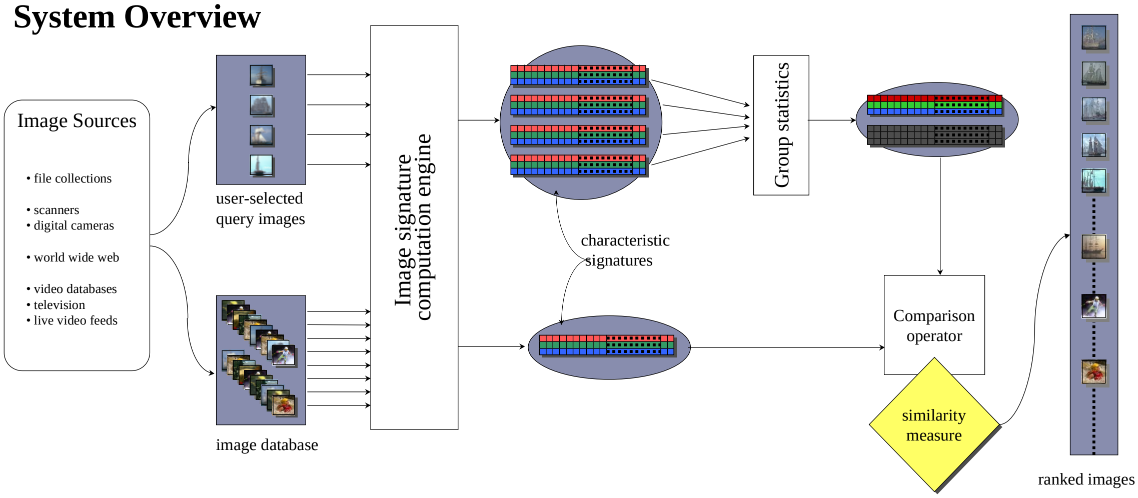
STRUCTURE DRIVEN IMAGE DATABASE RETRIEVAL

Paul Viola, Jeremy S. De Bonet, Kinh H. Tieu

Learning and Vision Group (http://www.ai.mit.edu/projects/lv), Artificial Intelligence Laboratory, Massachusetts Institute of Technology

System Overview



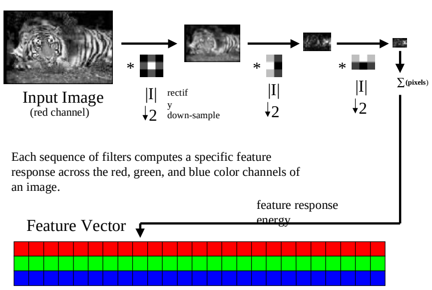
We present a system for example-based image database retrieval with automatically weighted, selective structural features and show results of the system on a database of 3000 images.

Images are transformed into a feature space which captures visual structure, texture and color using a tree of filters. Similarity is defined as the inverse of the distance in this "perceptual feature space."

Many image database systems use unselective global features such as color histograms. Queries are limited to a single positive example image and the user is required to hand-weigh the relative importance of features.

Our approach measures the statistical properties of an image such as response energies to local edge filters at various orientations. These responses are in turn used as input to another level of filtering. The resulting filtering sequence captures the response energy of the image to a particular selective structural feature. Our system allows the user to select both multiple positive and negative example images. Statistics computed from the set of query images are then used to automatically weigh the relative importance of features for that particular query.

Feature Computation



Each sequence of filters computes a specific feature response across the red, green, and blue color channels of an image.

Texture energy

$$E_i(I) = 2\downarrow\left[(F_i \otimes I)^2\right]$$

$$E_{i,j}(I) = 2\downarrow\left[\left(F_j \otimes E_i(I)\right)^2\right]$$

Textures-of-textures

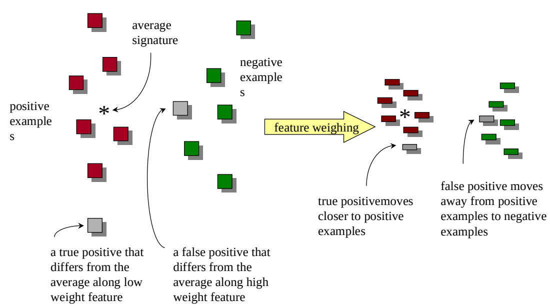
$$E_{i,j,k}(I) = 2\downarrow\left[\left(F_k \otimes E_{i,j}(I)\right)^2\right]$$

$$S_{i,j,k,c}(I) = \sum_{pixels} E_{i,j,k}(I_c)$$

Similarity Metric

$$Dist(t) = \sum_{e=0}^{25^3} \sum_{c=\{r,g,b\}} \left(\frac{\sigma^{PN}_{q_{e,c}}}{\sigma^{P}_{q_{e,c}}}\right)^2 \left(\frac{\overline{q_{e,c}} - t_{e,c}}{\sigma_{q_{e,c}}}\right)^2$$

Features which vary greatly across both the positive and negative examples but vary little across the positive images only are weighed more heavily. This is a diagonal approximation to Fisher's Linear Discriminant.

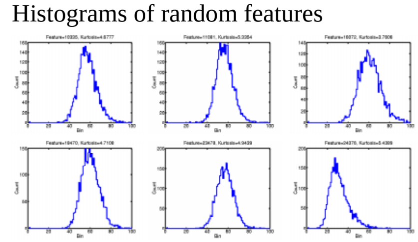


Feature Selection

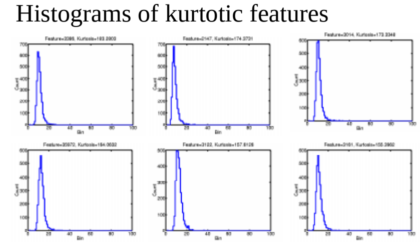
Although a large number of features provides a rich representation, it also requires a large amount of computation time and memory. In addition, some features may not lead to useful projections.

We've selected features with high kurtosis across a large sample of images. This chooses features which are highly selective and hence lead to more interesting non-Gaussian projections.

Histograms of random features



Histograms of kurtotic features



Results