



Special issue on content-based multimedia indexing in the era of artificial intelligence

Stevan Rudinac¹ · Jenny Benois-Pineau² · Stephane Marchand-Maillet³

Received: 8 March 2021 / Revised: 8 March 2021 / Accepted: 6 April 2021 /

Published online: 1 May 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Content-based approaches to multimedia analysis, indexing, search and exploration have long been a cornerstone of facilitating access to multimedia and multimodal collections. Modeling the properties of increasingly heterogeneous multimedia items and their complex relations in parallel to indexing is a challenging task. Recent advances in artificial intelligence (AI) and, in particular, deep learning have given a new boost to practically every aspect of multimedia research.

This special issue aims at gathering and presenting new advances in the tasks related to the applications of AI in content-based multimedia indexing. It follows the 17th International Conference on Content-Based Multimedia Indexing (CBMI '19), held on 4-6 September 2019 in Dublin, Ireland. Next to including revised and extended versions of some of the best papers presented at the conference, the special issue features a number of new contributions from multimedia community at large.

We have received 23 valid submissions, of which we finally accepted 8 to the special issue. While AI has been at the core of multimedia tools and applications from their early days, the papers accepted to our special issue show that this relationship has become even stronger. In addition, they demonstrate that the recent breakthroughs in deep learning have further reinforced synergistic collaboration between multimedia and related fields, from computer vision and information retrieval to computer graphics. Below we briefly describe the accepted papers.

Federico Magliani, Tomaso Fontanini and Andrea Prati present their work titled “Bag of indexes: a multi-index scheme for efficient approximate nearest neighbor search” (<https://doi.org/10.1007/s11042-020-10262-4>). The main contribution of the paper is a novel indexing

✉ Stevan Rudinac
s.rudinac@uva.nl

Jenny Benois-Pineau
jenny.benois-pineau@u-bordeaux.fr

Stephane Marchand-Maillet
stephane.marchand-maillet@unige.ch

¹ University of Amsterdam, Amsterdam, Netherlands

² University of Bordeaux, Bordeaux, France

³ University of Geneva, Geneva, Switzerland

approach for approximate nearest neighbour search, which yields a state of the art performance in terms of retrieval accuracy, while ensuring low response time.

Rahma Abed, Sahbi Bahroun and Ezzeddine Zagrouba present the paper titled “KeyFrame extraction based on face quality measurement and convolutional neural network for efficient face recognition in videos” (<https://doi.org/10.1007/s11042-020-09385-5>). In this paper the authors propose an approach to video keyframe extraction based on image suitability for reliable face recognition.

The paper “Average biased ReLU based CNN descriptor for improved face retrieval” is contributed by Shiv Ram Dubey and Soumendu Chakraborty (<https://doi.org/10.1007/s11042-020-10269-x>). The authors propose a novel rectifier function that improves image representation learnt in the last layers of a deep neural network by utilizing some of the information normally discarded by ReLU while, at the same time, discarding irrelevant positive information.

Adnan Muhammad Shah, Xiangbin Yan, Salim Khan, Waqas Khurram and Qasim Raza Khan contribute the paper titled “A multi-modal approach to predict the strength of doctor’s patient relationships” (<https://doi.org/10.1007/s11042-020-09596-w>). Through extensive evaluation of their proposed approach based on the data collected from a physician review website, the authors show that filtering irrelevant information and utilizing multiple modalities (i.e. text and images) leads to a more reliable estimation of doctor-patient relationship strength.

The special issue features an extended version of the best CBMI’19 paper titled “Learning accurate personal protective equipment detection from virtual worlds” by Marco Di Benedetto, Fabio Carrara, Enrico Meloni, Giuseppe Amato, Fabrizio Falchi and Claudio Gennaro (<https://doi.org/10.1007/s11042-020-09597-9>). The authors address a challenging problem of limited training data availability for recognizing appropriate use of personal safety equipment. They show the advantages of synthesizing large collections of photo-realistic images for training deep learning models and then performing domain adaptation with a few real-world images.

Youshan Zhang and Brian D. Davison present their work on “Domain adaptation for object recognition using subspace sampling demons” (<https://doi.org/10.1007/s11042-020-09336-0>). The authors tackle the problem of domain adaptation in object recognition tasks by proposing an efficient approach to learning and evaluating the quality of intermediate subspace features, designed to bridge the gap between the source and target domains.

The paper titled “Video object detection algorithm based on dynamic combination of sparse feature propagation and dense feature aggregation” is contributed by Danyang Cao, Jinfeng Ma and Zhixin Chen (<https://doi.org/10.1007/s11042-020-09827-0>). The proposed approach specifically focuses on efficient, real-time object detection, while also addressing the issues inherent to video content, such as motion blur, object deformation and occlusion.

Finally, Xin Shi, Huijuan Chen and Xueqing Zhao present their work on “REBOR: A New Sketch-based 3D Object Retrieval Framework Using Retina Inspired Features” (<https://doi.org/10.1007/s11042-021-10618-4>). The proposed approach, inspired by the properties of human vision, facilitates accurate retrieval of 3D models based on hand-drawn sketch queries.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.