Optimal Interactive Content-Based Image Retrieval

Nikolaos D. Doulamis, Anastasios D. Doulamis, and Klimis S. Ntalianis

Electrical and Computer Engineering Department
National Technical University of Athens
Heroon Polytechniou 9, Zografou 15773, Athens, Greece
E-mail: ndoulam@cs.ntua.gr

1. INTRODUCTION

No doubt, the performance of a Content-Based Image Retrieval (CBIR) system depends on a) how efficient the image visual content is represented and b) the degree of importance, which is assigned to each content-descriptor. In the first case, efficient visual representation is achieved, apart from the extraction of appropriate descriptors, through a proper organization of them [1]. The second case faces the problem arising from the subjectivity of humans, which often perceive the same visual content in a different way. This is, for example, the case when a person is interested in color information of an image, while another one, or even the same under different circumstances, in texture or motion information. Furthermore, even if both people are interested in the same type of indexing, e.g., color indexing, they may interpret the image content quite different. For instance, the first person may want to search for a particular object with specific color characteristics, while the second for the global brightness of an image. To overcome the aforementioned difficulty, the human can be considered as a part of the retrieval process resulting in an *interactive* CBIR scheme [2]. In particular, the retrieval results are evaluated by the user and a degree of relevance is assigned to some selected images so that the system response is adapted according to the user's demands. This interactive framework is usually called *relevance feedback*, similarly to the definition used in traditional text-based information systems [3].

Towards this direction, some CBIR systems incorporating relevance feedback algorithms have been proposed. In particular, a relevance feedback algorithm based on a probabilistic framework was reported in [4], while a relevance feedback algorithm, which exploits the variation of the feature elements to perform the weight updating, has been investigated in [2] and [5]. Extension of [5] to negative examples has been presented in [6] very recently, while in [7] the extension is performed to the interaction of different feature elements with each other. However, as mentioned in the conclusions of [5], there is a need for an optimal weight updating strategy. In this case, optimality means that the system should be adapted so that its response is as close as possible to all relevant selected images and as far as possible to all irrelevant images.

In this paper, an efficient optimal relevance feedback mechanism is presented to improve the performance of a content-based image retrieval system. In particular, the algorithm is based on an optimal weight updating strategy, which regulates the importance of each image descriptor to the retrieval response. The weights are updated so that the correlation between the query image and all selected images of high degree of relevance is maximized, while simultaneously the correlation over all selected images of low degree of relevance is minimized. The optimization results in an efficient scheme both in computational complexity and performance for the weight updating.

Convergence of the weight updating is achieved if the user submits "consistent" relevant or irrelevant images to the system, i.e., images of similar content. This means that the system response adapts its performance to the user's information needs in a stable manner.

2. INTERACTIVE CONTENT BASED IMAGE INDEXING (RELEVANCE FEEDBACK)

In the proposed relevance feedback mechanism, the following weighted similarity distance is used, to regulate the degree of importance of each image descriptor in the retrieval processes. Therefore, we have that

$$\rho_{\mathbf{w}}(\mathbf{f}_q,\mathbf{f}_i)=\frac{\sum_{k=0}^{P=Q^L-1} w_k\cdot f_{q,k}\cdot f_{i,k}}{\sqrt{\sum_{k=0}^{P=Q^L-1} w_k^2\cdot f_{q,k}^2}\cdot\sqrt{\sum_{k=0}^{P=Q^L-1} f_{i,k}^2}} \qquad (1)$$

where, vector $\mathbf{w}=[w_0\ w_1\cdots w_p]^T$ contains the weights w_k, k=0,1,…,P, each regulates the importance of the kth feature element to the similarity measure.

The set of weights w_k are selected according to the user's information needs, through a *relevance feedback mechanism*, where the user feeds back to the system the most relevant or irrelevant images. In the following of this paper, an optimal algorithm for weight updating is proposed.

2.1 Optimal Weight Updating

Let us assume that m out of the M best retrieved images are selected by a user. Let us denote as $\mathbf{y}_i$, i=1,…,m ($1 \le m \le M$) the feature vectors corresponding to these m selected images. Then, if we assume that all m images are relevant to the actual user's information needs, the weights w_k are estimated so that the following equation is maximized

$$C(\mathbf{w}) = \sum_{i=1}^{m} \rho_{\mathbf{w}}(\mathbf{f}_q, \mathbf{y}_i) = \\ = \sum_{i=1}^{m} \frac{\sum_{k=0}^{P=Q^L-1} w_k \cdot f_{q,k} \cdot y_{i,k}}{\sqrt{\sum_{k=0}^{P=Q^L-1} w_k^2 \cdot f_{q,k}^2} \cdot \sqrt{\sum_{k=0}^{P=Q^L-1} y_{i,k}^2}} \quad (2)$$

where $\mathbf{f}_q$ is the feature vector of the user's query, while $f_{q,k}$ and $y_{i,k}$ the kth element of $\mathbf{f}_q$ and $\mathbf{y}_i$ respectively. Different degree of relevance of the selected images can be also incorporated in the proposed scheme, by multiplying the feature elements of the selected images by a scalar $-1 \le \eta \le 1$, which expresses the image relevance. In our case, values of η close to -1 indicate irrelevant images, values of η close to 0 neutral images, while values of η close to 1 relevant images. In the following analysis, we are concentrated on the case of relevant images for simplicity.

The energy $E_{\mathbf{y}_i}$ of $\mathbf{y}_i$ is given as

$$E_{\mathbf{y}_i} = \sum_{k=0}^{P} y_{i,k}^2 \quad (3)$$

Thus, equation (2) can be written as

$$C(\mathbf{w}) = \sum_{i=1}^{m} \rho_{\mathbf{w}}(\mathbf{f}_q, \mathbf{y}_i) = \sum_{i=1}^{m} \frac{\sum_{k=0}^{P=Q^L-1} w_k \cdot f_{q,k} \cdot y_{i,k} / E_{\mathbf{y}_i}}{\sqrt{\sum_{k=0}^{P=Q^L-1} w_k^2 \cdot f_{q,k}^2}} \quad (4)$$

Maximization of $C(\mathbf{w})$ is performed by setting the derivatives $\partial C(\mathbf{w}) / \partial w_n = 0$, for all weights w_n n=0,…,P.

$$\sum_{i=1}^{m} \frac{y_{i,n}}{E_{\mathbf{y}_i}} \cdot \left(\sum_{k=0}^{P} w_k^2 f_{q,k}^2 \right) = \\ = \left(\sum_{i=1}^{m} \sum_{k=0}^{P} w_k f_{q,k} \frac{y_{i,k}}{E_{\mathbf{y}_i}} \right) \cdot w_n f_{q,n} \qquad \forall \; n=0,1,\ldots,P \quad (5)$$

As a result, (5) corresponds to a non-linear system of P+1 equations. By dividing two equations of the form shown in (5) one over the other, for example the ones corresponding to $\partial C(\mathbf{w}) / \partial w_n = 0$ and $\partial C(\mathbf{w}) / \partial w_l = 0$, $n \ne l$, the following relation of weights w_n and w_l, is obtained

$$w_n = w_l \cdot \frac{f_{q,l}}{f_{q,n}} \cdot \frac{\sum_{i=1}^{m} \frac{y_{i,n}}{E_{\mathbf{y}_i}}}{\sum_{i=1}^{m} \frac{y_{i,l}}{E_{\mathbf{y}_i}}} \quad \forall \, n \ne l \quad (6)$$

From (6) it is clear that, if one weight is known, the rest P weights can be easily estimated. However, substituting the weight ratio expressed in (6) to the system of (5), all the P+1 non-linear equations are satisfied. This means that equation (6) is the solution of the maximization problem of (4) and thus one weight out of P is a free variable. This can be explained by the properties of $C(\mathbf{w})$. An additional equation for estimating the weight vector $\mathbf{w}$ is to consider the L_2 -norm of the weights equal to one, i.e., $\|\mathbf{w}\|_2 = 1$

Let us assume, without loss of generality, that the free variable is the first weight w_0, i.e., l=0. Then, expressing all weights w_n, $n \ne 0$ with respect to w_0 and using equations (6) and the constrains $\|\mathbf{w}\|_2 = 1$, we can conclude that

$$w_0 = \sqrt{\frac{1}{1 + \sum_{k=1}^{P} A_k^2}} \;, \qquad A_k = \frac{f_{q,0}}{f_{q,k}} \cdot \frac{\sum_{i=1}^{m} \frac{y_{i,k}}{E_{\mathbf{y}_i}}}{\sum_{i=1}^{m} \frac{y_{i,0}}{E_{\mathbf{y}_i}}} \quad (7)$$

In the equation (7) we have selected the possible solution since this leads to the maximization of $C(\mathbf{w})$. In case that images of different degree of relevance are taken into consideration the negative solution may be also appropriate and the solution which maximizes the measure $C(\mathbf{w})$ is selected.

3. EXPERIMENTAL RESULTS

The proposed optimal relevance feedback scheme is evaluated both on synthetic and real life experiment. The first is performed for better comprehension of the proposed weight updating strategy.

In the synthetic experiment, the database consists of 51 frames, which depict a red circle moving in an elliptic trajectory in a static black background. The shot movement of the circle is depicted in Figure 1(a), where we have superposed in a single frame eight different characteristic position of the circle. Figure 1(b) illustrates the correspondence of the indices of the frames with the angle θ of the circle, which is defined as Figure 1(a) indicates. Angle θ provides a better representation of the circle trajectory.

Initially, we assume that the user submit as query to the database, the image of Figure 2(a). This image presents the circle at angle about θ=-180^o. The variation of the correlation measure over all frames of the database, (i.e., over all angles θ) is illustrated in Figure 2(b) ["before feedback"]. Let us, then, assume that the user is interested for a circle at position θ=0^o. For this reason, the user selects the image of Figure 2(a) as relevant. The system response over all angles θ after feedback iteration is depicted in Figure 2(b). It is observed that after the adaptation circles with angles θ close to zero are more relevant than circles with angles close to -180^o. The weight values before and after feedback iteration are also depicted in Figure 2(c). Another relevance feedback example is presented in Figure 2(c,d). In this Figure the user selects the images of angle about θ=0^o and θ=-90^{o,} as relevant. In this case the initial query is the same as in Figure 2(a). The system response before and after feedback iteration is illustrated in Figure 2(d). It is also observed, that the weight updating process correctly adapts the system to the user's needs.

In the following, the convergence of the optimal relevance feedback algorithm is presented. For this purpose, a multiple relevance feedback scheme is assumed in which only one image is selected as relevant at each iteration (feedback). Figure 3(a) shows the selected images versus the number of iteration.. As is observed, the user's target is the angle of 90^o. However, this is reached with an "inconsistency" which is decreasing with the number of iterations. The variation of two different weights versus the number of feedback iterations is shown in Figure 3(b,c). In this figure, three different values of factor λ have been used; λ=1, λ=0.9 and λ=0.1. As can be seen, small values of λ make the weights be sensitive to the current (maybe noisy) user's selection. On the contrary, as the values of λ increase, the weight variation becomes smoother, since the effect of images which have been selected as relevant at a previous iteration increases. Usually, values of λ around 0.8-0.9 provide a sufficient weight convergence and avoid abrupt weight variation from iteration to iteration.

In the following, the performance of the optimal relevance feedback mechanism is evaluated using a database of real-life images. Let us assume that the user initially submits the mage of Figure 4, which depicts a space shuttle (users' query). In this Figure, we present the system response before the feedback iteration. We assume that the user is actually interested in finding an image depicting the launching of a spacecraft. For this reason he selects the two last images as relevant to his needs. The system response after the relevance feedback is illustrated in Figure 4, where it can be seen that now most of images depict a spacecraft launching, which is in agreement with the actual user's information needs.

4. REFERENCES

[1] A. Doulamis, N. Doulamis, and S. Kollias, "A fuzzy video content representation for video summarization and content-based retrieval," *Signal Processing*, Vol. 80, pp. 1049-1067, June 2000

[2] Y. Rui, T. S. Huang and S. Mehrotra, "Relevance Feedback Techniques in Interactive Content-based Image Retrieval," in Proc. *IS&T and SPIE Storage and Retrieval of Image and Video Databases VI*, San Jose, CA, USA, Jan. 1998.

[3] C. Buckley and G.Salton, "Optimization of Relevance Feedback Weights," in Proc of *SIGIR* 1995.

[4] I. Cox, M. L. Miller, S. M. Omohundro and P. N. Yianilos, "Pichunter: Bayesian Relevance Feedback for Image Retrieval," in Proc. *Inter. Conf. Pattern Recognition*, Vol. 3, pp. 362-369, 1996.

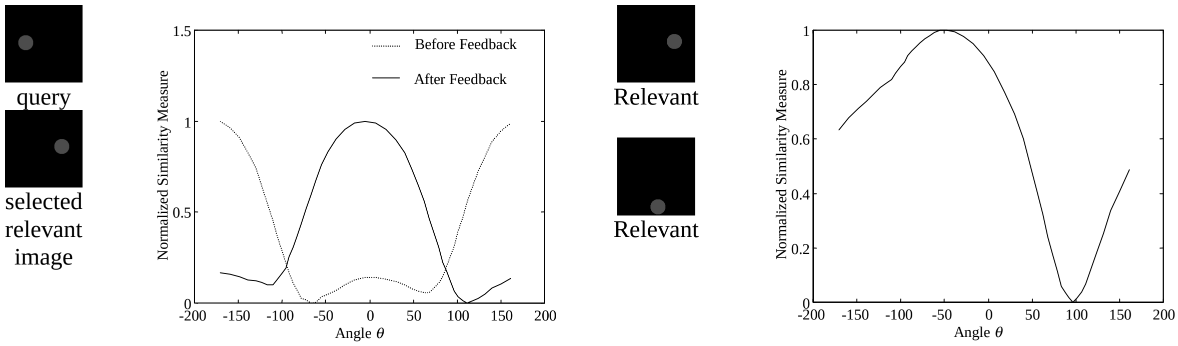
[5] Y. Rui, T. S. Huang, M. Ortega and S. Mehrotra, “Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval,” *IEEE Trans. Circuits. Systems for Video Technology*, Vol. 8, No. 5, pp. 644-655, Sept. 1998.

[6] Q. Tian, P. Hong and T. Huang, "Update Relevant Image Weights for Content-based Image Retrieval Using Support vector Machines," in Proc. *IEEE Inter. Conf. on Multimedia & Expo*, pp. 1199-1202, New York, USA, August 2000.

[7] Y. Choi, D. Kim, and R. Krishnapuram, "Relevance Feedback for Content-based Image Retrieval using Choquet Integral," in Proc. *IEEE Inter. Conf. on Multimedia & Expo*, pp. 1207-1210, New York, USA, August 2000.

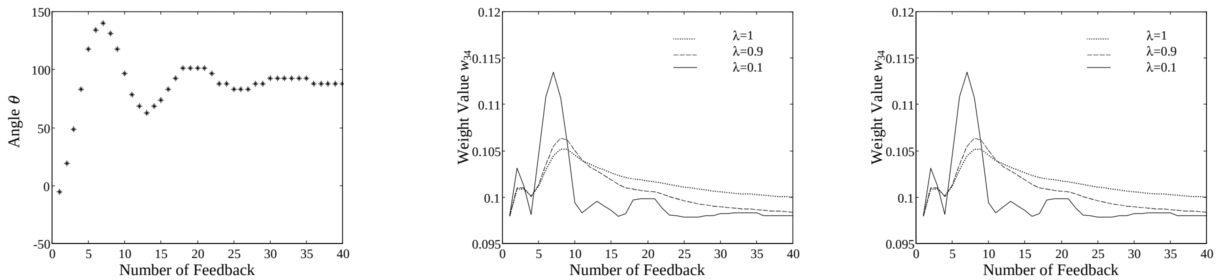
(a) (b)

Figure 1. (a) A synthetic database which represents a red circle moving in a elliptic trajectory. (d) The correspondence between frame number and angle θ of the circle position.



(a) (b) (c)

Figure 2. An example of the proposed optimal relevance feedback algorithm. (a) The query and the selected relevant image. (b) The similarity measure versus angle θ before and after relevance feedback. (c) The weight values before and after relevance feedback



(a) (b) (c)

Figure 3. Convergence of the proposed optimal relevance feedback scheme. (a) The selected relevant images. (b,c) The weight updating before and after the relevance feedback.

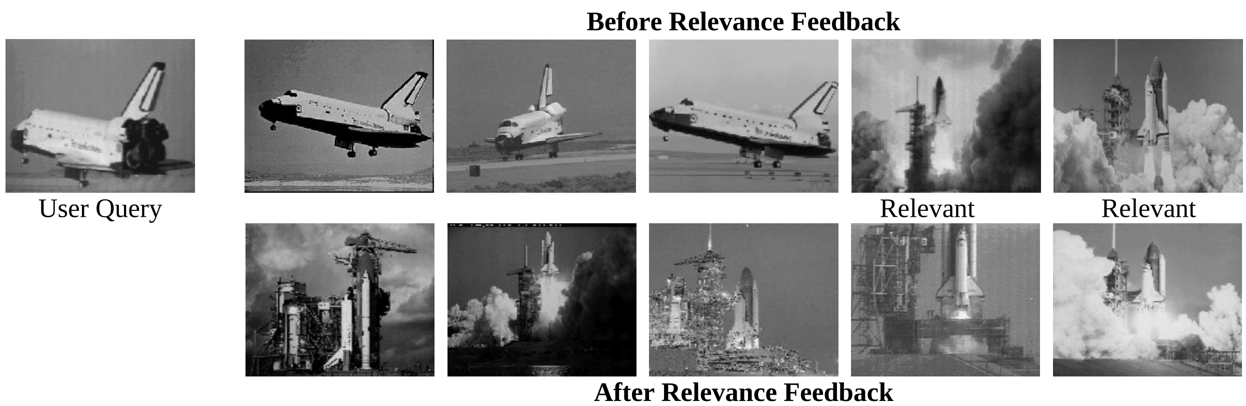


Figure 4. The optimal relevance feedback algorithm in a real experiment.