

Towards Gaze – Based Relevance Feedback in Image Retrieval

Horia Grecu, Ciprian Cudalbu, Vasile Buzuloiu
Image Processing and Analysis Laboratory,
Department of Applied Electronics and Information Engineering,
University “Politehnica” of Bucharest, Romania
E-mail: {hgrecu, ccudalbu, buzuloiu}@alpha.imag.pub.ro

Recently, relevance feedback (RF) algorithms have been shown to provide dramatic performance boost in content-based image retrieval (CBIR) systems. The purpose of relevance feedback is to provide ways to cope with the large variability in semantic interpretation of images across users. This is accomplished by allowing users to iteratively ‘guide’ the CBIR system, indicating the perceived relevance of retrieved results.

The problem of relevance feedback is usually cast as a user supervised classification problem where the task is to learn a discrimination function (over a certain image representation space), which separates the relevant image and non-relevant images. A fundamental difficulty for relevance feedback algorithms is the issue of small sample of training data: due to limitations in user’s willingness to cooperate, RF algorithm must deal with labeled data sets of few tenths of images, in representation spaces of hundredths dimensions.

We argue that the introduction of user’s eye-gaze measurements in the loop of relevance feedback can provide compelling improvements in retrieval performance.

The motivation is rooted in the long known evidence from cognitive sciences that eye-movements convey a great deal of information regarding the underlying mental process of subject.

The opportunity comes with the fact that state of the art eye tracking systems are able to provide accurate gaze data while being completely non-intrusive and almost transparent to user in terms of operation.

Our proposal is to preserve the basic relevance-feedback process - that consist of presenting the user a set of images and receiving an acceptance / rejection tag – while concomitantly gathering the eye-gaze data corresponding to the decision process.

While there is yet little knowledge on how to make the best use of the rich amount of information offered by eye-tracking systems, we show that the use of eye-fixations data can be exploited by providing the learning algorithm with importance weightings associated with each sample, leading to the following benefits:

- for the entire training set, it alleviates the limitation of dealing with global features only, by focusing the attention on the relevant parts based on the characteristics of fixation points in terms of duration and frequency;
- for positive samples, the analysis in terms of number of fixations and overall attendance time helps to identify the training samples containing a large number of distractors, which can be hence discarded or weighted accordingly;
- for negative samples, the analysis of fixations and overall attendance time can help to identify the negative samples that required the most attention, which informally are those images that are ‘similar but not quite’ the searched ones. These images are of high importance as they correspond to characteristic vectors on the border of relevant class, which, as known from the machine learning theory, are the most informative ones for the classification purpose.

We present results obtained by using the eye-gaze enhanced relevance-feedback on a simplified color-orientation image representation. The results compare positively with those obtained by the standard relevance-feedback using the same learning algorithm.