Improving Automatic Semantic Annotations of News Videos in Turkish through Web Alignment and Event Extraction

Dilek Küçük¹ and Adnan Yazıcı²

¹ TÜBİTAK - Uzay Institute, Ankara - Turkey
dilek.kucuk@uzay.tubitak.gov.tr

² Dept. of Computer Engineering, METU, Ankara - Turkey
yazici@ceng.metu.edu.tr
Outline

- Introduction
- Initial Semantic Annotation and Retrieval System
- New Features
  - Web Alignment
  - Event Extraction
- The Improved System
- The Prospective System
- Conclusion
Introduction

- In this study, we present our extension of a semantic video annotation and retrieval system for Turkish by incorporating two new features:
  - Web alignment
  - Event extraction

- Hence, the system is extended
  - to exploit related Web news articles as an additional information source, and
  - to extract common events mentioned in the news stories to use them as semantic annotations.
Initial Semantic Annotation and Retrieval System
Web Alignment [1]

- Web texts can be a useful information source
  - to compensate for the noise in the existing information sources like ASR outputs
  - to enrich the already existing semantic information.

- Web news articles from *Milliyet* are aligned with the corresponding news videos:
  - A search crawler is implemented.
  - Named entities extracted from the video texts are used as the search criteria.
  - Articles published on the broadcast dates of the news videos, with those on the previous and following days are considered.
Web Alignment [2]

- For each page \( (p) \) in which at least one of the named entities \( (NEs) \) of the story segment \( (s) \) is encountered, a confidence value for the page is calculated:

\[
confidence_{s,p} = \frac{\text{number of } NEs \text{ of } s \text{ encountered in } p}{\text{total number of } NEs \text{ in } s}
\]

- Pages with confidence less than 0.5 and less than the maximum confidence are eliminated.

- After sorting the remaining ones on their detection times, only the top-3 pages are retained.

- Resulting Web articles are associated with their news stories.

- NEs extracted from the Web articles are also associated with the stories.
Web Alignment [3]

Fig. 2. Web Alignment Evaluation.
Event Extraction [1]

- A bag-of-words approach is used.

- The video texts of a training data set are annotated with semantic events.
  - 35 videos broadcasted by TRT with a total duration of 4 hours
  - Story level annotation
  - One story can include zero to many events.
  - 463 events for the 341 news stories.
  - 69 distinct event types.
Event Extraction [2]

- The stories for each event type are juxtaposed to form (not disjoint) event texts.

- For the most frequent 10 events, the most frequent 10 keywords are determined in the event texts, each with a confidence value:

\[
c_{e,k} = \frac{t_{fe,k}}{\sum_{i=0}^{10} t_{fe,i}}
\]

- If an event keyword is encountered in a news story text, then the confidence value for that keyword is added to the particular event confidence score of that story.
Event Extraction [3]

Fig. 3. Event Extraction Evaluation.
The Improved System [1]

Fig. 5. The Architecture of the Improved System.
The Improved System [2]

Fig. 4. A Boolean Query Example Through the Semantic Retrieval Interface of the Improved System.
The Prospective System

Fig. 6. The Architecture of the Prospective System.
Conclusion

- We present our extension of a text-based semantic annotation and retrieval system for news videos in Turkish with two features:
  - Web alignment
  - Event extraction

- The prospective semantic annotation and retrieval system is envisioned to include other components for
  - topic segmentation
  - keyphrase extraction
  - news categorization
  - news summarization
Thank You