Project 3

Content-Based Image Retrieval

Peggy Gerardin
Digital Photography Course
Summer Semester 2002

1 Background.

More and more images are used in HTML documents on the WWW. Ideally, a tool should be available, i. e. an image search engine, that can retrieve relevant images quickly on demand. Effectively, digital image databases have seen an enormous growth over the last years. Since many image collections are purely indexed or annotated, there is a great need for automated, content-based methods that could help users retrieve or structure image databases. Given the current state-of-the-art in computer vision, no robust general purpose methods exist that could perform, for example, object recognition in a broad domain.

1.1 Applications of Content-Based retrieval.

Users tend to use three kinds of research strategies. One of them is *search by association* when they have no specific aim other than find interesting things. Another class of users aims to *search for a specific image*. Users may have a specific image in mind and the target is interactively specified as similar to a group of given examples, which are useful for art, industrial components or catalogues. The third class of applications may be the *category search*. Users aim to retrieve an image from a specific class (catalogues of varieties, for example).

1.2 The image domain.

In the image domain, there is a distinction between *narrow* and *broad* domains. The *narrow* domain possess images of limited variability. And when the object's appearance has limited variability, the semantic description is also well-defined. For example, if one looks for Tibetan Terrier picture, there are few chances to obtain something else than pictures about this type of dog. On the other hand, the *broad* domain has an unlimited variability in its appearance and in its semantic description. If one searches for water, for example, the query result's may contain a large variability of pictures with water (sea, ocean, drop, etc.). Then, objects in an image may have different meanings as well as the interpretation of the scene.

2 Image processing.

Two steps are important and sufficient to describe the content of an image. The first one is the image-content operations that transpose the data into another spatial data array, in three units: color, texture and geometry. They may be characterized by: $f(x) = g \circ i(x)$, where $i(x)$ is the image, g is an operator on images and the resulting image filed is $f(x)$. In that way, image retrieval with image processing must enhance aspects in the image data *relevant* to the query and to reduce the remaining aspects.

2.1 Color image processing.

Color has been an active area of research in image retrieval, more than in any other branch of computer vision. Color makes the image $i(x)$ take values in a color vector space. In fact, two aspects have a real contribution: one is that the recorded color *varies considerably* with the orientation of the surface, the viewpoint of the camera, the position of the illumination and which type of illumination, and the way the light interacts with the object. Second, the human perception of color has the ability to capture perceptual similarity.

2.2 RGB histograms.

To deal with the inequalities in observation due to surface reflection is to search for *clusters* in a color histogram of the image. In the RGB histogram, clusters of pixels from an object form streaks. Hence, a non-parametric cluster algorithm in RGB-space is used to identify which pixels in the image originate from one uniformly colored object.

2.3 RGB color representation.

RGB color representation is a good choice when there is no variation in the perception of an object. It describes the image in its literal color properties and it is sufficient in the case of art-paintings and color composition of photograps, for example. For retrieval applications, this space is helpful thanks to its opponent color representation, in using the opponent color axes (R-G, 2B-R-G, R+G+B). The advantage is that axes are invariant to changes in illumination intensity and shadows.

2.4 HSV representation.

The human system can ascribe fairly constant hues to surfaces viewed in different visual contexts. The HSV representation is often selected for its invariant properties. The hue is invariant under the orientation of an object with respect to the illumination and camera direction and hence more suited for object retrieval.

3 Task.

- To analyse images according to different color representations and to set-up a Color-Based Image retrieval:
  - Using color histograms: histograms display the number of pixels in the image that have the same characteristic. The image content $f(x)$ is extracted by the histogram g according to $f(x) = g \circ i(x)$. This content extraction is given by existing functions in Matlab. After normalisation, we will compute the intersecting area H of histograms R and M by $H(R, M) = \sum_{i=1}^{n} min(R_i, M_i)$. The intersecting area is used as the similarity measure between histograms.

– Using RGB color representation: the image is into RGB components, so there are three histograms for each image for R, G and B respectively. Use also opponent color representation (R-G, 2B-R-G, R+G+B) (see references below).

– Using a hue based representation with the help of HSV representation.

– Choosing the best strategy to retrieve images and improve it with the help of Finlayson's paper.

3.1 Tools:

- 20 images of jpeg format.
- colorspace.m file.

3.2 References:

- **Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A. and Jain, R.**, *Content-Based Image Retrieval at the End of the Early Years*, IEEE Transactions on pattern analysis and machine intelligence, Vol.22, NO.12, december 2000.

- **Gevers, T. and Smeulders, A.W.M.**, *A Comparative Study of Several Color Models for Color Image Invariant Retrieval*, Proceedings of the First International Workshop, IDB-MMS '96, Amsterdam, The Netherlands, pp. 17-26, August 1996.

- **Finlayson, G.D. and Schaefer, G.**, *Hue that is invariant to brightness and gamma*, British Machine Vision Conference, pp. 303-312, 2000.

3.3 Schedule:

Task (Weeks)	**Description and Questions**
T1 (1,2)	Generate the color histograms for each image; To obtain a fair comparison, normalise the histograms if needed; Overlapp the histograms and measure the similarity; Calculate the mean of each image and compare. Generate also RGB histograms for each image and compare the both methods. **Question 1: What other type of histogram would be useful (except RGB histogram)? Question 2: Which informations are helpful to retrieve color images?**
T2 (2,4)	Generate spatial histograms: divide the images in several squares and do the same way than previously. Calculate the distance between them and compare. **Question 1: What do spatial histograms bring more ?**
T3 (5,6)	Implement the RGB color representation technique to the given images and determine their color properties. Use the opponent color representation and see wether there is an improvement to applicate it in image retrieval. **Question 1: What is the advantage of RGB-color space with opponent color axes? Question 2: According to your results, is it the best way to match the human colour perception?**
T4 (7,8)	According to Finlayson's paper, define the hue of each image. Following their method, determine the brightness and gamma invariant method. **Question 1: What is the advantage of HSV representation? Question 2: Which method works the best?**
T5 (9)	Hand in your report (June 19, 2002)

4 Report and evaluation.

Here are a few recommendations for your report. First, your report should absolutely contain:

- The project title;
- A summary of the results (1/2 page max.);
- A short introduction: problem description, proposed method, etc.
- Your work and main results;
- A Conclusion: interpretation of the results, comments, future work, etc.
- A complete bibliography.

Your report should not exceed 8-10 pages. Include only salient results and developments. Use nicely done figures to illustrate a point. Use Appendix for relevant information, which is not necessary to make your point.

The project is worth 30/100 points for the digital photography course, including:
- 10 points for the Matlab code and demo (clarity, comments, etc.)
- 10 points for understanding and innovation
- 10 points for the report

Deadline: June 19, 2002.