A Watershed-based Social Events Detection Method with Support of External Data Sources

Minh-Son Dao, Giulia Boato, Francesco G.B. De Natale
mmLAB - DISI - University of Trento
Via Sommarive 5, 38123 Povo (TN), Italy
{dao, boato,denatale}@disi.unitn.it

Truc-Vien T. Nguyen
CIMeC - University of Trento
Corso Bettini 31, 38068 Rovereto (TN), Italy
trucvien.nguyenthi@unitn.it

ABSTRACT

In this paper, a watershed-based method with support from external data sources is proposed to detect Social Events defined by MediaEval 2012. This method is based on two main observations: (1) people cannot be involved in more than one event at the same time, and (2) people tend to introduce similar annotations for all images associated to the same event. Based on these observations, the whole metadata is turned to an image so that each row contains all records belonging to one user; and these records are sorted by time. Therefore, the social event detection is turned to watershed-based image segmentation, where Markers are generated by using (keywords, location) features with support of external data sources, and the Flood progress is carried on by taking into account (tags set, time) features. The high precision ($> 86\%$), and the acceptable recall ($\approx 50\%$) show the high effectiveness of the proposed method.

Categories and Subject Descriptors

H.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

1. INTRODUCTION

This paper presents the social events detection method that was specially built to meet challenge 1 and 3 of MediaEval 2012 [3]. We propose a watershed-based method for social events detection based on two main observations: (1) people cannot be involved in more than one event at the same time, and (2) people tend to use similar annotations (e.g. tags, description) for all images associated to an event. Details about the methodology are given in the next Section 2, while results on the MediaEval 2012 dataset are reported and discussed in Section 3.

2. METHODOLOGY

Starting from the assumptions reported in the Introduction, we define user-time images UT(i,j), by turning the whole metadata[1] to an image, so that each row of the UT image contains all records[2] belonging to one user and these records are sorted by time (i.e., date UT(i,j-1) < date UT(i,j)). In other words, pixel (i,j) of UT image (i.e., UT(i,j)) represents the j^{th} time-ordered record taken by the i^{th} user; and information such as description, tags, location, time written in j^{th} time-ordered record are considered as low-level features extracted from pixel UT(i,j).

The significant characteristic of UT image is that if pixel UT(i,j) belongs to event e, the left- or right-neighbour pixel of UT(i,j) must either belong to the same event e or another event. That leads to the idea of using watershed transform with markers to detect events, where **markers** are generated by using (*keywords, location*) features with support of *external data sources*, and the **flooding** progress is carried on by taking into account (*tags set, time*) features to build Merging-Condition (i.e., fill catchment basins). The Merging-Condition is built on the similarity of Tags set (using Jaccard index[3]) and time (using time-segmentation algorithm described in [1] to decide the time border of each event). Algorithms 1 and 2 explain how the proposed method runs.

Algorithm 1 watershed-based Social Events Detection

Input: UT image, Keywords, Locations
Output: Set of Events with Associated Images

1. Generate Set of Markers $\{m_i\}$ by calling **generateMarkers**(Keywords, Locations) function
2. FOR each Markers m_i MERGE left- and right-neighbours to the same Cluster c_i UNTIL **Merge-Condition** is not satisfied
3. FOR each Cluster c_i MERGE cluster c_j to c_i, $i \neq j$ IF it satisfies the **Merge-Condition**
4. Return set of remained clusters $\{c_i\}$. Each cluster represents one event.

2.1 Keywords and Locations Features

Keywords are considered as special "terms" that can model/represent the characteristics of events. It is important to notice that not all keywords may be considered equally

[1] $sed2012\_metadata.xml$, given by MediaEval 2012
[2] "photo" nodes of $sed2012\_metadata.xml$
[3] http://en.wikipedia.org/wiki/Jaccard index

MediaEval 2012 Workshop, October 4-5, 2012, Pisa, Italy

Algorithm 2 generateMarkers

Input: UT image, Keywords, Locations
Output: Set of Markers

1. USING tf-idf technique applied on *Keywords* to DETECT and RANK UT(i, j).
2. APPLY threshold to get the most related UT(i, j) to CREATE the CANDIDATE set $\{Can\}$ (we chose threshold as 0.5)
3. SELECT a subset $\{SCan\} \subset \{Can\}$ so that $SCan_k$ must contain *Locations*
4. Return $\{SCan\}$ as set of Markers

important: some of them perfectly represent an event (e.g., a "conference" is for sure a "technical event", thus this keyword is crucial in the recognition of technical events), while some other may refer to different type of events (e.g., the term "meeting" can refer to "technical event" but also to other types of events). Therefore, we need to assign a special weight for each keyword to increase the precision. In order to do this, we use the "semantic relatedness" technique [2] to measure the similarity between two terms by exploiting Wikipedia, a semi-structured, collaborative web-based resource.

Locations are used to determine where the event happened. It could be *name of cities, venues, public places*, or even corresponding (*latitude, longitude*) coordinates.

Unfortunately, it is very difficult to define generic markers that can run on all situations (e.g., MediaEval Challenges). Therefore, we build **generateMarkers** function separately for each Challenge, with support of External Data Sources.

2.2 Challenge 1

This challenge required to find all technical events organized in Germany. Therefore, we decide to build (keywords, locations) as follow:

Keywords: (1) define core keywords, such as "conference", "exhibition", and "workshop" and extend the set of keywords by collecting "synonym" defined by Oxford[4] and Macmillan dictionary[5]; (2) assign weights to keywords using semantic relatedness [2]; and (3) create the list of conferences (acronym, full name) by crawling data from some related websites such as www.allconferences.com, index.conferencesite.eu, www.tradeshowalerts.com, www.conferencealerts.com, www.ieee.org, www.acm.org, etc. We also considered some technical events only organized in Germany (e.g., FrosCon[6], CeBIT[7]). With this last group of keywords, the step 3 of algorithm 2 is not applied.

Locations: Create the list of cities in Germany (name, lat-long coordinates).

2.3 Challenge 3

This challenge required to find all Indignados movement events that took place in Madrid. Therefore, we decided to build (keywords, locations) as follow:

[4] www.oxforddictionaries.com
[5] www.macmillandictionary.com
[6] www.froscon.de
[7] www.cebit.de

Keywords: (1) define core keywords, such as "indignados", "demonstration", and "protest" and extend the set of keywords by collecting "synonym" defined by Oxford and Macmillan dictionary; (2) assign weights to keywords using semantic relatedness, as above.

Locations: Create the list of public places of Madrid (name, lat-long coordinates)

3. EXPERIMENTAL RESULTS AND DISCUSSION

In general, the effectiveness of the proposed method is quite good with precision being higher than 86%, and the recall being approximately 50%, as denoted in Table 1. Clearly, the markers generating progress plays a very important role in the proposed method, since it defines the precision rate. With 86% of precision, the generateMarker function works well in both challenges. Flooding progress influences the recall rate, especially the merging-condition. It decides how well we can detect images associated to the event detected by the markers generating progress. With average 50% of recall, there is the need to improve this process, especially with respect to the merging-condition.

For challenge 1, since we take into account acronyms of conferences, we can detect records that do not contain any keywords related to "technical event" as well as those that do not contain location information, such as events "designcamp", "CeBIT", etc. For challenge 3, since we set the geodistance sphere to 50 miles, and the set of public places of Madrid is not rich enough, we lost some significant markers and therefore lost images associated with these markers.

	Precision	Recall	F-Score	NMI
Challenge 1	86.23	59.13	70.15	0.6011
Challenge 3	86.15	47.17	60.96	0.4465

Table 1: Challenges Results

In the future, in order to increase the effectiveness of the proposed method, we will (1) take into account visual information, (2) define geodistance sphere for each location, (3) apply different threshold values in step 2 of algorithm 2, and (4) investigate Merging-condition w.r.t. features' influences (e.g. timestamp, tags set, visual similarity, etc.).

4. ACKNOWLEDGMENTS

This work has been partially supported by the EU Commission under the framework of the EU project Glocal, grant no. 248984.

5. REFERENCES

[1] A. C. Loui and A. Savakis. Automated event clustering and quality screening of consumer pictures for digital albuming. *IEEE TMM*, 5(3):309–402, September 2003.
[2] D. Milne and I. H. Witten. An effective,low-cost measure of semantic relatedness obtained from wikipedia links. In *Proceedings of the 22nd Conference on Artificial Intelligence*, 2008.
[3] S. Papadopoulos, E. Schinas, V. Mezaris, R. Troncy, and I. Kompatsiaris. Social Event Detection at MediaEval 2012: Challenges, Dataset and Evaluation. In *MediaEval 2012 Workshop*, Pisa, Italy, October 4-5 2012.