Image retrieval with SVM active learning embedding Euclidean search

Lei Wang
Nanyang Technological University, leiw@uow.edu.au

Kap Luk Chan
Nanyang Technological University

Yap Peng Tan
Nanyang Technological University

Publication Details
Image retrieval with SVM active learning embedding Euclidean search

Abstract
Image retrieval with relevance feedback suffers from the small sample problem. Recently, SVM active learning has been proposed to tackle this problem, showing promising results. However, a small but sufficient number of initially labelled samples are still required to ensure the subsequent active learning efficient and good retrieval performance. In the existing method, the user is asked to label more images before active learning starts. In this paper, a method of embedding Euclidean search into SVM active learning is proposed. With the help of Euclidean search, not only the adverse effect on retrieval performance due to lack of initially labelled samples can be reduced, the retrieval performance can be further enhanced when there is sufficient number of initially labelled samples. Experimental results demonstrate the improvement by the proposed method, especially when the number of initially labelled samples is small.

Keywords
euclidean, svm, image, active, embedding, learning, retrieval, search

Disciplines
Engineering | Science and Technology Studies

Publication Details

This conference paper is available at Research Online: http://ro.uow.edu.au/eispapers/783
Image retrieval with relevance feedback suffers from the small sample problem. Recently, SVM active learning has been proposed to tackle this problem, showing promising results. However, a small but sufficient number of initially labelled samples are still required to ensure the subsequent active learning efficient and good retrieval performance. In the existing method, the user is asked to label more images before active learning starts. In this paper, a method of embedding Euclidean search into SVM active learning is proposed. With the help of Euclidean search, not only the adverse effect on retrieval performance due to lack of initially labelled samples can be reduced, the retrieval performance can be further enhanced when there is sufficient number of initially labelled samples. Experimental results demonstrate the improvement by the proposed method, especially when the number of initially labelled samples is small.

1. INTRODUCTION

In the past few years, learning from user’s feedback was considered as an attempt to bridge the semantic gap in Content-Based Image Retrieval (CBIR) [4, 5, 9]. From relevance feedback, the user subjectively labels the retrieved images as “positive” or “negative”, and these labelled images are used to train a classifier that performs a bi-class classification on the image database. Those images with the larger scores or posterior probability values with respect to the positive image class are retrieved. However, the learning performance is often constrained due to the availability of a very small number of samples because a user is unwilling to label too many retrieved images for relevance feedback.

Recently, Support Vector Machines (SVM) based active learning has been proposed to tackle this problem by maximizing the learning efficiency while minimizing the number of labelled image samples [8]. In this approach, retrieval becomes a querying-learning cycle. In each cycle, the active learner queries the user about the labels of the unlabelled images that are carefully selected by itself. After the user labels these images, they are added into the current labelled data set, and an SVM classifier is retrained. Based on the retrained classifier, a new set of unlabelled images will be selected again, and a new cycle starts. When the user does not respond to more queries from the learner, the database images are sorted according to the distances to the learned decision boundary given by the final SVM classifier that separates the positive image class from the negative one. The images corresponding to larger positive distances are retrieved. This approach has shown promising retrieval performance. However, a small but sufficient number of initially labelled images are required to ensure subsequent learning efficient. Otherwise, the decision boundary given by the SVM classifier cannot tightly enclose the positive image class, and the retrieval performance will be poor. In [8], this problem is tackled by asking the user to label twenty more randomly selected images in the first round of relevance feedback and then starting active learning afterwards. Though by labelling more samples can get around this problem, it increases the burden on the user. On the other hand, heuristically labelling a fixed number of images is not in the spirit of active learning because this number may be different for different retrieval tasks and the image databases involved, and may not be known in advance.

In this paper, the above problem is solved by embedding Euclidean search into SVM active learning instead. In our method, an Euclidean search in the image database is performed by centering at a carefully selected positive image sample, and each database image is attached a distance from this positive image sample. When sorting the database images, both the distances from the learned decision boundary and those from Euclidean search are considered. This can effectively reduce the adverse effect of inaccurate decision boundary, and influence the retrieval to be from the region with high density of positive images. Experimental results demonstrate the overall improvement by the proposed method, especially when the number of initially labelled samples is small.

2. RELATED WORK

In [8], Tong et al. proposed SVM active learning and gave a clear analysis by using version space [2]. Given a set of
labelled samples, a version space will be formed. It can be viewed as a set including all hyperplanes that can correctly classify this labelled data set, and one hyperplane in it will be estimated as the current separating hyperplane. With the arrival of a new labelled sample, the size of version space will decrease if some hyperplanes in it cannot correctly classify this new labelled sample. Otherwise, it remains as before. Furthermore, the optimal separating hyperplane (OSH) will always lie in the series of version spaces because it can always correctly classify all the labelled data (Note that it is assumed that the labelled data are linearly separable in the kernel-induced feature space). Hence, when the size of version space becomes small enough, the optimal separating hyperplane can be well captured. Based on this idea, Tong et al. proposed to select the unlabelled sample that can reduce the size of the current version space as much as possible. Following this, the Simple method is proposed in which the unlabelled sample closest to the current separating hyperplane is selected. Compared to other selection methods, this method is much less computational intensive, and it may be the most suitable method for image retrieval where multiple samples are selected in each learning cycle. However, it was also found that, when only few initially labelled image samples are available, the poor learning of the SVM results in a poor retrieval.

In our view, this can be explained as follows. A version space with a large size will be formed in the small sample case with high probability. As a result, the estimated separating hyperplane may significantly deviate from the optimal one, OSH. In this case, the samples selected by the Simple method will not be those that should be selected, and it makes the subsequent learning inefficient and the size of version space cannot be quickly reduced. Hence, the next estimated separating hyperplane may still not reach the optimal one, even if several rounds of learning have been performed. Consequently, the decision boundary in the input space cannot tightly enclose the positive image class, and a poor retrieval performance will result. In [8], to avoid the poor performance, twenty more randomly selected images are labelled at the first round of relevance feedback before active learning starts.

3. SVM ACTIVE LEARNING EMBEDDING EUCLIDEAN SEARCH

3.1. The help of Euclidean Search

Euclidean search is a method commonly used in information retrieval. Given a query, Euclidean search calculates the Euclidean distances between the feature vectors corresponding the database images and the query, and the images with smaller distances are retrieved. The search scope of Euclidean search is spherical, and its effectiveness depends on the assumption that the positive images have a sphere-shape distribution around the query. Though this assumption may not be strictly true in reality, the Euclidean search seems to be a good heuristic when the labelled samples are so scarce that the estimation for the data distribution may be a bad disturbance rather than a helpful information.

Combining Euclidean search with SVM for image retrieval can also be found in [1]. Their work assumed that, in image retrieval, an accurate boundary enclosing the positive images could be found by SVM, and they used this boundary to restrict the Euclidean search to improve the search efficiency. However, this can be seldom achieved in practice because of the small sample problem. There are three differences between our work and that in [1]. Firstly, our concern is to improve the performance of SVM active learning instead of the Euclidean search. Secondly, and more importantly, our work focuses on the small sample problem actually happens in image retrieval, and tries to reduce the adverse effect of inaccurate boundary in SVM active learning caused by the small sample by using the Euclidean search. Finally, it will be shown later that our combination method can achieve better retrieval performance than the one proposed in [1].

---

**Fig. 1. Illustration of the help of Euclidean search**

Figure 1 illustrates the idea in this paper. The symbols “x” and “o” denote the positive and negative “image” samples, respectively, and the symbols in the bold face represent the labelled samples. Figure 1(a-1) shows the usual case at the beginning of retrieval where only one positive and one negative labelled images are available. In this case, the boundary (indicated by “OSH”) can be viewed as a hyperplane separating the two samples evenly. Calculating the distances from this hyperplane will make many negative im-
age samples be retrieved. In Figure 1(a-2), Euclidean search is incorporated, which is represented by a semi-transparent circular area centered at the labelled positive sample, \( x \). By sorting the database images with both the distances to the hyperplane and those from Euclidean search, the problem shown in Figure 1(a-1) can be avoided. Figure 1(b-1) shows the case in which more labelled image samples are available but the total number is still not sufficient. In this case, an inaccurate boundary can often be obtained. It can be seen that the region having larger positive distances to the boundary is in fact occupied by the negative images. In Figure 1(b-2), because multiple labelled positive images are available, the center of the Euclidean search is selected to be the labelled positive image corresponding to the largest positive distance to the current boundary. This selection borrows the estimation given by SVM active learning for the region of positive images. It avoids performing the Euclidean search in a region with low density of positive images, such as a region around a labelled sample near the boundary. However, the work in [1] treats the Euclidean search as the main body around a labelled sample near the boundary. However, the current boundary. This selection borrows the estimation given by SVM active learning for the region of positive images.

3.2. The Similarity Metric

Let \( d^i_{\text{SAL}} \) denote the distance of the database image \( i \) from the boundary given by SVM active learning, and

\[
d^i_{\text{SAL}} = \frac{f(x_i)}{||w||} = \left(\frac{w^T x_i + b}{||w||}\right)
\]

(1)

where \( w \) and \( b \) denote the normal vector and the bias of the separating hyperplane, respectively, and \( x_i \) is the feature vector representing the database image \( i \). Let \( d^i_{\text{EU}} \) denote the distance obtained by the Euclidean search, and

\[
d^i_{\text{EU}} = ||x_i - x_c||
\]

(2)

where \( x_c \) denotes the center of the Euclidean search, and

\[
x_c = \arg \max_{x_j \in D} d_j^{i_{\text{SAL}}}
\]

(3)

where \( D \) denotes the set of the labelled image samples. Differently, in [1], the search region is always around the user submitted query at the beginning of retrieval.

The similarity metric of our method for the \( i \)-th database image, \( d^i_{\text{SAL+EU}} \), can be defined as follows.

\[
d^i_{\text{SAL+EU}} = \begin{cases} 
  d^i_{\text{EU}} & \text{if } d^i_{\text{SAL}} \geq 0 \\
  \max_{d^j_{\text{SAL}} \geq 0} (d^j_{\text{EU}}) + |d^i_{\text{SAL}}| & \text{if } d^i_{\text{SAL}} < 0 
\end{cases}
\]

(4)

where \( \max_{d^j_{\text{SAL}} \geq 0} (d^j_{\text{EU}}) \) denotes the largest Euclidean distance among all the database images corresponding to the non-negative distances from the separating hyperplane.

The retrieval procedure is described as follows. (1) A retrieval is launched with one positive and one negative image examples provided by a user; (2) An SVM classifier is trained on the current labelled data set; (3) Following the Simple method, a given number of unlabelled images closest to the current separating hyperplane are selected to query the user; (4) After the user labels these images, add them into the labelled data set, and retrain the SVM classifier. Then go to step (3) if the user will carry out more feedbacks; (5) The SVM classifier is used to classify all database images, and an Euclidean search centered at the selected positive sample is also performed. Calculate the distance according to equation (4), and the database images corresponding to the smaller distances are shown as the retrieved results.

4. EXPERIMENTAL RESULTS

The experiments aim to evaluate the effectiveness of our method for the case of small sample. In the following, the methods of SVM active learning and Euclidean search are denoted as SAL and EU, respectively. The method proposed in this paper is denoted as SAL+EU(Proposed) while that in [1] is denoted as SAL+EU([1]).

An artificial database and a real color image database are used in the experiments. The artificial database, used in [10], includes seven 2D Gaussian-distributed classes. Each class has 100 samples representing 100 “images”. The real color image database includes 600 general color images composed from VisTex of MIT and Corel Stock Photos. Six image classes are defined based on high-level semantics (i.e. defined by a group of human observers), and each class includes 100 image samples. A perceptually uniform color space, CIE-\textit{Lab}, is used to represent general color images. Based on this color space, a feature vector of color moments [7] is defined for each image. The two pre-classified databases provide the ground truth for evaluation.

In the experiments, the number of selected unlabelled images in each learning cycle, \( k \), is set to 20 to be consistent with [8]. SVM\textit{light} [3] is used to perform SVM active learning. Gaussian kernel, \( k(x, y) = \exp(-\frac{||x-y||^2}{2\sigma^2}) \), is used, where the width, \( \sigma \), is set as the average of the Euclidean distances among the labelled samples.

Figure 2 shows the \textit{Precision} for retrieving top twenty images under different numbers of the initially labelled image samples (horizontal axis) available at the beginning of retrieval. The sub-figure(a) corresponds to the result after the first learning cycle. It can be seen that SAL+EU(Proposed) achieves the highest \textit{Precision} while SAL gives a lower one. The performance of EU is close to that of SAL+EU(Proposed). These results show that in this case SVM active learning provides much less helpful information for retrieval than the Euclidean search, and the Euclidean search is very helpful.
image database. Though the help of Euclidean search becomes less obvious because the real color image data lie in a higher dimensional space and may have a complex distribution, it can still be seen that SAL+EU(Proposed) achieves the best retrieval performance.

5. CONCLUSION REMARKS

We present a method that embeds Euclidean search into SVM active learning to counter the adverse effect of learning in the presence of the small sample problem. Our method reduces the burden on the user by avoiding requiring the user to label more images before active learning starts. With the help of Euclidean search, not only the adverse effect on retrieval performance due to lack of initially labelled samples can be reduced, the retrieval performance can be further enhanced when there is sufficient number of initially labelled samples. The experimental results demonstrate the improvement by our method over the existing methods, especially when the number of initially labelled samples is small.

6. REFERENCES


