

Haptic Telepresence System for Individuals with Visual Impairments

Yeongil Ryu and Eun-Seok Ryu*

Department of Computer Engineering, Gachon University,
1342 Seongnamdaero, Sujeong-gu, Seongnam, Gyeonggi, 13120, Korea

(Received April 1, 2016; accepted May 15, 2017)

Keywords: haptic telepresence, individuals with visual impairments, video streaming

In this paper, we propose a new haptic telepresence system for individuals with visual impairments (VIs) using a red-green-blue-depth (RGB-D) sensor and a haptic device. Recent improvements in RGB-D sensors have enabled real-time access to 3D spatial information. However, the real-time representation of the tangible haptic experience has not been sufficiently challenged. Thus, the proposed system addresses the telepresence of remote 3D information using an RGB-D sensor through video encoding and 3D depth-map enhancement. In the implemented system, the Kinect sensor from Microsoft is an RGB-D sensor that provides depth and 2D color images at a rate of approximately 30 fps. The Kinect depth data frame is buffered, projected into a 3D coordinate system with a resolution of 640×480 pixels, and then is transformed into a 3D map. To verify the benefits of the proposed video content adaptation method for individuals with VIs, in this study, we implemented a ‘2D plus depth map’-based haptic telepresence system; it conducts user experience experiments and presents the results of user response time.

1. Introduction

Over 39 million people are classified as legally blind,⁽¹⁾ but many devices and techniques to aid them still are only in the form of simple navigations and textual information for individuals with visual impairments (VIs). Recently, an art exhibition using 3D technology opened for individuals with VIs in the U.S. The art exhibition provided experiences in which individuals with VIs carved touch and feel 3D paintings, and audio described what feature was being touched on the painting. However, museum or art center exploration is still not easy for individuals with VIs, owing to constraints on their activity. Moreover, individuals with VIs cannot be provided with any video services [e.g., video conferencing, video on demand (VoD)] in contrast to people with unimpaired sight who can access video services anywhere. To address these problems, in this paper, we propose a new haptic telepresence system to deliver 3D spatial information on objects for individuals with VIs using an RGB-D (color and depth camera) sensor and a haptic device. The proposed system consists of (1) a 3D spatial information capture module, (2) an optimizing depth-map and real-time transmission module with multimedia encoder and decoder, and (3) a haptic interaction module. Figure 1 shows a conceptual diagram of the proposed haptic telepresence system that captures 2D true-color images and a depth map with Microsoft Kinect as the RGB-D sensor, optimizes the depth map, and transmits both the image and the map in real-time.

*Corresponding author: e-mail: esryu@gachon.ac.kr
<http://dx.doi.org/10.18494/SAM.2017.1594>

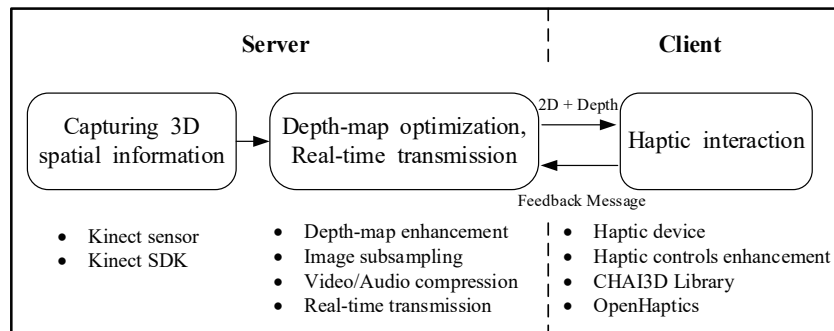


Fig. 1. Conceptual diagram of proposed haptic telepresence system.

2. Related Work

2.1 Prior work in assistive devices for VIs

Efforts to aid the daily lives of individuals with VIs have been made in various fields. Hong developed a car for blind drivers, which uses robotics, laser rangefinders, global positioning system (GPS), and smart feedback tools.⁽²⁾ The car for blind drivers enables a nonsighted driver to drive independently by informing the driver of the directions to turn the steering wheel as recognized by cameras and sensors. Martin *et al.* proposed a smart cane that can sense obstacles and guide individuals with VIs.⁽³⁾ The cane utilizes ultrasonic sensors with a detecting range of 45° forward, and it guides a user with vibration modules. However, systems for individuals with VIs recognize and analyze surrounding circumstances and deliver only the analyzed results to the user instead of assisting the user to perceive various information actively. Such systems differ only slightly from other conventional navigation systems for individuals with VIs.

Park *et al.* developed a telerobotic haptic exploration system for use in art galleries and museums by individuals with VIs.⁽⁴⁾ In Park *et al.*'s work, using 3D spatial information from an RDB-D sensor on a robot, individuals with VIs perceived environments around the telepresence robot, and controlled the robot using a haptic device. Park *et al.*'s work provided a new methodology of haptic telepresence for individuals with VIs by providing an enhanced interactive experience with which they can remotely access public places (e.g., art galleries and museums). It is meaningful because it enabled individuals with VIs to perceive 3D spatial information actively and not depend on the processed results from a computer. In this paper, we introduce the haptic telepresence project at Gachon University using “2D plus depth map” transmission, which extends our previous collaborative work with Park *et al.* Thus, it focuses on (i) the implementation of the system, (ii) an efficient video processing method, and (iii) the enhancement of “2D plus depth map” quality instead of robotic technologies.

2.2 Depth estimation

Depth estimation is an important field for many advanced video applications such as object reconstruction and human computer interface (HCI). To estimate depth, information has been developed in two directions: one is stereo/multiview matching, and the other is using a depth

sensor such as the Microsoft Kinect. Stereo/multiview systems calculate depth by measuring the disparity between two 2D images. Figure 2 shows an example of depth map estimation based on window-based stereo matching.

The Microsoft Kinect is an RGB-D sensor. Kinect v2, the latest Kinect sensor, estimates depth by measuring the time of flight (TOF) of infrared (IR) light. The TOF-based method estimates depth by measuring the phase difference between the emitted and received signals, and it enables the computation of the distance of the illuminated object from the sensor for each pixel. Kinect v2 provides a depth map with 512×424 resolution, and has an available depth sensing range from 0.5 to 4.5 m.

3. Proposed New Haptic Telepresence System

3.1 System architecture

Although many devices and techniques to aid individuals with VIs have been developed, they are limited to reading the information that has been transformed to voice or to using restricted tactile interaction. The architecture of the proposed system consists of a 3D spatial information capture module, a depth-map enhancement module, a video/audio real-time compression module, a real-time transmission module, and a haptic feedback/audio representation module. The rest of this section explains the modules in detail. In this study, we implemented the proposed system partially to verify its feasibility as shown in Fig. 3. The implemented modules enable individuals with VIs to feel the streamed 3D spatial information using a haptic device as shown in Fig. 4.

3.2 3D spatial information capture module

The 3D spatial information capture module was implemented using a Microsoft Kinect software development kit (SDK).⁽⁵⁾ The Kinect SDK allows capturing 2D true-color images by adjusting the resolution of the image to a size of up to 1920×1080 pixels and capturing depth-map functions up to a resolution of 640×480 pixels.⁽⁶⁾ The 3D spatial information capture module can be implemented using stereo/multiview matching or an IR sensor. Although the stereo/multiview matching provides a depth map of higher resolution than an IR sensor, its computational complexity is also much higher than that of the IR sensor. Therefore, the 3D spatial information capture module of the proposed system was implemented using IR sensors to support real-time service.

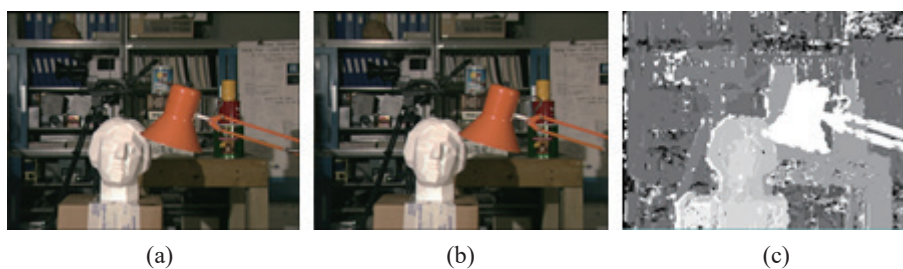


Fig. 2. (Color online) Depth map measured by traditional window-based stereo matching: (a) left image, (b) right image, and (c) measured depth map.

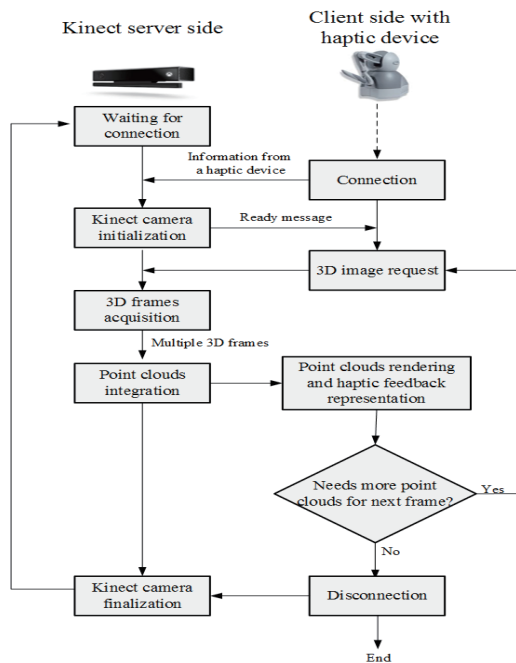
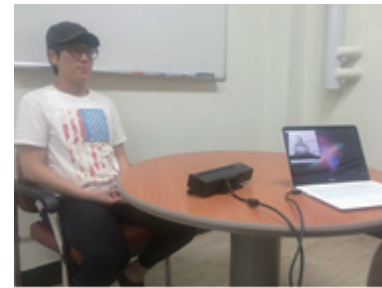
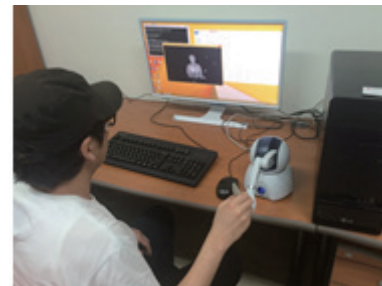


Fig. 3. Processing steps of the proposed system.



(a)



(b)

Fig. 4. (Color online) (a) Kinect server side and (b) client side using a haptic device.

3.3 Depth-map enhancement module

Depth maps from the IR sensor, the Microsoft Kinect, are generally noisy and not sufficiently accurate. Therefore, guided filtering is applied to the depth map enhancement module in this study.

Guided filtering smooths the image while preserving edges, and guided filtering is used to enhance a depth map like bilateral filtering, which is a popular smooth filter.⁽⁷⁾ However, the guided filtering is light-weight and provides better smoothing compared with bilateral filtering. Figure 5 shows the results of depth-map enhancement with guided filtering.

3.4 Video/audio real-time compression module

To transmit multimedia combined with 3D spatial information in the form of point clouds, 2D true color images, and audio, compression is essential in networks with limited bandwidth. High Efficiency Video Coding (HEVC), the newest video coding standard, was standardized by the Joint Collaborative Team on Video Coding (JCT-VC) group in partnership with the International Telecommunication Union Telecommunication Standardization Sector (ITU-T) Video Coding Expert Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG).⁽⁸⁾ HEVC enables significantly improved compression compared with existing major video coding standards in the range of 50% bit-rate reduction for equal perceptual video quality. Thus, a video/audio real-time compression module includes HEVC to transmit data seamlessly.

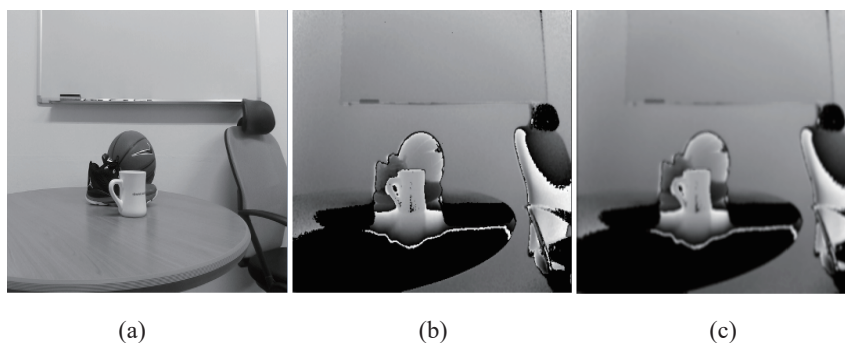


Fig. 5. Depth-map enhancement using guided filtering: (a) guidance image, (b) depth map, and (c) enhanced depth map using guided filtering.

3.5 Real-time transmission module

A haptic device for individuals with VIs is usually connected through a wired network, but a server system that includes a Kinect sensor can be connected through a wireless network for mobility. Thus, this module includes rate control, Raptor forward error correction (FEC), and unequal error protection (UEP) to improve the quality of service (QoS). These techniques are very useful in transforming visual perception (including depth) into nonvisual sensations over the network.^(9,10) Particularly considering the large bandwidth of data streams generated from a depth sensor (that includes both 2D color information as well as 3D point clouds), they enable the optimization of data size and reliable communication.

3.6 Haptic feedback/audio representation module

To optimize haptic feedback for individuals with VIs, a haptic interaction module is implemented using the CHAI3D library and OpenHaptics.^(11,12) The OpenHaptics tool kit is a framework to support low-level controls for a haptic device. The CHAI3D library is a powerful open source framework to support high-level access to a haptic device by providing OpenHaptics-based modules such as computer haptics, visualization, and interactive real-time simulation, and, as shown in Fig. 6, it involves audio, graphic, haptic effects, collision detection, and other features.⁽⁸⁾ The CHAI3D library has been applied to applications in automotive, aerospace, medical, entertainment, and industrial robotics areas.

4. Experimental Results and Discussion

Before implementing our proposed haptic telepresence system, we carried out feasibility tests to investigate the accuracy of 3D perception using the haptic device. The procedures of the feasibility test are: (1) a set of solid figures [(a) sphere, (b) cone, (c) cylinder, (d) pyramid, (e) cube] is given to 10 testers; (2) we randomly give one of the sets to testers; and (3) the testers choose a given solid figure through tactile cues from the haptic device, without using their vision. Figure 7 shows the box plot for time used to answer. The experimental results demonstrated that the 3D perception capabilities of each user have a certain gap, and considering that gap is the most

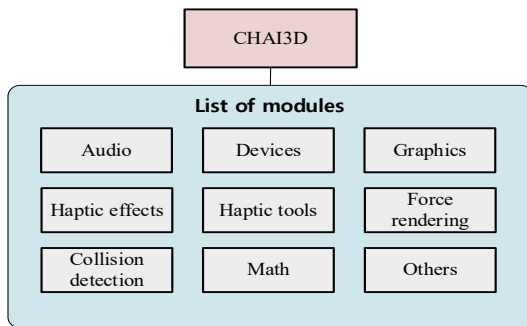


Fig. 6. (Color online) A list of CHAI3D modules.

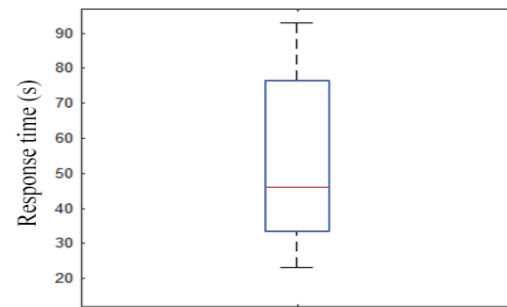


Fig. 7. (Color online) Box plot indicating response time.

important challenge in the proposed haptic telepresence system. The first step of the challenge is to develop a method of measuring the 3D perception capability of individuals with VIs. From the measurement results, the implemented system can provide the appropriate frame rate and the optimized smoothing level of 3D spatial information for each user.

5. Conclusions

In this paper, we propose a new haptic telepresence system for individuals with VIs as an extension of our collaboration with Park *et al.*⁽⁴⁾ It focuses on the enhanced 3D (2D + depth) video streaming technology using the HEVC standard, guided filtering, and the frame rate adaptation. The proposed system consists of (1) a 3D spatial information capture module, (2) optimizing the depth map and real-time transmission module with a multimedia encoder and decoder, and (3) a haptic feedback/audio representation module. The 3D spatial information is captured from a Kinect sensor and is transformed to the haptic feedback module through a haptic device. The system enables individuals with VIs to perceive the face of a family member in a remote place. Moreover, users can enjoy video contents using the system, although the video contents have a lower frame rate and resolution relative to ordinary video contents such as 1080p30 (1920 × 1080 resolution, 30 fps). The video frame rate changes according to predefined policies such as the perceived region of the 3D object and the haptic button event from users. In our current work, we carry out research and development on a 2D braille pad for individuals with VIs and implement an e-book reader application supporting Daisy and e-pub formats.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2015R1C1A1A02037743).

References

- 1 D. Pascolini and S. P. Mariotti: *Br. J. Ophthalmol.* Online **96** (2011) 614.
- 2 D. Hong: Making a car for blind drivers, http://www.ted.com/talks/dennis_hong_makin-g_a_car_for_blind_drivers (accessed June 2016).
- 3 W. Martin, K. Dancer, K. Rock, C. Zeleney, and K. Yalemirthi: *Proc. ASEE North Central Section Conf.* (2009) p. 1.
- 4 C. H. Park, E.-S. Ryu, and A. M. Howard: *IEEE Trans. Haptics* **8** (2015) 327.
- 5 Microsoft: Kinect tools and resources, <https://developer.microsoft.com/en-us/windows-/kinect/tools> (accessed June 2016).
- 6 Microsoft: Kinect hardware, <https://developer.microsoft.com/en-us/windows/kinect/hardware> (accessed June 2016).
- 7 K. He, J. Sun, and X. Tang: *European Conference on Computer Vision*, eds. K. Daniilidis, P. Maragos, and N. Paragios (Springer, Berlin, 2010) Chap. 1.
- 8 G. J. Sullivan, J.-R. Ohm, and W.-J. Han: *IEEE Trans. Circuits Syst. Video Technol.* **22** (2012) 1649.
- 9 A. Shokrollahi: *Trans. Inform. Theory* **51** (2006) 2551.
- 10 E.-S. Ryu and N. Jayant: *IEEE Trans. Consum. Electron.* **4** (2012) 1652.
- 11 CHAI3D: CHAI3D documentation, <http://www.chai3d.org/download/doc/html/wrapper-overview.html> (accessed June 2016).
- 12 Geomagic: OpenHatpics, <http://www.geomagic.com/en/products/open-haptics/overview> (accessed June 2016).