

The Xtrieval Framework at CLEF 2008: ImageCLEF photographic retrieval task

Thomas Wilhelm, Jens Kürsten, and Maximilian Eibl
Chemnitz University of Technology
Faculty of Computer Science, Dept. Computer Science and Media
09107 Chemnitz, Germany
[thomas.wilhelm | jens.kuersten | maximilian.eibl] at cs.tu-chemnitz.de

Abstract. This paper describes our participation at the ImageCLEF photographic retrieval task. We used our Xtrieval framework for the preparation and execution of the experiments. This year, we submitted 4 experiments in total. The experiments showed that our thesaurus based query expansions works well in improving the geometric mean average precision (GMAP) and binary preference (BPREF), but deteriorates the improvements gained by the addition of content-based image retrieval. The baseline (text-only) scored a mean average precision (MAP) of 0.0998. The combination of text and image retrieval gained a raise by 37 percent to a MAP of 0.1364. After applying the query expansion to both experiments the MAP for the text-only retrieval increased to 0.1081, but the MAP for the combined text and image retrieval decreased to 0.1140. By implementing an interface to the PostgreSQL database the retrieval speed and comparison operations for vectors could be speeded up.

Categories and Subject Descriptors

H.3 [Information Storage and Retrieval]: H.3.1 Content Analysis and Indexing; H.3.3 Information Search and Retrieval

Keywords

Evaluation, Content-based Image Retrieval, Query Expansion, Experimentation

1 Introduction

This year the data set was reduced again. All languages except English were removed. Only a subset of the topics of last year was supplied and the only available language again is English. So the multilingual character of this task got lost. In this year's experiments we actually repeated our monolingual runs of the last year with tweaked parameters and a new database backend for the storage of the MPEG-7 descriptors.

2 Experiment Setup

The base system from the last year (see [1], [2] and [3]) was reused with the following setup: Apache Lucene, customized analyzer with positional stopword¹ removal and Snowball stemmer². For the content-based image retrieval we used Caliph & Emir as described above.

Like last year a thesaurus was used for query expansion. The parameters of the last year were further tuned to reduce unfitting synonyms. As source for the thesauri we still use OpenOffice.org³.

¹ <http://members.unine.ch/jacques.savoy/clef/index.html>

² <http://snowball.tartarus.org/>

³ <http://wiki.services.openoffice.org/wiki/Dictionaries>

The MPEG-7 features were calculated by Caliph & Emir (see [4]). Contrary to our experiments of the last year the MPEG-7 descriptors were not stored as text representations in Lucene⁴, but as vectors in a PostgreSQL⁵ database. PostgreSQL was chosen because it supports arrays as data types. In fact it is not necessary to know the actual size of the arrays at design time. This approach is expected to achieve a much higher retrieval speed and it is possible to use descriptors of Caliph & Emir which has no string representation implemented (e.g. the dominant color descriptor).

The computations of the distance measures were externalized into the PostgreSQL database by implementing the algorithms as stored procedures in PL/pgSQL⁶. PL/pgSQL is an internal programming language of PostgreSQL which adds support for additional logic to SQL such as control structures. The following algorithms are implemented so far: cosine similarity, Dice coefficient, Euclidean metric, intersection, Jaccard similarity coefficient. The main advantage is the reduction of extra round trips between our application and the database server. On the other hand the speed could be reduced by the fact that PL/pgSQL is an interpreted language.

All topics were preprocessed ad-hoc to retrieve all needed resources to perform the experiments. Especially the example images were retrieved and analyzed in advance.

3 Results

Because of the before mentions reduction of data we only conducted four experiments. Considering the number of participants (25) and their submissions (100) this seems to be average. In the following table all four experiments are compared.

Table 1. Retrieval results

run	type	modality	feedback/ expansion	MAP	GMAP	BPREF
cut-txt	auto	txt	nofb	0.0998	0.0162	0.5008
cut-txt-qe	auto	txt	qe	0.1081	0.0300	0.5575
cut-mix	auto	txtimg	nofb	0.1364	0.0177	0.5008
cut-mix-qe	auto	txtimg	qe	0.1140	0.0275	0.5575

The results show that query expansion improves the mean average precision (MAP) only in text-only retrieval. In combination with content-based image retrieval the result gets even worse. But the geometric mean average precision (GMAP) and the binary preference (BPREF) improves in all cases. It is obvious that the thesaurus based query expansion improves the results in means of recall, but deteriorates them in means of precision.

Additional image information is only able to improve results if no query expansion is applied, but it is a quite high increase of the MAP by 37 percent.

References

- [1] T. Wilhelm, J. Kürsten, and M. Eibl, "Experiments for the ImageCLEF 2007 Photographic Retrieval Task"; http://clef-campaign.org/2007/working_notes/wilhelmCLEF2007.pdf.
- [2] J. Kürsten, T. Wilhelm, and M. Eibl, "The xtrieval framework at clef 2007: Domain-specific track," *LNCS - Advances in Multilingual and Multimodal Information Retrieval*, C. Peters et al., ed., Berlin: Springer Verlag, 2008.

⁴ <http://lucene.apache.org>

⁵ <http://www.postgresql.org>

⁶ <http://www.postgresql.org/docs/current/static/plpgsql.html>

- [3] T. Wilhelm, J. Kürsten, and M. Eibl, "Extensible retrieval and evaluation framework: Xtrieval," *LWA 2008: Lernen - Wissen - Adaption*, Würzburg: 2008.
- [4] M. Lux, W. Klieber, and M. Granitzer, "Caliph & Emir: Semantics in Multimedia Retrieval and Annotation," *19th International CODATA Conference*, 2004.