Live-Ticker Supported Sports Video Annotation

T. R. Mayer[2], D. Coquil[1], M. Döller[1], L. Limam[1], F. Stegmaier[1], and H. Kosch[1]

[1] Chair of Distributed Information Systems
University of Passau, Germany
[2] Laboratoire d'InfoRmatique en Image et Systèmes d'information (LIRIS)
INSA de Lyon, France

Abstract. Automatic semantic annotation of videos remains an open research problem. For soccer matches, live tickers, are very interesting sources of information. Indeed, being written by humans, they may contain pertinent semantic information that can be used to annotate the corresponding videos. This paper presents a prototype that processes live ticker texts and transforms them into machine-readable annotations. The analyzer produces semantic annotated actions as RDF-graphs. This annotation may be used for other purposes, such as tactical analysis orsemantic multimedia query processing.

Key words: Semantic video annotation, ontology, text analysis

1 Introduction

The analysis of sports videos is increasingly gaining in importance for professional and private consumption. Thanks to online video portals (e.g. YouTube), users can catch up with missed highlights of their favorite sport. To enable the retrieval of such data, annotation of this content is required. Several approaches have been proposed for automatic, semi-automatic or manual video annotation. In view of the amount of video material, automatic or at least semi-automatic techniques are preferable. As this annotator automatically processes live ticker texts written by human operators, it is a semi-automatic approach that benefits from human interpretation capabilities while still being able to cope with the mass of videos. Moreover, live tickers are available for free on the internet. They are thus a very good source of semantic annotations for sport videos. This annotation can be used for queries or also as support for video annotation systems. The remainder of the paper is structured as follows. After an overview of related work, the prototype architecture is introduced. This is followed by a summary of the challenges that have been faced. Finally, a conclusion is given.

2 Related Work

Soccer video analysis is an emerging research field in the domain of multimedia annotation. Early works [1] focused on the recognition of the edges of the soccer

field (e.g. penalty box). Later approaches dealt with for instance shot boundary detections with respect to events such as goals [7]. Besides, techniques have been introduced for player tracking [5] or player identification [2]. The trajectory of players and the ball has been analyzed in [3] to generate tactic annotations (e.g. "attack from left"). A first work targeting live ticker texts has been proposed in [6]. There, the individual entries in the text are assigned with the comments and timestamps to the video source.

3 Prototype Architecture

3.1 Current Prototype

The prototype consists of four modules: *Preprocessor*, *Text Analysis*, *Postprocessor* and *Annotation Engine* (see figure 1). A *Domain-Specific Web Ontology Language (OWL)* is used by the Annotation Engine. The text processing components, which have already been implemented, are described here in more detail. The Annotation Engine, which has not yet been realized is briefly outlined in section 3.2. The first three modules form the ticker text processing components, which create an informational structure for use by the Annotation Engine. The goal is to transform the raw ticker text into machine understandable data, and to enrich it by lexical and semantic information.

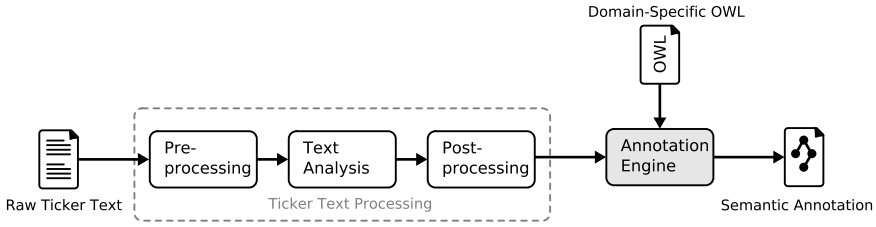

Fig. 1. Overview of the system architecture

Beginning with the *Preprocessor*, the raw ticker text, provided as a text file, is read. This character stream is then split into events. Each event represents a message submitted by the ticker author; it may consist of multiple sentences. The events are separated using timestamps provided in the original text, resulting in a list of tuples <*timestamp,event-text*>, which we call *event list*. In addition to In-Game (IG) events, the ticker text also contains Out-of-Game (OoG) events such as announcements. OoG events can be distinguished from IG event due to differences in the timestamp format; they are thus identified and removed using regular expressions. The result is an event list in which each entity is assigned to a phase of the match (first half-time, extra time of second half-time etc.) based on the timestamp and game rules. The last action of the preprocessor is to correct wrongly separated words, which usually occur for "-"-composed words with

special character. This most likely occurs because the ticker text is in German, whereas we used text analysis libraries dedicated to English.
The *Text Analysis* is the second part of the text preprocessing. It enriches the ticker events with lexical and semantic information. This is done in two steps. First, the Stanford Parser[3] performs lexical analysis and Part-of-Speech-Tagging (PoS-Tagging). Basing on statistical models, this parser tries to identify the lexical category (PoS-Tag) of a word in a sentence (e.g. verb, adjective). The library provides a model for German language. In the second step, the Stanford Named Entity Recognizer[3] or NE-Tagger identifies names in the event texts. As no German model is provided, training data have been used to let the NE-Tagger learn such a model. To this end, the ticker text has been annotated manually using domain-specific tags: PERSON, TEAM, LOC (locations, e.g. places or cities) and OTHER (for a word that is not a name). Using this as training data, a general model for the German language has been produced. Both, the PoS- and NE-Tagger use Sequence Labeling algorithms (see e.g. [4]) and calculate labels that are added to each word of an event text.

In the Text Analysis step, the PoS- and NE-Tagger split the event text into tokens and add their calculated labels. Due to different splitting rules, they may generate a different count of tokens for an event. The *Postprocessor* corrects this by merging multi-token entries, e.g. "Tim" and "Borowski" to "Tim Borowski". Furthermore, names are often written only as fore-/surname rather than completely. The Postprocessor corrects this by extending all names to their full version. This is done by comparing them with a lexicon, which has been developed by parsing a soccer website[4].

3.2 Future Work

An ontology has been created manually for the soccer domain to be used by the *Annotation Engine* (see figure 1), which has not yet been developed. This component will process text in two steps. The first step is similar to the present work: labels, which are references to OWL entities, will be assigned to tokens of the event text. A relation to a semantic context will then be assigned to each token. As a second step, the Annotation Engine will create an RDF graph using the domain-specific OWL and annotated event text tokens of the previous step. This graph can then be used for additional processing such as a tactical analysis.

4 Challenges

We had to face several challenges when developing our live ticker analysis prototype. As the ticker text is not machine understandable, it must first be transformed to a normalized format. This is a hard task due to different character encodings, different timestamp formats or incorrect punctuation. Writing styles

[3] http://nlp.stanford.edu/software
[4] http://www.fussball-wm-statistik.de

(e.g. jargon language) must be adapted. In the end, such a normalization engine is hard-linked to one specific ticker text provider. Another problem was that the analysis tools sometimes have trouble with special cases of word/sentence splitting. Finally, differences in timestamp formats complicated the assignment of an event to a phase of the game. As a summary, it seems that a tool with standardized output format would avoid many problems and enhance the ticker text production process.

5 Conclusion

This paper presents the current state of a live ticker analyzer for automatic generation of semantic annotations. A difficult problem is the lack of a standardized format for the ticker text. For now, texts must be normalized to a processable format for each provider, which is a hard task only valid until structural changes. Other issues such as unreliable punctuation complicate the semantic analysis. Once the text has been normalized and preprocessed, the annotation generation itself can take place. The result can be used for semantic multimedia queries and supporting additional annotation. In the future, the prototype will output an RDF graph including timestamps and representing game events. This will enable additional benefits such as the analysis of game tactics.

References

1. Gong, Y., Sin, L. T., Chuan, C. H., Zhang H., Sakauchi, M.: Automatic Parsing of TV Soccer Programs. Proc. of the International Conference on Multimedia Computing and Systems, 167–174 (1995)
2. Bertini, M., Del Bimbo, A., Nunziati, W.: Player identification in soccer videos. Proc. of the 7th ACM International workshop on Multimedia information retrieval. 25–32 (2005)
3. Zhu, G., Huang, Q., Xu, C., Rui, Y., Jiang, S., Gao, W., Yao, H.: Trajectory based event tactics analysis in broadcast sports video. Proc. of the 15th international conference on Multimedia, 58–67 (2007)
4. Nguyen, N., Guo, Y.: Comparisons of sequence labeling algorithms and extensions. Proc. of the 24th international conference on Machine learning (ICML '07), 681–688 (2007)
5. Miura, J., Kubo, H.: Tracking players in highly complex scenes in broadcast soccer video using a constraint satisfaction approach. Proc. of the International conference on Content-based image and video retrieval, 505–514 (2008)
6. Xu, C., Wang, J., Lu, H., Zhang, Y.: A Novel Framework for Semantic Annotation and Personalized Retrieval of Sports Video. IEEE Transactions on Multimedia, Vol. 10, Issue 3, 421–436 (2008)
7. Saric, M., Dujmic, H., Baricevic, D.: Shot Boundary Detection in Soccer Video using Twin-comparison Algorithm and Dominant Color Region. Journal of Information and Organizational Sciences, Vol. 32, Issue 1, 67–73 (2008)