April 22nd 2021

High Efficiency Light Field Image Compression - Hierarchical Bit Allocation and Shearlet-based View Interpolation

Author doctoral thesis: **Waqas Ahmad**
STC Research Centre
Mid Sweden University

Abstract

Over the years, the pursuit of capturing the precise visual information of a scene has resulted in various enhancements in digital camera technology, such as high dynamic range, extended depth of field, and high resolution. However, traditional digital cameras only capture the spatial information of the scene and cannot provide an immersive presentation of it. Light field (LF) capturing is a new-generation imaging technology that records the spatial and angular information of the scene. In recent years, LF imaging has become increasingly popular among the industry and research community mainly for two reasons: (1) the advancements made in optical and computational technology have facilitated the process of capturing and processing LF information and (2) LF data have the potential to offer various post-processing applications, such as refocusing at different depth planes, synthetic aperture, 3D scene reconstruction, and novel view generation. Generally, LF-capturing devices acquire large amounts of data, which poses a challenge for storage and transmission resources. Off-the-shelf image and video compression schemes, built on assumptions drawn from natural images and video, tend to exploit spatial and temporal correlations. However, 4D LF data inherit different properties, and hence there is a need to advance the current compression methods to efficiently address the correlation present in LF data.

In this thesis, compression of LF data captured using a plenoptic camera and multi-camera system (MCS) is considered. Perspective views of a scene captured from different positions are interpreted as a frame of multiple pseudo-video sequences and given as an input to a multi-view extension of high-efficiency video coding (MV-HEVC). A 2D prediction and hierarchical coding scheme is proposed in MV-HEVC to improve the compression efficiency of LF data. To further increase the compression efficiency of views captured using an MCS, an LF reconstruction scheme based on shearlet transform is introduced in LF compression. A sparse set of views is coded using MV-HEVC and later used to predict the remaining views by applying shearlet transform. The prediction error is also coded to further increase the compression efficiency. Publicly available LF datasets are used to benchmark the proposed compression schemes. The anchor scheme specified in the JPEG Pleno common test conditions is used to evaluate the performance of the proposed scheme. Objective evaluations show that the proposed scheme outperforms state-of-the-art schemes in the compression of LF data captured using a plenoptic camera and an MCS. Moreover, the introduction of shearlet transform in LF compression further improves the compression efficiency at low bitrates, at which the human vision system is sensitive to the perceived quality.

The work presented in this thesis has been published in four peer-reviewed conference proceedings and two scientific journals. The proposed compression solutions outlined in this thesis significantly improve the rate-distortion efficiency for LF content, which reduces the transmission and storage resources. The MV-HEVC-based LF coding scheme is made publicly available, which can help researchers to test novel compression tools and it can serve as an anchor scheme for future research studies. The shearlet-transform-based LF compression scheme presents a comprehensive framework for testing LF reconstruction methods in the context of LF compression.

Sensible Things that Communicate

PART OF MID SWEDEN UNIVERSITY