EPFL/STI/IEL/GR-EB, Station 11, CH-1015 Lausanne, Switzerland
(+41) 76 276 0454, pavel.korshunov@epfl.ch, http://mmspg.epfl.ch/page-73810-en.html

My research activities and contributions cover a wide area of multimedia research, including video streaming, video and image analysis, subjective, objective, and crowdsourcing assessments of image and video, video capture and representation, image and video compression, immersive technologies such as High Dynamic Range (HDR) imaging, Ultra High Definition (UHD), and 3D video, video capturing with mini-drones, media security and privacy issues, privacy protection and evaluation.

During my PhD years, I was developing distributed multimedia-rich applications focusing on such issues as scalability, usability, efficiency and robustness. For instance, one of my PhD projects and my internship focused on developing systems for control, capture, and layout composition of live video streams. Such systems can be used in video surveillance, video conferencing, and for recording video presentations. I was an active participant in development of Plasma, a scripting language and a tool for managing and layout composition of several video streams. Also, I developed a web-based system with multiple video sensors used for presentation capture, control and playback. The system is based on NCast Telepresenter box[1] and comprises of three projectors, as well as four PTZ and four panoramic video cameras.

In my PhD dissertation, I have studied the bandwidth minimization problem in the context of automated distributed systems, which stream video over the IP-network. The examples of such systems include video surveillance, sensor networks, remote autonomous vehicles, social web-based applications, and mobile and biometrics applications. Focusing on video analysis algorithms used in such applications, I have studied the relationship between video quality and algorithms' accuracy. I have demonstrated that conventional human-oriented requirements to video quality, its assessment and measurement, and video encoders do not suit video analysis algorithms well. It is evident that, with regards to video quality, computer vision is different from human vision.

I have also extended the work comparing computer vision and human vision during my postdoctorate, where I studied the difference between how people and video analytics are affected by visual privacy protection filters and, in more general, how visual information is perceived and analyzed by humans and computers. Based on these studies, I proposed and validated several privacy protection filters (warping-, morphing-, and false color-based filters) that can be used for reversible obfuscation of personal visual information. To be able to evaluate these methods, I have created image and video datasets, which were used in several MediaEval grand challenges that I also helped to organize.

A large amount of postdoctorate work was also focused on various aspects of subjective perception, specifically, how the perception and focus of attention are affected by new immersive technologies, such as HDR, UHD, and 3DTV. I also participated in several studies of the new standards and compression algorithms, including HEVC and VP9 video codecs, as well as HDR image and video compression. Taking into account the evaluation results, I also proposed a backward-compatible compression scheme for HDR images, which is partially implemented in the new JPEG XT standard for HDR images, and participated in development of backward-compatible compression for 3D images.

As a senior postdoctoral researcher, I led several research activities of the MMSPG lab, represented EPFL in FP7 NoE VideoSense project, and actively contributed to COST actions Qualinet (IC1003), HDRi (IC1005), and De-identification for privacy protection in multimedia content (IC1206). I have also regularly participated in preparing project proposals, supervising students, and teaching.

Below I discuss my research goals and research philosophy and propose possible future directions that I would like to work in.

[1] http://www.ncast.com/

Research Philosophy

I view computer science, if simply put, as a "middleware" between science and engineering. Hence, I am interested in practical problems that can be structured and formalized in such a way that can lead to general implications in science and engineering. I use the following approach when attacking a problem: identify problem's type, propose a solution, extend it into a formal framework, and then apply the framework to a similar class of problems in several joint areas. I also consider computer science as a discipline that grew out of mathematics, thus, I believe that solutions to the research problems in computing should be based on fundamental mathematical knowledge and traditions; one should have a concise, logical, and simple attitude when proposing a solution.

Research Goal

My long-term research goal is to make a notable contribution in multimedia research that would make video and image data as common and ordinary as text data. Today, text is used as the main medium for automated representation of information. Sophisticated tools are developed for analysis, search, transmission, and manipulation of textual information. However, tools for video and images are not yet available at a similar level of sophistication and accuracy. Therefore, developing more reliable and accurate analysis tools for video and images, as well as advancing the understanding of visual perception, intuitive integration of visual medium into existing systems (web applications, games, collaborative tools, etc.), tools for security and privacy protection of visual information, and seamless accessibility and transmission of video are challenging long-term research problems.

Currently, integration of different cross-domain multimedia services in a single system is receiving attention from researchers (video/image search, automated surveillance, smart environments, etc.) and companies (Google, Netflix, Facebook, and others). Still, multimedia research is mostly focused on several fairly disintegrated areas, including, computer vision, networking, multimedia presentation and interaction, 3D modeling, etc. However, I believe that the cross-domain problems are important, challenging, and are the most practical. Until now, my work focused on these types of problems, and I would like to continue finding and solving cross-domain challenges.

Future Work

I am interested in developing applications of computer vision and video analysis, video capture, compression, and streaming, wearables their integration into multimedia-based system, as well as, studying problems with 'human in the loop', including visual perception, visual security and personal privacy, and crowdsourcing and subjective assessment. In particular, the following research directions I would be interested to work in:

1. Efficient and comprehensive analysis of large collections of video and image data is a very challenging yet an important task in video sharing, video surveillance, and video-on-demand systems. Automatic object detection and tracking, detection of similar videos, detection of video copyright violation, pornography and violence detection remain to be challenging problems in large-scale systems. Besides designing efficient analysis algorithms, an important challenge is to build an appropriate system infrastructure (cheap, distributed, robust, and reliable) that can be adaptive to variable demand for computational, storage, and network resources and protect personal information about people appearing in these video and image data.

2. Collaborative smart environments attract researchers from many inter-disciplinary fields of computer science and engineering. Typically, smart environments consist of various types of sensors, including multimedia sensors (videos, microphones, mini-drones), connected to IP-network. Analysis, behavior, and control decisions of such system are based on the information combined from different sources including video and images from the sensors, location information, and information from social networks and sharing platforms. Typical examples of smart environments include monitoring systems (multi-camera surveillance systems and environment monitoring systems), intelligent homes that can automatically control different systems and appliances in a house, and intelligent care systems for patients, elders, and people in need. Many interesting research problems are related to these systems, such as real-time data fusion from heterogeneous sensors, designing architecture that is general for

different configurations of sensors and devices, scalability problem (especially for systems with media sensors), and the problems related to information security and privacy that take into account both context and content.

3. I would also like to explore wearable sensors that measure heart rate, respiration, skin conductivity, muscle contraction, etc. These sensors can help approximate an emotional and physical state of a person, and together with the contextual information, such as recorded preferences, GPS, timer, accelerometer, and phone camera, they can be used to guide and adapt TV and music preferences, smart home systems, fitness routines, etc. Principles of privacy by design will also be interesting to explore in relation to the personal sensors, as well as, main use cases, applications of these devices, and users' subjective perception.