AI4TV 2019: 1st International Workshop on AI for Smart TV Content Production, Access and Delivery

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

Jorma Laaksonen
jorma.laaksonen@aalto.fi
Aalto University
Espoo, Finland

Hamed R.-Tavakoli
hamed.r-tavakoli@aalto.fi
Nokia Technologies
Espoo, Finland

Lyndon Nixon
lyndon.nixon@modul.ac.at
MODUL Technology
Vienna, Austria

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

ABSTRACT

Technological developments in comprehensive video understanding – detecting and identifying visual elements of a scene, combined with audio understanding (music, speech), as well as aligned with textual information such as captions, subtitles, etc. and background knowledge – have been undergoing a significant revolution during recent years. The workshop brings together experts from academia and industry in order to discuss the latest progress in artificial intelligence research in topics related to multimodal information analysis, and in particular, semantic analysis of video, audio, and textual information for smart digital TV content production, access and delivery.

Related Workshop Proceedings are available in the ACM DL at: https://dl.acm.org/citation.cfm?id=3347449

CCS CONCEPTS

• **Information systems → Multimedia information systems**; **Multimedia databases**; **Multimedia content creation**; • **Computing methodologies → Artificial intelligence**.

KEYWORDS

Artificial Intelligence; TV content production; TV content delivery; TV content analysis; TV content annotation; intelligent multimedia

1 INTRODUCTION

New scientific breakthroughs in video understanding through the application of AI techniques along with the increase in the volume of multimedia content and more computational power have led to significant improvements in automated video description and have opened fresh avenues for the seamless combination of multiple modalities' analysis. The main goal of the workshop is to promote AI techniques for multimedia analysis to enable smarter content

production, access and delivery with the emphasis on large TV and radio program archives.

2 WORKSHOP SCOPE

The goal of this workshop is to bring together experts from academia and industry in order to discuss the latest progress in artificial intelligence research for multimodal information analysis, and in particular, semantic analysis of video, audio, and textual information for smart digital TV content production, access and delivery. Such topics include, but are not limited to, the following multimedia analysis techniques for streamed TV and radio programmes as well as TV archives (recorded content):

- Multimodal content analysis: scene segmentation, people and concept recognition, topic identification using video, audio, and/or (textual) metadata
- Embeddings for Multimedia Knowledge Graph
- Use or adaptation of multimedia description models or vocabularies for machine learning / neural networks
- Combination of AI and external knowledge (graphs) for improved multimedia analysis
- Automatic multimedia summarization and remixing
- Automatic deep captioning
- Interactive multimodal search and browsing in archives
- Hyperlinking and enrichment of TV content
- Breaking the language barrier of TV content using multimodal translation
- Comparative evaluations of AI techniques for multimodal analysis tasks
- Creation of multimedia benchmarks for AI evaluations
- Gender studies on TV and Radio programmes

3 WORKSHOP PROGRAMME

The workshop programme includes two keynote talks, 6 full papers and 4 demo papers.

The two keynote talks are:

(1) *Annotation automation to support dynamic exploration and creative retrieval of audiovisual archives*, delivered by Dr. Johan Oomen (NISV, The Netherlands).
(2) *AI gets creative*, delivered by Dr. Marta Mrak (BBC, UK).

The oral presentation sessions include six full papers:

(1) *L-STAP: Learned Spatio-Temporal Adaptive Pooling for Video Captioning* proposes a method that combines frame-by-frame spatial CNN processing and attention-based temporal analysis for automatic video captioning [4].

(2) *A Stepwise, Label-based Approach for Improving the Adversarial Training in Unsupervised Video Summarization* presents improvements to the SUM-GAN model, and shows excellent results on the SumMe and TVSum datasets [1].

(3) *On the Robustness of Deep Learning Based Face Recognition* first analyses the impact of image degradations on face detection and then proposes approaches to improve open-set face recognition [2].

(4) *Gender Representation in French Broadcast Corpora and Its Impact on ASR Performance* studies the gender imbalance in TV and radio broadcasts and shows how the under-representation of women leads to performance to decrease on recognition of female speech [5].

(5) *AI for audience prediction and profiling to power innovative TV content recommendation services* presents results from experiments with AI techniques to improve the accuracy of audience prediction through the addition of viewer, event and content features into the prediction model [7].

(6) *Data-driven summarization and synchronized second-screen enrichment of cycling races* focuses on data-driven cycling race summarization, which allows end-users to query for personalized stories of a race, and display of cycling heritage multimedia in a second screen [10].

The demo session includes four demos:

(1) *Examples of uses of Artificial Intelligence in video archives* presents different applications of artificial intelligence techniques applied to video archives such as face recognition, visual search and classification based either on faces, objects orlandmarks [9].

(2) *Automatically adapting and publishing TV content for increased effectiveness and efficiency* presents demos from the ReTV project showing AI powered content adaptation and recommendation through a TV program chatbot (4u2) and dynamic in-stream personalized video insertion (content switch) [8].

(3) *Personalized Movie Trailer Using Thumbnail Containers* introduces a client-driven method to generate personalized movie trailers for each user based on their interest by selecting personalized actions [6].

(4) *A workstation for real-time processing of multi-channel TV* presents the architecture of a workstation for the real-time processing of up to 32 channels of TV and its application to video copy detection [3].

4 WORKSHOP COMMITTEES

General Chairs:
  Raphaël Troncy (EURECOM)
  Jorma Laaksonen (Aalto University)
  Hamed R.-Tavakoli (Nokia Technologies)
  Lyndon Nixon (MODUL Technology)
  Vasileios Mezaris (CERTH-ITI)

Program Committee Members:
  Olivier Aubert (University of Nantes)
  Werner Bailer (Joanneum Research)
  Jean Carrive (INA)
  Mikko Kurimo (Aalto University)
  Tiina Lindh-Knuutila (Lingsoft)
  Symeon Papadopoulos (CERTH)
  Jörg Tiedemann (University of Helsinki)
  Dieter Van Rijsselbergen (Limecraft)
  Miggi Zwicklbauer (Rundfunk Berlin-Brandenburg)

ACKNOWLEDGEMENTS

REFERENCES

[1] Evlampios Apostolidis, Alexandros Metsai, Eleni Adamantidou, Vasileios Mezaris, and Ioannis Patras. 2019. A Stepwise, Label-based Approach for Improving the Adversarial Training in Unsupervised Video Summarization. In *Proc. Int. Workshop on AI for Smart TV Content Production, Access and Delivery (AI4TV 2019) at ACM Multimedia 2019*. ACM, Nice, France.

[2] Werner Bailer and Martin Winter. 2019. On the Robustness of Deep Learning Based Face Recognition. In *Proc. Int. Workshop on AI for Smart TV Content Production, Access and Delivery (AI4TV 2019) at ACM Multimedia 2019*. ACM, Nice, France.

[3] Mathieu Delalandre. 2019. A workstation for real-time processing of multi-channel TV. In *Proc. Int. Workshop on AI for Smart TV Content Production, Access and Delivery (AI4TV 2019) at ACM Multimedia 2019*. ACM, Nice, France.

[4] Danny Francis and Benoit Huet. 2019. L-STAP: Learned Spatio-Temporal Adaptive Pooling for Video Captioning. In *Proc. Int. Workshop on AI for Smart TV Content Production, Access and Delivery (AI4TV 2019) at ACM Multimedia 2019*. ACM, Nice, France.

[5] Mahault Garnerin, Solange Rossato, and Laurent Besacier. 2019. Gender Representation in French Broadcast Corpora and Its Impact on ASR Performance. In *Proc. Int. Workshop on AI for Smart TV Content Production, Access and Delivery (AI4TV 2019) at ACM Multimedia 2019*. ACM, Nice, France.

[6] Ghulam Mujtaba and Eun-Seok Ryu. 2019. Personalized Movie Trailer Using Thumbnail Containers. In *Proc. Int. Workshop on AI for Smart TV Content Production, Access and Delivery (AI4TV 2019) at ACM Multimedia 2019*. ACM, Nice, France.

[7] Lyndon Nixon, Krzysztof Ciesielski, and Basil Philipp. 2019. AI for audience prediction and profiling to power innovative TV content recommendation services. In *Proc. Int. Workshop on AI for Smart TV Content Production, Access and Delivery (AI4TV 2019) at ACM Multimedia 2019*. ACM, Nice, France.

[8] Basil Philipp, Krzysztof Ciesielski, and Lyndon Nixon. 2019. Automatically adapting and publishing TV content for increased effectiveness and efficiency. In *Proc. Int. Workshop on AI for Smart TV Content Production, Access and Delivery (AI4TV 2019) at ACM Multimedia 2019*. ACM, Nice, France.

[9] Antoine Mercie rand Sébastien Ducret, Charlotte Bürki, and Léonard Bouchet. 2019. Examples of uses of Artificial Intelligence in video archives. In *Proc. Int. Workshop on AI for Smart TV Content Production, Access and Delivery (AI4TV 2019) at ACM Multimedia 2019*. ACM, Nice, France.

[10] Steven Verstockt, Erik Mannens, and Jelle De Bock. 2019. Data-driven summarization and synchronized second-screen enrichment of cycling races. In *Proc. Int. Workshop on AI for Smart TV Content Production, Access and Delivery (AI4TV 2019) at ACM Multimedia 2019*. ACM, Nice, France.