Introducing the Sky and the Social Eye

Alan Woodley
Queensland University of Technology
Brisbane, Australia
a.woodley@ qut.edu.au

Shlomo Geva
Queensland University of Technology
Brisbane, Australia
s.geva@qut.edu.au

Richi Nayak
Queensland University of Technology
Brisbane, Australia
r.nayak@qut.edu.au

Timothy Chappell
Queensland University of Technology
Brisbane, Australia
timothy.chappell@qut.edu.a

ABSTRACT
We introduce the Sky and the Social Eye task which was run for the first time as a Grand Challenge at the 2016 ACM Multimedia Conference and as a Task Force at the 2016 MediaEval Workshop. Participants combined satellite images with social media to create a richer user experience. For the first year, the task was exploratory allowing participants to produce their own system without traditional constraints such as shared datasets or metrics. Here, we describe the task, summarize the participants' approaches and propose some future directions.

1. INTRODUCTION
Remote sensing data, such as satellite images, have been used to explore environmental and social challenges for decades. For example, NASA's Landsat program has been used to identify the extent of forest fires, map land use change, assess carbon stock and analyze reef water quality [5]. Increasingly, social media is also being used for similar purposes, for example, mapping the extent of damage from natural disasters such as floods or earthquakes [4; 10; 11]. Despite this, there is a lack of research that explores how remote sensing data and social media can be combined.

The Sky and the Social Eye task was proposed to fill this gap by using social media to enrich satellite images. It ran as a Grand Challenge at the Association of Computing Machinery (ACM) Multimedia Conference and as a Task Force at the MediaEval Workshop for the first time in 2016. The aim of the task was for participants to develop algorithms and systems that combined satellite data with social media data. For the first year, the task was exploratory and creativity was encouraged. This allowed participants to develop their own systems, using their own datasets, case studies and metrics.

Here, we describe the inaugural Sky and the Social Eye task. We begin by detailing the motivations for the task and then outline the approaches undertaken by the participants. We conclude by discussing potential future directions for the task.

2. MOTIVATION
There are strong advantages for using satellite images to explore social and environmental events. For example, satellite data:

- Provides strong and compelling evidence about events on the ground;
- Is relatively inexpensive compared with field visits; and
- Allows for a large spatiotemporal area to be investigated.

Despite this, satellite data also has limitations, for example:

- Satellites tend to have a low temporal frequency (for example fortnightly) and so may miss an important event such as a flood or forest fire; and
- Clouds are present in some events, such as floods and snowstorms, obscuring the ability of satellites to capture images.

Examples of a satellite image is presented in Figure 1 [13]. Here, two images are provided. Both images were taken in Sri Lanka and show the impact of floods. The first image was taken on March 21st 2016 was taken prior to the flooding, while the second image was taken on May 31st 2016 after flooding. The second image shows enlarged waterways from the flooding as well as the presence of cloud cover, which obscures the land below.

In contrast, social media is continuously being generated and is unencumbered by issues such as cloud cover. Based on this, social media is increasingly being used to address problems that were traditionally addressed by satellite images analysis [4; 10; 11]. The Sky and the Social Eye task extends this research by linking satellite images and social media, such as text, images and videos

3. PARTICPANTS
Three groups participated in the inaugural Sky and the Social Eye.

Ahmad et al. [1] developed a system called JORD, which retrieved the names of events from the EM-DAT database [8] and searched Twitter, Flicker, YouTube and Google to retrieve and fuse information about the event. They tested JORD with 80 events including floods, landslides, cyclones and wildfires.

Bischke et al. [2] produced a system that fused satellite images with Twitter data. They used a case study of a wildfire in Fort McMurray, Canada. They collected Landsat 8 satellite images from Amazon Web Services. Tweets, including text and other images, were collected from Twitter's Historical Powertrack API. Then the tweets' text was analyzed to identify geolocation and sentiment and the tweets' images were analyzed with a convolutional neural network to remove near-duplicates. Finally, the data were fused and presented as a visualization.

MediaEval 2016, October 20–21, 2016, Hilversum, Netherlands.

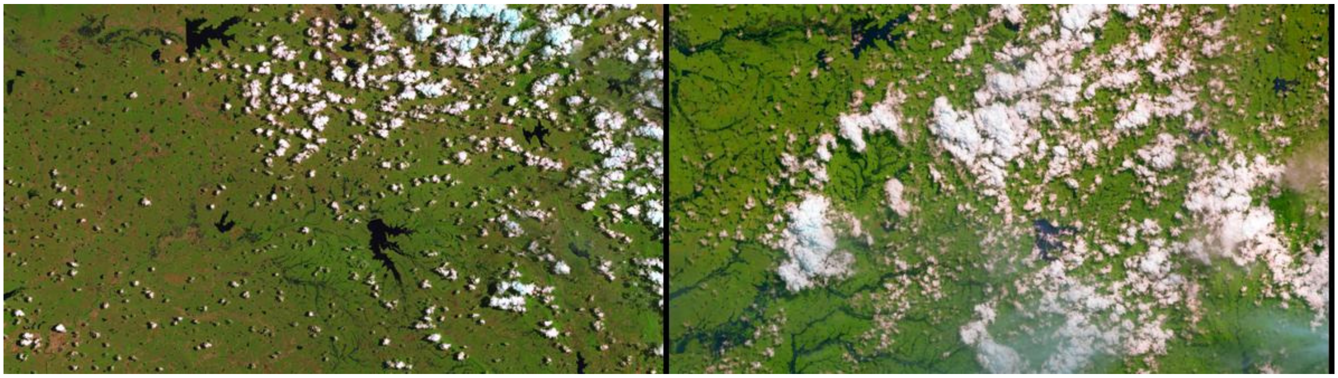

Figure 1. Satellite images which highlight the impact of a flood in Sri Lanka.

Crandall et al. [6] used satellite data as (noisy) ground truth to train two photo image classifiers: first, one that estimated if a photo contained evidence of an event and second, one that aggregated the estimates to produce an observation for given times and places. Satellite images were sourced from NASA's Terra satellite while photos were sourced from Flickr's public API. Satellite images where classified into bins according to their observed percentage of snow. These bins were then used to train the photo classifiers using both support vector machine on tagged text and a convolutional neural network on the photos themselves. Individual photos were then aggregated and a SVM trained over the aggregated data to estimate the probability of the actual environmental state. Finally, the results of the experiment were collected and presented as a visualization.

4. FUTURE DIRECTIONS

The major future direction of Sky and the Social Eye will be to transition it towards a traditional MediaEval Task. This will require the use of a shared set of documents, topics (queries) and metrics, enabling participants to better compare their systems to others. Moreover, the choice of the metrics is in itself a research problem as the metrics need to consider both multiple data types (that is: images and text) and draw together evaluations from the remote sensing community, whose metrics tend to be based on set retrieval (such as accuracy) [3], and the information retrieval communities, whose metrics tend to be list based (such as Mean Average Precision) [7].

In addition, the processing of remote sensing/social media on a cloud computing environment poses its own challenges. Traditionally, operational remote sensing analysis has been processed on a local server environment, often requiring the use of heavy sampling [14]. A cloud computer/supercomputer environment offers stronger computational capacity but has its own challenges, particularly regarding transfer of large datasets around the computing environment and the need to handle streaming data – challenges which may not be solved by current massive parallelization paradigm such as MapReduce or Mahoot [9; 12]. There is potential for the computer science community to play a leading role in this area of research.

5. SUMMARY

Here, we have described the Sky and the Social Eye task which was run as a Grand Challenge at the ACM Multimedia Conference and as a Task Force at the MediaEval Workshop. Participants in the task combined social media with satellite images, based upon on particular events. Sky and the Social Eye was run for the first time in 2016 and the organizers plan on transitioning it to a traditional MediaEval Task for future years.

6. ACKNOWLEDGMENT

The organizers of the Sky and the Social Eye task would like to thank the task participants and reviewers as well as the organizers of both the 2016 ACM Multimedia Conference and the 2016 MediaEval Workshop.

7. REFERENCES

[1] Ahmad, K., Riegler, M., Nguyen, D.T.D., Halvorsen, P., Lux, M., and De Natale, F., 2016. JORD - Linking sky and social multimedia data to events In *Proceedings of the 23rd International Conference on Multimedia Modeling* (Reykjavik, Iceland, 4-6 January 2016).

[2] Bischke, B., Borth, D., Schulze, C., and Dengel, A., 2016. Contextual enrichment of remote-sensed events with social media streams. In *Proceedings of the ACM Multimedia Conference* (Amsterdam, Netherlands 15-19 October 2016).

[3] Campbell, J.B. and Wynne, R.H., 2011. *Introduction to Remote Sensing*. Guilford Press.

[4] Cervone, G., Sava, E., Huang, Q., Schnebele, E., Harrison, J., and Waters, N., 2016. Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study. *International Journal of Remote Sensing 37*, 1, 100-124.

[5] Cohen, W.B. and Goward, S.N., 2004. Landsat's Role in Ecological Applications of Remote Sensing. *BioScience 54*, 6 (June 1, 2004), 535-545. DOI= http://dx.doi.org/10.1641/0006-3568(2004)054[0535:lrieao]2.0.co;2.

[6] Crandall, D., Wang, J., and Korayem, M., 2016. Tracking natural events through social media and computer vision. In *Proceedings of the ACM Multimedia Conference* (Amsterdam, Netherlands 15-19 October 2016).

[7] Croft, W.B., Metzler, D., and Strohman, T., 2010. *Search Egnines: Information Retrieval in Practice*. Pearson, Washington D.C., United States.

[8] Guha-Sapir, D., Below, R., and Hoyois, P., 2015. *EM-DAT: International disaster database.* Catholic University of Louvain.

[9] Lv, Z., Hu, Y., Zhong, H., Wu, J., Li, B., and Zhao, H., 2010. Parallel K-means clustering of remote sensing images based on mapreduce. In *Proceedings of the Proceedings of the 2010 international conference on Web information systems and mining* (Sanya, China2010), Springer-Verlag, 1927687, 162-170.

[10] Middleton, S.E., Middleton, L., and Modafferi , S., 2014. Real-time crisis mapping of natural disasters using social media. *IEEE Intelligent Systems 29*, 2, 9-17.

[11] Schnebele, E., Cervone, G., Kumar, S., and Waters, N., 2014. Real time estimation of the calgary floods using limited remote sensing data. *Water 6*, 2, 381-398.

[12] Sun, Z., Chen, F., Chi, M., and Zhu, Y., 2015. A spark-based big data platform for massive remote sensing data processing. In *Proceedings of the Proceedings of the Second International Conference on Data Science* (Sydney, Australia, 8-9 August 2015), Springer International Publishing, 120-126. DOI= http://dx.doi.org/10.1007/978-3-319-24474-7_17.

[13] United States Geological Survey, 2016. Landsat 8 Imagery Reveals Heavy Flooding in Sri Lanka. http://remotesensing.usgs.gov/gallery/gallery.php?cat=2#525.

[14] Wedderburn-Bisshop, G., Walls, J., Senarath, U., and Stewart, A., 2002. Methodology for mapping change in woody landcover over Queensland from 1999 to 2001 using Landsat ETM+. In *Proceedings of the The 11th Australasian Remote Sensing and Photogrammetry Association Conference* (Brisbane, Australia, September 2-6 2002).