Today's imaging technologies can now capture texture maps (e.g., color images) and depth maps (per-pixel distance between objects in the 3D scene and camera) from multiple viewpoints simultaneously. Given this rich set of input images, the objective is to develop key technologies that enable a range of interactive 3D imaging applications. In particular, we focus on three aspects of an end-to-end visual communication system: i) coding of 3D images, ii) transmission of 3D data over data networks, and iii) restoration of virtual views given transmitted 3D data. For coding of 3D data, we have developed a multi-resolution edge-adaptive image codec based on graph Fourier transform (GFT). For transmission of 3D data, we have developed a strategy to exploit the data redundancy inherent in the representation of multiview texture-plus-depth format for loss recovery. For restoration of 3D images, we have studied the use of graph Laplacian regularizer as signal prior for regularization.
The advance of image sensing technologies means that one can now readily capture texture maps (e.g., color images) and depth maps (per-pixel distance between objects in the 3D scene and camera) from multiple viewpoints simultaneously. Unlike conventional single-view video capture where the decoder can only play back video exactly as it was recorded at the encoder with no interactivity, having this rich set of multi-view input images enables a wide range of interactive 3D imaging applications at the decoder; e.g., free viewpoint TV (where observation viewpoints of an interested 3D scene can be freely selected), immersive visual communication (where a user’s head movements would induce corresponding change in the rendered viewpoint), etc.

The objective of this research is thus to explore signal processing technologies that can help bring these 3D imaging applications to reality.

The technical goal of this research is to develop key technologies that enable interactive 3D imaging applications. In particular, we focus on three important aspects of an end-to-end visual communication system: 1) representation and coding of 3D images, 2) robust transmission of 3D data over heterogeneous networks, and 3) signal restoration and rendering of virtual views given transmitted 3D data.

For representation and coding of 3D data, we have focused on the compression of depth images. The main technique we have developed is a multi-resolution edge-adaptive image codec based on graph Fourier transform (GFT). The basic idea is to first encode detected edges in a depth map in high resolution using arithmetic edge coding (AEC), then perform edge-adaptive low-pass filtering and down-sampling, and encode the low-resolution block using GFT.

For transmission of 3D data over networks, we have developed a strategy to exploit the data redundancy inherent in the representation of multiview texture-plus-depth format for loss recovery, resulting in a higher synthesized view quality at the decoder.

For restoration of 3D images (e.g., image interpolation, disocclusion hole filling, etc.), we have studied the use of graph Laplacian regularizer as signal prior for regularization. Together with a signal fidelity term in $L_2$-norm, we have shown that this problem can be solved efficiently via iterative reweighted least square (IRLS), where each iteration is an unconstrained quadratic programming problem with a closed form solution.

For coding of depth images, we have shown that our multi-resolution edge-adaptive image codec can outperform H.264 intra coding by up to 6.8dB in average PSNR. An example of our coding results is shown in the following PSNR versus bitrate plot for Teddy sequence.

For robust transmission of multiview texture-plus-depth video over lossy wireless links, we have shown that our proposed unequal forward error correction (FEC) protection scheme, together with decoder loss recovery that exploits inter-view redundancy for effective concealment (marked “patch-based” in the figure), can outperform other robust streaming proposals by up to 6.2dB.

For coding of depth images, we have shown that our multi-resolution edge-adaptive image codec can outperform H.264 intra coding by up to 6.8dB in average PSNR. An example of our coding results is shown in the following PSNR versus bitrate plot for Teddy sequence.

For robust transmission of multiview texture-plus-depth video over lossy wireless links, we have shown that our proposed unequal forward error correction (FEC) protection scheme, together with decoder loss recovery that exploits inter-view redundancy for effective concealment (marked “patch-based” in the figure), can outperform other robust streaming proposals by up to 6.2dB.
different transmission / loss recovery

(3) For image interpolation of DIBR-synthesized view, our devised local graph-based interpolation method using a graph Laplacian regularizer has shown remarkably natural image quality compared to naïve methods such as linear and bi-cubic interpolation. See Fig.3 for an example of actual DIBR-synthesized image and corresponding interpolated image.

Fig.3: (left) DIBR-synthesized image for virtual viewpoint closer to the 3D scene, (right) interpolated image using our proposed graph-based algorithm.


DOI: 10.1109/TMM.2015.2389714


DOI: 10.1109/JSTSP.2014.2330332


DOI: 10.1109/TIP.2014.2378055


DOI: 10.1109/TMM.2014.2332139


DOI: 10.1016/j.image.2014.03.007


DOI: 10.1109/TMM.2014.2299768


DOI: 10.1109/TMM.2013.2281024


DOI: 10.1109/TMM.2013.2272918


DOI: 10.1109/LSP.2013.2277595


DOI: 10.1109/TCSVT.2013.2269019

Thomas Maugey, Ismael Daribo, Gene Cheung, Pascal Frossard, "Navigation Domain Partitioning for Interactive

5. 主な発表論文等
（研究代表者、研究分担者及び連携研究者には下線）
6. 研究組織

研究代表者

チェン ジー（国立情報学研究所 コンテンツ科学研究系 准教授）

研究分担者

（ ）

研究者番号：

連携研究者

（ ）

研究者番号：