Warped Metrics for Image Similarity Retrieval

R. Brunelli, **O. Mich**
Istituto per la Ricerca Scientifica e Tecnologica
I-38050 Povo, Trento, ITALY

June 7, 2001

Keywords: image similarity, image retrieval by content, relevance feedback, negative examples

Finding images in large repositories is largely an exploratory activity during which the user needs appropriate tools supporting progressive goal refinements and (inter)active exploration. The tool investigated in this paper is a system for image retrieval by example based on a user warpable image similarity metric, supporting positive and negative examples as well as relevance feedback.
The user may compose her query to the system in a visual way by providing images that are similar to those searched for (positive examples) as well as not-wanted images (negative examples). Besides this very intuitive way of formulating the query the user can modify the system response in three important ways each of them related to (and controlled by) a single scalar value:

1. by varying the strength with which negative examples modify the metric properties of the multidimensional image space (α);
2. by selecting the amount of adaptivity of the system in modeling the concept of similarity implied by the positive examples of the query: the strength of the relevance feedback can be explicitily set by the user (β);
3. by changing the way the *distance* of a database image from each single image in the query set is used to construct the resulting overall similarity: both disjunctive and conjuctive searches with the possibility of negated terms can then be used, providing the fuzzy counterpart of a subset of boolean searches (γ).

All of the above mentioned effects can be achieved by suitable modifications of a basic L_1 metric structure imposed to the space of the image descriptors (in our case vectors beloging to $\boldsymbol{R}^{288}$ quantifying the distribution of luminance, hue, etc. for several image subregions). These modifications can most appropriately be seen as warping (or stretching) the original space geometry by changing in a continuous way the parameters $\{\alpha, \beta, \gamma\}$ defining the following family of image similarity functions:

$$D'(\mathcal{I}, \boldsymbol{E}_R, \boldsymbol{E}_{\hat{R}}, \alpha, \beta, \gamma) = D(\mathcal{I}, \boldsymbol{E}_R, \beta, \gamma) \left[\frac{D(\mathcal{I}, \boldsymbol{E}_R, \beta, \gamma)}{D(\mathcal{I}, \boldsymbol{E}_{\hat{R}}, \beta, \gamma)} \right]^{\alpha} \tag{1}$$

where $\boldsymbol{E}_R$ represent the set of positive examples, $\boldsymbol{E}_{\hat{R}}$ that of negative examples, and the similarity of an image $\mathcal{I}$ to a given query set $\boldsymbol{E}$ is given by

$$D(\mathcal{I}, \boldsymbol{E}, \beta, \gamma) = \left[\frac{1}{n} \sum_{\mathcal{I}' \in \boldsymbol{E}} d(\mathcal{I}, \mathcal{I}', \beta)^{\gamma} \right]^{1/\gamma} \tag{2}$$

where n is the number of images in $\boldsymbol{E}$ and

$$d(\mathcal{I}, \mathcal{I}', \beta) = k_\beta \sum_i \frac{1}{\sigma_i^\beta} |F_i - F_i'| \tag{3}$$

where $k_\beta = \left(\sum_{lm} \sigma_{lm}^{-\beta}\right)^{-1}$ is a normalizing factor.
The warping behaviour necessary to change system response can be obtained through the modification of the parameter set $\{\alpha, \beta, \gamma\}$:

$\alpha \in [0, \infty)$: the larger the value the larger the impact of negative examples on the metric structure of the image space: larger and more repulsive volume of space are defined around the negative examples;

$\beta \in [0, \infty)$: given a set of images the system increases the importance of the coordinates along which the positive examples are more similar to each others: the lower the value of β, the smaller the relevance feedback effect;

$\gamma \in (-\infty, \infty)$: the lower the value the more disjunctive the query is (the minimum dissimilarity value gets more weight), while a value of $\gamma = 1$ represents the simple average.

By interactive modification of the control parameters the user can then modify the response of the system specifying in a intuitive way complex queries, resembling fuzzy equivalents of the more traditional boolean textual queries. The system is currently being tested on several databases with different visual contents (images of fonts, drawings and environment photos) and sizes ranging from a few thousand images to more than twenty thousands items.