Coding for improved perceived quality of 2D and 3D video over heterogeneous networks.

Linda S. Karlsson

Department of Information Technology and Media
Mid Sweden University

Thesis No. 87
Sundsvall, Sweden
2010

Abstract

The rapid development of video applications for TV, the internet and mobile phones is being taken one step further in 2010 with the introduction of stereo 3D TV. The 3D experience can be further improved using multiple views in the visualization. The transmission of 2D and 3D video at a sufficiently perceived quality is a challenge considering the diversity in content, the resources of the network and the end-users. Two problems are addressed in this thesis. Firstly, how to improve the perceived quality for an application with a limited bit rate. Secondly, how to ensure the best perceived quality for all end-users in a heterogeneous network.

A solution to the first problem is region-of-interest (ROI) video coding, which adapts the coding to provide a better quality in regions of interest to the viewer. A spatio-temporal filter is proposed to provide codec and standard independent ROI video coding. The filter reduces the number of bits necessary to encode the background and successfully re-allocate these bits to the ROI. The temporal part of the filter reduces the complexity compared to only using a spatial filter. Adaption to the requirements of the transmission channel is possible by controlling the standard deviation of the filter. The filter has also been successfully applied to 3D video in the form of 2D-plus-depth, where the depth data was used in the detection of the ROI.

The second problem can be solved by providing a video sequence that has the best overall quality. Hence, the best quality for each part of the network and for each 2D and 3D visualization system over time. Scalable video coding enables the extraction of the parts of the data to adapt to the requirements of the network and the end-user. A scheme is proposed in this thesis that provides scalability in the depth and view domain of multi-view plus depth video. The data are divided into enhancement layers depending on the content's distance to the camera. Schemes to divide the data into layers within a view and between adjacent views have been analysed. The quality evaluation indicates that the position of the layers in depth as well as the number of layers should be determined by analysing the depth distribution. The front-most layers in adjacent views should be given priority over the others unless the application requires a high quality of the center views.