https://doi.org/10.2352/ISSN.2470-1173.2018.18.3DIPM-425

Holostream: High-Accuracy, High-Speed 3D Range Video Encoding and Streaming Across Standard Wireless Networks

Tyler Bell[1,2,*]***, Jan P. Allebach***[2]***, Song Zhang***[1]

[1] ***School of Mechanical Engineering, Purdue University; West Lafayette, Indiana 47907, USA***
[2] ***School of Electrical and Computer Engineering, Purdue University; West Lafayette, Indiana 47907, USA***
[*] ***tylerbell@purdue.edu***

Abstract

Holostream is a novel platform which enables high-quality 3D video communication on mobile devices (e.g., iPhones, iPads) using existing standard wireless networks. The major contributions are: (1) a novel high-quality 3D video compression method that drastically reduces both 3D geometry and color texture data sizes in order to transmit them within the bandwidths provided by existing wireless networks; (2) a novel pipeline for 3D video recording, encoding, compression, decompression, visualization and interaction; and (3) a demonstration system which successfully delivered video-rate, photorealistic 3D video content through a standard wireless network to mobile devices. The novel platform improves the quality and expands upon the capabilities of popular applications already utilizing real-time 3D data delivery, such as teleconferencing and telepresence. This technology could also enable emerging applications which may require high-resolution, high-accuracy 3D video data delivery, such as remote robotic surgery and telemedicine.

Introduction

Telecommunication methods evolved from telegraphs, telephones, and 2D video conferences with each advancement bringing an enhanced end user experience through more realistic interactions. None of these techniques are advantageous for truly natural interactions as users do not have context about the 3D world in which we live. Therefore, 3D video communication emerges as an interesting research subject because of the increased availability of novel 3D sensing technologies, enhanced network bandwidths, and increased computational power, even on a mobile device (e.g., tablets and smart phones).

Existing techniques for 3D video communications primarily focus on the application of teleconferencing and immersive telepresence. In each, an object at one location is captured in 3D and transmitted to a remote location for redisplay. Although these techniques may successfully accomplish their goals, compared with their mature 2D video communication counterparts, the state-of-the-art 3D video communication platforms are constrained by three major factors: (1) the availability of commercial 3D sensors that can provide equivalent high quality and resolution; (2) the drastically increased 3D data sizes that exceed the bandwidths of existing wireless networks; and (3) the limited real-time 3D data compression platforms which can efficiently reduce data sizes whilst maintaining data quality. Due to these limitations, high-quality 3D video communication across standard wireless networks is still in its infancy.

This paper proposes a novel modular platform, dubbed *Holostream*, which enables high-quality 3D video communications across existing standard wireless networks and existing mobile hardware devices (e.g., iPhones and iPads). Our novel platform advances the quality and capabilities of applications already utilizing real-time 3D data delivery (e.g., teleconferencing, telepresence). Further, the proposed platform could also enable applications where real-time delivery of high-resolution, high-accuracy 3D video data is especially critical, such as collaborative design and online facial behavior analysis.

Contributions

- We developed a novel 3D video compression method that can drastically reduce 3D video data size without substantially sacrificing data quality. With lossless frame by frame storage, the compression ratio is approximately 112:1 when compared with standard OBJ files. Lossy frame by frame storage can achieve a ratio of 518:1 without substantial quality reduction. When paired with the H.264 video codec, our method achieved a compression ratio of 1,602:1, with a bitrate of only 4.8 megabits per second (Mbps), while maintaining high-quality 3D video representations.
- We developed a novel and complete framework for 3D video recording, encoding, compression, decompression and visualization. The entire framework was tested using standard medium bandwidth networks to simultaneously deliver high-quality 3D videos to multiple mobile devices.
- We developed a demonstration system that achieved high-quality 3D sensing, compression, transmission, and visualization across standard wireless networks. The system wirelessly delivered precisely aligned coordinate and color data, consisting of over 300,000 vertices per frame, at 30 Hz to mobile phones and tablets.

Related Work

Much work has been done to enable realistic 3D telecommunications in the past several decades. Most recently, the Holoportation system [1] impressively acquires a colored 3D mesh of a user or object in one location and transmits it to remote users for visualization on augmented reality headsets. Although this system enables very natural user interactions, 1-2 Gbps of data must be transferred to realize communications at 30 Hz. This data rate greatly exceeds the bandwidth availabilities of today's existing wireless networks. In fact, the majority of the work in this area does not effectively address one of the major issues: how can 3D

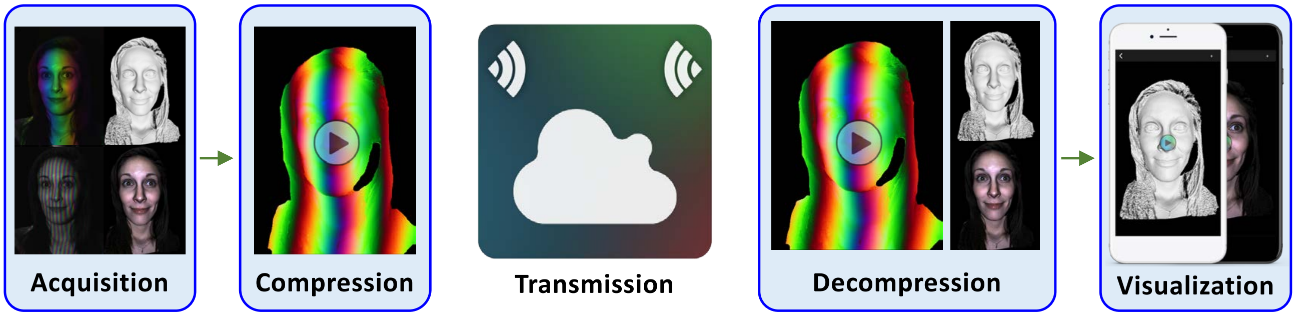

Figure 1. *Holostream pipeline starts with the* Acquisition Module *that captures 3D geometry and color texture in real time. The* Compression Module *encodes 3D data into 2D images that are packed into a video stream using a standard 2D video codec (e.g., H.264). The compressed 2D video is transmitted over existing standard wireless networks through the* Transmission Module*. The* Decompression Module *receives the compressed 2D video and decompresses it to recover the original 3D geometry and color texture. The* Visualization Module *on mobile devices (e.g., iPads, iPhones) visualizes 3D video in real time and allows the users to interact with the video instantaneously.*

data sizes be drastically reduced? Instead, it is typically assumed that the network bandwidth will simply be available. Of course, this assumption does not hold true in practice and thus it is still very difficult to realize high-quality, real-time 3D telecommunications over standard wireless networks.

In a tangential community, researchers have attempted to compress 3D data, both within the mesh format [2, 3] and the point cloud format [4, 5, 6, 7, 8]. As pointed out by the authors of the Holoportation system [1], however, although both dynamic mesh [3] and point cloud compressions [4] are being actively explored, it is still currently challenging to achieve real-time compression at the lowest bitrates (less than low tens of Mbps).

In contrast, standard 2D image and video compression techniques are quite mature and enable today's modern 2D video communications over standard wireless networks. If 3D geometry can be efficiently and precisely converted into standard 2D images, existing 2D video communication platforms can be immediately leveraged for low bandwidth 3D video communications. Given this, there has been effort devoted to encode 3D geometry information within regular 2D images [9, 10, 11, 12]. Although this direction is promising, current encoding methods may not be resilient—both in terms of accuracy and efficiency—to direct storage within a video codec. Further, current methods often only focus on encoding 3D geometry information, leaving other attributes (e.g., texture) to be stored or transmitted separately. This paper presents a novel method for the efficient and precise encoding of 3D video data and color texture into a regular 2D video format. This paper also presents how existing 2D video communication platforms may be leveraged for high quality 3D video communication in real-time over standard wireless networks.

System Overview

As discussed above, existing 3D video communication technologies have limited applications and potentially limited adoption partially because of (1) the required specialized and often expensive hardware; (2) inflexible and complex system setup; (3) highly demanding computational resources for real-time realization; (4) low-resolution and low-accuracy 3D sensing; and/or (5) the required high speed network bandwidths for 3D content transmission due to inefficient 3D video data compression methods.

Our proposed Holostream platform achieved high-resolution and high-accuracy 3D video communication by developing: (1) high-accuracy and high-resolution 3D video capture hardware system; (2) a novel 3D video compression technique; (3) a novel computational framework for 3D video streaming and decompression; and (4) a 3D video visualization application for mobile devices. To our knowledge, this system is the first of its kind which can deliver dense and accurate 3D video content in real time across standard wireless networks to a remote mobile devices (e.g., iPhones and iPads).

Figure 1 shows the overall pipeline of the proposed Holostream system. A set of structured patterns are projected and captured by a structured light system to reconstruct dense and accurate 3D video data, including both geometry and color texture information (*Acquisition Module*); 3D video data is encoded frame by frame into a standard 2D image format that is further compressed using a standard 2D video compression technique (*Compression Module*); the compressed video is delivered across standard wireless networks to remote mobile devices (e.g., iPhones) (*Transmission Module*); the mobile device decompresses the 2D video and decodes the original 3D data frame by frame (*Decompression Module*); and finally the mobile app visualizes 3D video contents in real time allowing the user to instantaneously interact with the 3D video data (*Visualization Module*). The rest of this section describes each individual module developed for the entire pipeline.

Acquisition Module

The Acquisition Module is a structured light scanner that acquires high-resolution 3D video data including 3D geometry and color texture in real time. The scanner consists of a single camera and a single projector, and uses the principle of triangulation for 3D shape reconstruction [13]. Each 3D reconstruction requires a sequence of structured patterns that vary sinusoidally along one direction and remain constant along the other direction. These types of sinusoidal patterns are often called fringe patterns; and this type of structured light technique is called digital fringe projection (DFP). Instead of directly using intensity information for 3D shape reconstruction, the DFP technique extracts the phase of phase-shifted fringe patterns for 3D shape reconstruction. Compared with conventional structured light techniques, the DFP technique has the advantage of simultaneously achieving high resolution (i.e., at camera pixel level) and high speed [14].

Our DFP system uses two sets of patterns with different frequencies, each having three phase-shifted patterns described as,

$$I_k^h(x,y) = I'(x,y) + I''(x,y)\cos\left(\phi^h + 2k\pi/3\right), \quad (1)$$

$$I_k^l(x,y) = I'(x,y) + I''(x,y)\cos\left(\phi^l + 2k\pi/3\right). \quad (2)$$

Here, $k = \{1,2,3\}$, $I'(x,y)$ is the average intensity, $I''(x,y)$ is the intensity modulation, and $\phi^h(x,y)$ and $\phi^l(x,y)$ is the high frequency phase and low frequency phase, respectively. Simultaneously solving three equations with the same frequency phase leads to

$$I'(x,y) = I_t^g(x,y) = (I_1^h + I_2^h + I_3^h)/3, \quad (3)$$

$$I''(x,y) = \sqrt{3(I_2^h - I_1^h)^2 + (2I_3^h - I_2^h - I_1^h)^2}/3, \quad (4)$$

$$\phi^h(x,y) = \tan^{-1}\left[\sqrt{3}\left(I_2^h - I_1^h\right)\right] / \left[2I_3^h - I_2^h - I_1^h\right], \quad (5)$$

$$\phi^l(x,y) = \tan^{-1}\left[\sqrt{3}\left(I_2^l - I_1^l\right)\right] / \left[2I_3^l - I_2^l - I_1^l\right]. \quad (6)$$

Both low and high frequency phase maps obtained here are called wrapped phase maps since they are bounded by $(-\pi, \pi]$ due to the use of an inverse tangent function. To recover the original phase value, the 2π discontinuous locations need to be identified and corrected; this is done through a step called phase unwrapping. Our system combined these two frequency phase maps and used a two-frequency phase unwrapping method [15] to obtain an unwrapped phase map $\Phi(x,y)$ for the high-frequency fringe patterns. The unwrapped phase is further used to reconstruct (x,y,z) coordinates per pixel once the system parameters (e.g., focal lengths of the camera and the projector, the transformation from the projector coordinate system to the camera coordinate system) are pre-calibrated [16]. Besides acquiring 3D geometry, DFP techniques naturally come with a texture image $I_t^g(x,y)$, and if a single sensor color camera is used, $I_t^g(x,y)$ can be converted to a color image through the demosaicing process. Figure 2 shows an example of reconstructing a single 3D frame for a camera resolution of 480×640.

Figure 2. Example of reconstructing 3D geometry and color texture from the Acquisition Module. Top row left to right: low frequency fringe patterns, high-frequency fringe patterns, low frequency wrapped phase ϕ^l, high-frequency wrapped phase ϕ^h; bottom row left to right: b/w texture I_t^g, unwrapped phase Φ, 3D geometry, color texture after demosaicing I_t^c.

Compression Module

The goal of our 3D video communication is to deliver dense and accurate 3D range geometry information, over existing standard wireless networks, at the video rate the Acquisition Module captures (e.g., 30 Hz). To accomplish this goal, the raw data must be substantially compressed for efficient transmission. Our proposed 3D video compression includes two steps: the first step is to encode 3D geometry and color texture into standard 2D images frame by frame. The second step is then to compress the 2D image sequence using standard video compression codecs.

3D Data Encoding. As previously discussed, (x,y,z) coordinates of a given pixel (i,j) are recovered directly from the unwrapped phase $\Phi(i,j)$ if the system parameters of the scanner are pre-calibrated [16]. In other words, there exists a one-to-one mapping between the unwrapped phase $\Phi(i,j)$ of a point and its recovered (x,y,z) coordinate. Therefore, if 3D coordinates of a pixel are known, we can determine its phase value with known system parameters. This tells us that we can directly use 2D data to represent 3D geometry. For each pixel (i,j) within the unwrapped phase map, a scaled corresponding phase value $\Phi'(i,j)$ is encoded as

$$E_r(i,j) = 0.5 + 0.5\sin\Phi'(i,j), \quad (7)$$

$$E_g(i,j) = 0.5 + 0.5\cos\Phi'(i,j). \quad (8)$$

The third color channel, in this case the blue channel, is used to store the natural texture value, $E_b = I_t^g$. Once the unwrapped phase and texture data are encoded and stored into the output 2D image, E, it can be further compressed frame-by-frame via a lossless (e.g., PNG), or via various levels of lossy (e.g., JPEG), image encoding. The final result of this novel approach is a compressed 2D RGB image from which both 3D coordinates and color texture can be later recovered.

The left image of Fig. 3 shows an example of an image, E, which encodes the same 3D geometry and color texture shown in Fig. 2. If the 480×640 image is stored with lossless PNG, the file size of the mesh is reduced from 32 MB to 288 KB, achieving a compression ratio of approximately 112:1 versus storing the same information within the OBJ format. As shown in the last image of Fig. 3, the difference between the original geometry and the geometry reconstructed from the PNG image, E, appears to be random noise which may be caused by small amounts of quantization.

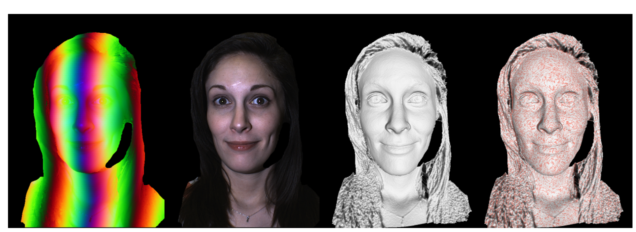

Figure 3. Encoding and decoding 3D geometry and color texture. A lossless PNG format results in a compression ratio of approximately 112:1. Images from left to right respectively show the encoded RGB PNG image, the recovered color texture, the 3D reconstructed geometry, and an overlay of the reconstructed 3D geometry on top of the original 3D geometry (gray color represents recovered geometry and red represents the original geometry).

Figure 4. Compressing encoded 3D data with the H.264 video codec at different qualities. First row: four frames decoded from the video stored with lossless H.264 (compression ratio 129:1); second row: four frames decoded from H.264 video using 4:2:0 subsampling and a constant rate factor of 6 (compression ratio 551:1); and third row: the corresponding frames using the same encoding instead with a constant rate factor of 12 (compression ratio 1,602:1).

If a lossy JPEG encoder is used, the file sizes can be reduced even further. Compression ratios of 107:1, 267:1, 406:1, and 518:1 were achieved when using JPEG 100%, 95%, 90%, and 85%, respectively, to compress the geometry within a 480×640 image. The 3D reconstructions from this lossy encoded image incur some reduction in measurement accuracy; however, for visual-based applications (e.g., telepresence), there may not be a distinguishable difference.

3D Video Compression. The compression ratios offered by PNG or JPEG may be high enough to stream E directly frame-by-frame over most medium bandwidth networks, however, even higher compression ratios can be achieved if the frames are video encoded. This is especially advantageous in terms of preparing E for delivery to mobile devices. In our platform, we encode E into an H.264 video stream using the x264 library [17]; this stream is then segmented into shorter transport stream files for delivery via HTTP Live Streaming (HLS). Figure 4 shows some example reconstructions decoded from the data stored using various H.264 qualities. When losslessly encoding a 10 second 3D video sequence consisting of 300 frames, a compression ratio of 129:1 was achieved resulting in 75 MB of data. When using H.264 to lossy encode video at various levels, the color texture was placed in a mosaic fashion to the right of the encoded 3D image, E. Encoding these same 300 3D video frames using 4:2:0 subsampling and a constant rate factor (CRF) of 6 provided a 551:1 compression ratio and a file size of 17.5 MB. Live streaming this encoded video sequence only required a 14 Mbps connection. A CRF value of 12 encoded the sequence into 6 MB of data giving a 1,602:1 compression ratio. As can be seen in Fig. 4, even at the very low bitrate of 4.8 Mbps the resulting 3D reconstructions are still of great quality in both geometry and color texture.

Transmission Module

The Transmission Module was implemented within an intermediary web server—built on top of HTTP and WebSocket technologies—which acts as a middleware between the Acquisition Module and Visualization Module(s). When a client connects to the server, an initialization message, M, is constructed and sent over a WebSocket. This message contains a few important parameters, such as the camera's demosaicing format and resolution, the phase scaling factor, the system's capture volume, and the DFP system's calibration data. Encoded live video streams are delivered over HTTP to connected clients within small video segments (transport streams) via HTTP Live Streaming (HLS).

Decompression Module

To highlight the efficiency of the proposed 3D video encoding method, our primary Decompression Module was implemented within a native iOS application responsible for receiving compressed 2D video data, decompressing it, and then reconstructing 3D geometry and color texture.

Initialization. When a Decompression Module client first connects to the Transmission Module's server, it receives properties about the incoming video stream, including the DFP system's calibration parameters and the video frame dimensions. Using these dimensions, a 2D mesh plane of 3D coordinates is initialized. For example, if the frame dimensions are 480×640, a regular mesh plane consisting of 307,200 vertices will be constructed. This initialization ensures that any pixel within the incoming encoded image, E, corresponds to a single 3D vertex within the mesh. The mesh is also created such that it is modifiable via vertex and fragment GPU shaders. All other properties within the initialization message, M, are sent to the shaders as uniforms.

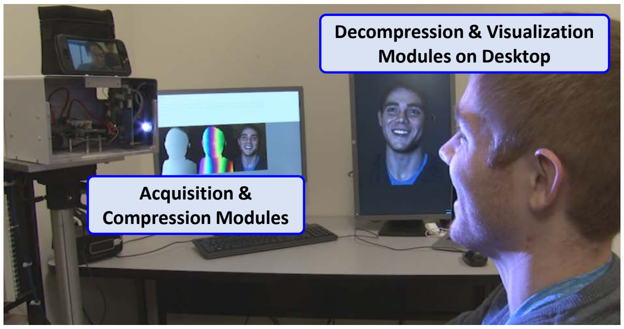

Figure 5. *Successful implementation of our proposed Holostream platform. The left image shows the Acquisition and Compression Modules. The right image shows multiple users on their mobile devices receiving and interacting with the live 3D video stream delivered across a standard wireless network. A video demonstration of our system can be found at* `http://www.xyztlab.com/holostream-ei-2018`.

Reconstruction. For each vertex (u,v) of the mesh, a wrapped phase value, ϕ_r, can be computed for the data stored within the two chosen color channels of E (e.g., red and green):

$$\phi_r(u,v) = \tan^{-1}\left[(E_r - 0.5)/(E_g - 0.5)\right]. \quad (9)$$

The wrapped phase is then unwrapped using the minimum phase unwrapping technology developed by An et al. [18]. This is done on the fly in the vertex shader using the DFP system's calibration parameters, the phase scaling factor, and the minimum z value of the capture volume—all of which were made available to the connected client upon initialization. Once the unwrapped phase $\Phi(u,v)$ has been computed it can be converted into a 3D coordinate (x,y,z) using the calibration parameters. Finally, the vertex (u,v) associated with the vertex shader updates its position attribute to the newly derived 3D coordinate.

The color texture image can also be decoded from the received encoded image, E, and this is done within the mesh's fragment shader. To recover a color texture value, we perform a demosaicing operation for each (u,v), using the grayscale texture encoded in E_b.

Visualization Module

To demonstrate the efficiency of our 3D video encoding method and the wireless nature of our platform, our main Visualization Module was implemented within a native iOS application. The app, upon connecting to a stream, receives an initialization message, M, from the Transmission Module over a WebSocket connection. Using this information, a 3D scene is initialized and a mesh is prepared for modifications, as described above. The H.264 encoded video stream from the Transmission Module is received via HLS. Geometry and color texture is then decompressed from encoded frames within custom Metal shaders. The end result is the ability to visualize and interact with live 3D video content on both iPhones and iPads.

Results

We developed a demonstration system to verify the performance of our proposed 3D video communication methods. The platform detailed above was implemented with a single DFP 3D capture system on a single PC. The PC, which ran the Acquisition, Compression, and Transmission Modules, had an Intel Core i7 (3.40 GHz) CPU, 16 GB of RAM, and a single NVIDIA GeForce GTX 980 Ti GPU. The DFP capture system consisted of a single camera (PointGrey Grasshopper3 GS3-U3-23S6C) and a single digital-light-processing (DLP) projector (Texas Instruments LightCrafter 4500). The resolution of the camera was set to 480 × 640 (allowing for up to 307,200 unique 3D coordinates) and the projector's resolution was 912 × 1140. We used the method developed by Zhang and Huang [16] to obtain the calibration parameters for our DFP system. The Acquisition and Compression Modules were implemented on the GPU within custom CUDA kernels in order to maintain a streamable frame rate of at least 30 (3D) Hz. Figure 5 shows two photographs of the successful implementation of our proposed Holostream platform. Note that the left image also shows an additional desktop implementation of the Decompression and Visualization Modules. A video demonstration of our system can be found at `http://www.xyztlab.com/holostream-ei-2018`.

To evaluate the proposed 3D range encoding method, the sequence of 300 3D face data frames (Fig. 4) were encoded and stored within H.264 video streams. CRF values of 0, 6, and 12 were used to encode the H.264 videos. Each frame of each video was then decoded using the proposed method and measured against its original reconstructed 3D frame. The mean error between the original and reconstructed geometries across the entire 10 second video sequence (representing the 300 encoded 3D frames) was 0.38 mm, 0.65 mm, and 0.69 mm for CRF values of 0, 6, and 12, respectively. The respective standard deviations were 0.50 mm, 0.51 mm, and 0.54 mm. It should be noted that these measurements excluded large boundary outliers.

To test the limits of our approach, the same 10 second video sequence was encoded with the H.264 codec using a CRF value of 25. Figure 6 shows two decoded video frames from this video stream. The 3D geometry data quality is noticeably worse, however, a compression ratio of 16,279:1 was achieved, resulting in only 0.59 MB of data for the 10 second video sequence. This means that only 0.48 Mbps of network bandwidth was required to deliver 300 3D frames at 30 Hz. Even at this heightened compression ratio, the color texture quality remains reasonably high (right two images). The apparent robust nature of the proposed 3D encoding method may make this technology valuable for 3D data delivery to remote locations or to throttled mobile devices.

Summary

This paper presented Holostream, a novel platform for high-quality 3D video communication across existing standard wire-

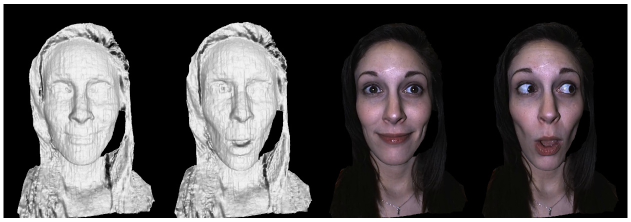

Figure 6. *Compressing encoded 3D data with the H.264 video codec at very lossy (CRF of 25) qualities achieving a 16,279:1 compression ratio.*

less networks and with existing mobile hardware devices (e.g., iPhones, iPads). Our novel 3D video compression method when paired with the H.264 codec achieved compression ratios of 1,602:1 while maintaining high-quality 3D video with color texture. This allowed for transmission of 3D video and color texture content over existing wireless networks using only 4.8 Mbps. We also developed an implementation of our entire 3D video acquisition, encoding, compression, decompression, and visualization framework. This system wirelessly delivered coordinate and color data, consisting of up to 307,200 vertices per frame, at 30 Hz to multiple mobile phones and tablets. We also showed that a bitrate of only 0.48 Mbps was sufficient to deliver lower quality 3D geometry while maintaining high quality color texture. The Holostream platform could accelerate development within areas which require high-resolution, high-accuracy 3D video data delivery in real time, such as video surveillance and biometrics.

References

[1] Sergio Orts-Escolano, Christoph Rhemann, Sean Fanello, Wayne Chang, Adarsh Kowdle, Yury Degtyarev, David Kim, Philip L. Davidson, Sameh Khamis, Mingsong Dou, Vladimir Tankovich, Charles Loop, Qin Cai, Philip A. Chou, Sarah Mennicken, Julien Valentin, Vivek Pradeep, Shenlong Wang, Sing Bing Kang, Pushmeet Kohli, Yuliya Lutchyn, Cem Keskin, and Shahram Izadi. Holoportation: Virtual 3d teleportation in real-time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16, pages 741–754, New York, NY, USA, 2016. ACM.

[2] R. Mekuria, M. Sanna, E. Izquierdo, D. C. A. Bulterman, and P. Cesar. Enabling geometry-based 3-d tele-immersion with fast mesh compression and linear rateless coding. *IEEE Transactions on Multimedia*, 16(7):1809–1820, Nov 2014.

[3] Alvaro Collet, Ming Chuang, Pat Sweeney, Don Gillett, Dennis Evseev, David Calabrese, Hugues Hoppe, Adam Kirk, and Steve Sullivan. High-quality streamable free-viewpoint video. *ACM Trans. Graph.*, 34(4):69:1–69:13, July 2015.

[4] R. L. de Queiroz and P. A. Chou. Compression of 3d point clouds using a region-adaptive hierarchical transform. *IEEE Transactions on Image Processing*, 25(8):3947–3956, Aug 2016.

[5] T. Golla and R. Klein. Real-time point cloud compression. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5087–5092, Sept 2015.

[6] J. Kammerl, N. Blodow, R. B. Rusu, S. Gedikli, M. Beetz, and E. Steinbach. Real-time compression of point cloud streams. In *2012 IEEE International Conference on Robotics and Automation*, pages 778–785, May 2012.

[7] Matthias Nießner, Michael Zollhöfer, Shahram Izadi, and Marc Stamminger. Real-time 3d reconstruction at scale using voxel hashing. *ACM Trans. Graph.*, 32(6):169:1–169:11, November 2013.

[8] A. Mossel and M. Kroeter. Streaming and exploration of dynamically changing dense 3d reconstructions in immersive virtual reality. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pages 43–48, Sept 2016.

[9] Nikolaus Karpinsky and Song Zhang. Composite phase-shifting algorithm for three-dimensional shape compression. *Opt. Eng.*, 49(6):063604, 2010.

[10] Zhiling Hou, Xianyu Su, and Qican Zhang. Virtual structured-light coding for three-dimensional shape data compression. *Opt. Laser Eng.*, 50(6):844–849, 2012.

[11] Yajun Wang, Lianxin Zhang, Sheng Yang, and Fang Ji. Two-channel high-accuracy holoimage technique for three-dimensional data compression. *Opt. Laser Eng.*, 85:48–52, 2016.

[12] Arnaud Bletterer, Frédéric Payan, Marc Antonini, and Anis Meftah. Point cloud compression using depth maps. In *IS&T International Symposium on Electronic Imaging 2016*, pages 3DIPM–397.1–3DIPM–397.6, 2016.

[13] Joaquim Salvi, Sergio Fernandez, Tomislav Pribanic, and Xavier Llado. A state of the art in structured light patterns for surface profilometry. *Patt. Recogn.*, 43(8):2666–2680, 2010.

[14] Jason Geng. Structured-light 3d surface imaging: a tutorial. *Advances in Opt. and Photonics*, 3(2):128–160, 2011.

[15] Jae-Sang Hyun and Song Zhang. Enhanced two-frequency phase-shifting method. *Appl. Opt.*, 55(16):4395–4401, 2016.

[16] Song Zhang and Peisen S. Huang. Novel method for structured light system calibration. *Opt. Eng.*, 45(8):083601, 2006.

[17] VideoLAN Organization. x264. `http://www.videolan.org/developers/x264.html`.

[18] Yatong An, Jae-Sang Hyun, and Song Zhang. Pixel-wise absolute phase unwrapping using geometric constraints of structured light system. *Opt. Express*, 24(15):18445–18459, 2016.

Author Biography

Tyler Bell received his BS in Computer Science (2012) and MS in Human Computer Interaction (2014) from Iowa State University. He is currently a PhD candidate in Computer Engineering at Purdue University, where he is co-advised by Prof. Song Zhang and Prof. Jan P. Allebach. Tyler's research interests include 3D communications/telepresence; 3D data/video compression; real-time 3D sensing systems; virtual/augmented reality; human computer interaction; and multimedia on mobile devices.

Jan P. Allebach is Hewlett-Packard Distinguished Professor of Electrical and Computer Engineering at Purdue University. Allebach is a Fellow of the IEEE, the National Academy of Inventors, the Society for Imaging Science and Technology (IS&T), and SPIE. He was named Electronic Imaging Scientist of the Year by IS&T and SPIE, and was named Honorary Member of IS&T, the highest award that IS&T bestows. He has received the IEEE Daniel E. Noble Award, the IS&T/OSA Edwin Land Medal, and is a member of the National Academy of Engineering. He currently serves as an IEEE Signal Processing Society Distinguished Lecturer (2016-2017).

Prof. Song Zhang is an associate professor of mechanical engineering at Purdue University. He received his Ph.D. degree in mechanical engineering from Stony Brook University in 2005. His major research interests are superfast 3-D optical sensing, 3D biophontic imaging, 3D geometry/video analysis, human and computer interaction, and virtual reality. Dr. Zhang has published over 100 research articles including more than 70 journal papers; edited one book; and holds 3 US patents.