# POTENTIAL PROJECTS

Dr. Wenbin Li
Department of Computer Science
University of Bath
wenbin.li AT bath.edu

Camera is fantastic sensor which brings human a rich set of information. Extracting and analyzing such information is the core issue of computer vision. My research interests have been in end-to-end pipeline for computational capture, as well as interior scene understanding and fabrication. I will briefly introduce some potential projects that are most interesting for me and potentially far-reaching:

**Multi-sensory Vision and Capture.** Many problems studied in computer vision have their roots in the desire to emulate certain cognitive capabilities of a human. However, those vision research mainly stands on a small amount of still images, which are inadequate to represent the biological vision process where multiple sense channels (intensity, voice, motion and depth) are involved for perception. With advances in electronics technology, this limited paradigm from single visual channel may no longer be the only source of inspiration for computer vision. Further investigation on other sense channels and the related electronic sensors/devices is promising because it would bring a rich set of information that describe images from different perspectives. Once such multi-sensory information is properly obtained, the fundamental question is then to identify or invent new intelligence and to explore other possibilities using a large collections of information from such multiple sense channels.

**Dynamic Interior Scene Understanding and Fabrication.** Interior scene is one of the most typical scenarios in daily life. The ability to understand such a scene depicted in a set of static images with other multi-sensory information has been an essential computer vision problem in practice. The dominant view in the community is to identify the semantic properties of still objects, as well as structures of scenes. However, this level of understanding is rather inadequate since the real-world interior scene is often dynamic where unknown number of objects (e.g. human, animals etc.) might be moving independently. If we could reason about the semantic coherence of those dynamic properties, it may enable richer applications with regards to longlife SLAM and natural interior scene digitalization.

**Human in Loop Vision System.** It is the fact that approaches of computer vision and graphics have been widely applied to some daily tasks of film and gaming companies. However, most of those tasks still mainly rely on human editing, although many related unsupervised algorithms have offered similar accuracy as what human can perform. It is because that industrial standard is more strict beyond the best performance the state-of-the-art unsupervised methods could give. Given this context, the further research efforts must go towards not only improving regular algorithmic metrics, but also the optimizing human operation and interaction to system. Such a research issue requires further investigation on user behaviors learning, as well as interdiscipline composed of semi-supervised computer vision and human-computer interaction.

For those projects mentioned above, an excellent mathematics background and good coding skills (e.g. in any of Python, C++ and Matlab etc.) are essential. A knowledge of neural networks (e.g. CNN, LSTM etc.) with experience in multi-sensory data presentation and analyses could be an important plus here.