

CUNI at TRECVID 2015 Video Hyperlinking Task Petra Galuščáková¹, Michal Batko³, Martin Kruliš², Jakub Lokoč², David Novák³, and Pavel Pecina¹

¹Institute of Formal and Applied Linguistics, Charles University in Prague ²SIRET Research Group, Charles University in Prague ³Faculty of Informatics, Masaryk University





Combination of:

Hyperlinking

- text-based similarity calculated on subtitles
 visual similarity between keyframes calculated using Feature Signatures
- preference based on whether the query and retrieved answer come from the same **TV serie**



- Find data segments which are similar to the query ones.
- Users can browse the archive using the links between query segments and related data segments.



1. Baseline

- All recordings are segmented into 60-second passages, new passage is created every 10 seconds.
 Terrier IR Platform
- Text-based retrieval using subtitles
 Metadata (title, description, broadcast channel)
 Context of the query (20 seconds before and after the query segment)
 Audio information from Doreso fingerprinting service
 Visual concepts recognized in each query keyframe with associated confidence score

4. Feature Signatures (FS)

- Feature Signatures approximate distribution of color and texture in the image.
- Identify keyframes with a similar background.
- Pre-calculated weights for each segment combined with the Baseline run.



5. FSSeriesRerank

• The results returned by the FS system were linearly combined with the same weights as those used in the Series run and reranked accordingly.

Run	Postfiltering	MAP	P10	MAP-bin	MAP-tol
Series	No	0.1312	0.2600	0.1443	0.1131
FSSeriesRerank	No	0.0987	0.1980	0.1094	0.0984
FS	No	0.1441	0.2750	0.1560	0.1234
Baseline	No	0.1405	0.2740	0.1536	0.1214
Series	Yes	0.1971	0.4313	0.2005	0.1718
FSSeriesRerank	Yes	0.1553	0.3414	0.1573	0.1340
FS	Yes	0.2131	0.4545	0.2138	0.1846
Baseline	Yes	0.2095	0.4495	0.2118	0.1826

Tab1. Performance comparison of results submitted to the Video Hyperlinking Task.

2. Postfiltering

• Remove retrieved segments occurring in the query video.

3. Series

• Information on whether the videos came from the same TV series

• Pre-calculated weights for each segment combined with the Baseline run.

Conclusion

• FS run outperformed the Baseline run - helpful in the retrieval of similar backgrounds and settings.

• Information about TV series did not outperform the Baseline run.

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.