

RESEARCH

Open Access

A semi-supervised learning algorithm for relevance feedback and collaborative image retrieval

Daniel Carlos Guimarães Pedronette^{1*}, Rodrigo T. Calumby^{2,3} and Ricardo da S. Torres²

Abstract

The interaction of users with search services has been recognized as an important mechanism for expressing and handling user information needs. One traditional approach for supporting such interactive search relies on exploiting relevance feedbacks (RF) in the searching process. For large-scale multimedia collections, however, the user efforts required in RF search sessions is considerable. In this paper, we address this issue by proposing a novel semi-supervised approach for implementing RF-based search services. In our approach, supervised learning is performed taking advantage of relevance labels provided by users. Later, an unsupervised learning step is performed with the objective of extracting useful information from the intrinsic dataset structure. Furthermore, our hybrid learning approach considers feedbacks of different users, in collaborative image retrieval (CIR) scenarios. In these scenarios, the relationships among the feedbacks provided by different users are exploited, further reducing the collective efforts. Conducted experiments involving shape, color, and texture datasets demonstrate the effectiveness of the proposed approach. Similar results are also observed in experiments considering multimodal image retrieval tasks.

Keywords: Content-based image retrieval; Semi-supervised learning; Relevance feedback; Collaborative image retrieval; Recommendation

1 Introduction

Image acquisition and sharing facilities have fostered the creation of huge image collections. This scenario has demanded the development of effective systems to support the search for relevant images, given users' information needs. One of the most promising approach for dealing with this challenge relies on the use of Content-Based Image Retrieval (CBIR) systems. The objective of CBIR systems is to provide relevant collection images by taking into account their similarity to user-defined query patterns (e.g., sketch, example image). In these systems, similarity computation is based on features that are associated with visual properties such as shape, texture, and color [1, 2]. The main challenge here consists in mapping low-level features to high-level concepts typically found within images, a problem named as *semantic gap*.

One suitable alternative to address this challenge consists in involving the user in the query processing loop, a procedure named as relevance feedback (RF). The objective is to refine the search systems based on relevance judgements provided by users. Along iterations, users assign labels to returned images (usually indicating if an image is relevant or not) [3], and the search system tunes itself in order to return more relevant images in the next iteration. The idea is to take advantage of the user perception in order to return more relevant images that better address her needs.

Typical RF approaches rely on the use of supervised mechanisms for learning from training sets composed of images labeled by users along iterations. At each iteration, the used machine learning method is retrained in order to define a novel ranked list containing potentially more relevant images. One important issue in the implementation of a RF approach concerns the number of images labeled at each iteration [4]. In fact, labeling a large number of images is a time-consuming and error-prone task. Therefore, RF approaches usually try

*Correspondence: daniel@rc.unesp.br

¹ Department of Statistics, Applied Mathematics and Computing - State University of São Paulo (UNESP), Rio Claro, SP, Brazil
Full list of author information is available at the end of the article

to minimize the number of iterations (and therefore the number of interactions) needed for providing relevant results.

Another issue concerns the imbalance problem. Usually, in a typical RF-based search scenario, only a few collection images are labeled. That opens a new area of investigation concerning the use of unsupervised approaches [5–7] that somehow can take advantage of the large number of unlabeled images available in a query session. Usually, these approaches benefit from contextual information defined in terms of information that can be extracted from the relationships among images (e.g., their distances and computed ranked lists).

In this paper, we propose a novel RF-approach based on semi-supervised learning mechanisms. It is semi-supervised in the sense that it learns from both labeled and unlabeled data [8]. Basically, the method combines labeled data available along RF iterations with contextual information provided by the large number of available unlabeled data. The use of these unlabeled data helps to minimize the user efforts, as potentially less labels need to be assigned to images along iterations.

In our semi-supervised method, we use the unsupervised Pairwise Recommendation [9] re-ranking algorithm, which has been demonstrated to yield effective results in image retrieval tasks. This algorithm exploits the relationships among images encoded in ranked lists. The relationships are modeled using unsupervised *recommendations* among images that are likely to be relevant. These unsupervised recommendations are later combined with supervised recommendations defined in terms of RF interactions.

This paper differs from our previous work [10] as it presents a novel collaborative image retrieval approach, based on the proposed semi-supervised learning algorithm. In collaborative image retrieval tasks, the semi-supervised learning algorithm benefits from feedbacks of different users. This scenario considers the feedback provided for different queries at the same time.

A large experimental evaluation was conducted considering several image descriptors and datasets, for both relevance feedback and collaborative image retrieval scenarios. Experiments were conducted on three image datasets, considering different visual descriptors (shape, color, and texture descriptors). The proposed approach was also evaluated on multimodal image retrieval tasks, which combine visual and textual descriptors. We also evaluated the proposed semi-supervised algorithm in comparison with a recently proposed genetic programming approach for relevance feedback. The experimental evaluation demonstrates that the proposed approach achieves significant effectiveness improvements in several image retrieval tasks by exploiting both supervised and unsupervised learning mechanisms.

The paper is organized as follows: Section 2 discusses related work, while Section 3 describes the problem formulation. In Section 4, we discuss the unsupervised Pairwise Recommendation [9] algorithm, while in Section 5 we present the proposed *Semi-supervised pairwise recommendation for relevance feedback* approach. Section 6 presents the *Semi-supervised pairwise recommendation for collaborative image retrieval*. Section 7 presents the experimental evaluation, and finally, Section 8 presents our conclusions and possible future work.

2 Related work

While huge volumes of imagery are generated daily, the task of finding the ones we want to see at a particular moment in time is becoming increasingly challenging. Additionally to the growing number of image collections, the *semantic gap* between low-level features and high-level semantic concepts often represents great obstacles for effective image retrieval. In this scenario, interactive retrieval methods have emerged as a promising solution, based on an interactive dialog between users and CBIR systems [3].

Despite its relatively short history, relevance feedback methods evolved consistently and it remains an active research topic [3, 11–13]. Initially, relevance feedback was developed along the path from heuristic-based techniques to optimal learning algorithms, with early works inspired by term-weighting and relevance feedback techniques for document retrieval. The main intuition behind heuristic-based methods, for instance, is to focus on the feature that can best cluster the positive examples and also separate the positive from the negative ones [11].

Different approaches and several machine learning techniques were used for relevance feedback in image retrieval tasks. In [14], relevance feedback was modeled as a Bayesian classification problem. The system analyzes the consistence among iterations: if the current feedback is consistent with the previous ones, higher probabilities are assigned to the images that are similar to the query target. Thus, the images similar to the user's interests are emphasized step by step [14].

A fuzzy approach [15] was also used for modeling relevance feedback tasks. A fuzzy set is defined, so that the degree of membership of each image to this fuzzy set is related to the user's interest in that image [15]. The user's feedback, both positive and negative, are then used to determine the degree of membership of each image to set being analyzed.

Various approaches focus on combining different features, using supervised approaches. In [16], a relevance feedback framework was proposed using genetic programming to find a function that combines non-linearly similarity values computed by different descriptors. The similarity functions defined for each available descriptor

are then used to compute the overall similarity between two images, and defining the retrieved results.

Another learning technique commonly used is Support Vector Machine (SVM) [17]. Basically, the problem is modeled as a binary classification problem, in which the goal of the SVM-based methods is to find a hyperplane that separates the relevant from the non-relevant images. The labeled images are usually the most ambiguous ones, using a principle called *active learning* [3]. For instance, those images are often selected by their proximity to the separation hyperplane.

Despite the success of supervised approaches, using unlabeled data to improve the retrieval results and consequently boost supervised learning has become a hot topic in machine learning [18]. An analysis on the value of unlabeled data is presented in [19]. The Manifold Ranking algorithm [20] was proposed aiming at ranking the objects with respect to the intrinsic data distribution. The unsupervised approach is different from distance-based ranking methods because it exploits the data distribution of all the samples for ranking rather than only considering the pairwise distances. This paradigm has been actively exploited in image retrieval systems in the past few years [6, 9, 21, 22].

Following this trend, the joint use of both labeled and unlabeled data led to development of various semi-supervised methods for relevance feedback. In [18], a semi-supervised approach attempts to enhance the performance of relevance feedback by exploiting unlabeled images, integrating semi-supervised learning and active learning. In each relevance feedback session, two simple learners are trained from the labeled data. Each learner then labels some unlabeled images in the database for the other learner. After re-training with the additional labeled data, the learners classify the images in the database again and then their classifications are merged.

The hypergraph-based transductive learning algorithm was also used [23] to learn beneficial information from both labeled and unlabeled data for image ranking. Images are taken as vertices in a weighted hypergraph and the task of image search is formulated as the problem of hypergraph ranking. The approach uses the similarity matrix computed from various feature descriptors and a probabilistic hypergraph, which assigns each vertex to a hyperedge in a probabilistic way.

SVM-based approaches were also adapted to semi-supervised learning in various ways. The performance of SVM is usually limited by the number of training data. Methods have been proposed for learning based on a kernel function from a mixture of labeled and unlabeled data [24], alleviating the problem of small-sized training data. The kernel uses a batch mode active learning method to identify the most informative and diverse examples via a min-max framework. In another SVM approach, the

information of unlabeled samples is integrated by introducing a Laplacian regularizer. The problem is formulated into a general subspace learning task, using an automatic approach for determining the dimensionality of the embedded subspace for relevance feedback [25].

Metric learning and rank-based approaches also have been exploited in relevance feedback tasks [3]. Based on a semi-supervised metric learning, a step-wise algorithm for boosting the retrieval performance of CBIR systems by incorporating relevance feedback information is proposed in [26]. In [27], a semi-supervised algorithm called ranking with Local Regression and Global Alignment (LRGA) is proposed. A Laplacian matrix for data ranking is employed, using a local linear regression model to predict the ranking scores of its neighboring points. A unified objective function is used to globally align the local models from all the data points, assigning a ranking score to each data point.

In order to further reduce the user efforts in the relevance feedback sessions, some approaches have been proposed for exploiting the feedback of various users in conjunction. In recent years, there is an emerging interest to analyze and exploit the historic data from different user interactions for improving the effectiveness of retrieval results considering multi-user collaborative environments [28]. This paradigm, commonly referred to as *Collaborative Image Retrieval* (CIR), has attracted a lot of attention [29–31]. In [30], a semi-supervised distance metric learning technique integrates both log data and unlabeled data information, using a graph approach. An approach for collaborative image retrieval using multi-class relevance feedback and Particle Swarm Optimization classifier is proposed in [31].

In this paper, we propose a novel RF-approach that combines various recent trends on interactive image retrieval systems, such as semi-supervised learning, ranked-based methods, and collaborative image retrieval. The proposed method is semi-supervised since it not only uses the supervised relevance feedback information but also exploits the unlabeled data. The method is inspired by the recent Pairwise Recommendation [9] algorithm, which considers the intrinsic dataset structure through a recommendation simulation model. Additionally, the approach presents other advantages, as the low computational cost and the use of an unified recommendation model for representing positive and negative feedback and collaborative image retrieval.

3 Problem formulation

This section discusses the problem formulation, which is divided into four main topics: (i) *image retrieval*: considering the retrieval process based only on image descriptors; (ii) *unsupervised learning*: performing a post-processing step after the initial retrieval; (iii) *semi-supervised*

relevance feedback: combining the post-processing step with information collected from user feedback interactions and; (iii) *collaborative image retrieval*: considering the feedback of various users.

3.1 Image retrieval

Let $\mathcal{C} = \{img_1, img_2, \dots, img_n\}$ be an *image collection* and \mathcal{D} be an *image descriptor* that defines a distance function between two images img_i and img_j as $\rho(img_i, img_j)$, or simply $\rho(i, j)$.

The distance $\rho(i, j)$ among all images $img_i, img_j \in \mathcal{C}$ can be computed to obtain an $N \times N$ distance matrix A , such that $A_{ij} = \rho(i, j)$.

The distance function ρ can be used to compute a ranked list τ_q given a query image img_q .

The ranked list $\tau_q = (img_1, img_2, \dots, img_{n_s})$ is defined as a permutation of the subset $\mathcal{C}_s \subset \mathcal{C}$, which contains the most similar images to the query image img_q , with $|\mathcal{C}_s| = n_s$. For a permutation τ_q , we interpret $\tau_q(i)$ as the position (or rank) of image img_i in the ranked list τ_q . In this sense, if img_i is ranked before img_j in the ranked list of img_q , that is, $\tau_q(i) < \tau_q(j)$, then $\rho(q, i) \leq \rho(q, j)$.

The same approach can be used for all images in the collection, i.e., we take each image $img_i \in \mathcal{C}$ as a query image img_q , in order to obtain a set $\mathcal{R} = \{\tau_1, \tau_2, \dots, \tau_n\}$ of ranked lists for each image of the collection \mathcal{C} .

3.2 Unsupervised learning

Our unsupervised learning approach relies on the use of an iterative function f_u that does not depend on any labeled training set. This iterative function takes an initial matrix $A^{(t)}$ and a set of ranked lists $\mathcal{R}^{(t)}$ (where t denotes the current iteration) as input and computes a novel distance matrix $A^{(t+1)}$ that is expected to be more effective. The new distance matrix $A^{(t+1)}$ is then used to compute a novel set of ranked lists, $\mathcal{R}^{(t+1)}$. More formally,

$$A^{(t+1)} = f_u(A^{(t)}, \mathcal{R}^{(t)}). \quad (1)$$

Several techniques like clustering [32], diffusion process [22, 33], graph models [34], and image re-ranking [9, 35] have been employed in unsupervised learning in image retrieval tasks. Most of these approaches share the objective of using contextual information encoded in the relationships among images. In this work, we use the Pairwise Recommendation [9] algorithm, discussed in Section 4, as the basis of our proposed semi-supervised approaches. This algorithm is used as the method that implements function f_u .

Alternatively, rank aggregation techniques have also been used in unsupervised approaches. The objective of these methods is to combine distance measures and ranked lists computed by different modalities/descriptors. In this case, the input of function f_u is a set of distance matrices $\{A_1, A_2, \dots, A_m\}$, where A_i is defined by the i th descriptor.

3.3 Semi-supervised relevance feedback

Given a query image img_q , the search process with relevance feedback is comprised of four main steps [16]: (i) presentation of a small number of retrieved images to the user; (ii) user indication of relevant and non-relevant images; (iii) learning the user needs by taking into account provided feedback; and (iv) the definition of a novel set of images to be presented in the next iteration.

Let $I_S^{(t)}$ be a set of images displayed to the user at each iteration t and L be the number of images displayed, such that $|I_S^{(t)}| = L$. Set $I_S^{(t)}$ contains both images labeled as relevant and non-relevant, i.e., $I_S^{(t)} = I_R^{(t)} \cup I_{NR}^{(t)}$, where $I_R^{(t)} = \{img_{R_1}, img_{R_1}, \dots, img_{R_r}\}$ is the set that contains the images labeled as *relevant* in a relevance feedback session, and $I_{NR}^{(t)} = \{img_{NR_1}, img_{NR_1}, \dots, img_{NR_{nr}}\}$ is the set that contains the images labeled as *non-relevant*.

The proposed semi-supervised relevance feedback approach is also defined in terms of an iterative function f_{rf} as following:

$$A^{(t+1)} = f_{rf}(A^{(t)}, \mathcal{R}^{(t)}, I_R^{(t)}, I_{NR}^{(t)}). \quad (2)$$

In our semi-supervised method, we use all information available on both unlabeled and labeled data. Therefore, function f_{rf} has as input the information used by the unsupervised methods (relationships among images encoded in a distance matrix and the ranked lists), as well as labeled data provided along the relevance feedback sessions.

3.4 Collaborative image retrieval

The recent collaborative image retrieval (CIR) paradigm is generally defined in the literature [30] as the use of the historical log data of user relevance feedback collected from CBIR systems over a long period of time. These approaches aim at avoiding the interaction overhead between systems and users. Generally, the main focus has been on exploiting information learned from previous user interactions as new query images are being processed [36].

In this work, we extend this definition for considering situations in which different users perform simultaneous queries and provides relevance feedback, during a single iteration. In fact, this extended definition aims at approximating a common real-world scenario, specially for web applications, where different users submit queries and interact with the retrieval system at the same time.

In our proposed approach, the information collected from interactions performed by different users is globally exploited, i.e., users can benefit from each other's interactions. The objective is to exploit the relationship among query images submitted by different users and propagate the collected feedback information in the proposed semi-supervised learning scheme. In summary, this strategy affects the processing of multiple queries, improving the effectiveness of provided results along iterations and reducing the users' efforts.

Formally, let $I = (img_q, I_R^{(t)}, I_{NR}^{(t)})$ be a tuple that defines a user interaction, where img_q is a query image and $I_R^{(t)}$ and $I_{NR}^{(t)}$ are the sets of relevant and non-relevant images, respectively. Let $\mathcal{S}^{(t)} = \{I_1, I_2, \dots, I_u\}$ be a set of users' queries and interactions provided for an iteration t , where u denotes the number of simultaneous queries at a given iteration t . The proposed semi-supervised learning for collaborative image retrieval is defined in terms of the function f_{cir} , which considers the set of interactions $\mathcal{S}^{(t)}$ in addition to the information analyzed by the unsupervised method. The function f_{cir} is defined in Eq. 3:

$$A^{(t+1)} = f_{cir}(A^{(t)}, \mathcal{R}^{(t)}, \mathcal{S}^{(t)}). \quad (3)$$

The relevance feedback considers a single query interaction, while the collaborative retrieval scenario exploits multiple interactions available in different queries submitted to the search system. The objective is to exploit the information inter-queries in such a way that the total collective effort is reduced.

4 Unsupervised pairwise recommendation

The *Pairwise Recommendation* [9] algorithm consists in an unsupervised image re-ranking method proposed for image retrieval tasks. The algorithm takes into account the relationships among images and information encoded in ranked lists with the objective of improving the effectiveness of CBIR systems. The algorithm is inspired by the concept of *recommendation*, originally proposed for reducing the information overload by selecting automatically items that match personal preferences.

In the case of the Pairwise Recommendation algorithm, images placed at top positions of ranked lists recommend *other* images to each other. The recommended images are expected to be similar to each other. In the algorithm, a recommendation indicates that the distance between two images should be reduced. Furthermore, weights are used to define how much distances should be decreased. These weights are defined based on the position of images in the ranked lists, and on the quality of ranked lists, which is defined in terms of a cohesion measure.

Once recommendations are performed, novel ranked lists $\mathcal{R}^{(t+1)}$ are computed based on the new distance matrix $A^{(t+1)}$. These steps are repeated over iterations until convergence.¹

4.1 Cohesion measure

A cohesion measure [9] is used for determining the quality of ranked lists. Images with better ranked lists, and therefore higher cohesion, have more authority for making recommendations. This measure assess the degree of agreement of ranked lists. If a ranked list is effective, i.e., has several similar images, then images at the top positions should refer to each other at the top positions of

their own ranked lists. Therefore, a perfect cohesion indicates that all considered images refer to each other at the first positions of their ranked lists [9].

4.2 Unsupervised recommendations

The unsupervised recommendation step relies on the analysis of the top- k positions of ranked lists. Let τ_i be the ranked list computed for image img_i and let img_x and img_y be images that are on the top- k positions of the ranked list τ_i . In this scenario, the image img_i recommends the img_y to img_x and vice versa. The recommendation decreases the distance between the images img_x and img_y , according to a weight w_r .

Algorithm 1 presents the method for performing recommendations considering a given ranked list τ_i . The weight w_r is computed in line 7 as a product of three factors: the cohesion c_i , and the partial weights w_x and w_y . While the cohesion c_i provides an estimation of the effectiveness of the ranked list τ_i , the weights w_x and w_y are computed based on the position of images in the ranked lists. For images at the first positions of the ranked list, a higher weight is assigned. The three variables are computed in the interval $[0,1]$.

Algorithm 1 Unsupervised recommendations based on ranked lists [9].

Require: Matrix A , Ranked list τ_i , Cohesion c_i , and Parameter L_c

Ensure: Updated matrix A

```

1:  $\tau_{ki} \leftarrow kNN(\tau_i)$ 
2: for all  $img_x \in \tau_{ki}$  do
3:    $w_x \leftarrow 1 - (\tau_{ki}(x)/k)$ 
4:   for all  $img_y \in \tau_{ki}$  do
5:      $w_y \leftarrow 1 - (\tau_{ki}(y)/k)$ 
6:      $w_r \leftarrow c_i \times w_x \times w_y$ 
7:      $\lambda \leftarrow 1 - \min(1, L_c \times w_r)$ 
8:      $A_{xy} \leftarrow \min(\lambda A_{xy}, A_{yx})$ 
9:   end for
10: end for
```

In Algorithm 1, line 8, a coefficient λ is computed based on the weight w_r of the recommendation and a constant L_c . The constant L_c determines the convergence speed. The higher the value of L_c , the faster the distances among images will decrease. Note that the value of λ is multiplied by the current distance matrix A_{xy} in order to updated it.

4.3 Rank aggregation

We also use the *Pairwise Recommendation* algorithm [9] in rank aggregation tasks. Our objective is to combine various descriptors so that retrieval results can be improved. We use a multiplicative approach to combine distance matrices defined by different descriptors. Later, the final matrix created is used as input of the Pairwise Recommendation algorithm. As this matrix combined contextual information provided by multiple descriptors,

the Pairwise Recommendation algorithm is expected to yield higher effectiveness gains.

5 Semi-supervised pairwise recommendation for relevance feedback

This section presents a novel *semi-supervised pairwise recommendation method for relevance feedback* scenarios. Since both the unsupervised learning as the relevance feedback procedures are intrinsically iterative, we propose an algorithm based on semi-supervised iterations. The objective of the proposed approach is to combine the unsupervised recommendations based on ranked lists with supervised recommendations based on user interactions obtained from relevance feedback sessions.

Algorithm 2 outlines the main steps of the proposed semi-supervised algorithm. Before each relevance feedback session, an unsupervised step is performed. The unsupervised recommendations (line 2) update the ranked lists aiming at improving the results showed to the user. In the following, the top-ranked images that were not labeled yet are displayed and the user indicates the relevant and non-relevant images. These two steps constitute a relevance feedback session, defined respectively in lines 3–4 of the algorithm. The last step (line 5) defines a set of supervised recommendations, which are performed based on the images labeled by the user. Finally, re-ranking based on these recommendations is performed.

Algorithm 2 Semi-supervised Pairwise Recommendation algorithm.

Require: Distance matrix A and set of ranked lists \mathcal{R}
Ensure: Processed distance matrix $A^{(t)}$ and set of ranked lists $\mathcal{R}^{(t)}$

```

1: for each iteration  $t$  do
2:   Perform Unsupervised Recommendations
3:   Display Images to the Users
4:   Collect Relevance Feedback
5:   Perform Supervised Recommendations
6: end for
```

Figure 1 illustrates the general workflow of the proposed semi-supervised algorithm. In other words, the information collected by each relevance feedback session is modeled as a set of recommendations that are later used by the unsupervised algorithm. The recommendations obtained from user interactions are also modeled as a distance updating among images, as discussed in the next section.

5.1 Semi-supervised recommendations

The supervised recommendations are defined in the same way as unsupervised recommendations: updating distances among images. In this scenario, the

supervised recommendations are defined in terms of variations of unsupervised recommendations, given by Algorithm 1. However, while the unsupervised recommendation exploits information from the ranked lists, the supervised recommendations uses the user feedback as input. Given a set of similar images $I_R^{(t)}$, labeled as “relevant” by the user, the supervised recommendations consider two different approaches when compared to the unsupervised recommendations:

- **Set of Relevant Images:** the unsupervised algorithm considers the ranked lists τ_{ki} as the source of the recommendations. For the supervised recommendations, the set of relevant images $I_R^{(t)}$ labeled by the user at iteration t is used instead of τ_{ki} .
- **Recommendation Weight:** on the unsupervised setting, the recommendation weight w_r is defined by three factors: w_x , w_y , and c_i . The terms w_x and w_y are computed based on the position of the images img_x and img_y involved in the recommendation (respectively, $\tau_{ki}(x)$ and $\tau_{ki}(y)$). The term c_i is computed according to the cohesion of the ranked list τ_{ki} . These factors aims at approximating the confidence of the unsupervised recommendation. For the supervised recommendations, on the other hand, the confidence is maximum, since they are obtained from the user feedback. Therefore, for representing the maximum confidence, we consider both the positions $\tau_{ki}(x) = \tau_{ki}(y) = 1$ and the cohesion $c_i = 1$. These values lead to a single recommendation weight w_r for all relevant images involved in the relevance feedback session.

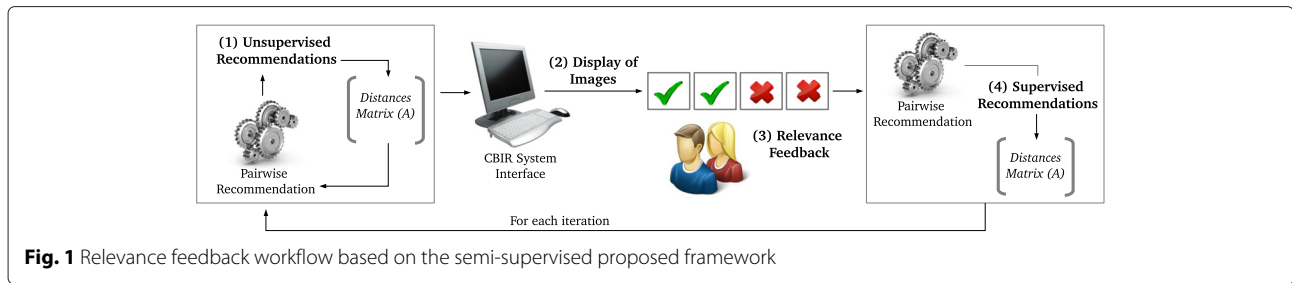
Based on these differences, Algorithm 3 defines the proposed supervised recommendation approach. In fact, all

Algorithm 3 Supervised recommendations based on relevance feedback.

Require: Set of Relevant Images $I_R^{(t)}$, Set of Non-Relevant Images $I_{NR}^{(t)}$, Parameter L_c
Ensure: Updated matrix A

```

1:  $w_x \leftarrow w_y \leftarrow 1 - (1/k)$ 
2:  $w_r \leftarrow w_x \times w_y$ 
3: { Positive Recommendation - Relevant Images}
4: for all  $img_x \in I_R^{(t)}$  do
5:   for all  $img_y \in I_R^{(t)}$  do
6:      $\lambda \leftarrow 1 - \min(1, L_c \times w_r)$ 
7:      $A_{xy} \leftarrow \min(\lambda A_{xy}, A_{yx})$ 
8:   end for
9: end for
10: { Negative Recommendation - Non-Relevant Images}
11: for all  $img_x \in I_R^{(t)}$  do
12:   for all  $img_y \in I_{NR}^{(t)}$  do
13:      $\lambda \leftarrow 1 + \max(1, L_c \times w_r)$ 
14:      $A_{xy} \leftarrow \max(\lambda A_{xy}, A_{yx})$ 
15:   end for
16: end for
```



pairwise distances among labeled images (relevant and non-relevant) are updated according to a single factor λ .

Notice that the supervised approach also exploits information provided by the set of non-relevant images, defining a set of *negative supervised recommendations* (line 11 of the Algorithm 3). The motivation consists in increasing the distances among similar images (defined by the set $I_R^{(t)}$) and non-similar images (defined by the set $I_{NR}^{(t)}$).

The negative recommendations use the same recommendation weight, which indicates maximum confidence ($\tau_{ki}(x)$, $\tau_{ki}(y)$, and $c_i = 1$). For negative recommendations, the differences regarding unsupervised recommendations are as follows:

- *Set of Images*: the negative recommendations replace the set of images involved in the recommendation. While the ranked lists τ_{ki} is used, the negative recommendations use $I_R^{(t)}$ and $I_{NR}^{(t)}$, aiming at increasing the distance among similar and non-similar images.
- *Negative Recommendation*: for increasing (instead of reducing) the distance among images, the λ factor is defined greater than 1 (as $\lambda = 1 + \min(1, L_c \times w_r)$) and the min operation is replaced by a max operation.

6 Semi-supervised pairwise recommendation for collaborative image retrieval

The collaborative image retrieval scenario aims at modeling a real-world situation in which various users submit simultaneous queries to a retrieval system and provide their relevance feedback. At each iteration, the relevance feedback information provided for a set of queries is collected and used by the semi-supervised approach.

For strict RF scenarios, the feedback information is processed in isolation for each user and affects only the results for that specific user. On the other hand, for collaborative image retrieval scenarios, the user feedback may be used for improving the effectiveness of several queries submitted by other users. In fact, the main principle of the Pairwise Recommendation [9] algorithm is based on exploiting the relationships among all images in a given dataset. Therefore, instead of processing the recommendations (both supervised and unsupervised) in isolation

for each user, they are processed considering the same distance matrix. As a result, the effectiveness improvements obtained by an issued query are propagated to various related queries, through the recommendations. Therefore, the efforts needed for obtaining effectiveness gains in the different query sessions are drastically reduced in comparison with typical RF scenarios, as discussed in Section 7.1.

The semi-supervised approach is modeled in the same way for the strict RF scenarios. Both supervised and unsupervised recommendations are used as defined in the previous section. However, the interaction workflow is different for collaborative image retrieval scenarios:

1. *Unsupervised Recommendations*: the unsupervised Pairwise Recommendation [9] algorithm is performed, updating the ranked lists for being showed to the users;
2. *Display of Images for Various Users*: for each simultaneous query, the top-ranked images that were not labeled yet by any users are displayed;
3. *Set of Relevance Feedback Interactions*: each user informs the relevant and non-relevant images from the set of images displayed for her corresponding query;
4. *Supervised Recommendations*: based on labels provided by various users, a set of recommendations are performed. The ranked lists are re-ranked based on this recommendations;

Figure 2 illustrates the general workflow of the proposed semi-supervised algorithm for collaborative image retrieval scenarios.

7 Experimental evaluation

This section discusses a set of experiments conducted for assessing the effectiveness of our method. We analyzed and compared the proposed method under several aspects, considering different datasets and descriptors. Section 7.1 discusses the experimental setup. Sections 7.2, 7.3, and 7.4 present the experimental results considering various shape, color, and texture descriptors, respectively. Section 7.5 in turn presents the results for

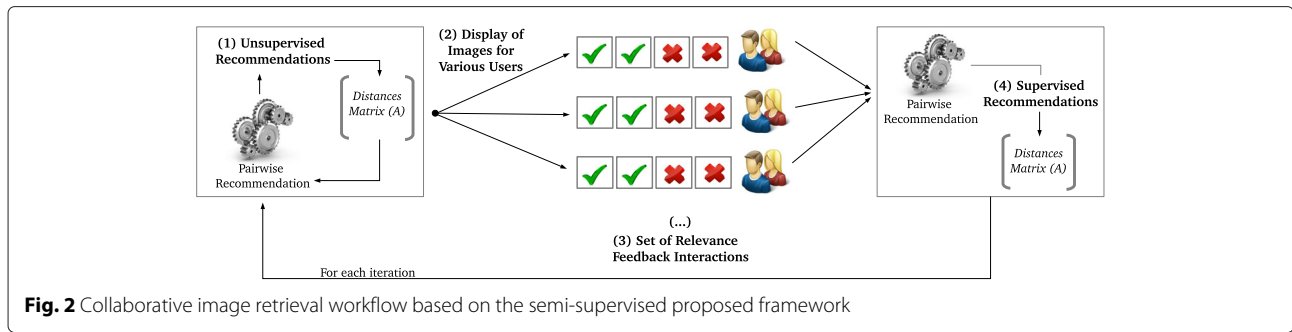


Fig. 2 Collaborative image retrieval workflow based on the semi-supervised proposed framework

multimodal image retrieval tasks. The main objective of these sections consists in assessing the improvements obtained by the proposed method along the relevance feedback sessions, evaluating the increase of effectiveness results. The goal of using various datasets and descriptors is to demonstrate that the proposed method can achieve significant gains regardless the considered description scenario. Experiments aiming at comparing the obtained results with related methods are presented in Section 7.6. Finally, Section 7.7 analyzes the impact of the number of users on the effectiveness of retrieval results.

7.1 Experimental setup

The conducted experiments aimed at evaluating the effectiveness of the proposed method along the iterations, considering relevance feedback and collaborative image retrieval scenarios. Experiments considered that 20 images are shown to the user at each iteration, along 10 iterations. In the experiments, the presence of users is simulated, such that all images belonging to the same class of the query image are considered relevant. For all experiments, we consider all collection images as queries and report the average effectiveness results for the whole dataset. The parameters settings used for the semi-supervised pairwise recommendation evaluation are the same used in [9].

In most of the performed experiments, we use two measures to evaluate the effectiveness of the proposed method: (i) precision vs. recall curves ($P \times R$) before the first iteration ($t = 0$) and after the last iteration ($t = 10$) and; (ii) the precision of top 20 images retrieved ($P@20$) vs. the number of iteration ($P \times t$ curve). The $P \times R$ analysis aims at evaluating the final improving effectiveness performance provided by the proposed method after a given number of iterations, while, the $P \times t$ curves are used to analyze how the evolution of the precision at top positions of ranked lists along iterations.

For collaborative image retrieval scenarios, we consider a random set of queries. The number of simultaneous queries per iteration q_i considered for each experiment is proportional to the dataset. We considered the value

of q_i as only 5 % of the size of the dataset. As discussed in the following sections, only this small subset is enough for obtaining similar results to relevance feedback considering all images as queries. In our experiments, we consider that the feedback of different users are processed in isolation at each iteration, and the improvements inter-queries can be observed at the next iteration.

7.2 Shape-based experiments

For the shape retrieval experiments, we consider the MPEG-7 collection [37], a well-known shape database composed of 1400 shapes divided into 70 classes. We use three shape descriptors: Segment Saliences (SS) [38], Beam Angle Statistics (BAS) [39], and Inner Distance Shape Context (IDSC) [40].

7.2.1 Relevance feedback results

Figure 3 presents the evolution of $P@20$ measure along 10 iterations for the three descriptors evaluated. We can observe very significant precision gains: the precision of the SS [38] descriptor, for example, goes from 36 % before the first iteration to 76 % after the last iteration. Figure 4 illustrates the precision vs. recall curves for the three

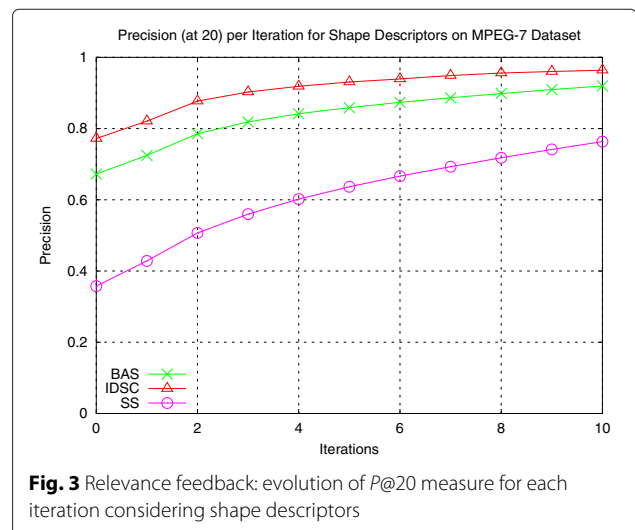
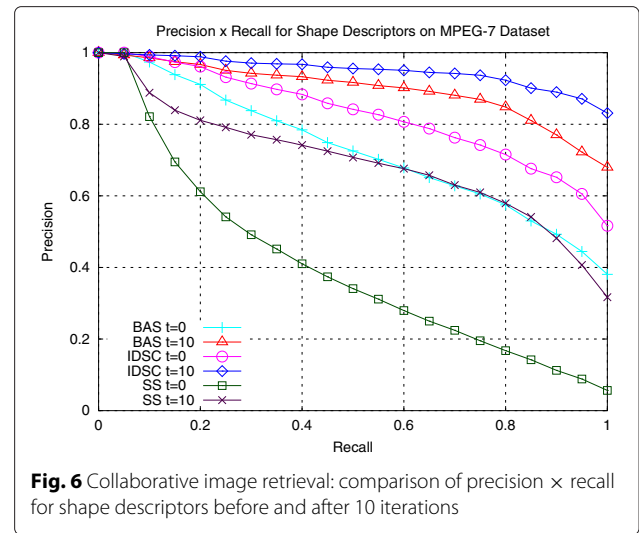
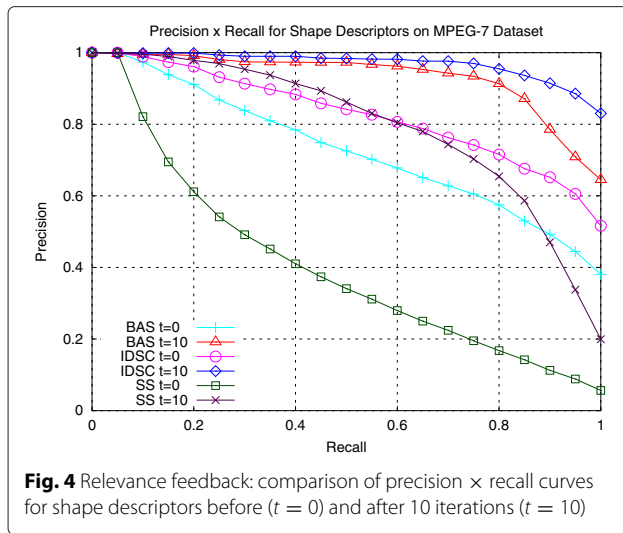


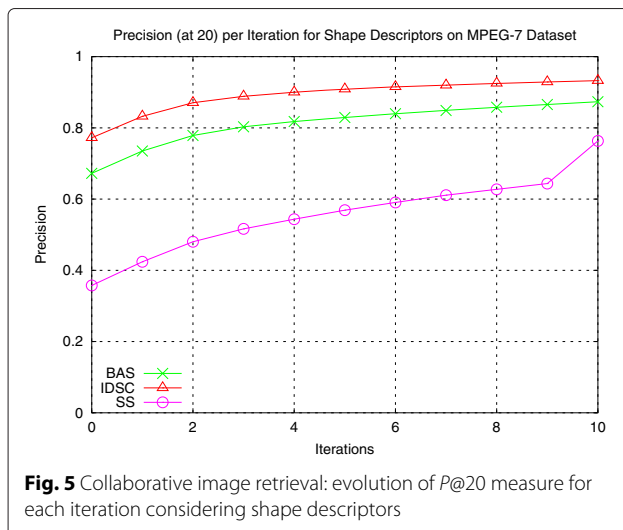
Fig. 3 Relevance feedback: evolution of $P@20$ measure for each iteration considering shape descriptors



descriptors considering the initial ($t = 0$) and final iterations ($t = 10$). As we can observe, the final curves are substantially superior for all descriptors.

7.2.2 Collaborative image retrieval results

The retrieval performance of the proposed framework on collaborative image retrieval based on shape descriptors is showed in Figs. 5 and 6. The evolution of $P@20$ measure along 10 iterations is illustrated in Fig. 5. We can observe very significant precision gains, even considering a small number of queries per iteration ($q_i = 70$). The precision of the SS [38] descriptor, for example, ranges from 36 % before the first iteration to 66 % after the last iteration. The precision vs. recall curves for the three descriptors considered is illustrated in Fig. 6. The final curves ($t = 10$) are substantially superior in comparison with initial curves ($t = 0$) for all descriptors.



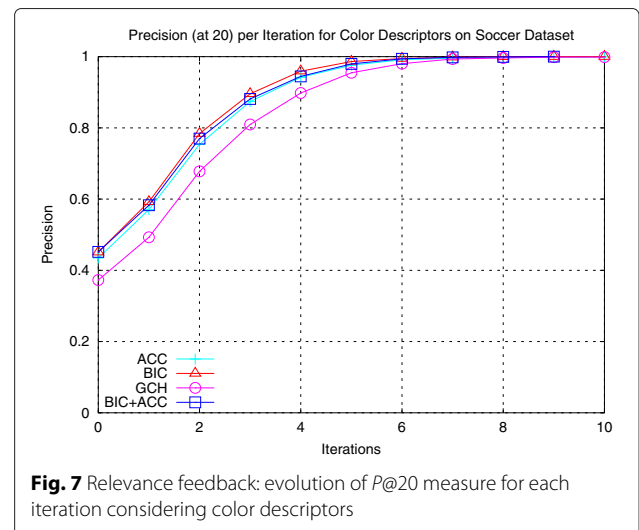
7.3 Color-based experiments

We evaluate our method for three color descriptors: Border/Interior Pixel Classification (BIC) [41], Auto Color Correlograms (ACC) [42], and Global Color Histogram (GCH) [43].

The dataset [44] used for color-based experiments is composed of 280 images illustrating soccer games collected in the Internet. The images are from 7 soccer teams containing 40 images per class. The size of images range from 198×148 to 537×672 pixels.

7.3.1 Relevance feedback results

Figure 7 illustrates how the $P@20$ measure evolves along the 10 iterations for the three color descriptors ($P \times t$ curve). We also consider the combination of ACC [42] + BIC [41] descriptors using rank aggregation. We can observe that the curve presents a remarkable ascendant



slope, indicating an increase of precision even greater than that obtained for shape descriptors. Figure 8 illustrates the precision vs. recall curves before and after the use of the proposed method (after 10 iterations). Again, the final curves are substantially superior for all descriptors.

7.3.2 Collaborative image retrieval results

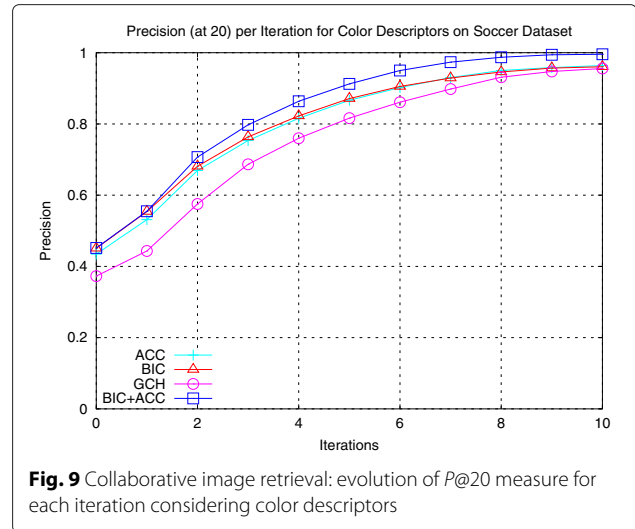
Figures 9 and 10 present the results of the proposed approach for collaborative image retrieval tasks considering color descriptors. Figure 9 shows the $P \times t$ curves considering 10 iterations for the three color descriptors. The combination of ACC [42] + BIC [41] descriptors using rank aggregation is also considered. Again, despite the small number of the queries per iteration ($q_i = 14$), we can observe very significant precision gains. Figure 10 illustrates the $P \times R$ curves before and after the use of the collaborative approaches. Notice that the final curves are superior for all descriptors and slightly superior to the traditional RF method, in this case.

7.4 Texture-based experiments

The texture experiments consider three well-known texture descriptors: Local Binary Patterns (LBP) [45], Color Co-Occurrence Matrix (CCOM) [46], and Local Activity Spectrum (LAS) [47]. We used the Brodatz dataset [48], which is composed of 111 different textures. Each texture is divided into 16 blocks, such that 1776 images are considered.

7.4.1 Relevance feedback results

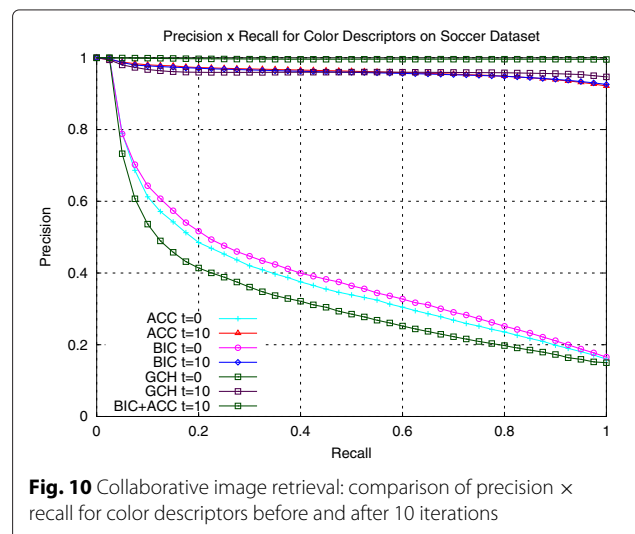
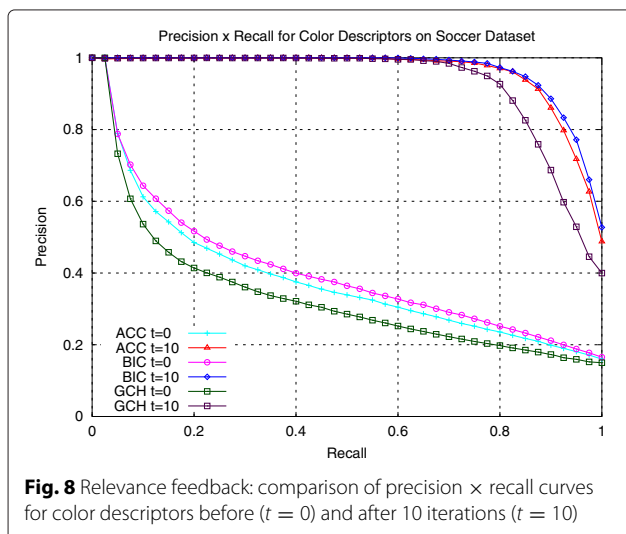
Figure 11 presents the $P \times t$ curves for the three descriptors and for the combination of LAS [47] + CCOM [46] descriptors. The results are consistent with shape and color descriptors, indicating the robustness of the proposed method for different visual properties. Figure 12

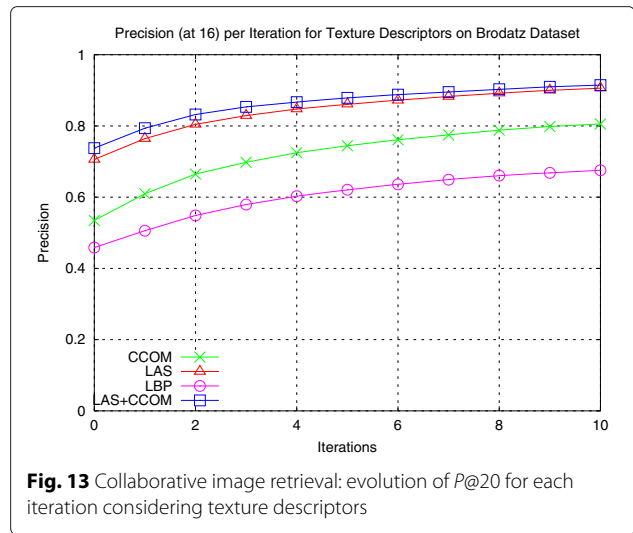
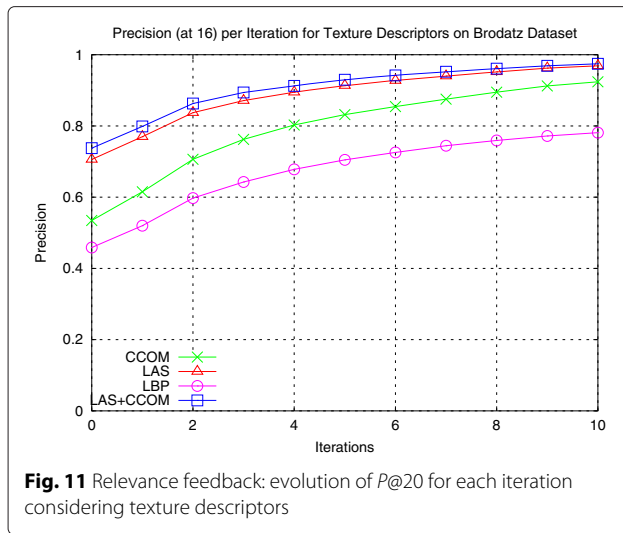


illustrates the precision vs. recall curves considering the initial ($t = 0$) and final iterations ($t = 10$). Again, a remarkable improvement in terms of effectiveness is observed for all descriptors.

7.4.2 Collaborative image retrieval results

The experiments for evaluating the collaborative approach on texture descriptors used the number of queries per iteration as $q_i = 88$. Figure 13 illustrates the $P@20$ measure evolution along 10 iterations ($P \times t$ curve). Figure 14 illustrates the precision vs. recall curves before and after the use of the proposed method (after 10 iterations). Considering both the $P@20$ measure and the $P \times R$ curves, we can observe results similar to the relevance feedback results, despite the small number of queries and user interactions required.





7.5 Multimodal retrieval

We also evaluated the proposed method considering a multimodal retrieval scenario, considering visual and textual descriptors. We used the UW dataset [49] created at the University of Washington. The dataset consists of a roughly categorized collection of 1109 images. The images include vacation pictures from various locations. These images are partly annotated using keywords. On the average, for each image the annotation contains 6 words. The maximum number of words per image is 22 and the minimum is 1. There are 18 categories, ranging from 22 images to 255 images per category.

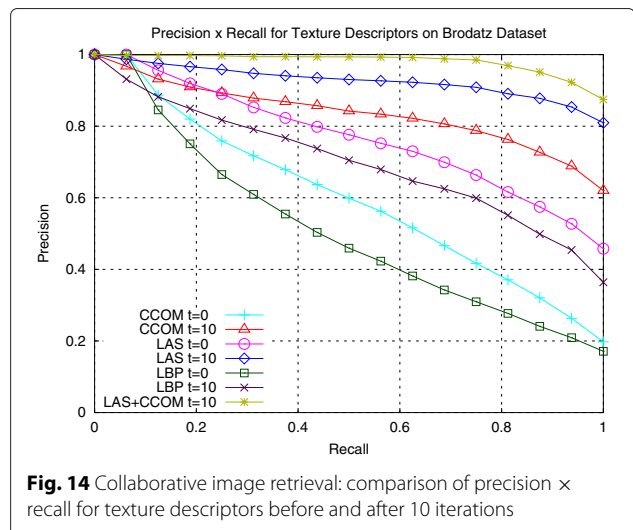
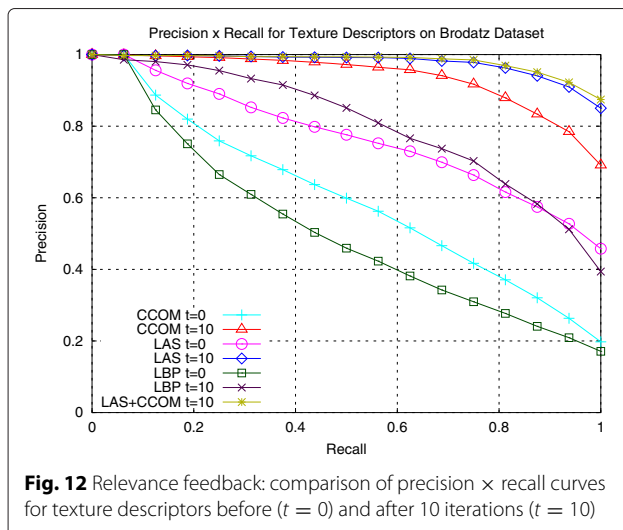
The experiments considered eleven descriptors:

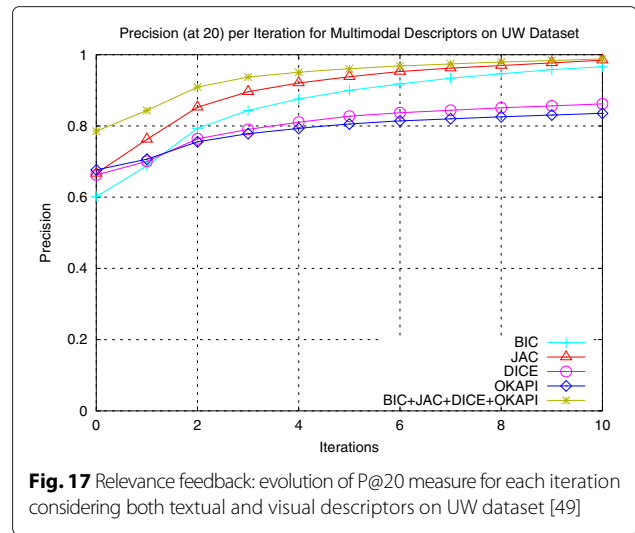
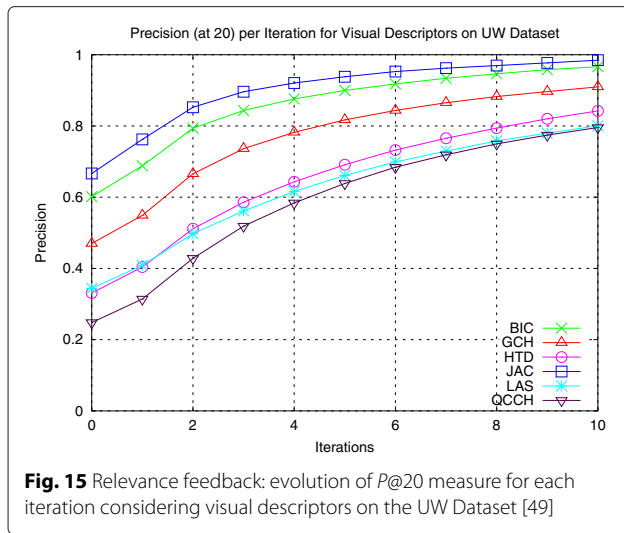
- **Visual Color Descriptors:** Border/Interior Pixel Classification (BIC) [41]; Global Color Histogram (GCH) [43] (both already used in Section 7.3); and the Joint Correlogram (JAC) [50].

- **Visual Texture Descriptors:** Homogeneous Texture Descriptor (HTD) [51]; Quantized Compound Change Histogram (QCCH) [52]; and Local Activity Spectrum (LAS) [47] (the last also considered in Section 7.4).
- **Textual Descriptors:** five well-known textual similarity measures [53] were considered for textual retrieval: the Cosine similarity measure (COS), Term Frequency - Inverse Term Frequency (TF-IDF), and the Dice coefficient (DICE), Jackard coefficient (JACKARD), and Okapi BM25 (OKAPI).

7.5.1 Relevance feedback results

We evaluated the $P@20$ measure along the 10 iterations for visual, textual, and multimodal retrieval. Figure 15 illustrates the $P \times t$ curves for the 6 visual descriptors considered, while Fig. 16 illustrates the $P \times t$ curves for





the textual descriptors. We can observe an increasing precision score for descriptors of both modalities along iterations. We can also observe that the precision gains are still more significant for visual descriptors.

We also evaluated the use of the proposed method for multimodal retrieval, considering two visual descriptors and two textual descriptors. Figure 17 presents the evolution of $P@20$ along iterations for visual (BIC and JAC), textual (DICE and OKAPI), and combined (BIC+JAC+DICE+OKAPI) descriptors. We can observe that the multimodal combination achieves a very high precision score.

7.5.2 Collaborative image retrieval results

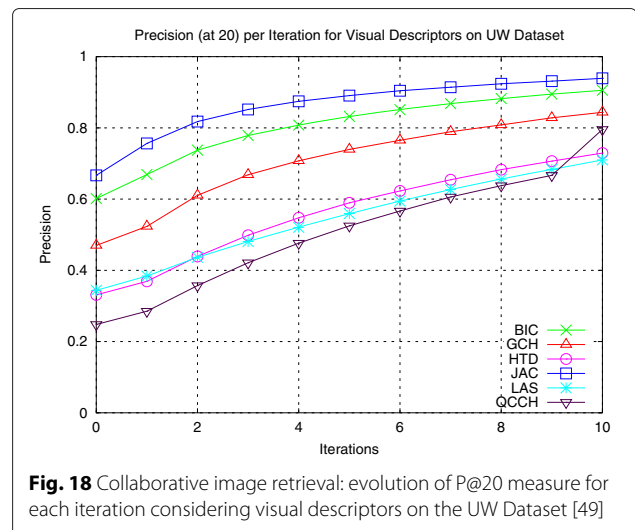
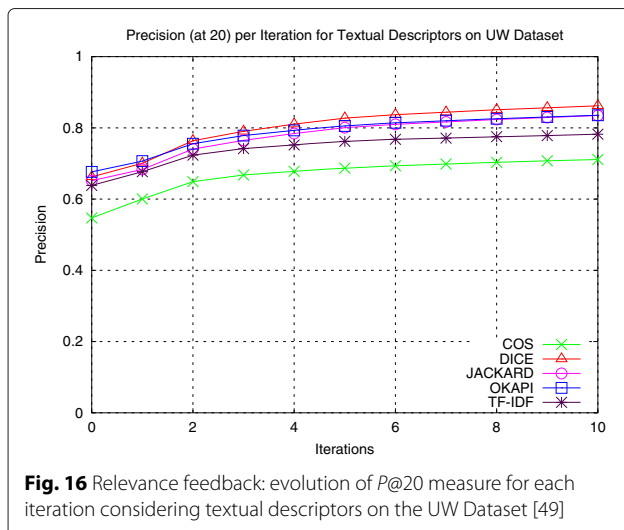
The collaborative image retrieval approach was evaluated using the same experimental setup of the relevance feedback. Figures 18 and 19 illustrates the $P \times t$ curves for the

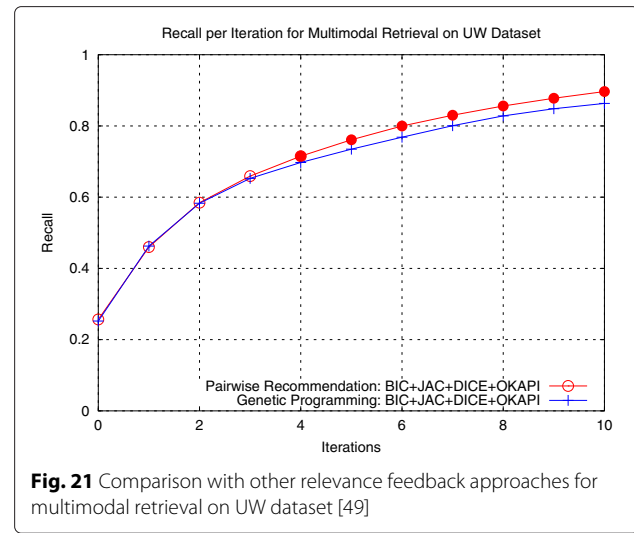
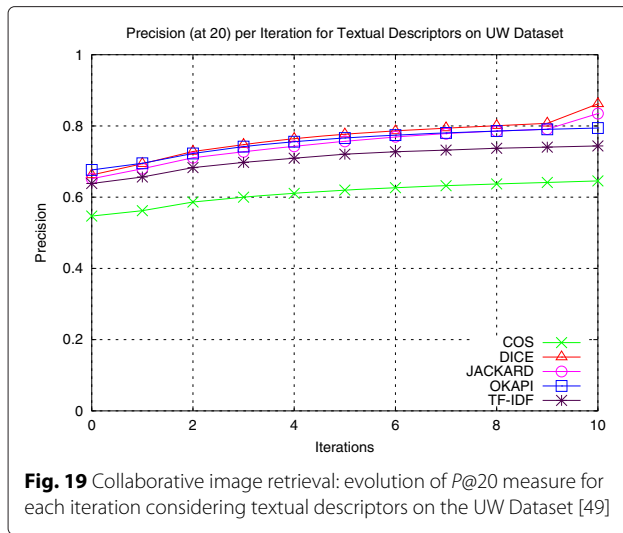
visual and textual descriptors, respectively. We can also notice an increasing precision score for all evaluated. We used the number of queries per iteration as $q_i = 55$.

The collaborative approach was also evaluated for multimodal retrieval, considering two visual descriptors and two textual descriptors. The results are illustrated in Fig. 20, which presents the evolution of $P@20$ along iterations for visual (BIC and JAC), textual (DICE and OKAPI), and the combination of the four descriptors.

7.6 Comparison with other approaches

Finally, we also evaluated our method in comparison with other multimodal relevance feedback approach. We considered a recently proposed relevance feedback approach based on Genetic Programming [53]. The UW dataset [49] was used considering a multimodal





retrieval scenario, combining textual and visual descriptors (BIC+JAC+DICE+OKAPI).

For effectiveness evaluation, we computed the evolution of recall along iterations ($R \times t$ curve) on the set of images actually observed by the user (20 per iteration). The goal is to analyze the percentage of relevant images retrieved given a number of relevance feedback iterations, which gives an approximation of the user effort on discovering new relevant images. Since the objective is to discover other relevant images, we increased the neighborhood size parameter to $k = 45$ on the Pairwise Recommendation method.

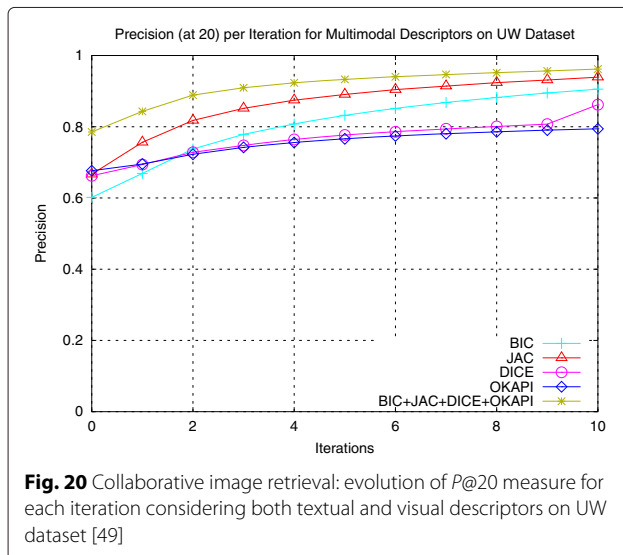
Figure 21 illustrates the $R \times t$ curve considering the proposed *semi-supervised pairwise recommendation for relevance feedback* and the Genetic Programming [53] approach as baseline. We can observe that, despite the

high effectiveness results of the baseline, the proposed method achieves comparable or better results for this task. We also used paired statistical significance t test to determine statistically significant differences in effectiveness. Filled red circles on the graph indicate iterations where the differences are statistically significant with a confidence level over 95 %.

7.7 Impact of collaborative users on effectiveness

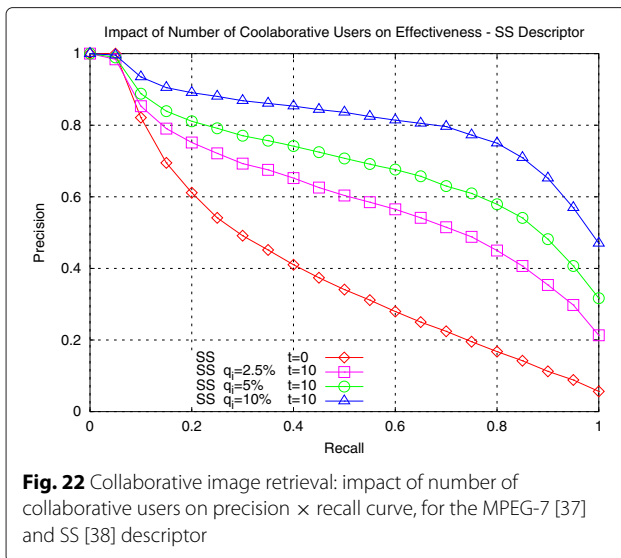
We also conducted an experiment for measuring the impact of the number of users on the effectiveness of retrieval results. The number of simultaneous queries per iteration q_i considered for each experiment is proportional to the dataset. We varied q_i between 2.5 and 10 %, computing the evaluation measures for each scenario. The MPEG-7 [37] dataset was considered for the experiment, due to existence of descriptors with lower effectiveness scores (SS and BAS shape descriptors were considered) allowing a better observation of the evolution of effectiveness according to the number of users. Figure 22 illustrates the results for the SS [38] descriptor while Fig. 23 considers the BAS [39] descriptor.

As we can observe for both descriptors, the more users are considered, the more effective are the results (better precision \times recall curves). We can also observe that the highest difference between curves is observed when we compare the original descriptor ($t = 0$) with the approach that considers the execution with 2.5 % of the dataset. That demonstrates the robustness of our method when dealing with a small number of users.



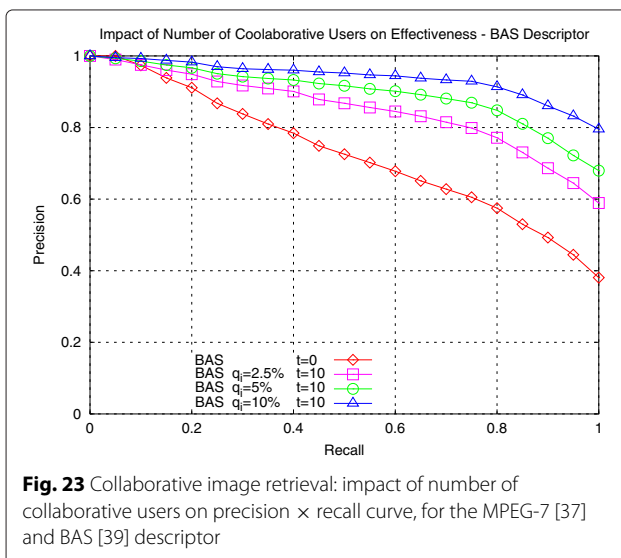
8 Conclusions

We have presented a novel semi-supervised learning algorithm for relevance feedback and collaborative image retrieval tasks. The proposed algorithm exploits both



labeled and unlabeled data aiming at improving the effectiveness of image retrieval tasks considering supervised and unsupervised steps. While the labeled data is obtained by user feedbacks, the unlabeled data is obtained from information encoded in ranked lists.

Various experiments considering several descriptors on four image collections demonstrated the effectiveness of the proposed approach. Experimental results showed the significant impact of the proposed semi-supervised algorithm on the quality of retrieved results along iterations. In diverse experiments, a drastically change of precision \times recall curve can be observed, illustrating the high effectiveness gains. The proposed was also evaluated considering statistical tests in comparison with a recently proposed genetic programming approach.



Future work focuses mainly on performing evaluations involving real users and the use of parallel programming and heterogeneous computing for computing simultaneous queries in collaborative image retrieval scenarios. We also plan to test the use of our collaborating retrieval approach in real-world search scenarios involving the access of multiple real users to publicly available image collections.

The joint use of active learning approaches with relevance feedback and collaborative image retrieval is also a promising research area. We intend to investigate new rank-based active learning methods for selecting the images showed in first user interactions.

Endnote

¹In this work, we do not use the clustering step proposed in the Pairwise Recommendation algorithm [9].

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

The authors are grateful to São Paulo Research Foundation - FAPESP (grants 2013/08645-0 and 2013/50169-1), CNPq (grants 306580/2012-8 and 484254/2012-0), CAPES, AMD, and Microsoft Research.

Author details

¹Department of Statistics, Applied Mathematics and Computing - State University of São Paulo (UNESP), Rio Claro, SP, Brazil. ²Recod Lab - Institute of Computing, University of Campinas (UNICAMP), Campinas, SP, Brazil. ³Department of Exact Sciences, University of Feira de Santana (UEFS), Feira de Santana, BA, Brazil.

Received: 30 January 2015 Accepted: 21 July 2015

Published online: 10 August 2015

References

1. RDS Torres, Falcão, MA Gonçalves, JP Papa, B Zhang, W Fan, EA Fox, A genetic programming framework for content-based image retrieval. *Pattern Recognit.* **42**(2), 283–292 (2009)
2. ATd Silva, AX Falcão, LP Magalhães, Active learning paradigms for cbir systems based on optimum-path forest classification. *Pattern Recognit.* **44**(12), 2971–2978 (2011)
3. B Thomee, M Lew, Interactive search in image retrieval: a survey. *Int. J. Multimedia Inf. Retrieval.* **1**(2), 71–86 (2012)
4. B Thomee, M Lew, in *13th Conference of the Advanced School for Computing and Imaging, Netherlands*. Relevance feedback in content-based image retrieval: promising directions (Heijten, The Netherlands, 2007), pp. 450–456
5. H Jegou, C Schmid, H Harzallah, J Verbeek, Accurate image search using the contextual dissimilarity measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(1), 2–11 (2010)
6. J Jiang, B Wang, Z Tu, in *IEEE International Conference on Computer Vision (ICCV'2011)*. Unsupervised metric learning by self-smoothing operator (IEEE, 2011), pp. 794–801
7. DCG Pedronette, Rds Torres, Image re-ranking and rank aggregation based on similarity of ranked lists. *Pattern Recognit.* **46**(8), 2350–2360 (2013)
8. I Cohen, FG Cozman, N Sebe, MC Cirelo, TS Huang, Semisupervised learning of classifiers: theory, algorithms, and their application to human-computer interaction. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(12), 1553–1566 (2004)
9. DCG Pedronette, Rds Torres, Exploiting pairwise recommendation and clustering strategies for image re-ranking. *Inform. Sci.* **207**, 19–34 (2012)

10. DCG Pedronette, RT Calumby, RdS Torres, in *27th SIBGRAPI Conference on Graphics, Patterns and Images*. Semi-supervised learning for relevance feedback on image retrieval tasks (Rio de Janeiro, 2014), pp. 243–250
11. TS Huang, XS Zhou, in *Proceeding of International Conference on Image Processing (ICIP)*. Image retrieval with relevance feedback: From heuristic weight adjustment to optimal learning methods (Thessaloniki, 2001)
12. XS Zhou, TS Huang, Relevance feedback in image retrieval: A comprehensive review. *Multimedia Syst.* **8**(6), 536–544 (2003)
13. J Li, NM Allinson, in *Handbook on Neural Information Processing*. Relevance feedback in content-based image retrieval: A survey, vol. 49, (2013). Online ISBN: 978-3-642-36657-4
14. L Duan, W Gao, W Zeng, D Zhao, Adaptive relevance feedback based on bayesian inference for image retrieval. *Signal Process.* **85**(2), 395–399 (2005)
15. M Arevalillo-Herráez, M Zacarés, X Benavent, Ves E de, A relevance feedback cbir algorithm based on fuzzy sets. *Image Commun.* **23**(7), 490–504 (2008)
16. CD Ferreira, JAdS Santo, RdS Torres, MA Gonçalves, RC Rezende, W Fan, Relevance feedback based on genetic programming for image retrieval. *Pattern Recognit. Lett.* **32**(1), 27–37 (2011)
17. S Tong, E Chang, in *Proceedings of the Ninth ACM International Conference on Multimedia*. Support vector machine active learning for image retrieval (ACM New York, USA, 2001)
18. Z-H Zhou, K-J Chen, H-B Dai, Enhancing relevance feedback in image retrieval using unlabeled data. *ACM Trans. Inform. Syst.* **24**(2), 219–244 (2006)
19. Q Tian, J Yu, Q Xue, N Sebe, in *IEEE International Conference on Multimedia and Expo (ICME '04)*. A new analysis of the value of unlabeled data in semi-supervised learning for image retrieval, vol. 2 (Taipei, 2004), pp. 1019–10222
20. D Zhou, J Weston, A Gretton, O Bousquet, B Schölkopf, in *Advances in Neural Information Processing Systems*. Ranking on data manifolds (MIT Press, 2004)
21. DCG Pedronette, OAB Penatti, RDS Torres, Unsupervised manifold learning using reciprocal knn graphs in image re-ranking and rank aggregation tasks. *Image Vis. Comput.* **32**(2), 120–130 (2014)
22. X Yang, L Prasad, LJ Latecki, Affinity learning with diffusion on tensor product graph. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(1), 28–38 (2013)
23. Y Huang, Q Liu, S Zhang, DN Metaxas, in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference On*. Image retrieval via probabilistic hypergraph ranking (San Francisco, CA, 2010), pp. 3376–3383
24. SCH Hoi, R Jin, J Zhu, MR Lyu, Semi-supervised SVM batch mode active learning with applications to image retrieval. *ACM Trans. Inform. Syst.* **27**(3), 16–11629 (2009)
25. L Zhang, L Wang, W Lin, Semisupervised biased maximum margin analysis for interactive image retrieval. *IEEE Trans. Image Process.* **21**(4), 2294–2308 (2012)
26. H Chang, D Yeung, in *Proceedings of the British Machine Vision Conference 2005*. Stepwise metric adaptation based on semi-supervised learning for boosting image retrieval performance (Oxford, UK, 2005)
27. Y Yang, F Nie, D Xu, J Luo, Y Zhuang, Y Pan, A multimedia retrieval framework based on semi-supervised ranking and relevance feedback. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(4), 723–742 (2012)
28. M Halvey, in *Proceedings of the 3rd International Workshop on Collaborative Information Retrieval*. Preference based feedback for collaborative image retrieval (ACM New York, USA, 2011), pp. 19–22
29. S Yan, W Zheng-xuan, W Dong-mei, in *International Conference on Biomedical Engineering and Computer Science (ICBECS)*. An image retrieval method based on relevance feedback and collaborative filtering (Wuhan, 2010), pp. 1–5
30. SC Hoi, W Liu, S-F Chang, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2008)*. Semi-supervised distance metric learning for collaborative image retrieval (Anchorage, AK, 2008), pp. 1–7
31. K Chandramouli, E Izquierdo, in *10th Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS'09)*. Multi-class relevance feedback for collaborative image retrieval (London, 2009), pp. 214–217
32. DCG Pedronette, EdS Torres, Exploiting clustering approaches for image re-ranking. *J. Vis. Lang. Comput.* **22**(6), 453–466 (2011)
33. X Yang, S Koknar-Tezel, LJ Latecki, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2009)*. Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval (Miami, FL, 2009), pp. 357–364
34. J Wang, Y Li, X Bai, Y Zhang, C Wang, N Tang, Learning context-sensitive similarity by shortest path propagation. *Pattern Recognit.* **44**(10–11), 2367–2374 (2011)
35. D Qin, S Gammeter, L Bossard, T Quack, Lvan Gool, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2011)*. Hello neighbor: Accurate object retrieval with k-reciprocal nearest neighbors (Providence, RI, 2011), pp. 777–784
36. P Tandon, CV Jawahar, in *National Conference on Communications (NCC '09)*. Long term learning for content extraction in image retrieval (Guwahati, India, 2009), pp. 1–5
37. LJ Latecki, R Lakmper, U Eckhardt, in *Conference on Computer Vision and Pattern Recognition (CVPR)*. Shape descriptors for non-rigid shapes with a single closed contour (Hilton Head Island, SC, 2000), pp. 424–429
38. RdS Torres, AX Falcão, Contour salience descriptors for effective image retrieval and analysis. *Image Vis. Comput.* **25**(1), 3–13 (2007)
39. N Arica, FTY Vural, BAS: a perceptual shape descriptor based on the beam angle statistics. *Pattern Recognit. Lett.* **24**(9–10), 1627–1639 (2003)
40. H Ling, DW Jacobs, Shape classification using the inner-distance. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(2), 286–299 (2007)
41. RO Stehling, MA Nascimento, AX Falcão, in *ACM Conference on Information and Knowledge Management (CIKM'2002)*. A compact and efficient image retrieval approach based on border/interior pixel classification (ACM New York, USA, 2002), pp. 102–109
42. J Huang, SR Kumar, M Mitra, W-J Zhu, R Zabih, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'97)*. Image indexing using color correlograms (San Juan, 1997), pp. 762–768
43. MJ Swain, DH Ballard, Color indexing. *Int. J. Comput. Vis.* **7**(1), 11–32 (1991)
44. JV de Weijer, C Schmid, in *European Conference on Computer Vision (ECCV'2006)*, Lecture Notes in Computer Science. Coloring local feature extraction, vol. 3952, (2006), pp. 334–348. doi:10.1007/11744047_26
45. T Ojala, M Pietikäinen, Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
46. V Kovalev, S Volmer, in *International Conference on Multimedia Modeling (MMM'98)*. Color co-occurrence descriptors for querying-by-example (Lausanne, 1998), pp. 32–38
47. B Tao, BW Dickinson, Texture recognition and image retrieval using gradient indexing. *J. Vis. Commun. Image Representation.* **11**(3), 327–342 (2000)
48. P Brodatz, *Textures: A Photographic Album for Artists and Designers*. (Dover Publications, Mineola, New York, 1966)
49. T Deselaers, D Keyers, H Ney, Features for image retrieval: an experimental comparison. *Inform. Retrieval.* **11**(2), 77–107 (2008)
50. A Williams, P Yoon, Content-based image retrieval using joint correlograms. *Multimedia Tools Appl.* **34**(2), 239–248 (2007)
51. P Wu, BS Manjunath, SD Newsam, HD Shin, in *IEEE Workshop on Content-Based Access of Image and Video Libraries*. A texture descriptor for image retrieval and browsing (Fort Collins, CO, 1999), pp. 3–7
52. C-B Huang, Q Liu, in *International Conference on Communications, Circuits and Systems (ICCCAS 2007)*. An orientation independent texture descriptor for image retrieval (Kokura, 2007), pp. 772–776
53. RT Calumby, RdS Torres, MA Gonçalves, Multimodal retrieval with relevance feedback based on genetic programming. *Multimedia Tools Appl.* **69**(3), 991–1019 (2014)