User-driven Nearest Neighbour Exploration of Image Archives

Luca Piras, Deiv Furcas and Giorgio Giacinto

Department of Electrical and Electronic Engineering, University of Cagliari, Piazza D'armi, Cagliari, 09123, Italy

{luca.piras, giacinto}@diee.unica.it, d.furcas@gmail.com

Keywords: Feature Space Exploration, Nearest Neighbour, Relevance Feedback, Query Shifting, Image Retrieval

Abstract: Learning what a specific user is exactly looking for, during a session of image search and retrieval, is a problem that has been mainly approached with "classification" or "exploration" techniques. Classification techniques follow the assumption that the images in the archive are statically subdivided into classes. Exploration approaches, on the other hand, are more focused on following the varying needs of the user. It turns out that image retrieval techniques based on classification approaches, though often showing good performances, are not prone to adapt to different users' goals. In this paper we propose a relevance feedback mechanism that drives the search into promising regions of the feature space according to the Nearest Neighbor paradigm. In particular, each image labelled as being relevant by the user, is used as a "seed" for an exploration of the space based on the Nearest Neighbors paradigm. Reported results show that this technique allows attaining higher recall and average precision performances than other state-of-the-art relevance feedback approaches.

1 INTRODUCTION

Nowadays, the high availability of pictures that digital cameras, tablets and smart-phones allows us to quickly capture, makes more and more pressing the need for systems that categorize and label our image archives in a "smart" way. While search engines on the internet such as Google and Bing play this role very well for images published on the web, effective approaches for personal and professional image archive search still require further investigation (Sivic and Zisserman, 2008).

Over the years, Content Based Image Retrieval (CBIR) techniques proved to be a good choice. Users query the system using a sample image, and expect that the system returns a set of images of the same *category* of the query. To perform this task, images are described through low-level features such as, for example, color, texture, shapes, or characteristic points. It is easily understood that the way in which these characteristics are represented inevitably constrains the results that can be obtained (Datta et al., 2008; Thomee and Lew, 2012). In addition, these approaches are always dependent on the choice of the low-level features and the used metrics (Lew et al., 2006; Pavlidis, 2008).

One of the main problems to face when a user is interested in performing a so-called "category" search is that different users have different perceptions of similarity and, often, at the beginning of the search process, the user may not have a clear idea of the images she is looking for. It is a common experience that at the time the user begins the search, she has in mind a rough idea of what she wants, and only after having seen several examples, and having explored part of the archive, she can focus her search more precisely.

In order to "help" Image Retrieval Systems to follow the user in this path, it is necessary to provide the system with a mechanism that interprets the will of the user and adapt *itself* to it. In the past years, several Relevance Feedback (RF) mechanisms have been proposed for this task, where the user can judge the images that the system returns as being relevant or not w.r.t. the user's query, and label them accordingly (Zhou and Huang, 2003). Over the years, the problem of learning what a specific user is exactly looking for has been mainly approached in two different ways, i.e., by "classification" or "exploration" approaches. The first approach is essentially based on training a pattern classifier using the set of images that the user, at each relevance feedback iteration, labels as being relevant or not (Thomee and Lew, 2012). In this way it is possible to incrementally create a training set that allows the classifier to "understand" the user's tastes. Several approaches follow this line of thinking and, as in other fields of Pattern Recognition and Machine Learning, Support Vector Machines (SVM) have been widely employed (Rao et al., 2006; Chen et al., 2001;

Hoi et al., 2009; Zhang et al., 2001; Tong and Chang, 2001). Even if SVMs are often used for Image Retrieval tasks, reported results often do not disclose the fact that good performance is strictly dependent on the choice of the most appropriate SVM kernel and the associated parameters. In addition, classification approaches, due to their own characteristics, tend to be static and not prone to adapt to the fickle needs of the user, because the underlying assumption is that the images in the archive can be thought as being statically subdivided into classes, and user's feedback is used to sample the class distribution of images.

On the other hand, approaches based on "exploration" paradigms aim to explore the feature space not only in the area of the initial query image, but also in different neighborhoods computed according to relevance information. Thanks to this prerogative, explorative approaches tend to be highly responsive because they are explicitly designed to follow the user's needs. For example, approaches based on the Nearest Neighbor (NN) paradigm can be used to easily implement explorative approaches (Piras et al., 2012), thanks to the very limited number of parameters to be set (Boiman et al., 2008).

Another issue that has been investigated in the past in the Relevance Feedback field, is related to the way the images are presented to the user. Often, the first n best ranked images are shown to the user, and, usually, these images are located in a limited area of the feature space quite close to the initial query. In this way, after the first few iterations, the system might not be able to find new relevant images to present to the user if the search converges towards a local optimum (Piras et al., 2012).

The above considerations motivate the proposal in this paper, i.e., to exploit the simplicity of the NN paradigm, based on the concept that similar images are located in adjacent areas of the feature space. In particular, we introduce the concept of *"transitive similarity"*, where two patterns I_1 and I_3 can be considered similar if I_2 is in the neighborhood of I_1 and I_3 is in the neighborhood of I_2. This concept is not new, and it has been inspired by the notion of data point $k-NN$ consistency for data clustering (Ding and He, 2004). We used this concept for computing, at each iteration, an exploration seed point that takes into account the set of relevant and not relevant images retrieved so far. Then, we evaluate the neighborhood of this seed point, and, for each neighbor, we consider its nearest neighbors. In this way, we avoid to focus on a limited area of the feature space by considering a large number of neighbors of the initial seed, as it may contain a large fraction of non-relevant images. On the other hand, the proposed mechanism allows exploring a larger number of search directions of the representation space, thus driving the search into "new" regions of the feature space where to find relevant images.

To illustrate in detail the proposed mechanism, this paper is organized as follows. Section 2 briefly reviews the related works on relevance feedback. Section 3 describes the proposed relevance feedback technique, that we named *"Nearest Neighbour Exploration Path"*. Experimental results are reported in Section 4. Conclusions are drawn in Section 5.

2 RELEVANCE FEEDBACK AND EXPLORATION OF THE FEATURE SPACE

The problem of finding and showing to the user new relevant images during her exploration of image archives has been addressed in the field of CBIR in different ways. One of the first techniques used to perform relevance feedback, that is still used in a number of image retrieval applications, is based on the query shifting paradigm. Originally, the query shifting mechanism has been developed in the text retrieval field, and based on the Rocchio formula (Rocchio, 1971). This formula has been then proposed for relevance feedback for CBIR tasks in (Rui et al., 1997):

$$Q_{opt} = \frac{1}{N_R} \sum_{i \in D_R} D_i - \frac{1}{N_T - N_R} \sum_{i \in D_N} D_i \tag{1}$$

Where D_R and D_N are the sets of relevant and non relevant images respectively, N_R is the number of images in D_R, N_T the number of the total documents, and D_i is the representation of an image in the feature space. This approach is motivated by the assumption that the query may lie in a region of the feature space that is in some way "far" from the images that are relevant to the user. On the contrary, according to the Eq.(1), the optimal query should lie near to the euclidean center of the relevant images and 'far" from the non relevant images. The same line of thinking has been also followed in (Giacinto and Roli, 2004a) where a Bayesian model for estimating the decision boundary between relevant and non-relevant images has been employed (see Section 3.1).

Relevance Feedback has been also formulated in terms of a pattern classification task using neural networks, self-organizing maps (SOMs) (Laaksonen et al., 2002) or approaches based on SVM. The latter have been widely used to model the concepts behind the set of relevant images, and adjust the search accordingly (Zhang et al., 2001; Chen et al., 2001). In these cases, it is usually difficult to produce a high-level generalization of a "class" of objects as it is difficult to provide a general model that can be adapted to represent different concepts of similarity. In addition, the number of available cases may be too small to estimate the optimal set of parameters for such a general model. This kind of problems have been partially mitigated thanks to the use of the *active learning* paradigm (Cohn et al., 1994), where the system is trained not only with the most relevant images according to the user judgement, but also with the most informative images that allow driving the search into more promising regions of the feature space. One of the approaches used to select informative images is based on choosing the patterns closest to the decision boundary, as described in (Hoi et al., 2009; Tong and Chang, 2001) where SVM based on active learning are used. In addition, Nearest Neighbor techniques have been used in the context of the *active learning* paradigm: in (Lindenbaum et al., 2004) the authors proposed to perform selective sampling for Nearest Neighbor classifiers. In order to choose the most informative patterns, the authors suggest to consider the effect of its classification on the remaining unlabeled points. Their algorithm is based on sampling sequences of neighboring patterns of length k, and selects an example that leads to the best sequence. The best sequence is the one whose samples have the highest conditional class probabilities.

The Nearest Neighbor paradigm over the years has been adapted in several forms for the exploitation of relevance feedback. One of these forms exploits relevance feedback by comparing all the images of the database against relevant and non-relevant images, and assigns to each image a *Relevance Score* (Giacinto, 2007) as follows:

$$rel_{NN}(I) = \frac{\|I - NN^{nr}(I)\|}{\|I - NN^{r}(I)\| + \|I - NN^{nr}(I)\|} \quad (2)$$

where $NN^{r}(\cdot)$ and $NN^{nr}(\cdot)$ denote the nearest relevant and non relevant image of the image I respectively, and $\|\cdot\|$ is the metric defined in the feature space at hand. In (Arevalillo-Herráez and Ferri, 2010) the authors propose to modify that formulation introducing a smoothed NN estimate (SNN) in order to increase the importance of the images more relevant to the user query. In (Arevalillo-Herráez and Ferri, 2013) instead, an improved score using a reliability estimate has been proposed.

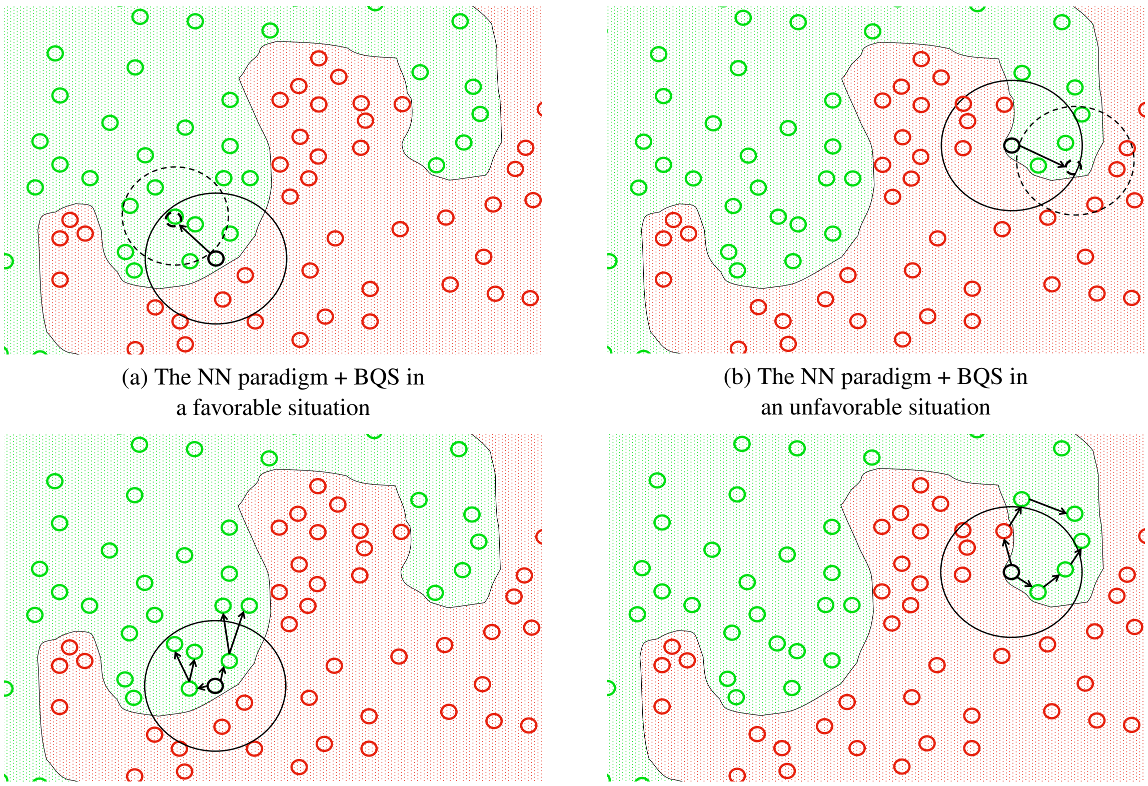
Apart from the techniques based on active learning, that are, however, based on a classification approach, there are not many papers focused on the exploration of the feature space. It is worth to note that also the approaches based on the Nearest Neighbor paradigm, that have a more clear explorative vocation, have been usually focused on maximizing the retrieval precision rather than on the exploration of the feature space, thus maximizing the recall. Our work aims to fill this gap.

3 NEAREST NEIGHBOR EXPLORATION PATH

Let us assume that the set of low-level features that we have extracted from each image of an archive, are such that a pair of images judged by the user as being similar to each other is represented by two near points in the feature space. While often CBIR tasks have been formulated in terms of a user that is interested in retrieving images belonging to a specific "category", we formulate the problem in terms of a user that is looking for "something similar" to the submitted image query, without any clear specification of a "category" the images should belong to. According to the first assumption, the images the user is interested in lie in a neighborhood of the query. If this assumption turns out to be true, i.e., the query lies in a region of the feature space where other similar images surround it, an isotropic search based on the Nearest Neighbor paradigm would allow retrieving a large number of relevant images.

Unfortunately, this configuration of the initial image query does not occur very frequently, and, in any case, being particularly easy to deal with, does not deserve further investigation. Much more interesting are the cases in which the initial query is close to regions containing images that are not relevant to the user's needs. In these cases we can distinguish between two possible configurations that are depicted in Fig. 1: a more favorable one in which the boundary between relevant and non-relevant images can be approximated as a convex hull within the area of influence of the query (Fig. 1(a)), and another one in which the separation between relevant and non-relevant images is not so clear (Fig. 1(b)). In the last case, an approach that just explores the feature space in the neighborhood of the query according to an isotropic NN search, is not effective. On the other hand, a technique that better explores the features space where the query lies, and is able to find more "interesting" regions where to perform the search for relevant images, is highly desirable.

In this section we provide the details of the two exploration methods that we propose in this paper. Both methods are based on the same *anisotropic* approach, that exploits the Nearest Neighbor paradigm in two different ways. The underlying rationale is the concept of *"transitive similarity"*, where two patterns I_1 and I_3 can be considered similar if I_2 is in the neighborhood of I_1 and I_3 is in the neighborhood of I_2. This concept is inspired by the notion of data point $k-NN$



(a) The NN paradigm + BQS in a favorable situation

(b) The NN paradigm + BQS in an unfavorable situation

(c) The NN paradigm + "*NN Exploration path*" in a favorable situation

(d) The NN paradigm + "*NN Exploration path*" in an unfavorable situation

Figure 1: The proposed Nearest Neighbor Exploration path algorithm in different situation compared to the BQS technique. The black circle represents the query, the green circles the images relevant to the query, the red circles the non relevant ones. Figure (c) shows the NN Exploration path through the $N+N\cdot M$ nearest points where $N=2$, $M=2$, and $k=6$. Figure (d) shows the NN Exploration path through the $N\cdot M$ nearest points where $N=2$, $M=3$, and $k=6$.

consistency for data clustering (Ding and He, 2004). In particular, if we are interested in retrieving k images relevant to the query (Q), instead of extracting the k nearest neighbors of Q, we use $N<k$ nearest neighbors of Q, and then, for each neighbor, we compute its nearest neighbors, so that the total number of images is k (see Fig. 1(c)). In this way, we consider the closest neighbors of Q, that are the most similar to Q by definition, and then we consider the most similar patterns to the neighbors of Q. Thus we do not take into account those images that may be loosely related to the query as their distance from it is larger than the distance of the nearest neighbors of each image in the neighborhood of Q.

The proposed methods are based on the computation of a reference point that we will call the "seed" in the following. Basically, we explore the feature space starting from the nearest point of the current seed. At the first step the role of seed is assigned to the query image, that the system receives as an input data. From the second iteration onwards, the role of the seed is taken by the query shifting mechanism computed according to (Giacinto and Roli, 2004a) that is briefly reviewed in the next subsection. Then, distances from the seed and each other image are calculated.

3.1 Bayesian Query Shifting (BQS)

In order to limit the exploration in regions not too far away from the region where known relevant images lie, the exploration approaches, that we show in the next sections, are seeded by a query point movement strategy (QPM) (Rocchio, 1971). In particular, in this paper at each iteration the role of the seed is played by a modified query vector computed according to the Bayes decision theory (Bayes Query Shifting, BQS) (Giacinto and Roli, 2004a):

$$Q_{BQS} = \mathbf{m}_R + \frac{\sigma}{\|\mathbf{m}_R - \mathbf{m}_N\|}\left(1 - \frac{k_R - k_N}{\max(k_R, k_N)}\right)(\mathbf{m}_R - \mathbf{m}_N) \tag{3}$$

where $\mathbf{m}_R$ and $\mathbf{m}_N$ are the mean vectors of relevant and non-relevant images respectively, σ is the standard deviation of the images belonging to the neighborhood of the original query, and k_R and k_N are the number of relevant and non relevant images, respectively. The new query Q_{BQS} lies on the line connecting the two means, in the $\mathbf{m}_R\mathbf{m}_N$ direction, the magnitude of the shift depending on the proportion of relevant and non-relevant images retrieved.

3.2 NN Exploration Path through the $N + N \cdot M$ Nearest Points

In order to explore the feature space in different directions, the first method we propose begins the exploration from N different points around the current seed Q (i.e., the initial query, or the BQS) that we can indicate as belonging to the set $S^0 = \{NN_i(Q) \mid i = 1,\ldots,N\}$ where $NN_i(Q)$ is the i^{th} nearest point of the seed Q. With the purpose of maximizing the exploration area, this algorithm is designed to avoid overlaps between the portions of space explored by different seeds. From these new seed points the algorithm continues to explore considering their M nearest images in $S' = \{NN_j(x) \mid x \in S^0, j = 1,\ldots,M\}$. The values of N and M are chosen such that $N + N \cdot M = k$ where k is the number of images to show to the user (e.g., Fig. 1(c)). The algorithm can be summarized in the following steps:

1. Let Q the first seed point and $NN(\cdot)$ the function that denote the nearest image, for $i = 1,\ldots,N$ evaluate $NN_i(Q)$.
2. Given the set $S^0 = \{NN_i(Q) \mid i = 1,\ldots,N\}$, $\forall x \in S^0$ evaluate $NN_j(x)$ for $j = 1,\ldots,M$.
3. Given the set $S' = \{NN_j(x) \mid x \in S^0, j = 1,\ldots,M\}$ the set of images to be shown to the user will be $S^k = S^0 \cup S'$

3.3 NN Exploration Path through the $N \cdot M$ Nearest Points

This second method performs the exploration of the feature space beginning from the N nearest images to the current seed Q (i.e., the initial query, or the BQS). As we showed in Section 3.2, the set of the new seed points is $S^0 = \{NN_i(Q) \mid i = 1,\ldots,N\}$, and, for each neighbour, we select its nearest neighbour that will play the role of a seed in the following phases of exploration. Accordingly, the set of seed points will be $S^j = \{NN(x) \mid x \in S^{j-1}, j = 1,\ldots,M\}$. The values of N and M are chosen such that $N \cdot M = k$ where k is the number of images to show to the user (e.g., Fig. 1(d)).

The algorithm can be summarized in the following steps:

1. Let Q the first seed point and $NN(\cdot)$ the function that denote the nearest image, for $i = 1,\ldots,N$ evaluate $NN_i(Q)$.
2. Given the set $S^0 = \{NN_i(Q) \mid i = 1,\ldots,N\}$, $\forall x \in S^0$ evaluate $NN(x)$.
3. Given the set $S^j = \{NN(x) \mid x \in S^{j-1}, j = 1,\ldots,M\}$ the set of images to be shown to the user will be $S^k = S^0 \cup S' \cup \cdots \cup S^M$

Summing up, this technique differs from the previous one in the number of neighbors considered. While here we consider just the nearest neighbor point for each of the N points in the neighborhood of the current seed, the former technique takes into account M points for each of the N points in the neighborhood of the current seed.

Figures 1(c) and 1(d) depict two examples of the use of the two proposed approaches. The first one shows the NN Exploration path through the $N + N \cdot M$ nearest points where $N = 2$, $M = 2$, and $k = 6$. Fig. 1(d) shows the second approach, i.e., the $N \cdot M$ nearest points, where $N = 2$, $M = 3$, and $k = 6$. The black circle represents the initial query, the green circles the images relevant to the query, the red circles the non-relevant ones. It is possible to see how in a favorable situation (e.g., Fig. 1(a) and 1(c)) both a query point movement strategy (such as the BQS), and the use of one of the proposed approaches are able to find images that are relevant to the query. On the other hand, in an unfavorable situation, an isotropic NN search would retrieve the k nearest images to the query, disregarding the fact that the volume that contains these images could have a large radius, and thus incorporate a large number of images that are non-relevant to the query. Our approach, instead, explores the feature space in the neighborhood of the query "step by step", through images close to each other, and thus it is able to find more "interesting" regions where to perform the search for relevant images. From Fig. 1(d) it is also possible to observe that even if one of the nearest images is non-relevant to the user query, our method is able to "correct" the path. This behavior could be explained by the assumption that the extracted features are such that a pair of images judged by the user as being similar to each other is represented by two near points in the feature space. In this situation, is thus likely that the retrieved non-relevant images are in some-way similar to the relevant ones and near to other relevant images.

4 EXPERIMENTAL RESULTS

4.1 Dataset

Experiments have been carried out using a subset of the Corel dataset obtained from the UCI KDD repository[1]. The dataset consists of 30,000 images manually subdivided into 71 semantic classes (Giacinto and Roli, 2004b). Images have been represented using the four features vectors available at the UCI web

[1] http://kdd.ics.uci.edu/databases/CorelFeatures/CorelFeatures.html

site: *Color Histogram*, *Color Histogram Layout*, *Co-Occurrence Texture* and *Color Moments*. Distances between features have been evaluated using the histogram intersection (Swain and Ballard, 1991) on the color histograms and the Euclidean distance for the other descriptors, they have been normalized in the range [0,1], and then summed up (Arevalillo-Herráez and Ferri, 2013) in order to obtain a unique value.

4.2 Experimental Setup

In order to test the performance of the proposed approaches, 500 query images from the dataset have been randomly extracted, so that they cover all the semantic classes. Relevance feedback is performed by marking images belonging to the same class of the query as relevant, and all other images in the pool of k to-be-labelled images as non-relevant. Performance is evaluated in terms of Precision, Recall, and Average Precision (Wang et al., 2010) that measures the average value of precision for each different recall value:

$$AP = \frac{1}{R}\sum_{i=1}^{n} rel(\tau(i))\frac{\sum_{j=1}^{i} rel(\tau(j))}{i} \quad (4)$$

where R is the number of relevant images, n is the number of images in the dataset, $\tau(i)$ is the image at the rank i, and $rel(\tau(i))$ is the associated binary relevance label equal to 1 if $\tau(i)$ is relevant w.r.t. the query, and 0 otherwise. The higher the value of AP, the better the ranking. To measure the Recall, the images that have been already labelled in a previous iteration are not considered as candidate for the next iterations. On the contrary, in measuring the Precision, all the images are considered as candidate in each iteration (Arevalillo-Herráez and Ferri, 2013).

For comparison purposes, the proposed approach has been compared against four approaches based on the NN paradigm: a NN technique enhanced with a smoothed estimator (SSN) as in (Arevalillo-Herráez and Ferri, 2010); a distance based approach where the image score is improved using a reliability estimate (Distance Based) (Arevalillo-Herráez and Ferri, 2013); an approach where the image relevance is estimated using a Relevance Score that takes into account the position in the feature space of the known relevant and non-relevant images (NN + BQS) (Giacinto, 2007), and an extension of the previous work where an exploration component has been introduced (NN + BQS + EE) (Piras et al., 2012). In the latter approach the parameters have been set according to the results obtained by the authors, in particular the parameter "α" has been set equal to 25%.

User's feedback has been also used to build the training set for an active SVM classifier (Tong and Chang, 2001). The choice of an active approach is due to its good performance in image retrieval tasks and in order to compare exploration techniques with a classification approach at the state of the art under the best possible setting. As SVM training requires choosing the kernel and the kernel parameters, a number of experiments have been performed using different kernels and different kernel parameters. Reported results are related to a Gaussian kernel as described in the original publication.

In order to provide the reader with a broader comparison, other relevance feedback algorithms have been considered: a query point movement approach (QPM) as the one described in Section 3.1; a probabilistic framework presented in (Arevalillo-Herráez et al., 2010) (Probabilistic); and the self-organizing map (SOM) method introduced in (Laaksonen et al., 2002).

4.3 Results

Figure 2 shows the performance for the methods proposed in Sections 3.2 (*NN-E* ($N+N\cdot M$)) and 3.3 (*NN-E* ($N\cdot M$)) using $N=5$, $M=4$, and $k=20$ for the latter and $N=5$, $M=3$, and $k=20$ for the first one. In order to choose the parameter that allowed attaining the highest performance, a number of preliminary experiments have been performed on a small subset of data.

It is easy to see how the proposed approaches provide better performance in terms of Precision and Recall than all the other methods. By considering the Average Precision, the proposed methods exhibit a higher performance than all the other approaches until the fifth/sixth iteration. Although the Average Precision may be of less interest for an approach focused on CBIR, it is interesting to see how the proposed algorithms work very well in the first few iterations, that are the ones performed by the vast majority of users, as typically just a tiny fraction of users go on after the forth/fifth iteration (Tronci et al., 2013).

This behavior can be explained by considering that the proposed approaches, after the first few iterations, explore regions where the number of relevant images is very small and the proposed algorithms are still able to find some images similar to the query, even if with a smaller increase than at the first few iterations.

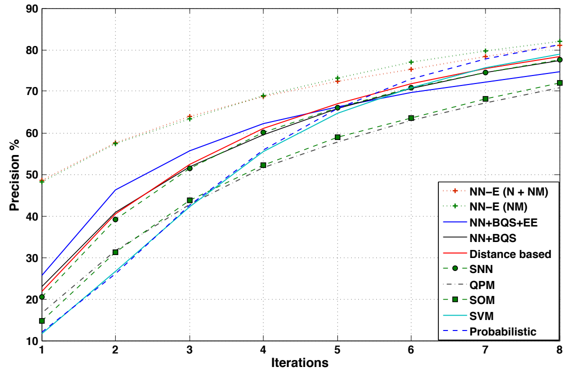
The proposed techniques are also able to perform better in Recall than the *NN + BQS + EE* approach, that, as expected, outperformed the *NN + BQS* mechanism thanks to the exploitation of the relevance feedback for exploring the feature space. In this case too, all the relevance feedback mechanisms based on *Exploration* approaches, work well till the fifth iteration in terms of Precision and Average Precision.

The results attained using the *SVM Active* approach show that the performance in the first few iterations is very low w.r.t. the other approaches. This is due to the too small number of samples to learn a model, and this is a problem that we observed when using approaches based on the "classification" paradigm, when no constraint is put on the minimum number of training images.

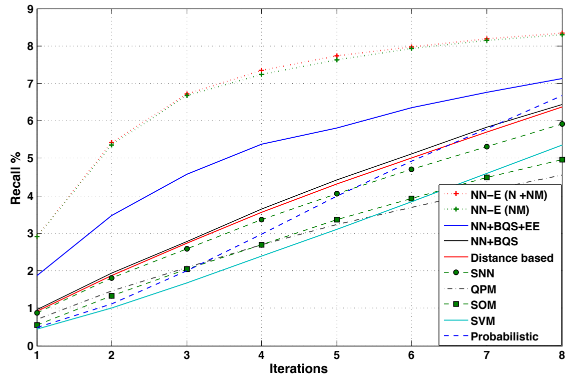
If we compare the performance attained by the two proposed approaches, the best results have been obtained by the exploration through $N+N \cdot M$ nearest points. The main reason behind this result is the capability of this approach to expand the search while remaining close to the area where relevant images lie. On the other hand, a drawback that can arise when using the $N \cdot M$ nearest points technique is that the approach could not be able to correct the "path" if among the nearest N images there are too many non-relevant images.

The visual inspection of the retrieval results confirmed the rationale behind the "classification" approaches, and the "exploration" approaches. Classification mechanisms provided good performance for those cases in which similar images can be considered as "near duplicates". The proposed exploration mechanisms exhibited better performance when a chain of similarities can be built among images bearing the same concept, because only small subsets of them can be considered as "near duplicates", the intersection of subsets providing the link between different images with the same concept. Thus, the proposed approaches proved to be more effective in concept retrieval thanks to their exploration capabilities.

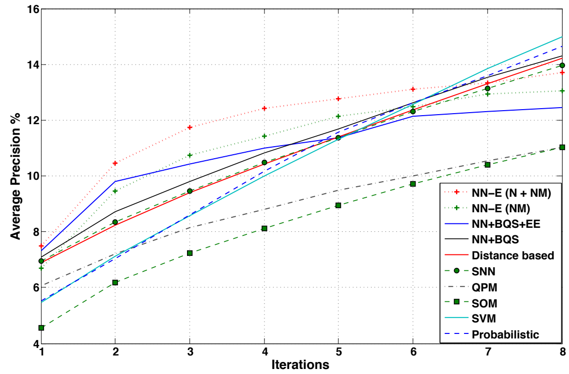
In order to test the significance of our results, the Friedman test (García et al.,) has been performed for each measure and query. This test demonstrated that there is statistically significant difference in precision, recall, and average precision among the proposed approaches and the best of the other methods (i.e., the *NN + BQS + EE* technique) according to a post-hoc Holm test at significance level $\alpha = 0.05$. A fortiori, it is possible to deduce the same for all the other methods. The only case in which the difference is not statistically significant, is the comparison between the average precision obtained by (*NN + BQS*) and (*NN-E* $(N \cdot M)$) where, as it is possible to see from the Figure 2(c), the two lines are quite close.



(a) Precision



(b) Recall



(c) Average Precision

Figure 2: Precision, Recall, and Average Precision for eight rounds of relevance feedback. The proposed methods are labelled $NN-E(N+NM)$ and $NN-E(NM)$, respectively.

5 CONCLUSIONS

In this paper we proposed two exploration approaches based on the query reformulation and the Nearest Neighbor paradigms. The main goal attained by the proposed mechanisms is to be able to explore the feature space around the images labeled as being relevant by the user, thus following the user's explorative behavior during a session of image search into a visual database. Reported results show that the proposed approach succeeded in showing the user a greater number of new relevant images during the first few iterations, in comparison with other techniques either based on "classification", or "exploration" approaches. We believe that the effectiveness of CBIR systems strongly depends on its adaptive behavior in response to relevance feedback. Accordingly, further experiments aimed at testing the system with real users are needed in order to assess the effectiveness of the proposed approach compared to other state-of-the-art relevance feedback mechanisms.

ACKNOWLEDGEMENTS

We would like to thank Dr. M. Arevalillo-Herrez and Dr. F. J. Ferri for providing us with the details of the results reported in (Arevalillo- Herra ez and Ferri, 2013). This work has been partly supported by the Regional Administration of Sardinia (RAS), Italy, within the project Advanced and secure sharing of multimedia data over social networks in the future Internet (CRP 17555, L.R. 7/2007, Bando 2009).

REFERENCES

Arevalillo-Herráez, M. and Ferri, F. J. (2010). Interactive image retrieval using smoothed nearest neighbor estimates. In Hancock, E. R., Wilson, R. C., Windeatt, T., Ulusoy, I., and Escolano, F., editors, *SSPR/SPR*, *LNCS* 6218, pages 708–717. Springer.

Arevalillo-Herráez, M. and Ferri, F. J. (2013). An improved distance-based relevance feedback strategy for image retrieval. *Image Vision Comput.*, 31(10):704–713.

Arevalillo-Herráez, M., Ferri, F. J., and Domingo, J. (2010). A naive relevance feedback model for content-based image retrieval using multiple similarity measures. *Pattern Recognition*, 43(3):619–629.

Boiman, O., Shechtman, E., and Irani, M. (2008). In defense of nearest-neighbor based image classification. In *CVPR 2008*. IEEE Computer Society.

Chen, Y., Zhou, X. S., and Huang, T. (2001). One-class svm for learning in image retrieval. In *ICIP 2001*, volume 1, pages 34 –37.

Cohn, D. A., Atlas, L. E., and Ladner, R. E. (1994). Improving generalization with active learning. *Machine Learning*, 15(2):201–221.

Datta, R., Joshi, D., Li, J., and Wang, J. Z. (2008). Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys*, 40(2):1–60.

Ding, C. H. Q. and He, X. (2004). K-nearest-neighbor consistency in data clustering: incorporating local information into global optimization. In Haddad, H., Omicini, A., Wainwright, R. L., and Liebrock, L. M., editors, *SAC 2004*, pages 584–589. ACM.

García, S., Herrera, F., and Shawe-taylor, J. An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons. *Journal of Machine Learning Research*, pages 2677–2694.

Giacinto, G. (2007). A nearest-neighbor approach to relevance feedback in content based image retrieval. In *CIVR '07: Proceedings of the 6th ACM international conference on Image and video retrieval*, pages 456–463, New York, NY, USA. ACM.

Giacinto, G. and Roli, F. (2004a). Bayesian relevance feedback for content-based image retrieval. *Pattern Recognition*, 37(7):1499–1508.

Giacinto, G. and Roli, F. (2004b). Nearest-prototype relevance feedback for content based image retrieval. In *ICPR 2004*, pages 989–992.

Hoi, S. C. H., Jin, R., Zhu, J., and Lyu, M. R. (2009). Semisupervised svm batch mode active learning with applications to image retrieval. *ACM Trans. Inf. Syst.*, 27(3):16:1–16:29.

Laaksonen, J., Koskela, M., and Oja, E. (2002). PicSOM-self-organizing image retrieval with MPEG-7 content descriptors. *IEEE Transactions on Neural Networks*, 13(4):841–853.

Lew, M. S., Sebe, N., Djeraba, C., and Jain, R. (2006). Content-based multimedia information retrieval: State of the art and challenges. *ACM Trans. Multimedia Comput. Commun. Appl.*, 2(1):1–19.

Lindenbaum, M., Markovitch, S., and Rusakov, D. (2004). Selective sampling for nearest neighbor classifiers. *Machine Learning*, 54(2):125–152.

Pavlidis, T. (2008). Limitations of content-based image retrieval. Technical report, Stony Brook University.

Piras, L., Giacinto, G., and Paredes, R. (2012). Enhancing image retrieval by an exploration-exploitation approach. In Perner, P., editor, *MLDM 2012*, *LNCS* 7376, pages 355–365. Springer.

Rao, Y., Mundur, P., and Yesha, Y. (2006). Fuzzy svm ensembles for relevance feedback in image retrieval. In Sundaram, H., Naphade, M. R., Smith, J. R., and Rui, Y., editors, *CIVR 2006*, *LNCS* 4071, pages 350–359. Springer.

Rocchio, J. J. (1971). Relevance feedback in information retrieval. In Salton, G., editor, *The SMART Retrieval System - Experiments in Automatic Document Processing*, pages 313–323. Prentice Hall, Englewood, Cliffs, New Jersey.

Rui, Y., Huang, T. S., and Mehrotra, S. (1997). Content-Based image retrieval with relevance feedback in MARS. In *International Conference on Image Processing Proceedings*, pages 815–818.

Sivic, J. and Zisserman, A. (2008). Efficient visual search for objects in videos. *Proceedings of the IEEE*, 96(4):548 –566.

Swain, M. J. and Ballard, D. H. (1991). Color indexing. *International Journal of Computer Vision*, 7(1):11–32.

Thomee, B. and Lew, M. S. (2012). Interactive search in image retrieval: a survey. *IJMIR*, 1(1):71–86.

Tong, S. and Chang, E. Y. (2001). Support vector machine active learning for image retrieval. In *ACM Multimedia*, pages 107–118.

Tronci, R., Murgia, G., Pili, M., Piras, L., and Giacinto, G. (2013). Imagehunter: A novel tool for relevance feedback in content based image retrieval. In Lai, C., Semeraro, G., and Vargiu, E., editors, *New Challenges in Distributed Information Filtering and Retrieval*, pages 53–70. Springer Berlin Heidelberg.

Wang, M., Yang, K., Hua, X.-S., and Zhang, H. (2010). Towards a relevant and diverse search of social images. *IEEE Transactions on Multimedia*, 12(8):829–842.

Zhang, L., Lin, F., and Zhang, B. (2001). Support vector machine learning for image retrieval. In *ICIP 2001*, pages 721–724.

Zhou, X. S. and Huang, T. S. (2003). Relevance feedback in image retrieval: A comprehensive review. *Multimedia Syst.*, 8(6):536–544.