Programming Project 1 – Content Based Image Retrieval Using Global and Local Features

CAP 5415, Principles and Algorithms of Computer Vision, Fall 2014
Department of Computer Science, Florida State University

Points: 100
Due: Monday, October 20, 2014
Maximum Team Size: 2

Submission: You can submit a hardcopy of your report with source code attached or email your report along with source code as compressed tar archive to liux@cs.fsu.edu (with subject line starting with "Programming Project #1 Submission for CAP5415). Only one copy from each team is required.

Purpose: To know how to use global image features and local image features to characterize image contents and use them for content-based image retrieval.

Background: As imaging devices (e.g., cameras) become ubiquitous (mainly due to built-in cameras in smart phones), public large image databases are established to document events in many aspects of people's life. These images, in turn, provide helpful information in inferring relationships among people, leading to useful knowledge for many applications. A fundamental and enabling algorithm is to be able to reliably compare image contents and therefore index, organize, and retrieve images based on contents.

Assignment: Given a dataset of images: 1) Implement a basic content-based image retrieval system using histograms of images (essentially what you have done in homework #1), 2) improve the performance of your system by using histograms of filtered images (i.e., using spectral image representations), and 3) further (hopefully) improve the performance of your system by using SIFT features and you can use available programs to compute SIFT features.

After you are done, apply your three methods on Corel-1000 image dataset; the images are available at http://www.cs.fsu.edu/~liux/courses/cap5415-2014/class-only/test1.zip (these images are also available on any computer science department server at ~liux/CAP5415/Corel-1000). In the dataset, there are 10 categories and each category has 100 images (first class: 0.jpg-99.jpg, second class: 100.jpg-199.jpg and so on). For each method, show the precision-recall for a good case and bad case and then generate the average precision and average rank (as shown in Figure 15 (the plots on the left and the middle) for color histograms and the SIMPLICITY method in the paper "SIMPLICITY: Semantics-sensitive integrated matching for picture libraries" (J. Wang et. al., *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 9, pp. 947—963, 2001, available at http://ieeexplore.ieee.org/iel5/34/20661/00955109.pdf?isnumber=20661&arnumber=955109 (from any computer on campus)). Rigorously speaking, the average precision and average rank plots show the class average of average precisions defined in the paper and class average of average ranks also defined in the paper. Note that for the second method, you need to try and choose different filters and for the third method, you need to work out how many SIFT features to use from each image and how the SIFT features of two images will be compared.

Submission:

- **Report** – You need to turn in a report, summarizing the algorithms you have used/developed, showing significant intermediate results to demonstrate the correctness of your algorithms and

programs, documenting your findings, and analyzing the accuracy of your methods based on average prevision and average rank.

- **Source code** – You need to attach all the source programs you developed for this programming assignment. Note that if you have used other's programs, you need to give proper references and credits, and you will be graded based on the additional work you have done. For SIFT features, you can use any implementation that is available to you.

- **Results** – Include in the report the precision-recall curve of a good case and a bad case from each of the three methods you have developed and the average precision and average rank plots for each of the three methods. Note that the results you submit must be obtained from your own programs.

Grading

- **Report** – 20 points.
  - Description, analysis, and justification –10 points.
  - Computation time comparison of the three methods – 5 points.
  - Performance comparison of the three methods – 5 points.

- **Correct implementation** – 50 points.
  - Program organization and supporting functions – 15 points.
  - Basic color histogram method – 15 points.
  - Spectral histogram method – 10 points.
  - SIFT feature method - 10 points.

- **Results** – 30 points.
  - Precision-recall curves **–** 10 points.
  - Average precision plots– 10 points.
  - Average rank plots – 10 points.

- **Documented better performance than SIMPLICITY –** 15 points.

Extra Credit: Please state clearly in your report if you have implemented the following extra credit option.

Documented better performance than SIMPLICITY: Optimize the parameter choices of your methods (e.g., choice of different filters, the number of SIFT features, and how images are compared). For each category, if your best performance (either in average precision or average rank) is better than the corresponding one given by SIMPLICITY, you will get 1 extra credit point for a maximum of 15 points. Note that there are 10 categories and two performance criteria and therefore there are 20 potential extra credit points.

Additional Information

The following paper uses SIFT features for image retrieval and it may give you some ideas.

Yushi Jing and Shumeet Baluja, "VisualRank: Applying PageRank to Large-Scale Image Search," *IEEE Transactions On Pattern Analysis And Machine Intelligence*, Vol. 30, No. 11, pp. 1877-1890, November 2008.

The average precision in the paper is defined as the average of the precision measurements when we return 1, 2, and until 100 images. In other words, it is defined as

$$\text{Average Precision} = \frac{1}{100}\sum_{l=1}^{100}\frac{m_l}{l}$$

, where m_l is the number of relevant images among the l images returned.

In the paper, it appears the rank starts from 0 instead of 1. In other words, the rank of the first image is 0, rather than 1 as I mentioned. (Note that paper states: "An ideal CBIR system should demonstrate an average p of 1 and an average r of 50." (p. 959), which is not quite correct. If the rank starts from 0, then an ideal CBIR system should produce an average rank of 49.5 (mean(0:99) in Matlab); if the rank starts from 1, then an ideal CBIR system should produce an average rank of 50.5 (mean(1:100) in Matlab)).

The average precision for the color histogram 2 method in the SIMPLICITY paper is given here:
0.28850 0.28632 0.23257 0.26635 0.91377 0.38402 0.41633 0.38620 0.21731 0.20715
and the corresponding average rank is
312.2110 332.2742 332.2713 284.3071 54.0995 187.5079 234.6007 278.2043 324.4203 427.3270

Similarly, for the SIMIPILITY method, the average precision is
0.47477 0.32446 0.33027 0.36296 0.98117 0.39964 0.40218 0.71858 0.34188 0.33971
and the corresponding average rank is
178.3529 242.0187 261.6305 260.7511 49.3074 197.1079 298.6917 91.5890 230.2441 271.2211

The above numbers were estimated from Figure 15 in the paper using linear interpolation.

Using histograms of four filters, I was able to obtain the following average precision
0.4667 0.3267 0.3078 0.6075 0.9914 0.3967 0.8038 0.4826 0.2358 0.3578

and the following average rank:

294.7775 356.1816 367.0497 173.8379 51.3128 309.7915 134.9138 242.4062 365.6872 310.5087 .

This means that using four filters, I was able to get 7 extra credit points, five from the average precision and two from the average rank.

I also did experiments using color histograms; while the results depend on the choice of the number of bins of the histograms, by using 56 bins for each channel, I was able to obtain the following average precision:
0.6585 0.3012 0.3166 0.3969 0.9876 0.4851 0.5461 0.6388 0.2228 0.4890
and the following average rank:
183.6762 359.4082 346.1323 208.4408 55.0364 238.9864 274.5182 312.1988 387.1628 270.5475

which means I was able to get 9 extra credit points (six from the average precision and three from the average rank).

In any case, I included these results to give you an idea what you should expect.