first 40 training digits together with their labels, arranged in a 4×10 array.

   (d) Compute the prior probability of each digit in the training set. Is it uniform?

2. Measuring Similarity/Dissimilarity.

   Let $\boldsymbol{x}^i \in \mathbb{R}^d$ (with $d = 28^2$) denote the ith training example concatenated as a column vector.

   (a) Implement the following pairwise comparison functions of the form $\mathcal{D}(\boldsymbol{x}^i, \boldsymbol{x}^j)$:

   - L_p norm: $\left(\sum_{k=1}^{d} |x_k^i - x_k^j|^p\right)^{1/p}$
   - Inner product: $(\boldsymbol{x}^i)^\top \boldsymbol{x}^j$
   - Normalized inner product: $(\boldsymbol{x}^i)^\top \boldsymbol{x}^j / \|\boldsymbol{x}^i\| \|\boldsymbol{x}^j\|$
   - χ^2 distance: $\frac{1}{2} \sum_{k=1}^{d} (x_k^i - x_k^j)^2 / (x_k^i + x_k^j)$

   Each is defined for $\boldsymbol{x} \in \mathbb{R}^d$ except χ^2, which requires $\boldsymbol{x}$ to be nonnegative and sum to 1.

   (b) Compute and display the best match (using max or min as appropriate) for the first 10 training digits (excluding self matches) vs. all M training digits using L_1, L_2, L_∞, and inner product (both normalized and raw). Use an asterisk to indicate errors.

   (c) Which choice of $\mathcal{D}(\cdot, \cdot)$ gave the fewest errors? Which gave the most?

3. Confusion Matrices and ROC Curves.

   (a) Compute the L_2 distance from all N testing digits to all M training digits.

   (b) Assuming a 1-nearest neighbor classifier, compute the 10×10 confusion matrix for this experiment. Display it as an image and comment on what it reveals about the classification behavior for digits such as 5 and 8.

   (c) Compute the histogram of distances for genuine matches and for impostors. Use bins of size 10 on the range 0 to 250, and normalize the histograms to sum to 1. Plot the two histograms on the same set of axes.

   (d) Plot the ROC curve for this experiment. What is the equal error rate?

4. Color Histogram Matching.

   (a) Select 10 objects from the Amsterdam Library of Object Images (ALOI) at http://staff.science.uva.nl/~aloi. For each object, download two images captured by the same camera under different illumination directions; call the resulting two sets of images $\mathcal{A}$ and $\mathcal{B}$. The preview thumbnail resolution of 154×115 is sufficient for this exercise.

   (b) For each of the 20 downloaded images, compute the color histogram using a color space of your choice with 15 equally spaced bins per channel.

   (c) Compute the 10×10 matrix of χ^2 distances between the color histograms from $\mathcal{A}$ to those of $\mathcal{B}$. Display the distance matrix, indicating the best matching entry in each row. Comment on the performance you observe, highlighting interesting successes or failures.