

Textures of Textures: Querying Image Databases by Example

Jeremy S. De Bonet

The Problem:

We are developing new techniques for querying, sorting, and measuring similarity between images. The technology behind this is based up the computation of image characteristic signatures. These signatures are used to as the basis of an image database paradigm based on querying by example.

Motivation:

There exists no way to directly measure the similarity between the content of images. Without the ability to measure similarity, it is impossible to treat images as queryable, searchable, or sortable data. As a result queries for images are typically satisfied by manually searching through all images in the entire image database. To automate this process, new techniques are needed to extract from an image qualities which can be used to make measurements of similarity.

Previous Work:

There are several organizations which support projects to develop new computer-vision algorithms to provide an automatic or semiautomatic method for performing such tasks.

To ask if one image is similar to another, one must specify, in some way, what criterion are to be used to make such a comparison. The principal developments to date, have come from the Query By Image Content (QBIC) project at IBM [1], the Visual Information Retrieval project at Virage [2], and the PhotoBook project in the MIT Media Lab [3]. All these projects form an image query using a single example image, and the selection of weights which determine the relative importance of each global image feature in measuring similarity. These projects have concentrated their effort on extracting a small number (fewer than ten) global image features. These features are essentially a collection of independently developed techniques such as color histograms, texture histograms, shape boundary descriptors, and eigenimages.

Approach:

In our image query paradigm, we describe similarity in terms of the difference between a test-image and a group of example-images. Thus, our methodology can be described as "a query by image example" where a query is formed by giving several examples which are indicative of the images wanted. We determine the relative importance of each global image feature in measuring similarity by computing the importance of each image feature in indicating membership in the group of example-images. Thus, the image features we use to measure similarity vary from query to query, depending on the group of example-images – for example, in one query chromatic-content may be the primary measure, while in another, spatial-arrangement might be penultimate. Because the selection of the relative importance of each image feature is derived from the set of example images (as opposed to being directly chosen by the user of the system) we can use a very large set of measurements of mathematical characteristics to describe an image; as opposed to limiting ourselves to using only a few, very specific image features, which correspond to those perceived by humans as salient.

The large set of measurements we make of each image are based on the responses of a collection of non-linear filter-networks. The responses of each filter-network in the collection are combined to form a characteristic signature for each image. These signatures used to measure the similarity between the images in the database and the group of example-images. The similarity measure of each image in the database is then used to rank and sort the images, satisfying the query.

Difficulty:

This query paradigm requires the extraction of a large, general and robust set of image features. Such a set must be complete enough to incorporate all the characteristics of an image that could potentially be needed criterion to measure similarity.

To meet this requirement, we use as our image features the responses of a large class of filter-networks. The numerical response of each filter network in this class is becomes element in what we call the characteristic signature of the image. It is this characteristic signature which is then used as the basis for image comparison.

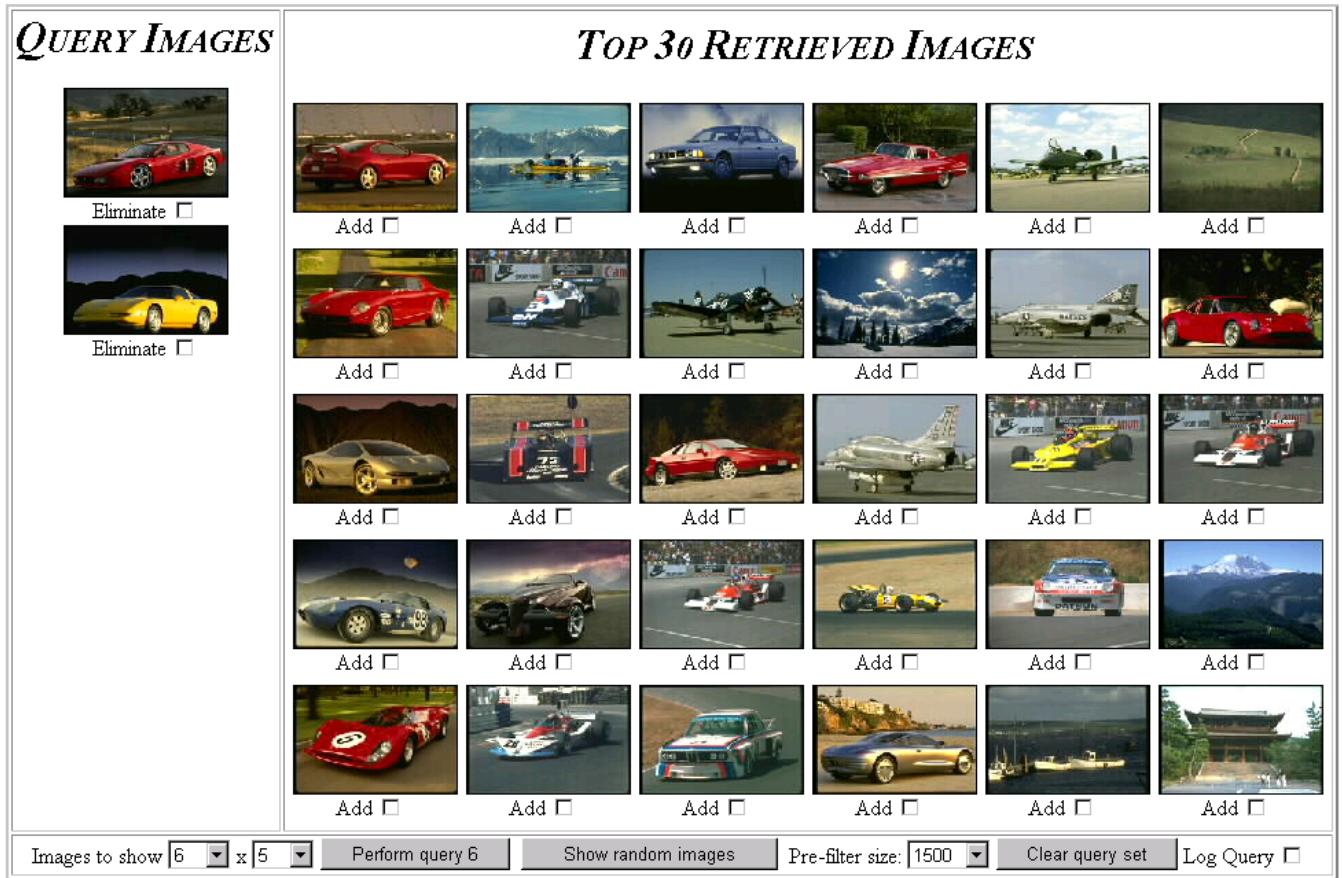


Figure 1: An example of this systems response to a two image query into a diverse 3000 image database.

Each filter network consists of several repetitions of a linear convolution operation, followed by a non-linear operation. The normalized sum of the result of these operations is (i.e. the response of the network) is a single element of the characteristic signature.

Impact:

With a functional metric over images, computers will be able to manipulate image data in the same way as they can currently manipulate text and numerical information. That is to say, they will be able to retrieve and sort information based on its visual content. This ability will cause a drastic shift in the tasks for which we use computers; a host of new applications will become possible.

Future work:

Future work in this area includes increasing the robustness of the underlying image representation, and building models based upon these techniques which will allow for unconstrained image recognition.

Research Support:

Research supported in part by DARPA under contract No. N00014-94-1-0128.

References:

- [1] QBIC Project, IBM Research, <http://www.qbic.almaden.ibm.com>
- [2] Virage, <http://www.virage.com>
- [3] Vision and Modeling Research Group, MIT Media Lab, Photobook Project, <http://www-white.media.mit.edu/vismod/der>