IEEE Transactions on Circuits and Systems for Video Technology

Special Issue on Segment Anything for Videos and Beyond

Summary

Segment Anything Model (SAM) has achieved remarkable progress in advancing semantic segmentation into the era of foundation models. By training on >1 billion ground-truth segmentation masks in >11 million natural images, SAM can produce high-quality segmentation results based on different types of prompts (e.g., bounding boxes, points, texts) or in a fully automatic manner. More importantly, SAM shows the promise of a wide applicability to various image segmentation tasks, without the previously-needed re-training or fine-tuning. Thus, the emergence of SAM has generated extremely active discussions in the computer vision community; it offers a paradigm shift comparable to chatGPT for natural language processing. Though showing great potential across diverse nature image tasks, applying SAM to video tasks poses new challenges and opportunities. For example, how can SAM ensure temporal consistency and coherence of the masks across video frames? How can SAM exploit cross-frame context to improve segmentation performance? How to evaluate the effectiveness and usefulness of SAM for downstream video applications? How to improve the efficiency and scalability of SAM for large-scale video data?

The aforementioned discussion merely scratches the surface of the innovative opportunities available in developing foundation models in the video domain. This special issue on "**Segment Anything for Videos and Beyond**" aims at promoting cutting-edge research in exploring the potential of SAM in video tasks and even beyond, and offers a timely collection of works to benefit researchers and practitioners.

Scope

This special issue welcomes original and high-quality submissions that present novel methods, benchmarks, applications, or surveys; the topics of interest include but are not limited to:
 • Video segmentation with SAM
 • Video object tracking with SAM
 • Video understanding with SAM
 • Video generation with SAM
 • Video compression with SAM
 • Video editing with SAM

- Video captioning with SAM
- Video processing with SAM

Important Dates:

Submission deadline: December 1, 2023
First review notification: January 20, 2024
Revision submission due: March 1, 2024
Second round review: March 1 to June 20, 2024
Notification of acceptance/rejection: June 20, 2024

Guest Editors:

Wenguan Wang, Zhejiang University, China (Email: wenguanwang.ai@gmail.com)
Hengshuang Zhao, The University of Hong Kong, China (Email: hszhao@cs.hku.hk)
Xinggang Wang, Huazhong University of Science and Technology, China (Email: xgwang@hust.edu.cn)
Fisher Yu, ETH Zurich, Switzerland (Email: i@yf.io)
David Crandall, Indiana University, USA (Email: djcran@indiana.edu)