

Combined Content based and Semantic Image Retrieval

Ioannis Boutsis, Theodore Kalamboukis
Department of Informatics
Athens University of Economics and Business
Athens 104 34

Abstract

i-score (Image Semantic and Content based Retrieval system) [1] developed at the Information Processing Laboratory combines two open source software libraries, Lire [2] and Lucene [3], with the aim to investigate the impact of images text-description in the quality and the effectiveness of image retrieval. In our runs for the ImagrCLEF2009 track the default Lucene's text analysis (stopword removal and stemming) was performed and the default Lucene's score function was used to evaluate the queries. Also all the duplicate descriptions of the images were removed from the database and a link was added to each record instead referring to a unique text. 39310 unique texts were remained in the database.

In both tasks Ad-Hoc and Case-based the semantic retrieval outperformed by far the visual and consequently the mixed retrieval. This is sensible for at least in our case we have used a naïve visual retrieval procedure. However give us promising evidence that techniques from textual retrieval can improve image retrieval in both the performance and efficiency

Construction of the Indexes

Two indexes were created automatically, one for the database of the images for visual retrieval and one for their descriptions for semantic retrieval.

For the images' data-base the index was created using Lire's DefaultDocumentBuilder and as an Analyzer Lire's SimpleAnalyzer. As a result the low level characteristics that we keep for each image are ScalableColor, ColorLayout and EdgeHistogram as they are defined at mpeg7.

For the texts' data-base firstly the HTML tags were removed. Then all the duplicate texts were removed and a link was added to each record instead referring to a unique text. 39310 unique texts were remained in the database. The index was based on the Lucene library and for each field of the images' records, that we want to be able to search, the following analysis was performed: The LowerCaseTokenizer was used and we have tokenized wherever the character is not a letter. Lucene's standard stop-words list was used and Porter's stemming algorithm applied on the remaining terms. Finally the filter (LengthFilter) was used to remove the terms that are either very small or very big to enter in a java stream.

Searching

For the visual retrieval the WeightedSearcher of Lire was used. The similarity measure was a weighted sum of the partial similarities due to each low level characteristic. In all our runs these values were set to: colorHistogramWeight=0.5 (ScalableColor), colorDistributionWeight=1.0 (ColorLayout) and textureWeight=0.7 (EdgeHistogram) wherever the input image was coloured and colorHistogramWeight=0.3, colorDistributionWeight=0.3 and textureWeight=1.0 for the black and white ones.

For the semantic retrieval, queries were subject to the same analysis as described in the indexing procedure (stopword removal, stemming) and the resulting sequence of terms was passed to Lucene. The default Lucene's score function was used to evaluate the queries.

Finally, at the combined retrieval the scoring function was defined as a linear combination of both, the image search and the corresponding text search, i.e.

$$\text{Mixed_score} = 0.8 * \text{textScore} + 0.2 * \text{imageScore}$$

Runs for CLEF-2009

The results for Image CLEF have been created off-line in order to be saved in trec_eval format. Thus, we have created seven runs as follows.

A. Topics

Run 1 – Visual: A visual search was performed for each of the images given as input and the results were combined, taking the average score. If the visual query contains p images then the score of a retrieved image, i , is given by

$$\text{visual_SCORE}(i) = \frac{1}{p} \sum_{k=1}^p \text{visual_sim}(i, i_k)$$

Run 2 – Semantic: For each topic the semantic index was searched.

Run 3 – Mixed: The semantic query was performed first. The visual retrieval was restricted to the images whose corresponding text was retrieved by the semantic query. If S is the set of images returned by the semantic query then

$$\text{mixed_SCORE}(i) = 0,8 * \text{semantic_SCORE}(i) + 0.2 * \text{visual_SCORE}(i)_{i \in S}$$

Run 4 : For each topic the retrieval was based on the type - semantic, visual, mixed - described in the topics2009.xml file.

B. Case Based Topics

Run 5 – Visual: A visual retrieval is applied to each topic and the score of an article is defined as the average of all visual similarities between the query-image and all the images in that article.

Run 6 – Semantic: The score of an article is defined as the average of all the semantic similarities between the semantic query and all the captions in that article retrieved by the query.

Run 7 – Mixed: A visual and a semantic retrieval are applied as in the runs 5 and 6 and the results are combined using percentages(80-20) for the text retrieval and image retrieval.

Results and Concluding Remarks

The results of our runs in ImageCLEFmed are summarized the table 1. In both tasks Ad-Hoc and Case-based the semantic retrieval outperformed by far the visual and consequently the mixed retrieval. This is sensible for at least in our case we have used a naïve visual retrieval procedure. However, this is a general remark in the results of all the participants, which give us promising evidence that techniques from textual retrieval may improve image retrieval in both the performance and efficiency. Indeed there is a lot of space for improvements techniques using relevance feedback techniques and domain ontologies or categorizing the images are currently under investigation. Definitely the availability of very large image

collections accompanied with descriptions, like the one in CLEFmed track, will contribute positively in this direction.

Task		num rel retr	MAP	P5	P10	P20	P30	P100	P1000
AdHoc Ret	Textual	1803/2362	0,3362	0,672	0,604	0,554	0,504	0,3152	0,0721
	Mixed	1381/2362	0,1466	0,376	0,328	0,28	0,2587	0,1696	0,0552
	Not applicable	824/2362	0,1255	0,328	0,288	0,246	0,224	0,1328	0,033
	Visual	242/2362	0,0048	0,024	0,028	0,026	0,024	0,024	0,0097
Case based Ret	Textual	93/95	0,1912	0,32	0,24	0,21	0,1867	0,116	0,0186
	Mixed	57/95	0,0159	0	0	0	0,02	0,016	0,0114
	Visual	39/95	0,0085	0	0	0	0,0133	0,01	0,0078

Table 1. Performance results of IPL laboratory for the ImageCLEFmed track.

References

1. <http://www.medas.gr:8084/iscore/>
2. <http://www.semanticmetadata.net/lire/>
3. <http://lucene.apache.org/java/docs/>