Social Event Detection at MediaEval 2012: Challenges, Dataset and Evaluation

Symeon Papadopoulos
CERTH-ITI
57001 Thermi, Greece
papadop@iti.gr

Emmanouil Schinas
CERTH-ITI
57001 Thermi, Greece
manosetro@iti.gr

Vasileios Mezaris
CERTH-ITI
57001 Thermi, Greece
bmezaris@iti.gr

Raphaël Troncy
EURECOM
Sophia Antipolis, France
raphael.troncy@eurecom.fr

Ioannis Kompatsiaris
CERTH-ITI
57001 Thermi, Greece
ikom@iti.gr

ABSTRACT

This paper provides an overview of the Social Event Detection (SED) task, which is organized as part of the MediaEval 2012 benchmarking activity and examines how social events can be detected by automatically analyzing social multimedia content. The 2012 edition of the SED task follows the successful completion of the first edition of SED, in 2011. This paper discusses the challenges set as part of SED 2012, the dataset, and the process of evaluating the submissions.

Categories and Subject Descriptors

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

General Terms

Experimentation

1. INTRODUCTION

The modeling, detection, and processing of events is an area that has started to receive considerable attention by the multimedia and semantic web communities[1] [2]. The Social Event Detection (SED) task of MediaEval 2012 requires participants to discover social events and detect related media items. By social events, we mean that the events are planned by people, attended by people and that the media illustrating the events are captured by people. A lot of multimedia content on the Internet was captured during such events or is otherwise related to events. Automatically establishing these underlying event-media associations represents a big step towards enabling multimedia browsing and search that is more natural to the users.

2. TASK OVERVIEW

The SED 2012 task is composed of three challenges and a common test dataset of images with their metadata (timestamps, tags, geotags for a small subset of them). Participants are free to use different approaches for addressing each

---

[1]See for example, the EiMM workshop series at ACM Multimedia or the DeRiVE workshop series at ISWC.

of the different challenges. For each challenge, up to five runs may be submitted, each run being a different set of results, which is produced by a different approach or approach variation, and is evaluated separately.

The evaluation criteria for each submission take into account the number of detected events (out of all relevant events in the testset) and the number of correct/incorrect media detected for these events. As a required (baseline) run for each challenge, the participants are asked to use any combination of the available image metadata they see fit, but no visual information, for finding the relevant events and images; the use of visual information in addition to the various provided image metadata is encouraged in subsequent runs.

Finding the events, in this task, does not mean finding some textual descriptions or metadata of the events. What we are looking for is a set of photo clusters, each cluster comprising only photos associated with a single event (thus, each cluster defining a retrieved event). The "photos associated with a single event" that we are looking for are all photos of the test collection that directly relate (in content, and also in terms of place/time) with the event of interest.

The image metadata that can be used by the participants for completing the SED challenges are only those provided to them as an XML file. The use of additional information (e.g. geotags) that may be available on the Internet for a given image of the dataset is not permitted. However, external resources (such as Wordnet, Wikipedia, or even visual concept detectors trained on external collections) can be employed, provided that they do not relate to specific images of the test dataset (or any images given for specifying the sought events), and that their development and use did not benefit from any knowledge of the task's dataset and challenge definitions.

3. CHALLENGES

3.1 Challenge 1

The first challenge reads: **Find technical events that took place in Germany in the test collection.**

Technical events, for the purpose of this task, are public technical events such as exhibitions and fairs. The annual CeBIT exhibition, taking place in Hannover, is a good (but of course, not the only) example of such an event.

Figure 1: Example images of (a) technical events (challenge 1), (b) soccer events (challenge 2), (c) Indignados demonstrations/protests (challenge 3)

3.2 Challenge 2

The second challenge reads: **Find all soccer events taking place in Hamburg (Germany) and Madrid (Spain) in the test collection.**

This challenge is very similar to the first of the two challenges defined in SED 2011 [3]. Similarly to 2011, soccer events, for the purpose of this task, are soccer games and social events centered around soccer such as the celebration of winning a cup. In contrast, a person playing with a soccer ball out in the street is not a soccer event under the task's definition. The reader is referred to [3] for more details on the meaning of a "soccer event".

3.3 Challenge 3

The third challenge reads: **Find demonstration and protest events of the Indignados movement occurring in public places in Madrid in the test collection.**

The Spanish Indignados movement centers around a series of demonstrations and other protests taking place all over Spain, which relate to the recent financial crisis outbreak as well as national politics in general. In contrast to the events that challenges one and two are concerned with, the events that are of interest to this third challenge are not scheduled, well-organized events (e.g., a technical fair or a soccer game, which are typically scheduled several months or days in advance, respectively). Instead, they are to a large extent spontaneous events, with any organization efforts related to them being typically centered around social media channels.

4. DATASET

A collection of 167.332 photos (more than twice as many as in the 2011 edition of this task) was created by issuing appropriate queries to the Flickr web service through its web-based API. The collected photos were all licensed under a Creative Commons licence, and were captured between the beginning of 2009 and the end of 2011 (specifically, 51.019 photos captured in 2009, 53.080 in 2010 and 63.233 in 2011) by 4,422 unique Flickr users. All photos were originally geo-tagged. However, before providing the XML photo metadata archive (including any tags, geotags, time-stamps, etc.) to the task participants, the geotags were removed for 80% of the photos in the collection (randomly selected). This was done for simulating the frequent lack of geotags in photo collections on the Internet (including the Flickr collection). The dataset and the ground truth will be made publicly available from the MediaEval website.

5. EVALUATION

The evaluation of the submissions to the SED task is performed with the use of ground truth that partially comes from the EventMedia dataset [1, 4] (for challenge 1), and in part as the result of a semi-automatic annotation process carried out with the help of the CrEve tool [5] (for all three challenges). Two evaluation measures are used:

- Harmonic mean (F-score) of Precision and Recall for the retrieved images. We use the macro version of F-score. This measures only the goodness of the retrieved photos but not the number of retrieved events, nor how accurate the correspondence between retrieved images and events is.

- Normalized Mutual Information (NMI). This compares two sets of photo clusters (where each cluster comprises the images of a single event), jointly considering the goodness of the retrieved photos and their assignment to different events.

Both evaluation measures receive values in the range [0, 1] with higher values indicating a better agreement with the ground truth results. These evaluation measures will be calculated both per challenge and on aggregate (for those teams that submit runs to all challenges).

6. CONCLUSIONS

The SED task gives its participants the opportunity to test and comparatively evaluate different approaches to the problem of social event detection in multimedia collections. The results of the submissions enrich the findings of the 2011 SED task results and give rise to interesting conclusions. Details on the methods and results of each individual participant can be found in the working notes papers of the MediaEval 2012 Workshop Proceedings.

Acknowledgments

7. REFERENCES

[1] H. Khrouf and R. Troncy. EventMedia: a LOD Dataset of Events Illustrated with Media. *Semantic Web - Interoperability, Usability, Applicability*, 2012.

[2] V. Mezaris, A. Scherp, R. Jain, and et.al. Modeling and Representing Events in Multimedia. In *19th ACM Int. Conf. on Multimedia*, Scottsdale, AZ, USA, 2011.

[3] S. Papadopoulos, R. Troncy, V. Mezaris, B. Huet, and I. Kompatsiaris. Social Event Detection at MediaEval 2011: Challenges, Dataset and Evaluation. In *Proc. MediaEval 2011 Workshop*, Pisa, Italy, September 2011.

[4] R. Troncy, B. Malocha, and A. Fialho. Linking Events with Media. In *6th Intern. Conference on Semantic Systems (I-SEMANTICS)*, Graz, Austria, 2010.

[5] C. Zigkolis, S. Papadopoulos, G. Filippou, Y. Kompatsiaris, and A. Vakali. Collaborative Event Annotation in Tagged Photo Collections. *Multimedia Tools and Applications*, 2012.