

CUNI at MediaEval 2015 Search and Anchoring in Video Archives: Anchoring via Information Retrieval

Petra Galuščáková and Pavel Pecina
Charles University in Prague
Faculty of Mathematics and Physics
Institute of Formal and Applied Linguistics
Prague, Czech Republic
{galuscakova,pecina}@ufal.mff.cuni.cz

ABSTRACT

In the paper we deal with automatic detection of anchoring segments in a collection of TV programmes. The anchoring segments are intended to be further used as a basis for subsequent hyperlinking to another related video segments. The anchoring segments are therefore supposed to be fetching for the users of the collection. Using the hyperlinks, the users can easily navigate through the collection and find more information about the topic of their interest.

We present two approaches, one based on metadata, the second one based on frequencies of proper names and numbers contained in the segments. Both approaches proved to be helpful for different aspects of anchoring problem: the segments which contain a large number of proper names and numbers are interesting for the users, while the segments most similar to the video description are highly informative.

1. INTRODUCTION

Thanks to a big rise of available multimedia data, the sizes of speech and video archives are recently growing rapidly. The amount of information stored in such archives is enormous, but the navigation in these archives is often tedious. Therefore, special attention should be paid to proposal and tuning of the methods for effective navigation in large multimedia archives. Effective navigation can not only help users to find required information easily but also to find additional information about the topic of their interest using an exploratory search.

The Search and Anchoring in Video Archives task organized in the MediaEval Benchmark 2015¹ explores effective navigation methods for large video archives. In this paper we deal with the Anchoring sub-task of the task. The main purpose of the sub-task is to automatically label anchoring segments. The anchoring segment should be in some way remarkable for the archive users so that the users may be interested in finding additional information about them. Hyperlinking [2, 3] can be further applied on the selected segments and more segments similar to the anchoring ones can be retrieved. The user can thus easily browse the archive using the links between anchoring segments and related data

¹<http://www.multimediaeval.org>

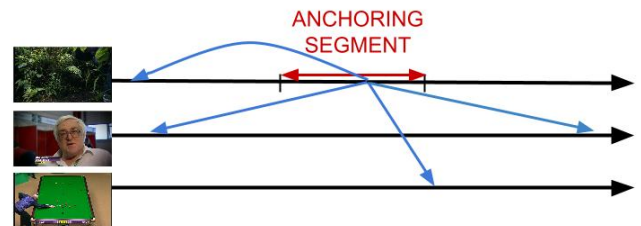


Figure 1: Hyperlinking of the anchoring segments and data segments.

segments and find information about related topics easily (see Figure 1).

Data used in the Anchoring sub-task consist of BBC TV programmes broadcast between 01.04.2008 and 31.07.2008. 38 programmes were used for the training and 34 programmes for the testing. 90 anchoring segments manually marked in the training data were available for training purposes. The Anchoring sub-task was evaluated using crowdsourcing campaign. The top 25 results returned by each task participant were manually marked as correct or wrong. Partially overlapping segments returned by the participants were joined before the annotation. We present Precision at 10 (P10), Recall and MRR scores. Details of the task, data, and evaluation process are given in the task description [1].

2. SYSTEM DESCRIPTION

We used subtitles available for each program in all our experiments. All videos were first segmented into shorter passages, which can possibly be the anchoring segments. The passages were either created regularly each 10 seconds (Fixed Segmentation)² or using machine learning-based methods which automatically detect probable segment ends (ML Segmentation) [4]. Afterwards, created segments were sorted according their likelihood of being the segment of interest (anchoring segment) for the archive users.

We used two methods for sorting created segments. In the first method, manually-created metadata available for each program were converted into the queries. The metadata consist of the program name and a short program description. Information retrieval (IR) system was then used to sort the created segments according to their similarity with metadata describing the program. The most similar segments

²Each segment was 60 seconds long.

Segmentation	Preference	P10	Recall	MRR
Fixed	—	0.27879	0.25119	0.92609
ML	—	0.30303	0.24995	0.90303
Fixed	Numbers	0.27879	0.24334	0.89091
Fixed	Names	0.31212	0.27030	0.86140
Fixed	Numbers and Names	0.31212	0.27061	0.83384

Table 1: Results of the anchoring task: a comparison of segmentation strategies and a preference based on frequencies of proper names and numbers contained in the segments. Best results for each measure are highlighted.

are expected to contain a relevant content and to describe the video recording well.

In the second method, the segments were sorted according to the frequency of numbers (words containing at least one digit) and proper names (words containing an upper case character while the previous word does not end with a dot) contained in each segment. Segments containing more proper names and numbers are intended to contain some kind of specific information.

In all experiments, we used the Terrier IR Framework version 3.5 [6], and its implementation of the Hiemstra Language Model [5] for retrieval, with its parameter set to 0.35. We used Porter stemmer and Terrier’s stopwords list.

Information retrieval system was also used for sorting the segments based on the proper name and number frequencies. As we further plan to combine the score from the retrieval based on metadata with the segment preference given by proper name and number frequencies, we first run the retrieval in the same way as in the case when only metadata were used. The weights corresponding to a frequency of numbers and names were precalculated. Finally, precalculated weights were linearly combined with the weight acquired as the output of information retrieval. However, in our preliminary experiments, the combination weight was set to highly prefer the ordering given by proper name and number frequencies.

3. RESULTS

The officially evaluated experiments are displayed in Table 1. The machine learning-based segmentation proved to increase Precision comparing to the fixed-length segmentation. Thanks to a large number of overlapping retrieved segments, which is even larger in the case of the machine learning-based segmentation, this Precision confirms the quality of the selection of the top-retrieved results. The fixed-length segmentation without any proper name and number preferences achieve the overall highest MRR score. This may confirm our assumption that highly ranked segments retrieved using metadata are the most informative for the recording.

The overall highest Precision and Recall scores are achieved when segments contain the largest number of proper names and numbers. This confirms the assumption that when the segment contains large number of proper names and numbers, it will be probably interesting for the archive user. Also, the overall largest number of segments marked by the annotators as relevant was retrieved when we sorted created segments according to how many proper names and numbers do they contain.

4. CONCLUSION AND FUTURE WORK

In the paper, we show several preliminary experiments with segmentation methods and selection process which can be used for anchor selection. High MRR score achieved by conversion of metadata into queries and their utilization for retrieving the most informative part of the document proved this approach to be convenient for the anchor selection.

Achieved Precision and Recall numbers show that the combination of this approach with the preference of the segment based on the frequencies of proper names and numbers may further improve the results. However, this combination would need some detailed tuning. The detection of proper names should be solved properly using the named entity recognition. Utilization of other information such as frequency of the content words, mentions of places and persons would also need further exploration.

5. ACKNOWLEDGMENTS

This research is supported by the Czech Science Foundation, grant number P103/12/G084, Charles University Grant Agency GA UK, grant number 920913, and by SVV project number 260 224.

6. REFERENCES

- [1] M. Eskevich, R. Aly, R. Ordelman, D. N. Racca, S. Chen, and G. J. F. Jones. SAVA at MediaEval 2015: Search and Anchoring in Video Archives. In *Proc. of MediaEval*, Wurzen, Germany, 2015.
- [2] M. Eskevich, R. Aly, D. N. Racca, R. Ordelman, S. Chen, and G. J. F. Jones. The Search and Hyperlinking Task at MediaEval 2014. In *Proc. of MediaEval*, Barcelona, Spain, 2014.
- [3] P. Galuščáková, M. Kruliš, J. Lokoč, and P. Pecina. CUNI at MediaEval 2014 Search and Hyperlinking Task: Visual and Prosodic Features in Hyperlinking. In *Proc. of MediaEval*, Barcelona, Spain, 2014.
- [4] P. Galuščáková and P. Pecina. Experiments with Segmentation Strategies for Passage Retrieval in Audio-Visual Documents. In *Proc. of ICMR*, pages 217–224, Glasgow, UK, 2014.
- [5] D. Hiemstra. *Using language models for information retrieval*. PhD thesis, University of Twente, Enschede, Netherlands, 2001.
- [6] I. Ounis, G. Amati, V. Plachouras, B. He, C. Macdonald, and C. Lioma. Terrier: A High Performance and Scalable Information Retrieval Platform. In *Proc. of SIGIR*, Seattle, Washington, USA, 2006.