MUSA2: First ACM Workshop on Multimodal Understanding of Social, Affective and Subjective Attributes

Xavier Alameda-Pineda, Miriam Redi, Mohammad Soleymani, Nicu Sebe, Shih-Fu Chang, Samuel Gosling

MUSA2 – First ACM Workshop on Multimodal Understanding of Social, Affective and Subjective Attributes.

Xavier Alameda-Pineda
Inria
Grenoble, France
xavier.alameda-pineda@inria.fr

Miriam Redi
Nokia Bell Labs
Cambridge, United Kingdom
miriam.redi@nokia.com

Mohammad Soleymani
University of Geneva
Geneva, Switzerland
Mohammad.Soleymani@unige.ch

Nicu Sebe
University of Trento
Trento, Italy
sebe@disi.unitn.it

Shih-Fu Chang
Columbia University
New York City, New York
sfchang@ee.columbia.edu

Samuel Gosling
University of Texas
Austin, Texas
samg@austin.utexas.edu

ABSTRACT

Multimedia scientists have largely focused their research on the recognition of *tangible* properties of data, such as objects and scenes.Recently, the field has started evolving towards the modeling of more complex properties. For example, the understanding of social, affective and subjective attributes of data has attracted the attention of many research teams at the crossroads of computer vision, multimedia, and social sciences. These *intangible* attributes include, for example, visual beauty, video popularity, or user behavior. Multiple, diverse challenges arise when modeling such properties from multimedia data. Issues concern technical aspects such as reliable groundtruth collection, the effective learning of subjective properties, or the impact of context in subjective perception. The first edition of the ACM MM'17 MUSA2 workshop has gathered together high-quality research works focusing on the computational understanding of intangible properties from multimodal data, including visual emotions, user intent, human relationships, and personality.

ACM Reference Format:
Xavier Alameda-Pineda, Miriam Redi, Mohammad Soleymani, Nicu Sebe, Shih-Fu Chang, and Samuel Gosling. 2017. MUSA2 – First ACM Workshop on Multimodal Understanding of Social, Affective and Subjective Attributes.. In *Proceedings of MM'17, October 23–27, 2017, Mountain View, CA, USA., ,* 2 pages.
https://doi.org/10.1145/3123266.3132057

1 INTRODUCTION

Traditionally, the recognition of tangible properties of data, such as objects and scenes [7], have overwhelmingly covered the spectra of applications in multimedia, computer vision and signal processing. In the recent past and partly fostered by social media, the understanding of subjective attributes (SA) of data has attracted the attention of many research teams at the crossroads of computer vision, multimedia, and social sciences. These subjective attributes include the ones assessed by *individuals* (e.g. safety [5], interestingness [2], evoked emotions [3], memorability [6]) as well as *aggregated* emergent properties (such as popularity [4] or virality [1]).

Given the inherent abstract nature of such concepts, many new challenges arise when attempting to automatically detect such properties in multimedia data, or to perform SA-based large-scale retrieval, including:

- Collecting huge amounts of annotations reflecting subjective judgments, as opposed to binary, objective annotations.
- Designing features or learning representations specifically tailored for subjective attribute recognition.
- Reliably evaluating the accuracy of detectors of subjective properties.
- Translating (social) psychology theories into computational approaches to systematically understand human subjective attribute perception.

This MUSA2 workshop aimed to gather high-quality contributions on the latest methodologies for understanding and recognizing intangible properties of multimodal data. In a nutshell, the focus of the workshop was on computational and experimental methods to learn, infer, or retrieve subjective attributes from multimodal data and their applications (e.g. subjective attribute-based advertising, retrieval and search), as well as to understand how and why humans perceive subjective attributes.

The scientific scope of the MUSA workshop stands on three areas of the ACMMM'17 program: Multimedia and Vision (Understanding), Multimedia Art, Entertainment and Culture (Experience), and Social Multimedia (Engagement). The first, because many recent and highly relevant studies to the proposed to topic appeared in the computer vision community (memorability, virality, etc); These studies would be helpful to inspire truly multimodal research in understanding subjective attributes. The second, because subjective attributes such as aesthetics or evoked emotions are a priori strongly tied to the culture, and hence it is worth boosting research in this direction. The third, not only because some of these subjective properties are required to be socially - as opposed to individually - perceived, but also because it is worth investigating the perception of intangible properties from a psychological and sociological perspectives.

MM'17, , October 23–27, 2017, Mountain View, CA, USA.

ACM ISBN 978-1-4503-4906-2/17/10.
https://doi.org/10.1145/3123266.3132057

Papers accepted at the MUSA workshop will be encouraged to submit and extended version to the associated ACM TOMM special issue[1].

2 TOPICS

This workshop is an opportunity for researchers and practitioners in the area of multimedia, computer vision, computational social science, and social-media analytics to present creative works in the area of subjective attribute understanding from multimedia data. More importantly, the multidisciplinary nature of this workshop allows submissions from research groups focusing on multimedia, computer vision, hci, experimental social and personality psychology, digital humanities, and cultural analytics.

We called for papers on multidisciplinary topics regarding subjective multimedia analytics. These topics included:

- *Data collection/annotation and evaluation* methods for subjective attribute studies, including active learning and crowdsourcing.
- *Learning* and inference techniques for individual and aggregated subjective attribute recognition in multimedia data, including beauty, sentiment, interestingness, memorability, creativity, ambiance, virality, popularity, engagement.
- *User diversity*-aware models for individual and collective subjective attribute detection and retrieval: systematic studies regarding the impact of user's demographics (e.g., gender, age, race) and psychological characteristics (e.g., personality, emotional state) in relation to subjective preferences are welcome.
- *Impact of Context:* Systematic analysis on the effect of the individual's social network on the perception of subjective attributes.
- *Human Understanding:* Experimental Psychology and Computational Social Science studies using multimedia data to understand human behavior.
- *Applications* of subjective attribute detection and retrieval methods, including advertising, retrieval, search and new contexts such as social good and urban spaces.

3 PROGRAM

The workshop program includes oral presentations, invited keynote presentations, and a poster session. We received 13 submissions, out of which we accepted 4 for oral (15 minutes) presentation (30% acceptance rate) and another 4 for spotlight (1 minute) presentation. All speakers will have the chance to further discuss with the audience during a 45-minute poster session. Each paper received 3 high-quality reviews.

Submissions span a variety of different topics related to the workshop, including, among others: the impact of personality on media consumption, how head pose affects photo aesthetics, how people caption images they wish to share, interpretable models for sentiment-biased concept detection. This exciting scientific program will be completed with an invited talk. Ayman Shamma,Senior Research Scientist at FX Palo Alto Laboratory, and Saeideh Bakhshi Quantitative UX Research at Facebook, joined our workshop as keynote speakers, sharing their knowledge at the intersection of human-computer interaction, user engagement, social media analytics, and multimedia modeling.

[1]http://tomm.acm.org/special-issues.cfm

4 FUTURE EDITIONS

In future editions, we would like to foster even more the multidisciplinary of our workshop. Welcoming such a wide spectrum of research fields could create an unprecedented opportunity to gather experts from domains that are traditionally disjoint, thus fostering interdisciplinary discussions and enabling the formation of solid collaborations between computer science and social science researchers. We will organize an interdisciplinary round-table where students and researchers will openly discuss the future of the rising field of intangible multimedia understanding, as: (1) How to actively engage social science experts in subjective attribute multimedia research, thus enriching the theoretical foundations of computational frameworks for subjective attribute detection. (2) How to design multimedia analysis tools and solutions that can support social psychology and computational social science studies.

5 ORGANIZERS

Our organization committee is very diverse in terms of expertise, topics, geographic location and seniority.

- Dr. Xavier Alameda-Pineda (Research Scientist, Perception Group, INRIA Grenoble)
- Dr. Miriam Redi (Research Scientist, Social Dynamics Team, Bell Labs Cambridge, UK)
- Dr. Mohammed Soleymani (Swiss NSF Ambizione Fellow, Multimedia Affective Analysis group, University of Geneva)
- Prof. Nicu Sebe (Professor and Head of Dept. of Information Engineering and Computer Science, University of Trento)
- Prof. Shih-Fu Chang (Richard Dicker Professor and senior Vice Dean of the School of Engineering and Applied Sciences at Columbia University.) - sfchang@cs.columbia.edu
- Prof. Samuel D. Gosling (Professor, Department of Psychology, Univeristy of Texas at Austin)

REFERENCES

[1] Xavier Alameda-Pineda, Andrea Pilzer, Dan Xu, Nicu Sebe, and Elisa Ricci. 2017. Viraliency: Pooling Local Viraliry. In *IEEE CVPR*.

[2] Michael Gygli, Helmut Grabner, Hayko Riemenschneider, and Luc Van Gool. 2013. The interestingness of images. In *ICCV*.

[3] Brendan Jou, Tao Chen, Nikolaos Pappas, Miriam Redi, Mercan Topkara, and Shih-Fu Chang. 2015. Visual Affect Around the World: A Large-scale Multilingual Visual Sentiment Ontology. (2015).

[4] Aditya Khosla, Atish Das Sarma, and Raffay Hamid. 2014. What Makes an Image Popular?. In *Proceedings of the 23rd International Conference on World Wide Web (WWW '14)*. ACM, New York, NY, USA, 867–876. https://doi.org/10.1145/2566486.2567996

[5] Lorenzo Porzi, Samuel Rota Bulò, Bruno Lepri, and Elisa Ricci. 2015. Predicting and understanding urban perception with convolutional neural networks. (2015), 139–148.

[6] Aliaksandr Siarohin, Gloria Zen, Cveta Majtanovic, Xavier Alameda-Pineda, Elisa Ricci, and Nicu Sebe. 2017. How to Make an Image More Memorable? A Deep Style Transfer Approach. In *ACM ICMR*.

[7] Bolei Zhou, Agata Lapedriza, Jianxiong Xiao, Antonio Torralba, and Aude Oliva. 2014. Learning deep features for scene recognition using places database. In *Advances in neural information processing systems*. 487–495.