Chapter VI
Active Learning for Relevance Feedback in Image Retrieval

Jian Cheng
National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China

Kongqiao Wang
Nokia Research Center, Beijing, China

Hanqing Lu
National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China

ABSTRACT

Relevance feedback is an effective approach to boost the performance of image retrieval. Labeling data is indispensable for relevance feedback, but it is also very tedious and time-consuming. How to alleviate users’ burden of labeling has been a crucial problem in relevance feedback. In recent years, active learning approaches have attracted more and more attention, such as query learning, selective sampling, multi-view learning, and so forth. The well-known examples include Co-training, Co-testing, SVM active, etc. In this literature, the authors will introduce some representative active learning methods in relevance feedback. Especially, they will present a new active learning algorithm based on multi-view learning, named Co-SVM. In Co-SVM algorithm, color and texture are naturally considered as sufficient and uncorrelated views of an image. SVM classifier is learned in color and texture feature subspaces, respectively. Then the two classifiers are used to classify the unlabeled data. These unlabeled samples that disagree in the two classifiers are chose to label. The extensive experiments show that the proposed algorithm is beneficial to image retrieval.
INTRODUCTION

Content-Based Image Retrieval (CBIR) has been one of the most active research topics in computer vision and pattern recognition fields since 1990’s (Rui, Huang, & Chang, 1999; Smeluders, Worring, Santini, & Gupta, 2000; Datta, Li, & Wang, 2005). Most of existing CBIR systems adopted low-level features (color, texture, shape, etc) to represent images. However, it is inadequate to describe the semantic concepts with the low-level features of images, which is named semantic gap. The concept of semantic gap has been extensively used in the CBIR research community to express the discrepancy between the description of the low-level features extracted from the images and the semantic understanding of human.

To narrow the semantic gap, a straightforward way is to take human in the loop. As one of important ways of human-in-the-loop, Relevance Feedback (RF) is a query modification technique, which was initially developed in document retrieval and then introduced into CBIR during mid 1990’s (Picard, Minka, & Szummer, 1996; Rui, Huang, Ortega, & Mehrotra, 1998). Relevance feedback attempts to capture the user’s preference through iterative feedback and query refinement. In each round, the user is requested to provide feedbacks regarding the relevance or irrelevance of the current retrieval results. Then the classifier will be refined based on the feedback results. The pool of unlabeled images is classified as relevance or irrelevance by the learned classifier, and the relevant images are ranked and returned to user for next round labeling. With the interactive labeling and learning procedures, systems can learn user’s preferences and improve the performance of image retrieval.

Many relevance feedback algorithms have been proposed for image retrieval in past years (Huang, & Zhou, 2001; Zhou, & Huang, 2003). The early work was mainly inspired by term-weighting and relevance feedback in document retrieval (Rocchio, 1971). Rui and Huang (1998) introduced the query refinement algorithm based on term-frequency and inverse-document-frequency in text retrieval into CBIR. Picard, et al. (1996) grouped the images or regions into hierarchical trees whose nodes were constructed through single-link clustering, and then weighted on grouping. These methods fall into heuristic-based formulation with empirical parameter adjustment. Later, there are more works focusing on learning-based strategy and many classic machine learning techniques are applied. Tieu and Viola (2000) assumed that an image was generated by a sparse set of visual cause and that images which were visually similar share causes. They proposed a mechanism for computing a very large number of highly selective features which captured some aspects of this causal structure, then used Boosting to learn a classification function which only relied on 20 features. In (Vasconcelos, & Lippman, 2000), Gaussian mixture model on DCT coefficient was used as image representation, then Bayesian inference was applied for image region matching and learning. Hong, et al. (2000) treated relevance feedback as a binary classification problem and incorporated Support Vector Machines (SVM) into the classification process. However, an inevitable issue in performing relevance feedback is the small sample size. The fact makes many learning methods inefficient, such as Bayesian, boosting, even SVM.

In general, more labeled data will result in higher accuracy. While the labeled data is limited, the unlabeled data is often abundant and easy to obtain. Therefore, the key issue addressing the small sample size is how to efficiently utilize the unlabeled data. Recently, the idea of semi-supervised learning (SSL), i.e. learning from both labeled and unlabeled data, has attracted much attention in information retrieval (Zhu, 2006). One recent approach is to apply Transductive SVM (TSVM) to exploit the unlabeled data (Joachims, 1999; Wang, Chan, & Zhang, 2003). However, some study challenged that TSVM might not be so helpful from unlabeled data in theory and in practice (Zhang, & Oles, 2000).
12 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage: www.igi-global.com/chapter/active-learning-relevance-feedback-image/28832?camid=4v1


Related Content

Incorporating Spatial Queries into Semantic Sensor Streams on the Internet of Things
www.igi-global.com/article/incorporating-spatial-queries-into-semantic-sensor-streams-on-the-internet-of-things/194998?camid=4v1a

Automatic Data Enrichment in GIS Through Condensate Textual Information
www.igi-global.com/chapter/automatic-data-enrichment-gis-through/20712?camid=4v1a

IT Value of Software Development: A Multi-Theoretic Perspective
www.igi-global.com/chapter/value-software-development/4294?camid=4v1a

www.igi-global.com/article/understanding-business-domain-models/49724?camid=4v1a