

Reducing Time and RAM Requirements in Content-Based Image Retrieval using Retrieval Filtering

Jens Forster and Thomas Deselaers

Human Language and Pattern Recognition Group, Computer Science Department,
RWTH Aachen University, D-52056 Aachen
{forster,deselaers}@i6.informatik.rwth-aachen.de

Abstract: In this paper we present and evaluate how computationally cheap comparison measures can be applied in content-based image retrieval applications to reduce the time- and memory requirements. The time requirements are reduced by applying filtering techniques and evaluating computationally costly distance measurements only to a suitably chosen subset of images. The RAM requirements are reduced by keeping only those features in memory that are absolutely required. It is shown that runtime- and memory efficiency is greatly improved with hardly any changes in retrieval quality.

Introduction

With the ubiquity of cameras and the ever-increasing necessity of digital images in medicine, the amount of images stored in databases is growing quickly. Access to these data is commonly achieved using textual meta data. Content-Based Image Retrieval (CBIR) systems are an alternative and allow accessing image databases by image content rather than by textual information. A problem with CBIR systems is that they require computationally expensive operations and large amounts of memory to allow for acceptable results. The aim of a CBIR system is to find visually similar images for a given query image.

To find visually similar images, typically features are extracted from each image in the database and from a given query image. Then, the features of the query image are compared to the features of each database image and thus the most similar images from the database can be determined. Some of these distance measures, e.g. the Image Distortion Model (IDM) [KGN04], provide good results in terms of error rates but have high computational costs which do not allow for interactive use in the context of large databases.

The concept of *filtered retrieval*, i.e. use a computationally cheap distance function to preselect images for the computationally more costly distance function, is well-known in the database and data exploration community, e.g. [FBF⁺94] and [SH94] propose to use a lower dimensional distance function as a filter for a higher dimensional quadratic distance function.

Filtered Retrieval in FIRE

We integrated the concept of filtered retrieval into the Flexible Image Retrieval Engine (FIRE)¹ [DKN04]. In FIRE, an *image* X is a set of *features* X_1, \dots, X_M representing certain aspects of the images. Normally, FIRE calculates the distance between a query image Q and a given image X contained in the database as a weighted sum over all individual feature distances. This calculated distance is used to assign a similarity score $S(Q, X)$ to

¹<http://www-i6.informatik.rwth-aachen.de/~deselaers/fire.html>

the database image.

$$S(Q, X) = \exp(-d(Q, X)) = \exp\left(-\sum_{m=1}^M w_i \cdot d_m(Q_m, X_m)\right) \quad (1)$$

where $d(Q, X)$ denotes the total distance between Q and X , $d_m(Q_m, X_m)$ the distance of Q and X wrt. feature m and w_i the weighting factor of feature m .

In filtered retrieval, we avoid calculating all distances, by first calculating the distance for one feature type for all images and keeping only the best matching images for subsequent steps in which the images are compared using the remaining features and distance functions. The order in which feature distances are evaluated is crucial since later steps are influenced by earlier steps. If several different features are used, the aim is to reduce the number of images to be considered in subsequent steps from feature to feature:

Let the database \mathcal{B} consist of N images: $\{X_1, \dots, X_n, \dots, X_N\}$, each image X_n is represented by M features $\{X_{n1}, \dots, X_{nm}, \dots, X_{nM}\}$. Then the distance according to the first feature is evaluated for all N database images. The N_1 best matching images are considered for the evaluation of the second feature. This is repeated with decreasing N_m until all M features were used. To be able to calculate scores (Eq 1), we set distances that were not evaluated to a value slightly higher than the highest calculated distance.

Now, not every feature for each image is required to process a query, e.g. if image X has been ruled out after the evaluation of feature m it is not necessary to keep features $m+1, \dots, M$. Therefore it is possible to remove some of the features from RAM and to load features only on-demand which may lead to a large reduction in memory requirements if it is possible to select small features first and load large ones only when requested. The gain in memory and speed is biggest if the features that are evaluated first are small features and allow for efficient distance calculation and the features evaluated last need more memory and computation time for distance calculation.

Databases

ImageCLEF² is an evaluation of CBIR techniques which is part of the Cross Language Evaluation Forum (CLEF)³ [DWK⁺05, DWN07, CMD⁺05]. In the context of this evaluation two medical image retrieval tasks are defined using the following databases:

The *IRMA 10,000* database was used in ImageCLEF 2005 and consists of 9,000 training images subdivided into 57 classes and 1,000 test images. The images have been chosen randomly from daily routine at the RWTH University Aachen Hospital.

The ImageCLEF 2005 and 2006 medical retrieval database consists of more than 50,000 images together with medical case descriptions and various other meta data

Our group participated in ImageCLEF 2005 and 2006 and obtained very good results. However, for the experiments, high computing times were required on machines with much RAM. Given the announcement that the databases are growing over the next years and that faster evaluation of the experiments is desired we use the proposed techniques for these tasks and present the outcomes in the following section.

Experimental results

First we evaluate the reduction of runtime on the IRMA task as here the image distortion model is required to obtain good results. Therefore we use Euclidean distance filtering

²<http://ir.shef.ac.uk/imageclef/>

³<http://clef.iei.pi.cnr.it/>

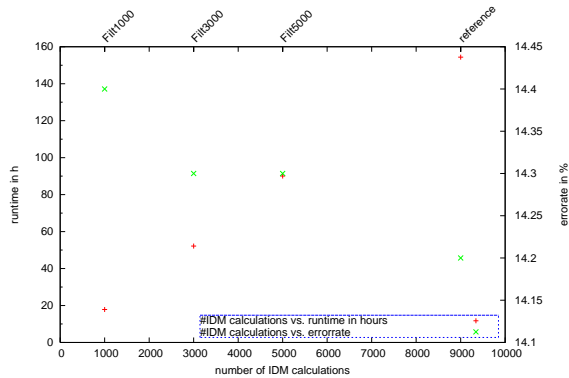


Figure 1: Time and error rate results on the IRMA 10,000 database

on 32×32 thumbnails of the images and apply the image distortion model distance to a subset only. The IDM distance function has been weighted 5 to 1. In the reference experiment filtered retrieval has not been used. In the following experiments the number of IDM distance calculations has been varied from 5,000 to 1,000. The achieved results depicted in Figure 1 show that the runtime of overall experiment has been reduced from about 154 hours to about 18 hours while increasing the error ate by only 0.2 %, i.e. two images more are classified incorrectly. On the ImageCLEF medical retrieval dataset, no class information is given but we have relevance judgements for a set of query images and measure the performance in mean average precision (MAP). Experimental results are given in Table 1. A reduction from 2.069GB to 0.7995GB peek RAM is observed while hardly changing the retrieval results.

Summary

The taken approach lead to very good results. In total a runtime reduction up to a factor of 8.6 and a reduction of peek RAM requirement up to 66% have been achieved with hardly any changes in retrieval performance. Hence it is now possible to apply CBIR to large databases in near real-time tasks. Nevertheless experiments show that the feature sequence needs to be well-chosen to keep retrieval performance. A more detailed description of the taken approaches and further results can be found in [For07].

Table 1: Experimental results to RAM reduction on the ImageCLEF2006 med database

run	reference	partial loading		
		FIL10P	FIL01P	FIL014P
32 x 32 image feat.	*	2.(15000)	1.(30000)	1.(30000)
color hist.	*	1.(30000)	2.(15000)	2.(15000)
32 x 32 image feat.	*	-	-	3.
peek memory usage in GB	2.069	0.4811	0.5404	0.7995
mean average precision	0.1470	0.0653	0.0746	0.1024

References

- [CMD⁺05] Paul Clough, Henning Mueller, Thomas Deselaers, Michael Grubinger, Thomas Lehmann, Jeffrey Jensen, and William Hersh. The CLEF 2005 Cross-Language Image Retrieval Track. In *Workshop of the Cross-Language Evaluation Forum (CLEF 2005)*, LNCS, page in press, Vienna, Austria, September 2005.
- [DKN04] Thomas Deselaers, Daniel Keysers, and Hermann Ney. Features for Image Retrieval – A Quantitative Comparison. In *DAGM 2004, Pattern Recognition, 26th DAGM Symposium*, number 3175 in LNCS, pages 228–236, Tübingen, Germany, September 2004.
- [DWK⁺05] Thomas Deselaers, Tobias Weyand, Daniel Keysers, Wolfgang Macherey, and Hermann Ney. FIRE in ImageCLEF 2005: Combining Content-based Image Retrieval with Textual Information Retrieval. In *Workshop of the Cross-Language Evaluation Forum (CLEF 2005)*, volume 4022 of LNCS, pages 652–661, Vienna, Austria, September 2005.
- [DWN07] Thomas Deselaers, Tobias Weyand, and Hermann Ney. Image Retrieval and Annotation Using Maximum Entropy. In *Evaluation of Multilingual and Multi-modal Information Retrieval – Seventh Workshop of the Cross-Language Evaluation Forum, CLEF 2006*, LNCS, page to appear, Alicante, Spain, September 2007.
- [FBF⁺94] C. Faloutsos, R. Barber, M. Flicker, J. Haffner, W. Niblack, D. Petkovic, and W. Equitz. Efficient and Effective Querying by Image Content. *IIIS*, 3(3–4):231–262, July 1994.
- [For07] Jens Forster. Reducing Time and RAM Requirements in Content-Based Image Retrieval using Retrieval Filtering. Studienarbeit, RWTH Aachen University, January 2007.
- [KGN04] Daniel Keysers, Christian Gollan, and Herman Ney. Classification of Medical Images using Non-linear Distortion Models. *BVM 2004, Bildverarbeitung fr die Medizin 2004*, pages 366–370, March 2004.
- [SH94] Harpreet S. Sawhney and James L. Hafner. Efficient Color Histogram Indexing. In *Image Processing 1994, ICIP-1994, IEEE International Conference*, volume 2, pages 66–70, November 1994.