Abstract

This paper examines the problem of image retrieval from large, heterogeneous image databases. We present a technique that fulfills several needs identified by recent research in the field. This technique fairly integrates diverse and expandable sets of image properties (for example, color, texture, and location) in a retrieval framework, and allows users to control the retrieval process. We propose a novel set of evaluation methods in addition to standard Performance Evaluation protocols for image retrieval; our technique proves competitive with state-of-the-art methods in these tests and does better on certain tasks. Furthermore, it improves on many standard image retrieval algorithms by supporting queries based on subsections of images. For certain queries, this capability significantly increases the relevance of the images retrieved, and further expands the user's control over the retrieval process.

The STAIRS* Engine

A series of transformations converts a raw image into a vector that captures the spatial layout of color and texture in the image. The image is described in terms of component image tokens, or small patches described by their color, texture, and location. The joint histogram of these token values forms the representation of the image, which is compared with other images using a modified cosine metric. Using a similarity matrix $S$ in the distance equation

$$D(h_i, h_j) = \cos^{-1} \left( \frac{h_i^T S h_j}{\sqrt{h_i^T S h_i} \sqrt{h_j^T S h_j}} \right)$$

Compare two images based on their histogram vectors, using a modified cosine metric. This method is then extended to include an additional function $m$ using a parameterized decomposition $m$. Further tuning can be obtained by normalizing the cosine to a projected space.

Region-Based Retrieval

Often users are interested in only a portion of an image, perhaps a particular object in a scene containing many others. In such cases, a query based on the full image will return many false hits due to spurious matches with irrelevant areas of the scene. To solve this problem, an image-retrieval system must retrieve images based upon a match of some region in the target image with a specified region of the query image. STAIRS supports a form of region matching as a special case of a more general capability: matching some image region at a distance more or less strictly than others. In this framework, a region query is formed by requiring a close match in the region of interest, while allowing the rest of the image to match anything. Because the image has already been segmented into tokens, any areas that potentially match the target can be identified easily.

Some examples of region-based retrieval: