

Diversity in User Relevance Feedback

Björn Þór Jónsson (bjorn@itu.dk)

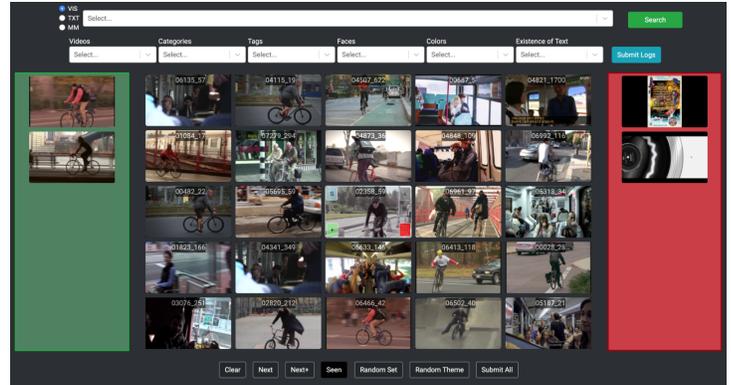
Martin Aumüller (maau@itu.dk)

Edgar Chavéz, CICESE, Mexico (elchavez@cicese.mx)

Project Outline

In user-relevance feedback systems, such as Exquisitor, the system presents potentially relevant images to users who label them as either relevant or irrelevant. Using the labeled examples, the system then refines the model of the user's need, until the user is satisfied [1].

The **choice of images** to present to the user is a difficult problem, as a naïve approach may present too many similar images, as shown in the image above. The challenge addressed in this project is to ensure **diversity** (aka “one of each”) as well as **relevance**.



Since Exquisitor uses a cluster-based index [2], the cluster structure can be employed directly to improve diversity. A more refined proposal is based on the Half-Space Proximity (HSP) graph [3], however, which is a sparse subgraph of the image similarity complete graph, where each node (that we call the *center*) in the HSP is connected to a natural number of similar neighbors (the *spike* of neighbors). Every node in the spike acts as a proxy of a *direction*. There is one spike for each object, and computing each spike is linear using a naïve algorithm, hence it has quadratic complexity for the entire collection. There are several challenges in this project, among them is computing efficiently the HSP, computing the spike of a query and embedding the resulting exploring mechanisms into Exquisitor.

The project is suitable for 2-3 well-qualified MSc students. The intention is to publish the results in international research venues, both as a conference paper and a journal paper. And the presentation of the MSc project should be exceptionally visual and interesting!

Skills Developed

The project provides excellent training for large-scale software development, including:

- C++ / backend programming and data structure development.
- Performance evaluation and analytic skills.

References

- [1] Omar Shahbaz Khan, Björn Þór Jónsson, Stevan Rudinac, Jan Zahálka, Hanna Ragnarsdóttir, Þórhildur Þorleiksdóttir, Gylfi Þór Guðmundsson, Laurent Amsaleg, Marcel Worring. Interactive Learning for Multimedia at Large. Proceedings of the European Conference on Information Retrieval (ECIR). Online, April 2020.
- [2] Gylfi Þór Guðmundsson, Björn Þór Jónsson, Laurent Amsaleg. A Large-Scale Performance Study of Cluster-Based High-Dimensional Indexing. Proceedings of the Workshop on Very-Large-Scale Multimedia Corpus, Mining and Retrieval, Firenze, Italy, October 2010.
- [3] Chavez E. et al. (2006) Half-Space Proximal: A New Local Test for Extracting a Bounded Dilation Spanner of a Unit Disk Graph. In: Anderson J.H., Prncipe G., Wattenhofer R. (eds) Principles of Distributed Systems. OPODIS 2005. Lecture Notes in Computer Science, vol 3974. Springer, Berlin, Heidelberg.