Master2/Engineering degree Internship: Intelligent interaction for deep learning-based video segmentation

Contact: Claire-Hélène Demarty (Claire-helene.demarty@interdigital.com) and Frédéric Lefevbre (Frederic.Lefevbre@interdigital.com)

Summary:

This internship proposal targets the development of deep learning approaches for high-end visual effects, and in particular in the context of rotoscoping.

Because the requested spatial accuracy and temporal consistency of the segmentation are extremely strong in the rotoscoping task, the interaction with the user (roto-artist) is essential to guide and refine the final object segmentation.

On the other hand, how to efficiently incorporate user interaction while learning deep models is still a non-solved problem. The internship will therefore consist in studying smart ways of taking the user interaction into account for the task of video segmentation.

Resulting algorithms might be integrated in a professional VFX software.

Dates: 6 months (Starting from Jan-Feb. 2020)

Skills: machine learning, deep learning, computer vision, video/image processing, PyTorch, TensorFlow or Keras deep learning frameworks, Python or C++

Keywords: machine learning (deep learning), video processing, computer vision, interaction, segmentation, tracking, rotoscoping.

Objective

In the past years deep learning has considerably gained in popularity and was applied to various tasks of multimedia analysis, processing and generation. The goal of this internship is to develop new deep learning-based approaches for high-end visual effects. More precisely, the internship will be held under the umbrella of a project targeting to develop tools that would highly simplify the video rotoscoping process. The rotoscoping process consists in extracting an object of interest from its background in order to replace this background by another one in final compositing, and it is usually performed manually by roto artists. Hence, rotoscoping is an interactive editing process which requires a large amount of processing time as the requested segmentation accuracy is critical. As such, the global goal of the Technicolor

project is to propose new interactive video editing tool that would help the artists to accelerate and improve the rotoscoping. In this context, the internship will focus on how to efficiently incorporate user interaction in the task of video segmentation.

Task Description

What is expected from this internship can be divided in:

- Studying relevant state of the art on interaction in the context of deep learning

- Designing and implementing deep learning architectures that leverage different types of user interaction for the task of segmentation.

- Evaluating the proposed deep learning methods on the relevant benchmarks and comparing them with other state of the art methods

- In case a successful solution is achieved, integrating the architecture into the corresponding VFX plugin within the project

As outputs of this internship, a paper submission to an international conference, and a potential integration of one final solution into the plugin would be expected.

Working environment

InterDigital is one of the world's largest pure research, innovation, and licensing companies, with more than 300 engineers around the world. Our focus is on research and development with pervasive impact: mobile technologies that underpin smartphones, networks and services via global standards, and video technologies that are the foundation for today's most popular products and services. Wireless and video – arguably the two most impactful technologies today.

On the wireless side, InterDigital has been a pioneer for four decades, with our engineers designing and developing a wide range of advanced technologies that are used in digital cellular and wireless products and networks, including 2G, 3G, 4G and IEEE 802-related products and networks. Today, we're a leader in 5G research and beyond, a thought leader in our industry and, over the course of the last two decades, the source of more than 30,000 contributions to key global standards.

In video, our existing efforts in key standards have been dramatically expanded with the addition of the Research & Innovation operation of global market leader Technicolor in 2019. With a video R&D heritage spanning decades, while yielding one of the leading innovation portfolios in the industry, what is now InterDigital R&I is a world leader in video research, pioneering new capabilities and making more than 100 contributions to key global video standards. In 2019, the company incorporated a world-class video and AI research team,

formerly Technicolor R&I, and an established portfolio of video expertise to expand our work in wireless and video technologies and consumer electronics.

Our track record of research & development is matched by our fair licensing practices, which are a model for the industry, and our willingness to partner with virtually anybody in developing new capabilities that will improve technology for consumers and businesses around the world.

InterDigital R&I France in Rennes is the largest of the InterDigital research centers all over the world with more than 160 employees and 25 nationalities. Formerly the biggest research center of Technicolor, it has been acquired in June 2019. Our innovations fuel Interdigital's technology, and at the same time our researchers publish papers in the top academic conferences. We actively collaborate with numerous universities around the globe. We welcome post-docs, visiting faculty, PhD students and numerous graduate student interns to come and spend time with us in our lab for periods of time from 3 to 36 months.

Since the acquisition by Interdigital, the Technical Area "Content Processing" in which the internship is proposed is continuing its developments and research with Technicolor as final client.

Technicolor is an industry leader in the production of video content for movies, TV, advertising, games and more. The company provides production, postproduction, and distribution services to content creators, network service providers and broadcasters. Technicolor is driving innovation in 3D movie making, 3D entertainment experiences for the home, automatic meta-data creation, second screen experiences, media management and the digital home of the future.

Context

Within the Image Science Laboratory at InterDigital R&I Rennes, France (former Technicolor R&I), this internship is proposed in the Technical Area "Content Processing" gathering more than 20 researchers, engineers, PhDs, and Postdocs coming from 10 different countries. The main line of research in the team is to understand, organize, and enhance content, both professional and user-generated.