Using Collection Information To Improve Low Level Feature Based Multimedia Retrieval

Reede Ren
Department of Computing Science
University of Glasgow
17 Lilybank Gardens, Glasgow, UK
reede@dcs.gla.ac.uk

Joemon M. Jose
Department of Computing Science
University of Glasgow
17 Lilybank Gardens, Glasgow, UK
jj@dcs.gla.ac.uk

ABSTRACT

We propose a statistical representation for media documents called feature terms. This approach normalises feature distributions in a collection and leads to a mixed query/document model for multimedia retrieval. Two related problems are addressed: (1) how to extract discrete feature terms from continuous low level features; (2) how to rank the relevance between documents. TRECVid 2006 video collection and its 24 content-based query topics are employed for evaluation.

1. INTRODUCTION

Document collection is one of the major facts affecting retrieval performance. Zhai et al. [7] assert that the retrieval is a statistical decision problem based on the variance of term distributions in both document collection and a query. For example, Salton et al. [5] optimise retrieval models by local statistics in a document collection. However, it remains a new research topic in multimedia retrieval (MIR), how to model a multimedia collection. The discussion on concept number in TRECVid [2][6] is a study on the collection nature, namely how many high level features are effective and efficient to represent a news video collection. These technical reports show that collection knowledge is essential for the implementation of search engines.

In this paper, we limit our scope on low level features and propose an accumulated representation for media documents. The solution involves two steps: (1) project continuous low-level features onto a set of discrete variants called *low level feature terms* or *feature terms* in short; (2) compute a statistic on *feature terms* to implement a query. The highlights of this new approach are: (1) a general efficient document representation based on collection nature; (2) an efficient framework for document similarity measurement in high-dimension feature space, which avoids the complexity in distance definition; (3) the facilitation of feature aggregation from multiple samples, which results in an efficient mixed query model. In addition, this representation allows the exploration of text retrieval models in the application of MIR.

SAMT 2008 Dec 3-5 2008, Koblenz, Germany

2. RELATED WORKS

A feature term denotes a range interval of a feature. Since our approach exploits the same principles employed in statistical text retrieval, we review related works such as the discussion about text term distribution.

As an important aspect in term weighting, text term distribution is well discussed for the justification of retrieval models [1]. Harter et al. [3] propose that a term follows a 2-Poisson distribution because term appearance is a Boolean random and sparsely distributed. Margulis et al. [4] extend this model to N-Poisson mixed distribution. Several class numbers from two to seven were evaluated on real document collections, but no specific class number of Poisson mixed model shows a significant out-performance. Amati et al. [1] simulate a retrieval process with a Bernoulli distribution. The authors suggest a uniform term distribution. This is because the joint probability of multiple terms is so small that a simple uniform distribution is good enough for modelling term distribution.

Both Poisson and uniform distribution hypothesis are acceptable for distribution modelling of feature terms. We employ a uniform distribution to extract feature terms. In [1], the uniform hypothesis of term distribution leads to a superior retrieval performance. Moreover, the computational cost of a uniform distribution is significantly lower than N-Poisson. This is essential in MIR.

3. FEATURE TERM EXTRACTION

A criterion is necessary to justify the projection from continuous features to discrete feature terms. Given the uniform assumption, we propose following three criteria, minimised χ^2 test, maximised entropy and minimised AC/DC rate.

3.1 χ^2 Test

χ^2 test evaluates the similarity of a sample sequence from the given distribution. The optimised term probability should be $\hat{p}(tf_i) = \frac{1}{M}$. We get

$$\chi^2(M) = \sum_{i=0}^{M-1} \frac{(Mp(tf_i) - 1)^2}{M} \tag{1}$$

Therefore, the criterion for an optimised solution is,

$$I_{\chi^2} = \arg\min_M \chi^2(M) \tag{2}$$

3.2 Maximum Entropy

To reduce the effect of term set size on entropy calculation, we compute the overall entropy for a given feature-term projection as:

$$Entropy_s(M) = \frac{1}{\sqrt{M-1}} Entropy(M) \tag{3}$$

A high entropy measurement $I_{entropy}$ indicates a good term selection.

3.3 AC/DC rate

For a term frequency sequence $tf_0, tf_1, \ldots, tf_{M-1}$, discrete Fourier transform is used to compute the DC and the first AC parameter. The rate of AC/DC reflects the bias of a term frequency sequence from the average. A low AC/DC rate is preferred.

4. EXPERIMENTS

The TRECVid 2006 collection is used for evaluation. Three MPEG-7 low level features are extracted, including colour layout (12 dims), dominant colour (7 dims), and edge histogram (88 dims).

We rank video documents by KL and $BM25$ model. Figures 1, 2, and 3 display the sum of *number of relevant document returned* (Num-rel-ret) for each topic with the feature of dominant colour, colour layout, and edge histogram, respectively. These experiments show: (1) that term-based queries can reach at least similar performance as direct comparison and (2) that term extraction keeps the effectiveness of low level features in visual similarity measurement.

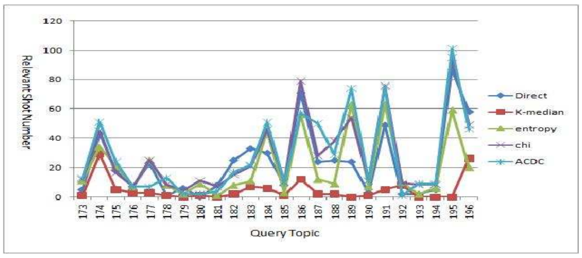

Figure 1: Dominant Colour Query By One Key Frame

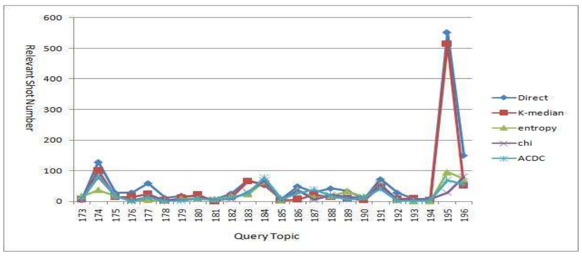

Figure 2: Colour Layout Query By One Key Frame

5. CONCLUSION

The major contribution of this paper is the exploration of a term generation approach and its application in video retrieval. Two closely associated research topics are explored:

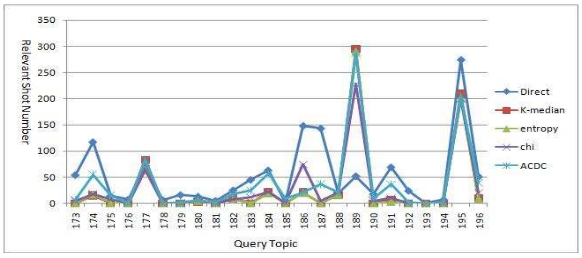

Figure 3: Edge Histogram Query By One Key Frame

(1) how to extract a term-like variant from continuous low level features; and (2) how to use these feature terms for a query. Such a projection is an important step towards an efficient video retrieval model. We suppose that the background knowledge from a video collection is of great importance for a retrieval: (1) the feature distribution in a document collection will hint what feature will be effective for a query topic; (2) the feature distribution difference between query samples and video collection can represent the query topic more accurate than a feature clustering. Moreover, our approach will make it possible to employ well developed text retrieval models, such as BM25, in the application of video retrieval. Note that these retrieval models are domain independent, which will lead to an efficient multimedia retrieval framework.

6. ACKNOWLEDGEMENT

The research leading to this paper was supported by European Commission under contract FP6-045032 (SEMEDIA).

7. REFERENCES

[1] G. Amati and C. J. V. Rijsbergen. Probabilistic models of information retrieval based on measuring the divergence from randomness. *ACM Transactions on Information Systems (TOIS)*, 20(4):357–389, 2002.

[2] M. Campbell, A. Haubold, M. Liu, A. Natsev, J. Smith, J. Tesic, L. Xie, R. Yan, and J. Yang. Ibm research trecvid2007 video retrieval system. In *TRECVid 2007*. NIST, 2007.

[3] S. Harter. A probabilistic approach to automatic keyword indexing, part i on the distribution of speciality words in a technical literature. *Journal of the ASIS*, 26:197–216, 1975.

[4] E. Margulis. N-poisson document modelling. In *the 15th international ACM SIGIR conference on Research and development in information retrieval*, pages 177–189. ACM, ACM Press, 1992.

[5] G. Salton and C.Buckley. Term-weighting approaches in automatic text retrieval. *Information Processing and Management*, (24):513–523, 1988.

[6] C. G. M. Snoek, B. Huurnink, L. Hollink, M. de Rijke, G. Schreiber, and M. Worring. Adding semantics to detectors for video retrieval. *IEEE Transactions on Multimedia*, 9(5):975–986, 2007.

[7] C. Zhai and J. Lafferty. A risk minimization framework for information retrieval. *Inf. Process. Manage.*, 42(1):31–55, 2006.