ATQAM/MAST'20: Workshop on Aesthetic and Technical Quality Assessment of Multimedia and Media Analytics for Societal Trends

Tanaya Guha
University of Warwick
Coventry, UK

Vlad Hosu
Dietmar Saupe
Bastian Goldlücke
University of Konstanz
Konstanz, Germany

Naveen Kumar
Disney Research
Los Angeles, US

Weisi Lin
Nanyang Technological University
Singapore

Victor Martinez
Krishna Somandepalli
Shrikanth Narayanan
University of Southern California
Los Angeles, US

Wen-Huang Cheng
National Chiao Tung University
Hsinchu, Taiwan

Kree McLaughlin
Hartwig Adam
Google
Los Angeles, US

John See
Lai-Kuan Wong
Multimedia University
Cyberjaya, Malaysia

ABSTRACT

The Joint Workshop on Aesthetic and Technical Quality Assessment of Multimedia and Media Analytics for Societal Trends (ATQAM/MAST) aims to bring together researchers and professionals working in fields ranging from computer vision, multimedia computing, multimodal signal processing to psychology and social sciences. It is divided into two tracks: ATQAM and MAST.

ATQAM track: Visual quality assessment techniques can be divided into image and video technical quality assessment (IQA and VQA, or broadly TQA) and aesthetics quality assessment (AQA). While TQA is a long-standing field, having its roots in media compression, AQA is relatively young. Both have received increased attention with developments in deep learning. The topics have mostly been studied separately, even though they deal with similar aspects of the underlying subjective experience of media. The aim is to bring together individuals in the two fields of TQA and AQA for sharing of ideas and discussions on current trends, developments, issues, and future directions.

MAST track: The research area of media content analytics has been traditionally used to refer to applications involving inference of higher level semantics from multimedia content. However, multimedia is typically created for human consumption, and we believe it is necessary to adopt a human-centered approach to this analysis, which would not only enable a better understanding of how viewers engage with content but also how they impact each other in the process.

MM '20, October 12–16, 2020, Seattle, WA, USA

ACM ISBN 978-1-4503-7988-5/20/10.
https://doi.org/10.1145/3394171.3421895

CCS CONCEPTS

• **Computing methodologies** → *Perception*; *Image manipulation*; **Computer vision tasks**; • **Information systems** → **Multimedia information systems**; *Data mining*.

KEYWORDS

Quality assessment, aesthetics assessment, visual enhancement, datasets, multimedia, societal impact, media consumption

ACM Reference Format:
Tanaya Guha, Vlad Hosu, Dietmar Saupe, Bastian Goldlücke, Naveen Kumar, Weisi Lin, Victor Martinez, Krishna Somandepalli, Shrikanth Narayanan, Wen-Huang Cheng, Kree McLaughlin, Hartwig Adam, John See, and Lai-Kuan Wong. 2020. ATQAM/MAST'20: Workshop on Aesthetic and Technical Quality Assessment of Multimedia and Media Analytics for Societal Trends. In *Proceedings of the 28th ACM International Conference on Multimedia (MM '20), October 12–16, 2020, Seattle, WA, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3394171.3421895

1 INTRODUCTION

The Joint Workshop on Aesthetic and Technical Quality Assessment of Multimedia and Media Analytics for Societal Trends (ATQAM/MAST) aims to bring together researchers and professionals working in fields ranging from computer vision, multimedia computing, multimodal signal processing to psychology and social sciences. It is divided into two tracks. The first track, ATQAM, presents contributions that address aspects of perceptual quality in multimedia, with a focus on technical and aesthetic quality. The second track, MAST, has the goal to understand different computational tools

and methodologies for the systematic study of the societal impact of media.

1.1 ATQAM

Every day over 450 million photos and videos are being uploaded to Facebook and Instagram. The exponential growth of visual media has made quality assessment become increasingly important for various applications, including image and video acquisition, synthesis, restoration, enhancement, search and retrieval, storage, and recognition.

Broadly, visual quality assessment techniques can be divided into image and video technical quality assessment (IQA and VQA, or broadly TQA) and aesthetics quality assessment (AQA). TQA focuses on the effect of image-level technical aspects of perceived quality, such as sharpness, noise, color reproduction, contrast, dynamic range, and others. On the other hand, AQA deals with more abstract aesthetics-related quality factors that capture the subjective aesthetics experience. Aesthetics judgments are associated with the adherence to established photographic rules encompassing lighting (emphasis, contrast), composition, colors, and more. Even though these topics have mostly been studied independently, they represent tightly related aspects of the same underlying subjective experience of media items, value judgments.

This workshop aims to bring together individuals in the two fields of TQA and AQA for sharing of ideas and discussions on current trends, developments, issues, and future directions, with the vision to accelerate the progress of research in both fields. Our hope is that bridging TQA and AQA, will result in a better understanding of quantitative measures of quality of experience in the broader context of multimedia applications.

1.2 MAST

Traditional multimedia content analytics and research usually deals with tasks such as indexing and summarization. In the context of understanding the impact of media on the society and shaping our experiences, we believe in the need for a holistic approach to quantify how people, places and topics are portrayed in media, complementing the traditional branches of media analytics research. The goal of our workshop is to close the loop with the audience's experience, and to analyze the impact of media in terms of ever-evolving societal patterns.

The workshop is strongly in line with the Experience theme of the ACM MM 2020. The focus of this workshop is on developing computational media intelligence, which deals with quantifying how people, places and topics are presented in media content to understand their impact on individual and society. The MAST workshop is unique in its approach to media. Since the very first MAST workshop in San Diego in conjunction with the International Conference of Multimedia and Expo (ICME 2018), our workshop has thrived to connect industry, academia and media stake-holder. MAST is truly interdisciplinary where computer scientists and engineers discuss open problems and challenges with social scientists and media experts.

2 SCOPE AND TOPICS

The topics of interest for this workshop include:

2.1 ATQAM

- **Analysis and prediction** of aesthetic and technical visual quality, encompassing absolute and comparative judgments about media:
  - Traditional and deep-learning-based approaches
  - Aesthetics and QoE related concepts such as interestingness, popularity, viralness
- **Datasets** for TQA and AQA, including:
  - New approaches to data collection procedures and sources
  - New data augmentation methods
- **Applications** of TQA and AQA in computer vision or image processing tasks:
  - Visual filtering and retrieval (recommendation, image gallery/video)
  - Visual editing (recomposition, retargeting, cropping)
  - Assessment guided visual enhancement
  - Real-world systems and applications
  - Applications to light fields, 360, stereo, point clouds.

2.2 MAST

Traditional research on different forms of media analysis has been primarily focused on the low and mid-level tasks, such as indexing and summarization. With the availability of big data and powerful learning techniques, the human factors in media, such as the affective content, insight generation, societal impact and trends. To this end, the workshop invites original papers related to the topics listed below.

- **Impact of Media**
  - Quantifying and analyzing media impact at individual, group and societal level
- **Affect in media**
  - Affect prediction analysis
  - Expressed vs. perceived affect in multimedia
  - Therapeutic uses of media forms
- **Computational narratology**
  - Understanding narratives, tropes and character portrayals
  - Computational analysis of storytelling
  - Interaction among characters in multimedia content
- **Methodologies**
  - Data mining for labeling media data
  - semi-supervised/ self-supervised learning

3 WORKSHOP ORGANIZATION

3.1 ATQAM

The ATQAM track takes place as in the first half-day of the ATQAM/MAST'20 joint workshop program. The program features a short overview talk provided by the workshop organisers on the challenges and potential of bridging the technical and aesthetics quality assessment, two invited keynote talks, two talks that present original work on databases for TQA and AQA, and an invited talk on "rating distribution and personality prediction for image aesthetics assessment".

Invited keynote speakers. **Alan Conrad Bovik** is an American engineer and vision scientist. He is a Professor at The University of Texas at Austin (UT-Austin), where he holds the Cockrell Family Regents Endowed Chair and is Director of the Laboratory for Image

and Video Engineering. He is a faculty member in the UT-Austin Department of Electrical and Computer Engineering, the Institute for Neuroscience, and the Wireless Networking and Communications Group. Bovik won a Primetime Emmy Award in 2015 for his development of video quality measurement tools that are now standards in television production. Two of Bovik's research publications in the area of visual image quality have been recognized as 2017 Google Scholar Classic Papers, which are selected for being highly-cited papers that have stood the test of time, and are among the ten most-cited articles in their area of research published ten years earlier.

Professor Bovik is giving a talk related to video quality assessment, "Speeding it Up: Perception of High-Frame Rate Videos".

James Z. Wang is a professor at Pennsylvania State University. Wang's research seeks to advance knowledge through modeling objects, concepts, aesthetics, and emotions in big visual data. He is well-known for his pioneering research in the field of aesthetics quality assessment. His research team have developed the ACQUINE aesthetic quality inference engine, SIMPLIcity semantics-sensitive image retrieval system, the ALIPR real-time computerized image tagging system, which are all widely cited. His research has been reported widely by significant media, including Discovery, Scientific American, MIT Tech Review, Public Radio, NPR, and CBS. Wang also received an NSF Career award and the endowed PNC Technologies Career Development Professorship.

Professor Wang's talk is on "Modeling Aesthetics and Emotions in Visual Content - From Vincent van Gogh to Robotics and Vision".

Invited talks. The first presentation "Going Big or Going Precise: Considerations in building the next-gen VQA Database" discusses choices that should be considered when scaling-up video TQA databases. It introduces the trade-off between rating precision and the total number of annotated videos when crowd-sourcing VQA databases, and its effect on VQA model performance. The second talk "Rating Distribution and Personality Prediction for Image Aesthetics Assessment" discusses two issues significant to image aesthetics assessment (IAA). Compared with technical image quality assessment, IAA is expected to be much more individualized, because 'beauty is in the eye of the beholder'. In this talk, the presenter will demonstrate the objective prediction of subjective rating distribution (rather just the mean score in most existing methods) in IAA; in addition, he will also explore the use of personality prediction in IAA.

Accepted paper. The ATQAM track includes the presentation of the "EVA: An Explainable Visual Aesthetics Dataset". The work is a promising contribution to the field of AQA. While EVA is not the largest in terms of number of images, it is the most comprehensive aesthetics dataset annotated under controlled conditions. Moreover, EVA includes annotations of other aesthetics-related image attributes, such as technical quality, and studies their connection to aesthetics.

Organizing/Program Committee. The ATQAM track organizers comprise researchers from both the TQA and AQA fields from several Universities spread across Europe and Asia, such as University of Konstanz (Germany), National Chiao Tung University (Taiwan), Nanyang Technological University (Singapore), and Multimedia University (Malaysia).

We appreciate the reviewers' efforts and would like to thank the program committee for their valuable support: **Seyed Ali Amirshahi** (NTNU, Norway), **Raouf Hamzaoui** (De Montfort University, UK), **Matthias Hirth** (TU Ilmenau, Germany), **Shujun Li** (University of Kent, UK), **Hanhe Lin** (University of Konstanz, Germany), **Yuen Peng Loh** (Multimedia University, Malaysia). All paper submissions were reviewed by three members of the program committee with respect to scientific quality and the suitability to the workshop's topic.

3.2 MAST

The MAST track will be held in the second half-day of the ATQAM/MAST'20 workshop. The program features a short introduction about the workshop's motivation and relevance by the workshop organisers, three invited talks, one panel discussion and two contributed talks that present original reasearch work on understanding societal trends from media data. All submitted papers were reviewed by at least two members of the program committee in terms of novelty, scientific quality and the suitability to the workshop's topic. Submissions were accepted with at least two clear accept ratings.

Accepted papers. The MAST track program contains 2 full-length accepted original research work with novel contributions.

First, the paper by **Kexin Feng** et al. (Texas A&M University) titled *Exploring Speech Cues in Web-mined COVID-19 Conversational Vlogs* reports on the impact of COVID-19 on the population in New York using Youtube conversational vlogs as a window into their mental state.

The second paper by **Danni Chen** et al. (UCLA) is titled *Understanding Gender Stereotypes and Electoral Success from Visual Self Presentations of Politicians in Social Media.* This work explores the role that gender stereotypes play in political campaigns primarily on social media.

Program committee. MAST track organizers sincerely appreciate the reviewers' efforts and would like to thank the members of the diverse program committee for their valuable support: **Abhinav Dhall**, Monash University, Australia ; **James Kennedy**, Disney Research, US; **Jangwon Kim**, Amazon Alexa Speech, US; **Makarand Tapaswi**, INRIA, Paris; **Subarna Tripathi**, Intel AI, US; **Samuel Kim**, Canary Speech, US

Organizing committee. The MAST track organizing committee shows a healthy partnership between academia and industry. Academics from University of Warwick and University of Southern California (USC) came together to host this workshop with the research scientists at Disney Research Los Angeles and Google Los Angeles.

ACKNOWLEDGMENT

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 251654672 – TRR 161 (Project A05).