

Balancing Exploration – Exploitation in Image Retrieval

Dorota Głowacka and Sayantan Hore

HIIT, Department of Computer Science,
University of Helsinki, Finland

Abstract. In recent years there has been an increased interest in developing exploration–exploitation algorithms for image search. However, little research has been done as to what type of image search such techniques might be most beneficial. We present an interactive image retrieval system that combines Reinforcement Learning with an interface designed to allow users to actively engage in directing the search. Reinforcement Learning is used to model the user interests by allowing the system to trade off between exploration (unseen types of image) and exploitation (images the system thinks are relevant). A task-based user study indicates that for certain types of searches a traditional exploitation-based system is more than adequate, while for others a more complex system trading off exploration and exploitation is more beneficial.

Image retrieval techniques operating on meta-data, such as textual annotations, have become the industry standard. However, with the explosive growth of image collections, tagging new images quickly is not always possible. Secondly, there are many instances where image search by query is problematic, e.g. finding an illustration for an article about “youth”. A solution to such a problem is content-based image retrieval (CBIR) [6]. Early experiments show that CBIR can be improved through relevance feedback by involving the user in the search loop [1]. However, relevance feedback can lead to a context trap, where users specify the context so strictly that they can only exploit a limited area of information space. Combining *exploration/exploitation* strategies with relevance feedback is a popular attempt at avoiding the context trap [2, 3, 7]. However, few studies have been done showing the advantages (and disadvantages) of exploratory image retrieval systems. We report preliminary studies showing under what conditions exploratory image search might be most beneficial and where exploratory search may actually hinder the search results. For this purpose, we built a query-less image search system incorporating state of the art reinforcement learning (RL) techniques to allow the system to efficiently balance between exploration and exploitation.

System Overview. The system assists users in finding images in a database of unannotated images without query typing. The RL methods and interactive interface allow users to direct the search according to their interests. The interface and an example search are presented in Figure 1. The search starts with a display of a collage of images. To ensure that the initial set is a good representation of the entire image space, we cluster all the images in k clusters, where k is the number of displayed images and then we sample an image from each cluster. Our pre-user study shows that this technique provides a good starting point for the search. When the mouse hovers over an image,

Open search - browsing a collection of images without knowing what the target may look like. The study included three conditions: 1) our Gaussian Process system (GP), 2) a version of our system that uses only exploitation (EXPLOIT), and 3) a system that presents random images at each iteration (RAND). In EXPLOIT, the exploration level was set to 0, which means that the system can only present images similar to the ones marked as relevant. The same interface was used in all settings. We used the MIRFLICKR-25000 dataset [4] consisting of 25000 images from the social photography site Flickr and commonly used in assessment of image retrieval and annotation tasks. We recruited 20 post-graduate students to run the experiments. Each participant was asked to perform three tasks for all three types of searches, i.e. each participant performed 9 searches. We counterbalanced between the tasks and the systems for each subject so that each task was performed the same number of times with each system. The participants were asked to finish the task when they find the target image (in target search) or when they feel they found the ideal image in category and open searches. In all the tasks, the search was limited to 25 iterations. In target search, participants were presented with an image and a short description of that image and then asked to look for that image. In category and open searches, no example images were provided and participants were only given a short description of what to look for, e.g. red rose or illustration for an article about gardening.

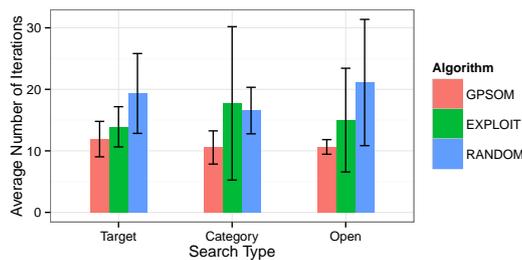


Fig. 2. Average number of iterations along with 95% confidence intervals.

We measured the average number of iterations to complete each task (Figure 2), which is a standard performance measure to evaluate CBIR systems [1]. GP-SOM outperforms EXPLOIT and RAND in all search types indicating that adding exploration to image search provides better support for user needs. There is little difference between GP-SOM and EXPLOIT in target search, indicating that when users have a specific image in mind from the onset, adding exploration makes little improvement. GP-SOM is more suitable for searches that are more exploratory in nature, such as category or open search, where the user first wants to browse the dataset before deciding what image they really want. We also counted the cumulative number of images that received positive feedback over search sessions in order to assess users' engagement in the search process (Figure 3). In target search, GP-SOM and EXPLOIT behave in a similar way. In category and open searches, GP-SOM displays relevant images throughout the search,

while EXPLOIT stops providing relevant images after about 10 iterations, which indicates that users “get stuck” in a very limited area of the image space.

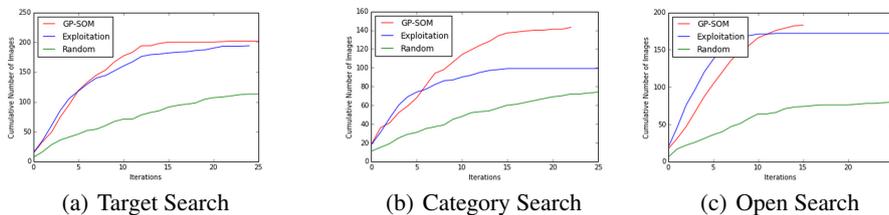


Fig. 3. Cumulative number of images marked as positive by user over iterations.

To summarize, GP-SOM exposes users to a higher number of relevant images in searches that are more vague in nature compared to EXPLOIT which narrows down the image space available to the user from the onset, which makes it more suited for target search. The results have significant implications for design of image retrieval system, where different strategies should be applied depending on the type of search, e.g. if we know that users will always make short searches then an EXPLOIT-type system will do a good job. However, if users have to perform longer open-ended searches (e.g. browsing a database of missing people), then a system based on exploration-exploitation might be more appropriate. In the future, we plan to run extensive user studies to get a better understanding of the relationship between search type and various levels of exploration.

Acknowledgements. The project was supported by The Finnish Funding Agency for Innovation (under projects Re:Know and D2I) and by the Academy of Finland (under the Finnish Centre of Excellence in Computational Inference).

References

1. I. Cox, M. Miller, T. Minka, T. Papatomas, and P. Yianilos. The bayesian image retrieval system, pichunter: theory, implementation, and psychophysical experiments. *Image Processing*, 9(1):20–37, 2000.
2. D. Głowacka, T. Ruotsalo, K. Konyushkova, K. Athukorala, S. Kaski, and G. Jacucci. Directing exploratory search: Reinforcement learning from user interactions with keywords. In *Proc. of IUI*, 2013.
3. D. Głowacka and J. Shawe-Taylor. Content-based image retrieval with multinomial relevance feedback. In *Proc. of ACML*, pages 111–125, 2010.
4. M. Huiskes and M. Lew. The mir flickr retrieval evaluation. In *Proc. of MIR*, 2008.
5. K. Konyushkova and D. Głowacka. Content-based image retrieval with hierarchical gaussian process bandits with self-organizing maps. In *Proc. of ESANN*, 2013.
6. H. Kosch and P. Maier. Content-based image retrieval systems-reviewing and benchmarking. *JDIM*, 8(1):54–64, 2010.
7. N. Suditu and F. Fleuret. Iterative relevance feedback with adaptive exploration/exploitation trade-off. In *Proc. of CIKM*, 2012.