FEATURE SPACE WARPING RELEVANCE FEEDBACK WITH TRANSDUCTIVE LEARNING

Daniele Borghesani, Dalia Coppi Costantino Grana, Simone Calderara, and Rita Cucchiara

University of Modena and Reggio Emilia, Modena, Italy
WWW home page:http://imagelab.ing.unimore.it

Abstract. Relevance feedback is a widely adopted approach to improve content-based information retrieval systems by keeping the user in the retrieval loop. Among the fundamental relevance feedback approaches, feature space warping has been proposed as an effective approach for bridging the gap between high-level semantics and the low-level features. Recently, combination of feature space warping and query point movement techniques has been proposed in contrast to learning based approaches, showing good performance under different data distributions. In this paper we propose to merge feature space warping and transductive learning, in order to benefit from both the ability of adapting data to the user hints and the information coming from unlabeled samples. Experimental results on an image retrieval task reveal significant performance improvements from the proposed method.

Keywords: Relevance feedback, covariance matrices, transductive learning, feature space warping

1 Introduction

The use of relevance feedback strategies in information retrieval and in particular in content-based image retrieval systems is widely considered a very precious (sometimes necessary) addition to the system itself. Actually it is the most effective way to capture user's information need and, more generally, user's search intention. The reason is pretty straightforward: the automatic association of low-level features to high-level semantics is still a very open problem, and the only practical way to identify what the user is looking for is by including him in the retrieval loop, with the input of feedbacks (positive, negative or both). The common scenario in which relevance feedback is used within content-based image retrieval systems is the following:

1. An initial query-by-keyword or query-by-example is performed, in the form of a list of results ranked with increasing distances from the query in the feature space;

2. The user provide some good (and bad, implicitly or explicitly) feedbacks given the displayed images, choosing in other words what is relevant and what is irrelevant;
3. An algorithm uses these information to change the displayed results in a "refinement" step to accommodate user's judgements;
4. Back to step 2 and loop until a certain condition (or satisfaction) is reached.

In this paper, we focus on the third step proposing a new and effective strategy for relevance feedback based on Transductive Learning. The main contribution of this work is the joint use of Transductive Learning (successfully applied to a wide variety of pattern recognition and classification tasks) with the Feature Space Warping, a widely used technique for relevance feedback which aims at modifying accordingly the relations between feature points: objects similar to positive feedbacks go closer to the query, while object dissimilar are pushed away. We show that the union of these two techniques overcome the respective limitations providing a significant boost in performance over the two techniques used alone.

2 Related work

The literature on this topic is countless [5, 22], since this problem can be faced from several point of view (computer vision, database management, human-computer interaction, artificial intelligence, even psychology). Moreover, aside the research on the algorithm for relevance feedback, there is a wide literature about the way in which the performance of a system with relevance feedback can be safely evaluated in order to provide fair comparison with different techniques. Regarding the algorithms, we can identify very generally three classes. In the first one (called Query Point Movement, QPM in short), we try to move the query point in order to create a more complete query (a fast technique to overcome to slow convergence is proposed by [10]). In the second one (called Feature Space Warping, FSW in short), we try instead to manipulating the feature space or the metric space, in order to shape it in the direction of the users' feedbacks [1, 12]. The third one applies some machine learning procedures (like SVM or Adaboost) to learn how to separate relevant samples from irrelevant ones [16,17]. Among the usual techniques, based on SVM or boosting, we preferred testing a transduction-based learning. Some author followed the same path (see [13–15, 20]). The idea is to take advantage both of the unlabeled and labeled samples in a transductive inference manner, learning from an incremental amount of training samples (feedbacks, in this case).

Given the algorithm, the problem of evaluation is controversy. Even back in the Seventies, Williamson [19] proposed an evaluation methodology to tackle the so called "ranking effect" in the "fluid" relevance feedback evaluation, i.e. the overestimated performance improvement (in terms of recall and precision) due to the reposition of positive feedback in the top of the rank, aside the under-estimated performance improvement of the "frozen" relevance feedback evaluation, which maintain the original ranks of documents along the sessions. In his "reranked original" ranking proposal, the best ranks are assigned to thy relevant documents and the worst ranks to the non-relevant documents; those documents not yet judged would remain in their original order, but with a rank decreased by the number of non-relevant documents identified. In [9] a comprehensive analysis tries to find out the reasons of relevance feedback evaluation problems, in particular problems with the dataset (characteristics and relative ground truth), problem with the comparison (different measures, different ranking approaches that make a comparison unfair, the need of rank normalization), and finally the problem of the parameter settings which can be impractical in real context. In our opinion, a quite fair and complete set of measures has been proposed in [11], where authors proposed:

- *actual* recall and precision (computed at each iteration and relative to the current set of retrieved images solely)
- *new* recall and precision (computed at each iteration and relative to the previous set of retrieved images solely)
- *cumulative* recall and precision (computed at each iteration and relative to the whole set iterations so far)

In this way, we can describe the behavior of the retrieval system both in terms of speed (how fast valuable images are retrieved over time) and in terms of completeness (how many good images the retrieval system finds out globally). Finally, as suggested in [8], we tried to concentrate the analysis on *feasible search tasks*, i.e. visual topics with a good number of representatives with a low degree of uncertainty in the evaluation, in order to assure a valuable reference ground truth.

3 Visual similarity using Covariance matrices

In order to accomplish an effective similarity retrieval upon these images, we relied on a simple yet effective feature which allows to consider both color and edge based information, that is covariance matrices. Computing the covariance region descriptor from multiple information sources yields a straightforward technique for a low-dimensional feature representation [18].

The covariance matrices do not form a vector space, but if we concentrate on nonsingular covariance matrices, we can observe that they are symmetric positive definite, and as such they lay on a Riemannian manifold.

For the image retrieval task, we use normalized pixel locations $(x/W, y/H)$, RGB values in the range $[0, 1]$ and the norm of the first derivatives of the intensity with respect to x and y, calculated through the filter $[-1\ \ 0\ \ 1]^T$. The covariance of a region is thus a 7×7 matrix.

In order to rank images by visual similarity to a given query, we need to measure the distance between covariance matrices. As already mentioned, the covariance matrices do not lie on Euclidean space, thus in [6] the following distance measure for positive definite symmetric matrices is proposed:

$$\rho(\mathbf{X}, \mathbf{Y}) = \sqrt{\sum_{i=1}^{d} \ln^2 \lambda_i(\mathbf{X}, \mathbf{Y})} \tag{1}$$

where $\{\lambda_i(\mathbf{X}, \mathbf{Y})\}_{i=1..d}$ are the generalized eigenvalues of two non singular covariance matrices $\mathbf{X}$ and $\mathbf{Y}$.

Unfortunately distance alone is not enough for our purposes. In fact to enable the user to provide relevance feedbacks, we need to work on an Euclidean space, which allows us to move the query and the other points with linear combinations. To this aim two steps are required: the projection on the tangent space, and the extraction of the orthonormal coordinates of the tangent vector. By combining these two steps, the projection $\mathbf{y}$ of a covariance matrix $\mathbf{Y}$ on the hyperplane tangent to covariance matrix $\mathbf{X}$ can be written as

$$\mathbf{y} = \mathrm{vec}\left(\log\left(\mathbf{X}^{-\frac{1}{2}}\mathbf{Y}\mathbf{X}^{-\frac{1}{2}}\right)\right) \tag{2}$$

where the vector operator is defined as

$$\mathrm{vec}(\mathbf{y}) = \left[y_{1,1}\ \sqrt{2}y_{1,2}\ \sqrt{2}y_{1,3} \ldots y_{2,2}\ \sqrt{2}y_{2,3} \ldots y_{d,d}\right] \tag{3}$$

In this way, after the selection of an appropriate projection origin, every covariance matrix gets projected to a 28-dimensional feature vector laying on an Euclidean space.

This process is easily invertible. We can compute the relative covariance matrix in the Riemannian Manifold starting from the 28-dimensional feature vector using the following formulation:

$$\mathbf{Y} = \mathbf{X}^{\frac{1}{2}} \exp\left(\mathrm{vec}^{-1}(\mathbf{y})\right) \mathbf{X}^{\frac{1}{2}} \tag{4}$$

4 Mean-Shift Feature Space Warping with Remapping

In this work, we started from the relevance feedback technique proposed by Chang *et al.* [4] called Mean Shift Feature Space Warping (MSFSW). Given a query point $\mathbf{q}$ in the feature vector space, k samples are retrieved by nearest neighbor search. By examining the results, the user provides his feedback by specifying the relevance of M of these samples, forming two sets: $\{\mathbf{f}_p\}$ and $\{\mathbf{f}_n\}$, the relevant and irrelevant sets respectively. These are employed to move all data samples $\{\mathbf{p}\}$ toward or away from the warping center $\mathbf{w}$. In particular, for each $\mathbf{p}$, its warped point $\mathbf{p}'$ is given by

$$\mathbf{p}' = \mathbf{p} + \lambda \sum_{j=1}^{M} u_j \exp\left(-c|\mathbf{p} - \mathbf{f}_j|\right)(\mathbf{w} - \mathbf{p}) \tag{5}$$

where the scalar value u_j is set to $+1$ if $\mathbf{f}_j \in \{\mathbf{f}_p\}$, and -1 if $\mathbf{f}_j \in \{\mathbf{f}_n\}$. Two global coefficients c and λ are required to control the influence of each feedback to each sample and the maximum moving factor of any point $\mathbf{p}$ toward or away from the warping center $\mathbf{w}$.

The original FSW algorithm fixes the warping center $\mathbf{w}$ to $\mathbf{q}$. Thus, the query point will always stay in its original position. Other points will move toward or far away from $\mathbf{q}$ based on its proximity to relevant and irrelevant sets. But, according to the analysis proposed in [4], FSW algorithm tends to perform poorly under Gaussian distributions when the query point is far away from the cluster center. For this reason, in the MSFSW, authors proposed to move the warping center instead of staying at $\mathbf{q}$. They suggest to adopt the Rocchio's query movement formula:

$$\mathbf{w}' = \alpha\mathbf{w} + \beta\overline{\mathbf{f}_p} - \gamma\overline{\mathbf{f}_n} \tag{6}$$

where $\mathbf{w}$ is the warping center (initially set to $\mathbf{q}$), $\overline{\mathbf{f}_p}$ and $\overline{\mathbf{f}_n}$ are the mean of the set $\{\mathbf{f}_p\}$ and $\{\mathbf{f}_n\}$. Another set of parameters α,β and γ is required, and must be tuned to optimize the performance.

With the above formulations, the MSFSW algorithm provides a flexible parameterization for switching between the two extreme algorithms: QPM by setting $\alpha = \gamma = \lambda = 0$ and $\beta = 1$, and FSW by setting $\alpha = 1$ and $\beta = \gamma = 0$. Given the final user target of our application, exposing the parameters configuration to the user was out of question. Thus, we determined the parameters configuration which provided best results on a small initial training set, using an automatic exhaustive search procedure.

From the above equations, it is clear that we need a way to compute a linear combination of the feature vectors. For this reason, we employed the projection of the covariance matrices on the tangent space previously described (Eq. 2). As mentioned before, the projection requires a point from which determine the orthonormal coordinates of the tangent vector (i.e. the vector in the Euclidean space). Our experiments confirm that the choice of this point is fundamental to guarantee an optimal correspondence between the distances computed on the Riemannian manifold and those computed on the tangent space.

Thus, when the user requires a refinement of a similarity search of a previously selected image, we project the whole feature space on the chosen query point (i.e. the covariance matrix of the selected image), then we rank the results and show them to the user in order to perform further refinements.

Differently from the original formulation, we observed that the use of weighted negative samples, as well as the the weighted shifting of the original query, can have a negative impact on the performance of the algorithm especially when the discriminative power of the feature used is lacking for a particular class of images, or when visually similar images containing semantically different content are discarded by the user contributing on pushing away positive samples (in visual terms) yet to retrieve. For this reason, we modified the Mean Shift part of the algorithm: at each step we perform a new map, from the Euclidean space back to the Riemannian Manifold (by means of Eq. 4), using the mean of the positive feedbacks collected so far as reference. The new query is then used to project the whole feature space, then we rank the results and show them to the user in order to perform further refinements using the same approach. In this way we can also simplify the system getting rid of three additional parameters (α,β and γ). Thus, the algorithm proceeds as follows:

1. Mapping of the entire feature set from Riemannian Manifold to Euclidean Space using the query as tangent point
2. Computation of the mean point of the positive feedbacks
3. Conversion of the mean point from the Euclidean Space back to the Riemannian Manifold
4. Remapping of the entire feature set using the mean of positive feedbacks as tangent point

5 Transductive relevance feedback

As mentioned in the Section 2, the relevance feedback problem can be analyzed as a semisupervised learning problem, in which the positive and the negative feedbacks given by the users constitute iteratively (and incrementally) the training set of the algorithm. In this paper, we propose a graph-based transductive learning method to tackle this purpose, defining a graph where the vertices represent the labeled and unlabeled images of the dataset, while the edges incorporate the similarity between them, in our case obtained from the distance between covariance matrices. Graph-based methods are nonparametric, discriminative, and transductive by definition [21], and labels can be assumed to be smooth over the graph. Starting from the whole dataset with n images, let's define a set $\mathcal{L}$ of labeled images $(x_1, y_1), \ldots, (x_l, y_l)$ in which C classes are defined, so $y_i \in 1 \ldots, C$. The other images belongs to a set U of u unlabeled images, with $n = l + u$. Now let's define a function $f : \mathbb{R}^n \to [0, 1]$ which denotes the confidence of each image to one class. Formally, we can define a cost function J on f as:

$$J(f) = \sum_{(i,j)=1}^{n} \|f(x_i) - f(x_j)\|^2 w_{ij} + \lambda \sum_{i=1}^{l} \|f(x_i) - y_i\|^2 \tag{7}$$

with λ as a regularization parameter (in our case, $\lambda = 1$). This equation, in a minimization process, tries to match the confidence w_{ij} between samples with the the true confidence y_i with respect of the current confidence. Once converted in matrix notation, Eq. 7 becomes:

$$J(f) = (f(X) - Y)^T (f(X) - Y) + \lambda f(X)^T L f(X) \tag{8}$$

where $L = D - W$ is the graph Laplacian. W is the weight matrix, while D is the matrix which represents the degree of vertices:

$$d_{ii} = \sum_{j=1}^{n} w_{ij} \tag{9}$$

The cost function minimization has a closed solution, that is:

$$f = (I - \lambda L)^{-}1Y \tag{10}$$

At the end of the computation, f contains the class confidence of each new sample.

In the transductive process, we want to transfer labels from labeled samples to unlabeled ones: in other words, we want the samples which are close in the feature space to share the same label. To satisfy this *local constraint*, we construct accordingly the weight matrix W, that is the matrix in which each element w_{ij} contains the relation between two vertices (thus two images) xi, xj. To move from distances to affinities, we use the following formulation:

$$w_{i,j} = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{11}$$

where σ is a bandwidth parameter to tune the relations between vertices, and the distance is computed as the L_2 norm after the conversion of covariance matrices in the Riemannian Manifold to vectors in the Euclidean Space. W can be subdivided into four submatrices:

$$W = \begin{pmatrix} W^{ll} & W^{lu} \\ W^{ul} & W^{uu} \end{pmatrix} \tag{12}$$

where W^{ll} (a full connected graph) denotes relations between labeled data, W^{uu} (a k-nearest neighbor graph) denotes relations between the candidate images yet to label, and the symmetric subgraphs W^{ul} and W^{lu} (still k-nearest neighbor) denote the relations between positive and candidate images. The relations used to compute values in these graphs are the following:

$$W^{ll}_{i,j} = \frac{1}{n} \tag{13}$$

$$W^{uu}_{i,j} = \begin{cases} \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2_{uu}}\right) & \text{if } x_i \in knn(x_j); \\ 0 & \text{otherwise;} \end{cases} \tag{14}$$

$$W^{lu}_{i,j} = W^{ul}_{i,j} = \begin{cases} \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2_{uu}}\right) & \text{if } x_i \in knn(x_j); \\ 0 & \text{otherwise;} \end{cases} \tag{15}$$

where $k = 10$ and $\sigma = 1$.

After the first ranking by similarity, the user selects the positive feedbacks while the unselected samples are considered as negative. Then the process described in this section is iterated following the user's need or until no more changes in the rank occurs.

6 Transductive relevance feedback with Feature Space Warping

The main benefit regarding the use of FSW is the possibility to move potentially irrelevant samples away from the query center, while attracting far away relevant samples toward the query center. This characteristic turns out to be very important in the the transductive learning context, especially when the samples retrieved by the system using similarity only are poor. Recalling that covariance matrices express points lying on a Riemannian Manifold and their distances, according to (1), are the geodesic distances on the manifold, we can motivate the choice of the graph Laplacian transductive approach observing that during feedback iterations we try to learn the underlying geometry of the manifold composed by positive query results points as realizations. Since the Laplacian is the discretized approximation of the Laplace-Beltrami operator on the manifold [3], we can learn the manifold structure directly from the analysis of the graph Laplacian itself.

The basic idea is to strengthen, in the Laplacian, the contribution of positive query points while weakening the contribution of negatives one. This is equivalent, under a continuous relaxation of the graph affinities, to adding and removing links and path in the graph. With this premises, the user intervention allows to enhance the geodesic distances to better represent the geometry of the manifold where the query results are lying on. Additionally, it is important to remark that for transductive methods, the ratio between the number of labeled and unlabeled samples is fundamental: if the number of training samples is low as well as the number of unlabeled data goes to infinity, the learning procedure leads to uninformative membership functions [2].

To overcome this problem, we introduced a further step in the transductive learning procedure, modifying the value of W^{lu} and W^{ul} elements in the affinity matrix in analogy with Eq. 5. In particular, given P the set of positive samples indexes and N the set of negative ones, for each element of the affinity matrix W^{lu}_{ij}, its warped version $W^{lu'}_{ij}$ is corrected by factors δ^p_j and δ^n_j:

$$\delta^p_j = -\sum_{i \in P} \exp\left(-cW_{ij}\right) \tag{16}$$

$$\delta^n_j = \sum_{i \in N} \exp\left(-cW_{ij}\right) \tag{17}$$

The warped elements $W^{lu'}_{ij}$ are finally described by the following equation:

$$W^{lu'}_{ij} = W^{lu}_{ij} + \lambda\left(\delta^p_j + \delta^n_j\right) \tag{18}$$

where the global coefficients c and λ assume the same tuning functionalities as described in Section 4.

7 Experimental results

In order to verify the effectiveness of the proposed approach, we report the results on a publicly available and challenging historical images dataset [1], created using the procedure described in [7], and composed of 2282 pictures. We performed an

[1] http://imagelab.ing.unimore.it/files/imp.zip

Fig. 1. Samples taken from the 6 query classes.

automatic simulation of relevance feedback interaction, in order to avoid human errors. We evaluate the techniques using the 171 visual queries provided with the dataset annotation, in which several prototypes of 6 object classes were retrieved (Fig. 1). We compared the following algorithms:

– Naive relevance feedback (actually no relevance feedback at all): the system discards the current set of n results and proposes to the user the next n, following the original rank given by the visual similarity;
– MSFSW: original Mean Shift Feature Space Warping proposal by [10], with an empirically optimized set of parameter $\alpha = 0.2$, $\beta = 0.5$ and $\gamma = 0.3$ for the means-shift part and $\lambda = 0.7$ and $c = 0.8$;
– MSFSW with remapping: our modification of the original MSFSW algorithm which performs a remapping of the entire feature set using the mean of positive feedbacks as tangent point for the conversion from Riemannian Manifold to Euclidean Space;
– TL with NN: the transductive learning approach which uses positive feedbacks as samples and defines the relevance feedback as a process to assign a label to unlabeled samples. The affinity matrix is filled only for the $k = 20$ nearest neighbors;
– TL with FSW: the same transductive learning approach jointly used with Feature Space Warping on the affinity matrix

Performance has been evaluated with a user-centric perspective: we choose to use metrics clearly easy to comprehend by a user in front of the application, and we included a fixed number $T = 10$ of iterations, to convey that the user will get bored and stop pursuing in the search at most after 10 refinements. We chose cumulative recall as representative metric, namely the recall provided by the system at each step $i = 1 \dots T$. While the first steps give an idea of the convergence capabilities of the algorithm, the last step give an overall evaluation of the algorithm itself. The results are presented in Table 1 and Fig. 2.

The first step in the chart is the retrieval by similarity, sorted by increasing distances. The naive procedure is the baseline for the comparison. The original MSFSW proposal has a good behavior in the first step, but the deformation after the first projection limits its effectiveness in the following steps, probably due to the side effects highlighted in Section 4. Much better performance are obtained with remapping on the mean of positive feedbacks, which continually improve in substantial way up to the forth step. In the next steps, the improvement remains marginal, mimicking the behavior of the original version. The

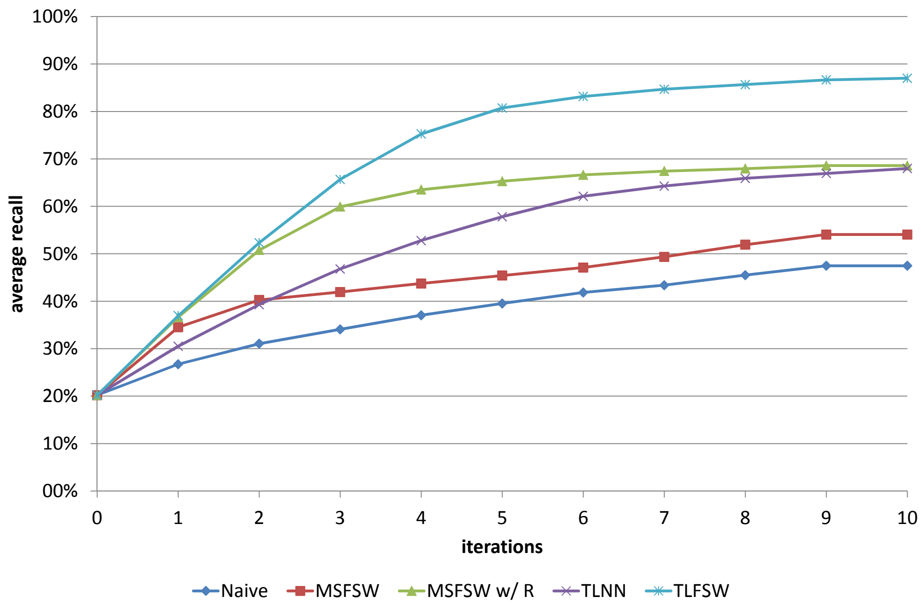

Fig. 2. Comparison of the proposed techniques in terms of recall at each iteration step.

Transductive Learning approach has a much slower gradient: the performance is comparable in the first steps, but the number of steps required to gather the same performance of the other technique is too high. The proposal of this paper, instead, shows the steepest gradient in the first steps and, from the fifth steps on, it maintains an considerable increasing improvement over the best so far (MSFSW with remapping) up to 19%.

Table 1. Recall values at different iteration steps.

method	1	2	3	4	5	10
Naive	26,7	31,1	34,1	37,1	39,5	47,5
MSFSW	34,5	40,3	41,9	43,7	45,4	54,1
MSFSW w R	36,6	50,8	59,9	63,5	65,3	68,6
TLNN	30,5	39,3	46,8	52,8	57,8	68,0
TLFSW	37,0	52,3	65,7	75,3	80,7	87,0

8 Conclusions

In this paper we presented an innovative relevance feedback strategy which merges a transduction-based relevance feedback with the advantages of the feature space warping. In our tests, this procedure appears to be extremely promising. In the future, we plan to improve this technique in the scalability side, in other words to test it on large-scale image collections.

References

1. H.Y. Bang and T.H. Chen. Feature space warping: an approach to relevance feedback. In *IEEE International Conference on Image Processing*, pages 968–971, 2002.
2. B. Belkin, N. Srebro, and X. Zhou. Statistical analysis of semi-supervised learning: The limit of infinite unlabelled data. In *Neural Information Processing Systems*, 2009.
3. Mikhail Belkin and Partha Niyogi. Semi-supervised learning on riemannian manifolds. *Machine Learning*, 56(1-3):209–239, June 2004.
4. Y.J. Chang, K. Kamataki, and T.H. Chen. Mean shift feature space warping for relevance feedback. In *IEEE International Conference on Image Processing*, pages 1849–1852, 2009.
5. Michel Crucianu, Marin Ferecatu, and Nozha Boujemaa. Relevance feedback for image retrieval: a short survey. In *In State of the Art in Audiovisual Content-Based Retrieval, Information Universal Access and Interaction including Datamodels and Languages (DELOS2 Report*, 2004.
6. Wolfgang Förstner and Boudewijn Moonen. A metric for covariance matrices. Technical report, Stuttgart University, 1999.
7. Costantino Grana, Daniele Borghesani, and Rita Cucchiara. Picture Extraction from Digitized Historical Manuscripts. In *ACM International Conference on Image and Video Retrieval*, July 2009.
8. Mark J. Huiskes and Michael S. Lew. Performance evaluation of relevance feedback methods. In *ACM International Conference on Image and Video Retrieval*, pages 239–248, 2008.
9. Xiangyu Jin, James French, and Jonathan Michel. Toward consistent evaluation of relevance feedback approaches in multimedia retrieval. In *International Workshop on Adaptive Multimedia Retrival*, July 2005.
10. Danzhou Liu, K.A. Hua, Khanh Vu, and Ning Yu. Fast query point movement techniques for large cbir systems. *IEEE Transactions on Knowledge and Data Engineering*, 21(5):729–743, May 2009.
11. Jie Luo and Mario A. Nascimento. Content-based sub-image retrieval using relevance feedback. In *ACM international workshop on Multimedia databases*, pages 2–9, 2004.
12. G.P. Nguyen, M. Worring, and A.W.M. Smeulders. Interactive search by direct manipulation of dissimilarity space. *Multimedia, IEEE Transactions on*, 9(7):1404–1415, November 2007.
13. V. Radosavljevic, N. Kojic, G. Zajic, and B. Reljin. The use of unlabeled data in image retrieval with relevance feedback. In *Symposium on Neural Network Applications in Electrical Engineering*, pages 21–26, September 2008.
14. H. Sahbi, J.-Y. Audibert, and R. Keriven. Graph-cut transducers for relevance feedback in content based image retrieval. In *IEEE International Conference on Computer Vision*, pages 1–8, October 2007.
15. H. Sahbi, P. Etyngier, J.-Y. Audibert, and R. Keriven. Manifold learning using robust graph laplacian for interactive image search. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2008.

16. Dacheng Tao, Xiaoou Tang, Xuelong Li, and Xindong Wu. Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(7):1088–1099, July 2006.
17. K. Tieu and P. Viola. Boosting image retrieval. In *IEEE International Conference on Computer Vision and Pattern Recognition*, volume 1, pages 228–235, 2000.
18. Oncel Tuzel, Fatih Porikli, and Peter Meer. Pedestrian Detection via Classification on Riemannian Manifolds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(10):1713–1727, 2008.
19. Robert E. Williamson. Does relevance feedback improve document retrieval performance? *SIGIR Forum*, 13:151–170, May 1978.
20. Ying Wu, Qi Tian, and T.S. Huang. Integrating unlabeled images for image retrieval based on relevance feedback. In *International Conference on Pattern Recognition*, volume 1, pages 21–24, 2000.
21. Yufei Zha, Yuan Yang, and Duyan Bi. Graph-based transductive learning for robust visual tracking. *Pattern Recognition*, 43:187–196, January 2010.
22. Xiang Sean Zhou and Thomas S. Huang. Relevance feedback in image retrieval: A comprehensive review. *Multimedia Syst.*, 8(6):536–544, 2003.