HAL open science

Active Boosting for interactive object retrieval.

Alexis Lechervy, Philippe-Henri Gosselin, Frédéric Precioso

To cite this version:

Alexis Lechervy, Philippe-Henri Gosselin, Frédéric Precioso. Active Boosting for interactive object retrieval.. International Conference on Pattern Recognition, Aug 2010, Turkey. pp.1. hal-00520294

HAL Id: hal-00520294

https://hal.science/hal-00520294v1

Submitted on 22 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Active Boosting for interactive object retrieval

Alexis Lechervy, Philippe-Henri Gosselin and Frédéric Precioso
ETIS, CNRS, ENSEA, Univ Cergy-Pontoise, F-95000 Cergy-Pontoise
{*alexis.lechervy,philippe-henri.gosselin,frederic.precioso*}*@ensea.fr*

Abstract

This paper presents a new algorithm based on boosting for interactive object retrieval in images. Recent works propose "online boosting" algorithms where weak classifier sets are iteratively trained from data. These algorithms are proposed for visual tracking in videos, and are not well adapted to "online boosting" for interactive retrieval. We propose in this paper to iteratively build weak classifiers from images, labeled as positive by the user during a retrieval session. A novel active learning strategy for the selection of images for user annotation is also proposed. This strategy is used to enhance the strong classifier resulting from "boosting" process, but also to build new weak classifiers. Experiments have been carried out on a generalist database in order to compare the proposed method to a SVM based reference approach.

1 Introduction

Recent years have seen exponential growth of image and multimedia data. The machine learning framework has shown to be powerful in this situation. The two most frequent forms are either supervised or unsupervised learning models. In model with supervision we see different approaches, we can annotate sets of images sharing common visual concepts or sets of equivalence constraints between pairs of images. In this paper we use interactive learning framework with binary annotations. We start with few labeled images and at each feedback loop the user enriches the training set.

Recent machine learning techniques have demonstrated their capability for identifying image categories from image features, like SVM or Boosting. On the one hand, SVM classifiers can be successfully applied to all datasets, thanks to kernel trick [11]. However, the performance of classification depends highly on kernel design. On the other hand, the boosting classifiers have good performance [13] and they can be tuned more easily by using weak classifiers based on simple features. In this paper we focus on boosting context.

In interactive search context, the users may be more interested in the most relevant items. For example, in web browsing search, the users often read only the first result page. Traditional approaches in this case are ranking methods. A major approach in boosting to learn good ranking functions is proposed by Freund et al. : RankBoost [6]. Many learning-to-rank algorithms have been proposed after in interactive learning context [4].

The learning process starts with few examples (for instance: only one positive label). Some boosting methods have been proposed to learn with a small dataset [3, 15]. One of the simplest but most popular method in interactive learning is the so called *relevance feedback*. Active learning can be used in an interactive loop to annotate samples. In boosting community, some methods have been proposed for active learning [2, 10, 9]. More recently, Vijayanarasimhan et al. [14] propose an active learning approach to select the best learning sample set depending on the precision of annotations.

Finally, computational efficiency is important in interactive learning context. Grabner and Bischof propose an online boosting method to reduce learning runtime in tracking context [8, 12]. They introduce dynamic pools of weak classifiers, whose size is much smaller than the complete set of all weak classifiers. In this paper, we also use this idea for computation saving, but in a different manner, in order to deal with the interactive context.

This paper is organized as follows: Section 2 motivates the choices of the proposed active learning framework and introduces our novel method. Section 3 present our weak classifiers. In Section 4, we report the experimental results and demonstrate the effectiveness of our method.

2 Proposal method

In this paper we use a RankBoost algorithm with a dynamic pool. This pool is built from active process. In our method we extract a weak classifier from the features of the images. The whole learning scheme is described in Algorithm.1.

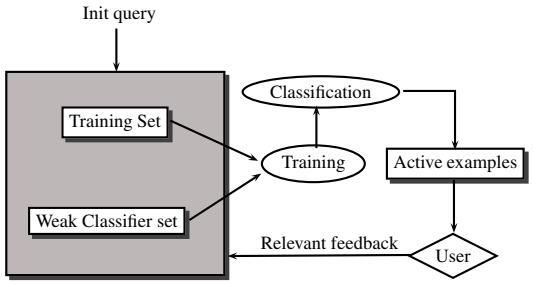


Figure 1. Interactive boosting framework

2.1 Learning Algorithm

This paper introduces a new boosting algorithm for object retrieval in interactive context. Fig. 1 present the stages of our method. Retrieval session starts with an empty set W_0 of weak classifiers. The user initializes a retrieval session with a query composed of a set X_0 of positive and negative labels. We propose to construct the set of weak classifiers W_j at relevance feedback iteration j, using the positives examples. For all positive examples, we add the weak classifiers built from the visual characteristics of the positive images (more details in section 3):

$$\begin{cases} W_0 & = \varnothing \\ W_{j+1} & = W_j \ \cup \mathcal{H}_{x_i} \end{cases} \tag{1}$$

With $\mathcal{H}_{x_i} = \{h_{k,x_i}\}_k$ the set of weak classifiers built using image x_i visual features. The algorithm trains the strong classifier using a RankBoost algorithm. This algorithm uses the previous annotations X_j and the set of weak classifier W_j. The strong classifier is used to rank the database. The system selects new images to annotate through an active process. The user then decides whether to stop or not.

2.2 Active Learning Strategy

We use an active method to build a weak classifier set. At each step, we rank the images looking for new images to annotate with the best classifier on images already annotated. We search the images in $x_{i^*} \notin X_p \cup X_n$ with:

$$i^* = \arg\max_i \left(\max_{h \in \mathcal{H}_{x_i}} r_0(h) \right) \tag{3}$$

If the weak classifier set is chosen randomly, the error computed on small datasets is unreliable and the classifier selection is not correct. Our method proposes for annotations the images with the classifier which maximizes r_0 (cf. Eq. 2 in algorithm description). This active process is illustrated on Fig.3. The active selection is made among the top 100 ranking images. In our experiments, this approximation gives similar result to a full active learning selection, but saves a lot of computational time.

Algorithm 1 after j step

Require: Given example image $x_i \in X_p \cup X_n$

Require: a base of Weak classifier W_{j-1} (starting with $\varnothing$)

1: Initialize weights

$$\nu_0(x_i) = \begin{cases} \dfrac{1}{|X_p|} \text{ if } x_i \text{ is positive } (x_i \in X_p) \\ \dfrac{1}{|X_n|} \text{ if } x_i \text{ is negative } (x_i \in X_n) \end{cases}$$

2: **if** $x_i \in X_p$ **then** $W_j = W_{j-1} \cup \mathcal{H}_{x_i}$

3: **else** $W_j = W_{j-1}$

4: **end if**

5: **for** $t = 0$ to T **do**

6: Choose $h_t \in W_j$ such that:

$$r_t(h) = \sum_{x_p \in X_p} \nu_t(x_p) h(x_p) - \sum_{x_n \in X_n} \nu_t(x_n) h(x_n) \tag{2}$$

7: Compute α_t:

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 + r_t}{1 - r_t} \right)$$

8: Update ν_t:

$$\nu_{t+1}(x_i) = \begin{cases} \dfrac{\nu_t(x_i) \mathrm{e}^{-\alpha_t h_t(x_i)}}{\sum_{x_p \in X_p} \nu_t(x_p) \mathrm{e}^{-\alpha_t h_t(x_p)}} \text{ if } x_i \in X_p \\ \\ \dfrac{\nu_t(x_i) \mathrm{e}^{\alpha_t h_t(x_i)}}{\sum_{x_n \in X_n} \nu_t(x_n) \mathrm{e}^{\alpha_t h_t(x_n)}} \text{ if } x_i \in X_n \end{cases}$$

9: **end for**

10: **return** The final strong classifier is:

$$H(x) = \sum_{t=0}^{T-1} \alpha_t h_t(x)$$

3 Weak Classifiers

For each image x_i, we define a set $\mathcal{H}_{x_i}$ of weak classifiers. We cut each image in 9 areas z_l according to a 3x3 grid. An histogram (color, texture,...) is computed in each area. We also use the regions $\rho_m = \bigcup_l z_l$. In this case, the histogram is the sum of each area histogram. For example, in Fig. 2, the areas 1,2,3,5,6 form the region $\rho_{12356} = z_1 \cup z_2 \cup z_3 \cup z_5 \cup z_6$.

3.1 The classifier of type 1

The type 1 classifier uses a region ρ_k and a reference image x_i. The classifier is defined by:

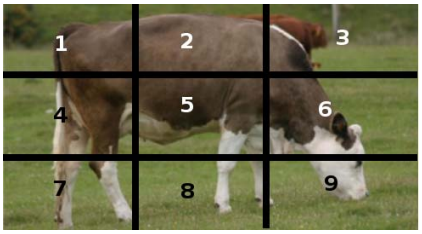


Figure 2. Examples of regions for a weak classifier

$$h^1_{k,x_i}(x_j) = 1 - d(histo_{\rho_k}(x_i), histo_{\rho_k}(x_j)) \quad (4)$$

It compares the histogram in region ρ_k in reference image x_i to the one in test image x_j. We use χ^1 distance for comparison:

$$d(g,x) = \frac{1}{M}\sum_{p=1}^{M}\frac{|g_p - x_p|}{g_p + x_p} \quad (5)$$

M is the size on histogram g_p and x_p is the bin of histograms. $\mathcal{H}^1_{x_i} = \{h^1_{k,x_i}\}_k$ is the set of all weak classifiers of type 1, using image x_i.

3.2 The classifiers of type 2

The classifier of type 2 uses an area in the reference image x_i. It is defined by:

$$h^2_{k,x_i}(x_j) = 1 - \min_{z_{k'} \in Z} d(histo_{z_k}(x_i), histo_{z_{k'}}(x_j)) \quad (6)$$

The comparison distance is also a χ^1 distance (Eq.5). $\mathcal{H}^2_{x_i} = \{h^2_{k,x_i}\}_k$ is set of all weak classifier of type 2.

4 Experiments and Results

4.1 Experimental procedure

We evaluate the performance of our algorithm on Visual Object Category (Voc)2006 dataset. This database contains 5304 images provided by Microsoft Research Cambridge and flickr. Voc2006 database contains ten categories (cat, car, motorbike, sheep ...). All images can belong to several categories. There are two distinct sets, one for training and one for test with 9507 annotations. To facilitate evaluation and comparison of search performance in this active learning context, we use a specific protocol which cannot be compared with voc2006 campaign protocol [5]. The simulation starts with a random positive image. At each step, five images provided by the active strategy are added. We repeat 10 times this relevance feedback loop. At the end, we get 51 labels and compute the results. This protocol is repeated 100 times for each category. We compare our method with a SVM based approach [7] with same 64-bin histograms: 32 for color CIE L*a*b, 32 for quaternion wavelets [1] (Qw) and same distance (χ^1).

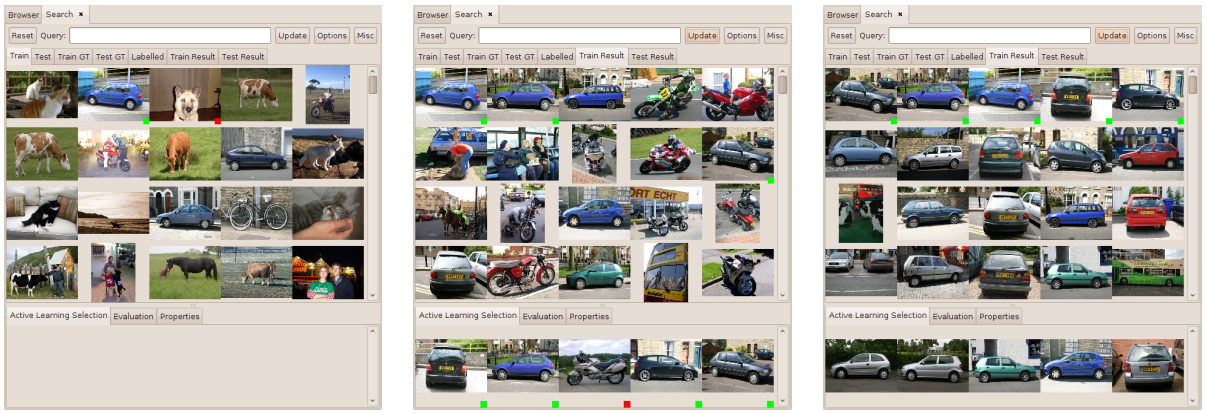
4.2 The experimental results and discussions

Fig.4 shows the result for with or without active strategy for RankBoost and SVM. Let us point out that scales on axis are different on the three figures. Without active learning, we get similar results whatever is the classifier (yellow and green curves). On the one hand, SVM computes optimal hyperplanes, but its performance is bounded by global histograms. On the other hand, boosting hyperplanes are perhaps less optimal, but its simplicity allows the use of more complex features/weak classifiers.

The active method in both RankBoost (blue) and SVM (red) performs better, but the evolution depends on the category. For the "car" category (*cf.* Fig. 4(a)), Average Precision increases a lot during the first feedback steps, but results are quite similar for 40 labels and 50 labels. Initially the cars are easily found in the foreground, however the algorithm does not find car in background. The feature are not suited for retrieving very small objects, and hence the active method has a small impact. Per contra, for the sheep category (*cf.* Fig. 4(b)), the active method improves significantly the ranking results. When considering all categories (*cf.* Fig. 4(c)), active RankBoost is more efficient than active SVM, certainly because we use richer features. Our method allows to easily and efficiently combine many features type and does not require laborious design such as kernel function building.

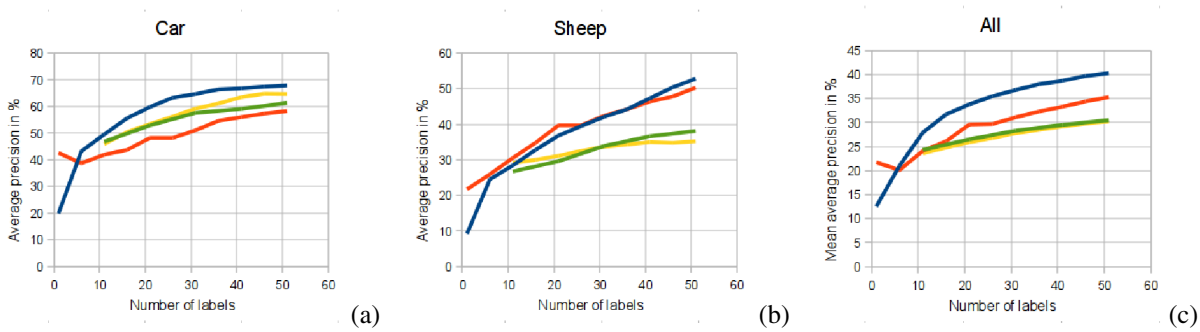
5 Conclusion

This paper propose a new active learning method based on RankBoost for interactive search. The proposed method build a set of weak classifiers and use an active method to select the best samples. This approach is based on an association weak classifiers/training samples. A boosting algorithm selects and weights the most interesting classifiers among those extracted from positive images. Finally, an innovative active method is proposed to select the images which, when annotated, increase the improve the database ranking a number of weak classifiers. The experimental results demonstrated, for a given small training sample set, that the proposed method can provide better classification performance than an active SVM method using global histograms. With this method, we showed that the boosting framework can deal with the specificity of interactive image retrieval: for instance the method can provide relevant results with only 2 examples. Furthermore, design of classification tools for complex features is easy in comparison to other frameworks such as kernel functions. The proposed method can effectively reduce the labeled required while maintaining a satisfactory and real time classification performance.



(a) (b) (c)

Figure 3. Illustration of active learning process: (a) Initial labels (green square= positive, red square= negative), (b) First ranking and Active selection, (c) Second ranking



(a) (b) (c)

Figure 4. Average precision in % (VOC2006). Blue is our algorithm, red is active SVM, yellow is RankBoost and green SVM without active learning

6. Acknowledgement

This work was supported by a grant from DGA.

References

[1] W. L. Chan, H. Choi, and R. Baraniuk. Quaternion wavelets for image analysis and processing. In *IEEE ICIP*, volume 5, pages 3057–3060, October 2004.

[2] B. Collins, J. Deng, K. Li, and L. Fei-Fei. Towards scalable dataset construction: An active learning approach. In *IEEE ECCV*, 2008.

[3] L. Diao, K. Hu, Y. Lu, and C. Shi. A method to boost support vector machines. In *PAKDD '02*, pages 463–468, London, UK, 2002. Springer-Verlag.

[4] P. Donmez and J. G. Carbonell. Optimizing estimated loss reduction for active sampling in rank learning. In *ICML '08*, New York, NY, USA, 2008.

[5] M. Everingham, A. Zisserman, C. K. I. Williams, and L. Van Gool. The PASCAL Visual Object Classes Challenge 2006 (VOC2006) Results. http://www.pascal-network.org/challenges/VOC/voc2006/results.pdf.

[6] Y. Freund, R. Iyer, R. E. Schapire, Y. Singer, and G. Dietterich. An efficient boosting algorithm for combining preferences. In *JMLR*, pages 170–178, 2003.

[7] P. Gosselin, M. Cord, and S. Philipp-Foliguet. Combining visual dictionary, kernel-based similarity and learning strategy for image category retrieval. *CVIU*, 110(3):403–417, 2008.

[8] H. Grabner and H. Bischof. On-line boosting and vision. In *CVPR*, pages 260–267, Washington, DC, USA, 2006. IEEE Computer Society.

[9] X. Li, L. Wang, and E. Sung. Improving adaboost for classification on small training sample sets with active learning. In *ACCV*, Korea, 2004.

[10] Y. Lu, Q. Tian, and T. Huan. Interactive boosting for image classification. In *MCAM*, pages 315–324, 2007.

[11] J. Shawe-Taylor and N. Cristianini. *Kernel methods for Pattern Analysis*. Cambridge University Press, 2004.

[12] N. Thuy, B. D. Nguyen, and B. Horst. Efficient boosting-based active learning for specific object detection problems. In *CVISP 2008*, 2008.

[13] K. Tieu and P. Viola. Boosting image retrieval. *CVPR*, 1:1228, 2000.

[14] S. Vijayanarasimhan and K. Grauman. Whats it going to cost you?: Predicting effort vs. informativeness for multi-label image annotations. In *IEEE CVPR*, 2009.

[15] L. Wolf and I. Martin. Robust boosting for learning from few examples. In *CVPR05*, pages I: 359–364, 2005.