Telerobotic Haptic Exploration in Art Galleries and Museums for Individuals with Visual Impairments

Chung Hyuk Park, *Member, IEEE*, Eun-Seok Ryu, *Member, IEEE*, and
Ayanna M. Howard, *Senior Member, IEEE*

Abstract—This paper presents a haptic telepresence system that enables visually impaired users to explore locations with rich visual observation such as art galleries and museums by using a telepresence robot, a RGB-D sensor (color and depth camera), and a haptic interface. The recent improvement on RGB-D sensors has enabled real-time access to 3D spatial information in the form of point clouds. However, the real-time representation of this data in the form of tangible haptic experience has not been challenged enough, especially in the case of telepresence for individuals with visual impairments. Thus, the proposed system addresses the real-time haptic exploration of remote 3D information through video encoding and real-time 3D haptic rendering of the remote real-world environment. This paper investigates two scenarios in haptic telepresence, i.e., mobile navigation and object exploration in a remote environment. Participants with and without visual impairments participated in our experiments based on the two scenarios, and the system performance was validated. In conclusion, the proposed framework provides a new methodology of haptic telepresence for individuals with visual impairments by providing an enhanced interactive experience where they can remotely access public places (art galleries and museums) with the aid of haptic modality and robotic telepresence.

Index Terms—Haptic telepresence, assistive robotics, 3D haptic rendering, visual impairment, video encoding, depth sensors, multimedia streaming

◆

1 INTRODUCTION

About 285 million people worldwide are visually impaired, and among them, over 39 million people are classified as legally blind [1]. Many assistive devices and techniques have been developed to aid in the navigation and textual information transfer for individuals with visual impairments (VIs) [2]. However, shopping and exploration in public places such as art galleries and museums are still hard challenges for individuals with VIs [3]. Observing that individuals with VIs incorporate many tactile [4], [5] and auditory cues [6], [7] to perceive the environment, feel objects, or interact with others, we developed an assistive robotic system for remote exploration. For providing a richer experience by allowing individuals with Vis to feel remote objects or gain perception over remote places, this paper proposes a novel haptic exploration framework with a telepresence robot [8], [9], [10], which is an advanced form of robotic teleoperation that grants controllability and reality over a remote environment through robotic perception and actuation systems, for enabling remote exploration in public places for individuals with visual impairments.

Despite the many assistive devices that have been developed to aid in daily living [2], [11], the most general and popular assistive means for individuals with VIs are still the walking cane and guide dogs as can be found in the current official handbooks and information material on assistive technology for VI [12], [13]. These aids are effective in assisting the user in navigating within an environment, only with the effective working range limited to the proximity of the user. Assistive methods for remote interaction for individuals with VIs, such as real-world haptic exploration for accessing remote environments rich with visual information, have not been studied in depth.

We envision that a telerobotic approach equipped with haptic display (as sketched in Fig. 1), which grants a person to gain 3D tactile feedback of the remote environment as well as controllability through a telepresence robot, can be a viable solution for the telepresence of individuals with VIs. Haptic and auditory representations of the environment [14], [15], [16] can be effective interactive modalities for assisting individuals with VI, and it will be able to increase the effectiveness more if the system can handle dynamic *"real-world"* renderings in *"real time"*.

As such, we employ a robotic agent equipped with RGB-D (color & depth sensing, e.g., [17]) sensors to collect 3D spatial and visual information of a remote environment and transfer them to the user through a bi-directional haptic interface, enabling both tele-operation and tele-perception.

- *C. H. Park is with the School of Engineering and Computing Sciences, New York Institute of Technology, New York, NY 10023. E-mail: chung.park@nyit.edu.*
- *E.-S. Ryu is with the Department of Computer Engineering, Gachon University, Seongnam-si, Gyeonggi-do 13120, Korea. E-mail: esryu@gachon.ac.kr.*
- *A. M. Howard is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332. E-mail: ayanna.howard@ece.gatech.edu.*

Manuscript received 1 Sept. 2014; revised 15 July 2015; accepted 15 July 2015. Date of publication 22 July 2015; date of current version 14 Sept. 2015.
Recommended for acceptance by N.G. Bourbakis, J.A. Gardner, N. Giudice, V. Hayward, M.A. Heller, and D. Pawluk.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TOH.2015.2460253

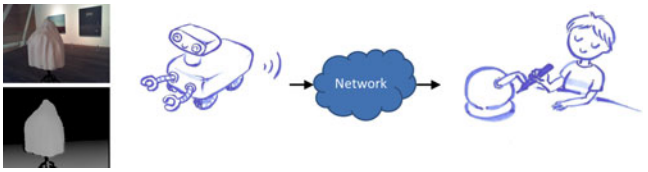


Fig. 1. Conceptual diagram of our telepresence robot based haptic exploration system for individuals with visual impairments.

Our framework enables a person to control a remotely located robot through a haptic interface while feeling the remote 3D space through haptic rendering, which transforms color and 3D spatial information through 3D haptic rendering and verbal feedback to the user, as the system architecture illustrated in Fig. 2.

In our previous studies, the haptic telepresence robotic system [18], [19], [20], [21] provided (i) sensory devices to perceive the environment and generate a 3D perception model of the environment, (ii) transmission of the environmental perception to the human user in a non-visual way, and (iii) transmission of human controls on the haptic interface to enable teleoperation and telepresence for the human user. However, the system was not directly designed for networked telepresence robotic system for individuals with VIs. Thus, this paper extends the previous research to a networked telepresence with a new video content adaptation method for the 3D visual information transmission in real scenarios of telepresence in museums and art galleries.

Experiments with three different control methods for robotic navigation are designed to evaluate the capability of remote navigation for users with and without VIs. Another set of experiments with several 3D scenes captured from real-world objects in museums and art galleries are also designed to validate how effective our system can be in transferring the haptic perception for remote object exploration.

The details of this approach are described in the following sections. Section 2 discusses previous and on-going research in the areas of assistive robotics, haptics, and networked multimedia systems. Section 3 presents the details of our system architecture, including our multi-modal signal handling and networked multimedia modules. Section 4 describes our experimental design and displays results acquired from participants, in which both individuals with and without visual impairments were recruited. Finally, we further analyze our results in Section 6 and conclude with future directions in Section 7.

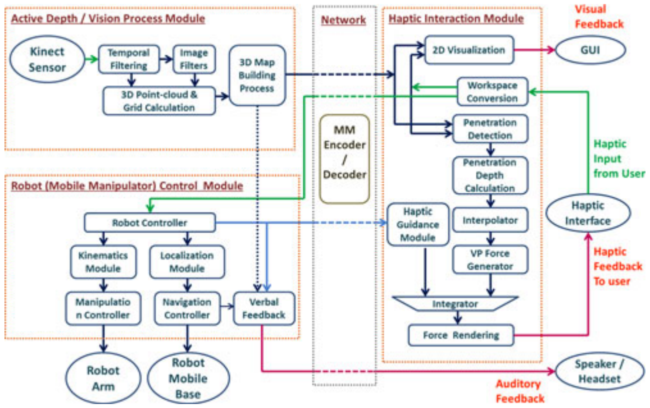


Fig. 2. System architecture for our telepresence robotic system.

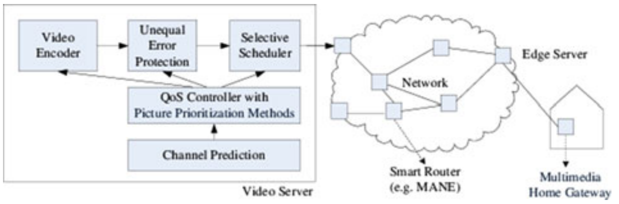


Fig. 3. The general architecture of telepresence system with video streaming technology.

2 RELATED WORK

2.1 Prior Work in Assistive Devices for VI

Efforts to aid in the daily living of individuals with VIs have been made in two directions: equipping the person with intelligent tools and changing the environment intelligently. Borenstein and Ulrich developed the GuideCane [22] that can sense obstacles in an environment and guide the individual through the environment with a robotic cane. Utilizing arrays of ultrasound sensors detecting full range of 180 degree forward, the GuideCane can detect obstacles and rotate its mobile base to adjust its path automatically, freeing the user from the efforts to steer around the walking cane to detect obstacles by physically making contacts. Jansson and Monaci developed haptic guidance system for outdoor navigation [23], Loomis et al. added approaches such as wayfinding and location identifiers [24], and Katz et al. developed vision-based navigation guidance system with sonification and text-to-speech (TTS) techniques for the visually impaired [25], which could guide a person with visual impairments with GPS navigation and this system. Recent improvements can be found from Ye [26], [27] who developed more smart cane that integrates computer vision, mapping, and object recognition technologies to generate verbal feedback of spatial information.

Approaches related to the second direction include the work from Kulyukin et al. who incorporated a mobile-robot based guidance approach with radio-frequency identification (RFID) devices to assist in navigation and provide richer information than the former robotic assistance [28]. By implanting RFID chips into objects and environmental structures and transferring each RFID information, the user can get both navigational guidance and more detailed object information. Using this concept, Kulyukin et al. developed a RoboCart that can assist individuals with VI in navigating and finding items in a grocery store that has RFID tags installed. However, this approach requires both equipping the person and the environment with assistive devices, requiring whole changes of the infrastructure.

2.2 Telepresence with Video Communications

Multimedia technology, on the other hand, has a long history of enhancements and optimization for providing rich and crisp visual information to the user as Fig. 3 [29]. This technological merit can be greatly useful in transforming visual perception (including depth) into non-visual sensation over the network. Especially, considering the large bandwidth of data stream generated from the depth sensor (that includes both 2D color information as well as 3D point clouds), the optimization of data size and network protocol is essential to achieving a reliable telerobotic communication.

Regarding the reliable telerobotic communication system, there have been several studies on telepresence using medical video communication technologies and haptic feedback. Examples of the telepresence technology include telementoring, telemedicine, and telesurgery. A telemedicine study was conducted over intercontinental cities even in the 1990's, and the first human telesurgery consultation was performed in 1996 [30]. In 2001, a telesurgery of Laparoscopic cholecystectory was demonstrated between New York and Strasbourg (France), with a network delay of 200 ms and a constant delay of 155 ms over a 10 Mbps network (3 Mbps guaranteed minimum bandwidth), achieving no packet loss in the application level and the overall subjective quality measured as 9.5 points (0:worst; 10:best) [31].

With these on-going studies, several commercialized surgical robotic systems were also developed. The da Vinci (Intuitive Surgical Inc) [32] and Zeus (Former Computer Motion) [33] are well-known systems [34], which are designed primarily for on-site surgery but also having some capabilities to provide HD (high definition) and 3D medical images and even a haptic feedback for aiding in surgery. Raven [35] is another surgical robot platform that has two robotic arms and two cameras, which is gaining more collaborating researchers through the Raven's open-source model [36].

For the reliable utilization of surgical robotic systems for telesurgery, a real-time transmission of data (e.g., medical images) is most important. However, the network delay of broadband internet is not quite predictable. For example, the telerobotic experiment Plugfest 2009 reported the delay for video transmission in inside of Unite States took from 21 to 112 ms, and the delay for worldwide transmission took from 115 to 305 ms [37]. The delays were distributed over a wide range due to many network routers between end-to-end nodes that cannot be controlled directly, which suggests that telesurgery systems without careful network-delay management could be very dangerous. Even though the study reported that most participants could adapt to the delays by maximum 500 ms, the maximum allowable delay is considered 200 ms in the area of real-time video communications. In another study, an experienced surgeon instructor demonstrated a virtual reality (VR)-based online training class (remote surgical master class) to teach the anatomic structures for surgical exercise over two cities 12,000 km apart (Stanford University in US and Canberra, Austria) [38]. This study also revealed the importance of haptic feedback as well as network quality and latency issues.

While most studies on telepresence used wired networks, some telemedicine studies are focusing on medical image transmission over integrated wireless and wired networks. One example is the medical video streaming system with a joint source channel coding and decoding (JSCC/D) technology [39]. This study introduced the JSCC technology for streaming ultrasound video, and the images and video streams acquired by OTELO (mobile-tele-Echography using an ultra-light-robot) system were transmitted. The studied system supported from 7.5 fps (frames per second) to 15 fps of video frame rate by using different network protocols according to the characteristics of data and network condition such as video, audio, frame rate, and network bandwidth. In their experiment, GOP (Group of Picture) size was from 8 to 15, and bitrate was from 201 to 384 kbps. They also protected video packet headers with robust header compression (RoHC) technology using LDPC codes, and provided unequal error protection (UEP). However, the study did not consider the delay issues in detail as well as a low resolution video (CIF) and a low network bandwidth were used for their experiments with ultrasound images instead of high quality and high resolution medical images.

2.3 Telepresence for the Visually Impaired through Haptic and Auditory Modalities

As mentioned in the Introduction, the exploration of individuals with VIs in public places, especially in spaces with high concentration on the needs of visual and tactile observation, is still a challenging problem. There have been many approaches such as exploring 3D arts in virtual reality [40] or Museum of Pure Form–a virtual reality tactile exploration system for art works at museums [41], and studies on perspectives on artists with VIs [42]. However, they required the haptic data to be prepared *a priori*, usually utilized only one modality, and the system was either complex or expensive for individual users. Thus, we set two objectives in our study: (i) combining multiple modalities for exploration and (ii) enabling the user to feel remote locations through a robot.

We chose haptic feedback and auditory signals as our main feedback modalities since haptic modality has been proven effective in extracting objective properties [43], understanding structural properties [7], cognitive mapping of unknown spaces [44], and even feeling graphical information [45] for individuals with VIs. Auditory signals are also found effective as shown in the examples of graphic exploration with simple sonic feedback [46], data sonification and visualization [47], and BATS the audio-tactile mapping system [48]. Furthermore, in alignment with prior discussion on tactile experience of art pieces [49], [50] and prior work on the Museum of Pure Form [41], this paper aimed at bringing the tactile experience of object and spatial exploration to the user through a robotic agent in a remote place.

We focus on utilizing non-visual modalities to create an interactive channel for communication between a robotic system and the user to aid in controlling the robot while achieving perception in a remote environment. To be more specific, this paper aims to provide a novel framework for a user with VI to expand the limitations of one's physical living area through a telepresence robotic assistant using haptic and auditory feedback modalities. The major contributions to this effort include: (i) the design and evaluation of a real-time interactive haptic telepresence robotic system for individuals with VI in navigation and object perception scenarios in remote places, (ii) the study and design of haptic and auditory modalities as non-visual feedback methods for in-situ perception in public places, and (iii) the design and evaluation of a vision-based haptic exploration framework that generates 3D haptic perception based on real-world 3D spatial information.

3 HAPTIC TELEPRESENCE

3.1 System Architecture

Fig. 2 illustrates the system architecture and functional modules of the haptic telepresence robotic system discussed


(a) (b)

Fig. 4. Our system platform: (a) Mobile manipulation robot with the Kinect sensor. (b) Graphic user interface and a haptic device.

in this paper. The hardware platform, as shown in Fig. 4, is composed of a robotic arm (Pioneer2 Arm) and a robotic mobile base (Pioneer 3AT) equipped with the Kinect depth sensor. Interaction with a user is achieved in three ways: (i) visual feedback through a graphic-user-interface (GUI) for typical users, (ii) haptic feedback through a haptic interface for users with VI, and (iii) verbal feedback through a speaker or a headset for users with VI.

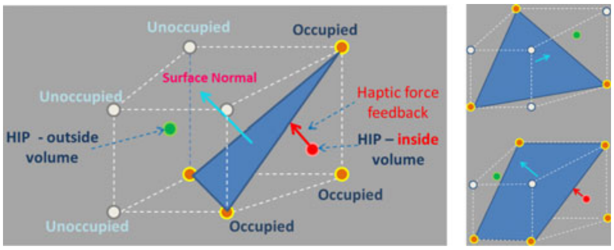
The system architecture is composed of four major functional blocks: (i) an active depth/vision perception process module, (ii) a haptic interaction module, (iii) a robot (mobile manipulator) control module, and (iv) a network module with multimedia encoder and decoder. The robot control module is described in our previous work [18], [19], [20], [21] and thus not discussed in this paper. The active depth/vision process module creates a 3D map of the environment and enables the creation of haptic/auditory feedback for the user in other modules. Haptic interaction is linked with the verbal feedback module, which synchronously transfers status on 3D map representation (depth and color) and robot navigation status to present haptic and verbal feedback to the user. Lastly, the network module with multimedia encoder and decoder handles the multimedia content adaptation and robust transmission technology for individuals with VIs by enhancing the sensed 3D visual information for the remote haptic device.

3.2 Haptic Perception of Remote Environment

We have developed a 3D haptic rendering algorithm based on 3D-point maps generated by either a stereo-vision process or a Kinect depth camera [19], [20], [21]. In this paper, we discuss the extension of the system with a networked configuration with multimedia compression and multi-modal interaction. To handle continuous depth and color image streams from the Kinect and transfer them over the network, a two-stage pipeline structure has been constructed [21].

The first stage of the pipeline projects the Kinect depth data frame from the buffer into a 3D coordinate system then transform it into a 3D occupancy map, as illustrated in Fig. 5. The occupancy map is used to represent a discretization of objects in the observed volumetric objects in the environment by marking occupied and unoccupied cells. The details of the process mainly involve image filtering and noise handling of the Kinect sensor, which are described in our previous work [21].

The second stage of the pipeline handles the haptic rendering from a 3D map while updating the 3D map as the depth data stream comes in from the Kinect sensor. First, the haptic interaction point (HIP) corresponding to the user's movement on the haptic interface (user's haptic workspace) is projected into the 3D map (virtual volumetric space of the remote environment). Then, the virtual proxy forces are calculated by finding the closest surface point of the object the HIP is penetrating from the 3D map. To expedite the calculation for the virtual-proxy algorithm, neighbor-based surface normals and penetration depth estimation process are incorporated as illustrated in Fig. 5.



(a) (b)

Fig. 5. Illustration of a virtual-proxy force calculation: (a) Virtual-proxy force estimation within neighbor points. (b) Example of the surface-normal estimation given neighboring 3D Points.

For rendering the virtual-proxy force for the haptic interface, a typical spring-damper model is used as in Fig. 6 and Eq. (1). The spring-damper model is widely adopted in haptic rendering [51], [52], [53] due to its fast computation with simple second-order systems, which are suitable for reliable haptic rendering with fast control cycles of the haptic interface. ($\vec{F}_{VP}$: virtual-proxy force-feedback, k: spring constant, b: damping constant, $\vec{P}_{proxy}$: position vector of the proxy, $\vec{P}_{HIP}$: position vector of the HIP, and $\vec{V}_{proxy}$ and $\vec{V}_{HIP}$: velocities of the proxy and the HIP)

$$\vec{F}_{VP} = k(\vec{P}_{proxy} - \vec{P}_{HIP}) + b(\vec{V}_{proxy} - \vec{V}_{HIP}). \tag{1}$$

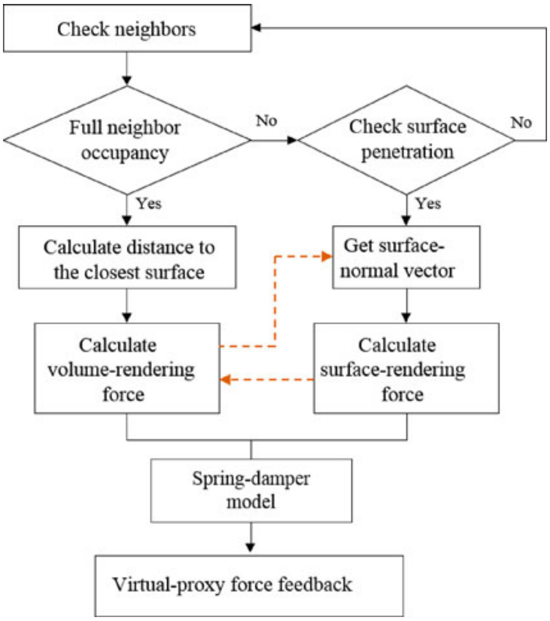
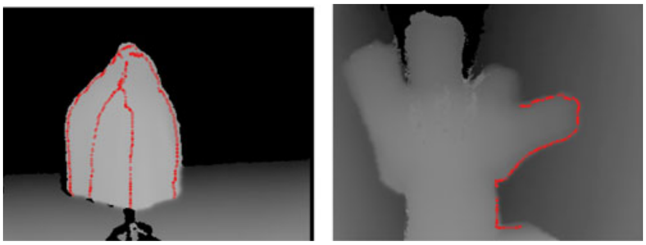


Fig. 6. Haptic virtual-proxy algorithm for calculating interaction force with a 3D map.



(a) (b)

Fig. 7. Example trajectories during 3D haptic exploration (points marked in 30 ms intervals).

This way, we achieve direct haptic rendering from RGB-D perception and real-time interaction with dynamic environments as shown in Fig. 7. These benefits encourage real-time interaction and exploration in the remote environment through the robotic platform, bringing the telepresence of the user closer into a realistic application.

3.3 Voice Feedback for Supplementary Modality

To supplement the user's perception over a remote environment, we utilize auditory feedback to translate color and distance information of the remote environment the user is exploring with the haptic interface remotely. Many individuals with VI have prior perception of colors, and even if a person was born blind the person learns about color through literatures and education. Our approach about color information was in fact due to a strong request from participants in our previous studies [19], [21], [54] and the above knowledge was gained through our interaction with them over more than three years. The use of color information can be also found in other studies on indoor navigation, object identification [55], and clothes matching for blind users [56].

The auditory feedback consists of verbal descriptions of color and distance information synchronized with an object the HIP is in contact with. The color and/or distance information is reported to the user when the HIP is interacting with an object in the 3D map. The color information is retrieved by (i) calculating the surface proxy point of the HIP in the 3D map (haptic perception process) and (ii) accessing the color information of the corresponding pixel in the color image stream. We have implemented our voice system to provide the color information in 65 colors, which can be reduced into fewer (simpler) color names or expanded into larger color sets (the list of colors recognized by our system are defined in Table 1). The auditory feedback also includes brief verbal reports on the status of robotic movements when the robot is navigating. The verbal description of the status of our robotic system consist of "forward", "left", "right", "backward", and "stop", and are reported only when the status changes or status is requested by the user.

3.4 Multimedia Transmission and Content Adaptation

The multimedia transmission and content adaptation method for individuals with VIs enhances the sensed 3D visual information for remote haptic exploration. The method consists of six sub-modules: an adaptive filter and a down-sampler, a 2D image and depth map encoder, an adaptive FEC module, a quality of service (QoS) controller in the server side, and a haptic device tracer with a network monitor in the client side as shown in Fig. 8.

TABLE 1
List of Color Names Differentiable with Verbal Feedback

#	Color	#	Color	#	Color	#	Color
1	Aqua	18	Gold	35	Midnight blue	52	Sea green
2	Aquamarine	19	Green	36	Navy blue	53	Sienna
3	Black	20	Green yellow	37	Neon blue	54	Sky blue
4	Blue	21	Grey	38	Olive	55	Slate blue
5	Blue violet	22	Hot pink	39	Olive green	56	Spring green
6	Bronze	23	Indian red	40	Orange	57	Steel blue
7	Brown	24	Ivory	41	Orange red	58	Summer sky
8	Cadet blue	25	Khaki	42	Orchid	59	Turquoise
9	Charteuse	26	Lawn green	43	Pale green	60	Violet
10	Chocolate	27	Light blue	44	Pale violet	61	Violet red
11	Coral	28	Light coral	45	Peach	62	White
12	Dark green	29	Light gold	46	Pink	63	Wood
13	Dark grey	30	Light sky blue	47	Plum	64	Yellow
14	Deep pink	31	Lime green	48	Purple	65	Yellow green
15	Deep sky blue	32	Magenta	49	Red		
16	Dodger blue	33	Maroon	50	Royal blue		
17	Forest green	34	Medium purple	51	Scarlet		

Once the Kinect sensory device generates 2D images with 3D depth data, the method first builds an enhanced 3D depth-map by using both original 2D image and depth data and then applies a low-pass filter to raw 2D image and downsamples the image and depth map temporally according to the decision of the QoS controller. For an individual with VIs who is using the haptic device, high resolution and temporal video may not be always required. Thus, the QoS controlling helps to reduce the required bandwidth as well as the computational complexity of the system. Furthermore, the method encodes the 2D images and depth maps with given encoding option by the QoS controller. For example, each pixel consists of luma (brightness) and chroma (color) components, and the detailed color information of the image is not needed. Thus, the QoS controller applies lower quantization-parameter (QP) values for the chroma components compared to luma components. In addition, the encoder applies different encoding technology

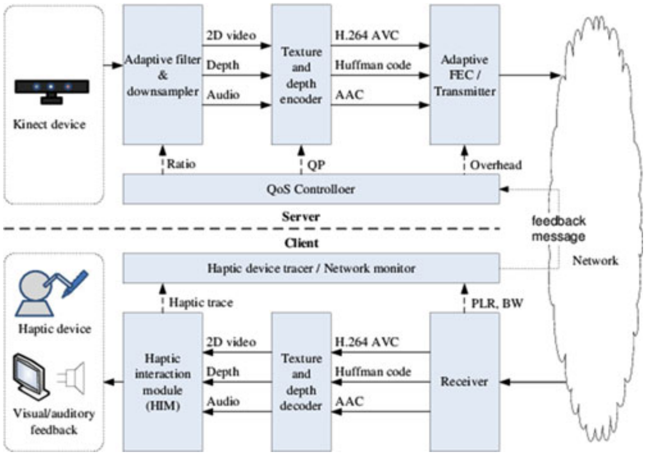


Fig. 8. The multimedia transmission and content adaptation module in the proposed system.

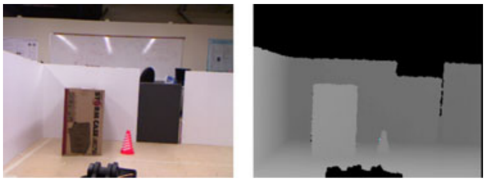

(a) Color view (b) Depth view

Fig. 9. Tutorial site.

to the 2D image and the 3D depth map [57] [58]. To overcome situations of degraded overall quality of experience (QoE) due to any loss in the depth data, the encoder uses H.264 AVC for the 2D image and Huffman coding method for the depth map, which is more important than the image quality for the user.

The FEC codes protect the encoded bitstream with some amount of overhead according to network conditions (e.g., packet loss rate and bandwidth). The QoS controller receives feedback messages from the haptic device tracer and the network monitor of client and makes a decision to control the other modules in the server. In real-time video/audio transmission, maximum of 200 ms delay in video and 70 ms delay in audio are preferable because human user can feel the time differences if the delay is more than those thresholds, which are basically from experiment-based psycho-visual and psycho-acoustic models. However, our proposed system design allows more than 200 ms of delay in video because individuals with VIs take some time to feel 3D image by exploring the surface of transmitted image (a video frame) with a haptic device. Besides, the exploring time depends on the complexity of an image and the sensitivity of each person. Thus, our designed system updates the image when the user explores certain amount of image surface areas with the haptic device (e.g., 80 percent). As an example, if a user moves the haptic interface to feel the surface of a remote object (e.g., a human face, an art work in a museum, etc.), the haptic device traces its movement and measures average elapsed time for exploration. Then, the module sends feedback information to the QoS controller in the server to adjust a video frame rate adaptively. This is for the purpose of our system design, and the system provides an adaptive frame updating rate.

The bitrate saving by the content adaptation is highly related to the perceptual sensitivity of individuals with VIs, and the bitrate savings could be different for each person. Thus, the perceptual sensitivity of individuals with VIs could be an important research topic to increase overall QoE. To increase the QoE of 3D perception, the texture and depth decoder module applies the bilateral filter [59] to original 2D images and acquires images with the edges smoothed. Then, the 3D depth-map enhancement method crops the edge image because the resolutions of the original 2D image and depth-map are different, and the method merges the cropped edge image to depth map. Additionally, the temporal median filter and histogram equalization could be applied for the depth-map to enhance the image. Thus, the proposed module provides the smoothed and emphasized 3D shapes with sharpened edges, which helps individuals with VIs to feel 3D images more precisely.

The proposed system architecture includes the Kinect sensor from Microsoft that provides depth and color images at a rate of approximately 30 fps. The Kinect depth data of the captured frame is buffered, projected into a 3D coordinate system with resolution 640 × 480, then transformed into a 3D map (640 × 480 × 1024) structure. By using open source *ffmpeg* library, H.264 AVC encoder and the Huffman coding encode 2D images and depth maps, respectively.

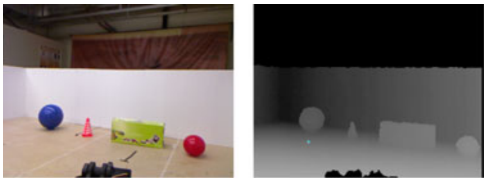

(a) Color view (b) Depth view

Fig. 10. Experimental setup.

4 EXPERIMENT DESIGN AND RESULTS

We implemented the system framework discussed in Section 3.1 using our mobile manipulation robot platform (Fig. 4), and designed a set of experiments to evaluate the system performance as well as the users' responses. The first experiment (Section 4.1) was designed to evaluate the haptic and auditory modalities in navigating the telepresence robot in public places such as museums or art galleries. For safety and permission issues, we setup a test site in our lab at this stage. We also designed the second experiment (Section 4.2) to evaluate the capability or our system in enabling users to feel objects and distinguish different collections in art galleries or museums. A total of 26 participants (five female and 21 male) joined our experiments. Among them, five had visual impairments (with VI: two were low-vision and three were blind) and 21 were sighted (without VI). The participants were divided into two groups: first group of 14 participants including five individuals with VI joined Experiment 1, and the other group of 12 sighted participants joined Experiment 2. There was no overlap in the user groups between two experiments.

4.1 Experiment 1: Robotic Navigation with Haptic and Auditory Exploration

This experiment was performed given a common haptic input method for haptic exploration and a headset for auditory feedback. Participants were allowed to use two distinctive control mechanisms: *keyboard control (KeyNav)* and *haptic click (AutoNav)*. The Keynav is designed to enable the user to control the robot using the keyboard, i.e., up arrow: move forward, left arrow: turn left, right arrow: turn right, down arrow: move backward, and shift key: stop. The AutoNav is designed to enable the user to just click on a 3D environment represented through the haptic interface to control the robot to autonomously reach the clicked goal point in the remote environment. Using these two control mechanisms, the user is asked to accomplish a goal, which was to "*find a blue/red ball and approach it with the robot*." Prior to the experiment, participants were given a one minute tutorial for using the haptic interface and understanding 3D scenes based on a tutorial site. The spatial setups of a tutorial site and the actual remote scene (5 × 6 m space) for the experiments were as shown in Figs. 9 and 10, respectively.

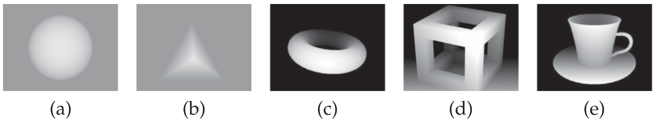

(a) (b) (c) (d) (e)

Fig. 11. Training models for haptic perception of object.

The procedures for our experiments are as follows:

1) Consents from adults or parents' consent along with verbal assents from minors are acquired prior to the experiment.
2) Explanation on the robotic platform, computer system, and the haptic interface is given.
3) Tutorial on the haptic exploration is given. The space contains two objects, a box and a cone, that have different sizes from objects to be placed in the actual experiment scene.
4) Tutorial on the control methods—KeyNav and AutoNav—are given to the participant in a random order.
5) Explanation on the auditory feedback—color, distance, and the status of the robot—are provided to the participant.
6) The participant starts controlling the robot and accomplishes the task of finding a specific-colored ball. Two balls, of different sizes and colors, are placed in different locations inside the test scene. To rule out the memorization effect, each participant is asked to find a ball with a random choice of color (between red and blue) in the first trial, and then asked to find the other colored ball in the second trial.

The hypothesis we wanted to study through the first experiment were as follows:

- Hypothesis 1: Linkage between control modality and haptic feedback can affect the performance of teleoperation.
- Hypothesis 2: Multi-modal (haptic and auditory) feedback for environmental perception can aid in teleperception of an individual with a visual impairment.

We determined the independent variables as the haptic and auditory feedback methods and the control method. The dependent variables we measured were the task completion time, the success rate, and the frequency of giving commands to the robot. As for the controlled variables, we fixed the number of target objects and object shapes, which were presented to the participant. During the experiments, we measured dependent variables on the following criteria: (i) Time taken for task completion, (ii) number of control sequences made by the participant, and (iii) the trajectory of the robot. After finishing the experiments, a questionnaire with eight items were provided for the user to complete. After the experiments, all participants were asked with a small set of questions for qualitative analysis:

- Q1: How easy was the haptic exploration in perceiving 3D shapes? *(1. Very Easy 2. Easy 3. Neutral 4. Somewhat difficult 5. Very difficult)*
- Q2: How helpful was the verbal feedback on colors in distinguishing the shapes? *(1. Very helpful 2. Helpful 3. Neutral 4. Somewhat helpful 5. Not helpful at all)*
- Q3: How realistic did you feel like being telepresent through the robotic system with haptic and auditory exploration? Please answer in the scale of 1~7 *(1: Not realistic, 7: Very realistic)*.

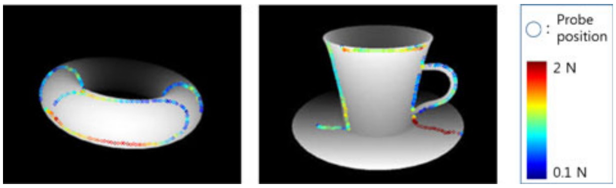


Fig. 12. Trajectories during 3D haptic exploration of the doughnut and tea cup shapes. The color represents the intensity of virtual-proxy force feedback.

4.2 Experiment 2: Object Perception with Haptic and Auditory Exploration

The second experiment was performed with same human interfaces (haptic and auditory interfaces) but without the telepresence. Since our telepresence robotic system is at a research stage and the safety and privacy issues have not been cleared yet, we only utilized Kinect sensor to capture 3D scenes of collections from art galleries and museums. The collected 3D scenes were later provided to users invited to our lab to participate in the experiment. Prior to the experiment, participants are given a one minute tutorial for using haptic interface and understanding 3D scenes based on simple geometric models of different shape and complexity. Then, the participants are allowed to use the same interface to explore objects from real 3D scenes. Three sets of three~four objects were provided to the participant, and then allowed to freely explore the object for 30 sec ~ 1 minute per each model. The participant was provided with haptic force feedback through the haptic interface, visual feedback of the object and HIP through a graphic user interface. Participants were divided into two groups (control and experimental group), only the experimental group was provided with the verbal feedback of color information through speakers or headsets. After feeling one set of objects, the participant was provided with a paper with the set of objects printed on, was provided with a randomly ordered 3D objects from the set while being blocked from using the GUI (i.e., the visual feedback was turned off), and was asked to tell which object the user was touching with the haptic interface. Images from the second experiment were shown below for training models (Fig. 11) and actual 3D scenes from museums and art galleries (Fig. 13).

The hypothesis we wanted to study through the second experiment were as below:

- Hypothesis 3: Verbal feedback on color information can aid haptic perception of complex 3D objects such as collections from art galleries and museums.

5 RESULTS

5.1 Experiment 1: Navigation and Haptic Exploration

The individuals without VI were blindfolded and individuals with low-vision were restricted to face the computer screen to simulate physically similar conditions among participants. All participants are safely seated in front of the desk as other participants and were guided through the experiment with same condition. The age group of participants in the

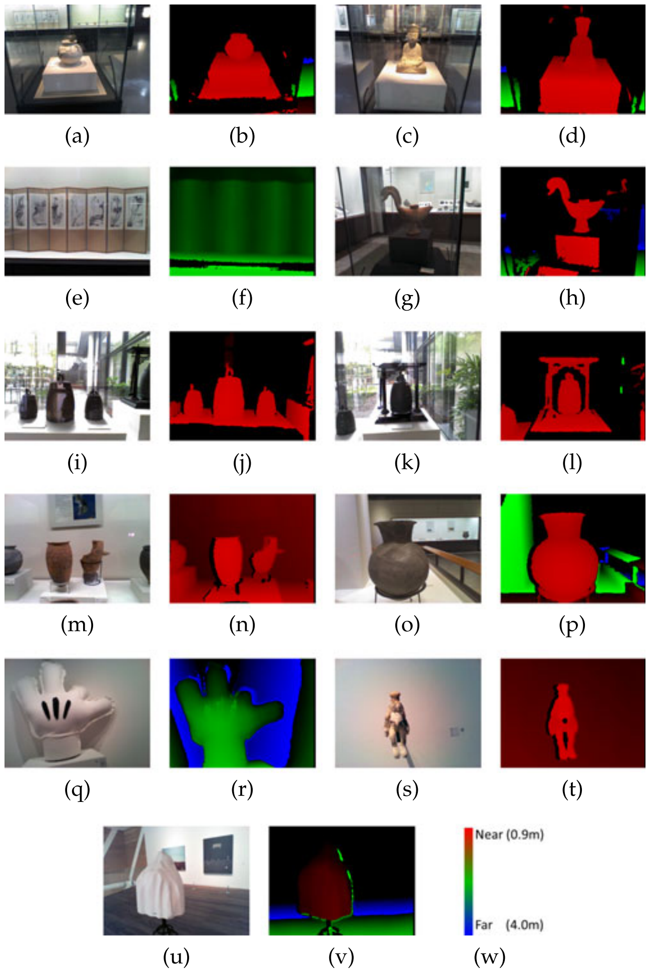


Fig. 13. 3D scenes for experiment 2 (color view and depth view pairs). (a∼h): Set 1 - Museum, (i∼p): Set 2 - Museum, (q∼v): Set 3 - Art Gallery, (w):color scale for depth representation. [Photos in courtesy of artists: (q) “EVOLG-Part2” by Mun Kyung Chung, (s) “YFOOG” by Mun Kyung Chung, and (u) “Someone We Know” by Min Sook Kang].

Experiment 1 was between 14 and 48 (average = 29.3, standard deviation (SD) = 9.16). Each participant was given a tutorial of the system (about 5-10 minutes) with instructions on the two control mechanisms in random order. Then, each participant was given two trials in a mixed order. We recorded log files to quantitatively evaluate the user’s performance in KeyNav and AutoNav trials. The experimental results and analysis are discussed in the following sections.

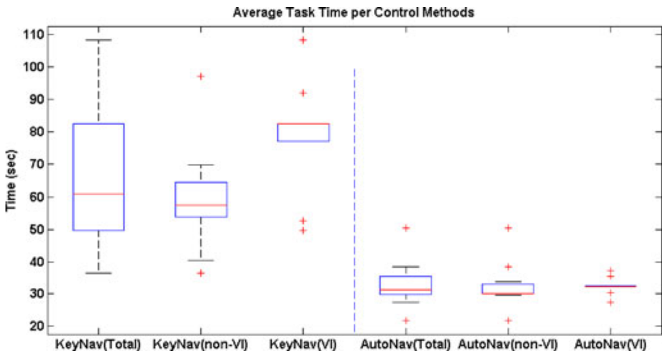


Fig. 14. Task time comparison between KeyNav and AutoNav.

TABLE 2
Success Rates of Participants in Achieving the Tasks Using Different Control Methods

Participants Control	1	2	3	4	5*	6	7*
KeyNav	O	O	O	X	O	O	X
AutoNav	O	X	X	O	O	X	O
Participants Control	8	9*	10	11	12	13*	14*
KeyNav	X	O	X	X	O	X	O
AutoNav	O	O	O	O	O	O	O

(: participants with VI, O: success, X: failure)*

5.1.1 *User Performance for Task Completion*

To evaluate user performance with our telepresence system, we measured the control time and the success rate for KeyNav and AutoNav. The average times taken for task completion were 65.52 sec (SD = 22.19 sec) for KeyNav and 32.81 sec (SD = 6.58 sec) for AutoNav. The median values and the variations are depicted as a bar plot in Fig. 14. This shows that, on average, users (with both VI and non-VI) were able to minimize the time to complete the ”find/approach ball” task when using the robot control method in AutoNav, as opposed to KeyNav.

The condition for goal achievement was classified as the point at which the telepresence robot had reached within a distance of 35 cm from the target, since the robot’s manipulator has a workspace range between 25 and 50 cm. When assessing how often users were successfully controlling the robot to achieve the task, we see that the success rates for KeyNav and AutoNav for the 14 participants were 57.1 and 78.6 percent respectively, as shown in Table 2. (The average success rate of individuals with VI were 60.0 and 100.0 percent respectively.)

We applied the repeated-measure ANOVA on the user performance results measured from task completion time and success rates. As Table 3 shows, there was a statistically significant effect of haptic modality on control with respect to task completion time, $F(1,13) = 27.18, p < 0.0002$.

However, Table 4 displays that there was not enough statistical significant effect of control modality on the success rate, $F(1,13) = 0.47066, p < 0.2178$.

5.1.2 *Control Load for the User*

Another important aspect in human-robot interaction is the factor of control load or amount of workload imposed on the user. We measure this term by counting the frequency of robot commands given by the human user, i.e. the number of keyboard strokes in KeyNav and the click counts in

TABLE 3
Repeated-Measure ANOVA on Task Completion Time (SS: Sum of Squares, df: Degree of Freedom, MS: Mean Squares)

Source	SS	df	MS	F-statistic
Time	7490.1	1	7490.07	27.18
Error	3582.3	13	275.56	

TABLE 4
Repeated Measure ANOVA on Success Rate
(success = 1, fail = 0)

Source	SS	df	MS	F-statistic
Time	0.32143	1	0.32143	0.3356
Error	4.17857	13	0.32143	

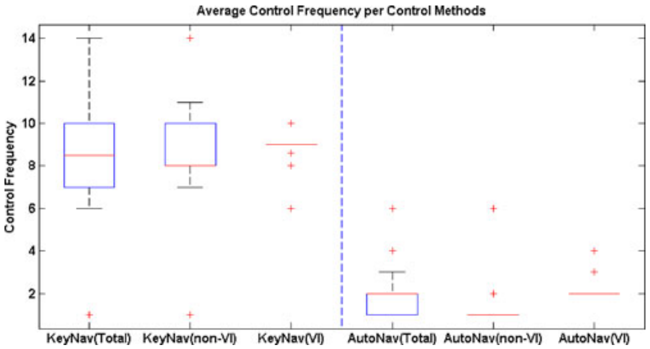


Fig. 15. Control frequency comparison between KeyNav and AutoNav.

AutoNav. The average control frequency taken for task completion was 8.5 (SD = 2.955) for KeyNav, and 2.1 (SD = 1.439) for AutoNav. The median values and the variations are as shown in Fig. 15. This shows that, on average, users (with both VI and non-VI) were able to accomplish the task with minimal effort to complete the "find/approach ball" task when using the robot control method in AutoNav, as opposed to KeyNav.

5.2 Experiment 2: Object Perception

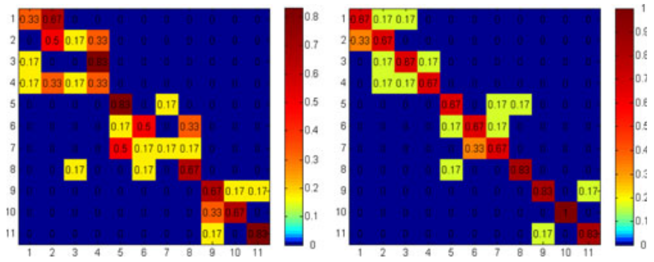
For evaluating haptic perception of real objects in museums and art galleries, we captured RGB-D scenes (color images + depth images) in the Museum of Art at Seoul National University (MoA-SNU, http://www.snumoa.org/) and Seoul National University Museum (http://museum.snu.ac.kr/) in Korea, as shown in Fig. 13. From the dataset, we selected different objects' scenes with the rule of the followings: (i) Whether it have single object or multiple objects in the scene; (ii) whether the objects are rather planar, circular, or rectangular, etc; (iii) whether the objects represent living things or inanimate objects; and finally (iv) whether the shape is simple (geometric) or more complex (or artistic). Total 10 scenes for three sets of experiments (Museum 1, Museum 2, and Gallery) were selected so we can show 3-4 objects per set to the participant. We provided three sets of 3D objects (3-4 objects in each set) to the user through haptic experience, and measured the correct matching rate of the participants' answers in blind haptic exploration. The recognition rate of participants are shown in Table 5, and the confusion matrix for the total 11 objects is visualized in Fig. 16. On average, the Experiment group who were given both haptic and auditory feedback showed about 33 percent higher recognition rage compared to the Control group who were only given the haptic feedback. These results clearly showed that the verbal feedback of color information during haptic exploration significantly increased the perception of remote objects.

Initial observation on the participants' trials displayed the tendency in users to make errors (get confused) due to the users' biased attention on verbal feedback on colors, or haptically exploring only on partial area in the workspace thus observing only part of the object. However, considering the users were all novice haptic device users (with less than 1 hour of haptic experience), we expect increased performance for more experienced users. We will evaluate this assumption in future studies.

TABLE 5
Average Recognition Rate for 3D Objects from
a Museum and an Art Gallery

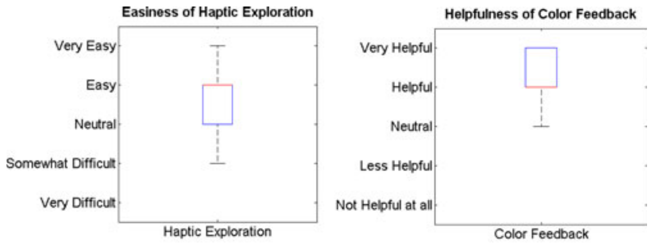
Object Sets	1 (Museum)	2 (Museum)	3 (Gallery)	Total
Control	29.2%	50.0%	72.2%	48.5%
Experiment	62.5%	70.8%	88.9%	72.7%

The verbal feedback on color information was provided only to the Experiment group, but not to the Control group.

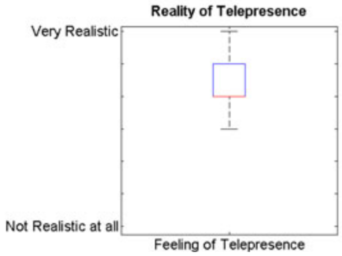


(a) Control group (b) Experiment group

Fig. 16. Confusion matrix on the participants' object recognition rates visualized with heatmap (1∼4: Set 1, 5∼8: Set 2, and 9∼11: Set 3). Control group was not given any verbal feedback, while Experiment group received verbal feedback of color information during haptic exploration.



(a) Responses to Q1 (b) Responses to Q2



(c) Responses to Q3

Fig. 17. Responses to the questionnaire.

5.3 Questionnaire Results

The participant's responses to the questionnaire were collected to get qualitative measure on the users' experiences. As shown in Fig. 17, the participants found the haptic modality fairly easy to work with and the voice feedback on the color information quite helpful for exploration. Also, the users reported that the feeling of telepresence was fairly realistic, which suggests that our system is feasible to transfer the feeling of "telepresence" with haptic and auditory modality to the user.

6 DISCUSSION

The experimental results demonstrated that our robotic telepresence system is capable of enabling the user to explore remote environment and interact with the system through haptic and auditory modalities on 3D perception of remote environment.

To evaluate Hypothesis 1 ("Whether providing linkage between control modality and haptic feedback can affect the performance of tele-operation"), we applied the repeated-measure ANOVA on the user performance results measured from task completion time and success rates. There was a statistically significant effect of haptic modality on control with respect to task completion time, however, not with the success rates. The latter analysis can be explained by our observation that individuals without VI, although blindfolded, showed tendency to make decisions only with color information. To develop a more quantitative answer, we are expanding the number of participants, both with VI and without VI, for further analysis.

To evaluate Hypothesis 2 ("Whether providing multimodal (haptic and auditory) feedback for environmental perception can aid in tele-perception of an individual with a visual impairment"), we designed the experiment to include two balls with different colors (blue and red) and a box with multiple colors (shapes printed on its surface). The results from the experiments showed that participants distinguished between boxes and balls easily, and mostly succeeded in finding the right colored balls according to the given task. This implies that our system achieved tele-perception through multi-modal means.

Finally, to evaluate Hypothesis 3 ("Verbal feedback on color information can aid haptic perception of complex 3D objects such as collections from art galleries and museums"), we have provided the individuals with 11 haptic 3D models from real collections in a museum and an art gallery to two user groups with and without the verbal feedback. For the control group, we have provided only haptic feedback (and visual feedback for initial explorations) to the user. For the experimental group, we also provided verbal feedback on color information, and we studied the effectiveness of auditory feedback as a supplementary modality to haptic modality.

For follow-up studies, we are conducting more evaluations with participants with visual impairments. We are using 3D printed models from the scenes we have acquired depth data, with which the participants are asked to first feel virtual shapes with our haptic interface and then to match from the 3D printed models. We will include this process for a prospective study with more extensive lists of objects in diverse environments.

Lastly, we want to add that a number of studies exists that show the importance of distributed information during haptic exploration [60] and the effectiveness of other haptic contacts besides point-based haptic interaction, such as rolling and sliding [61] or thimble-based interaction for cutaneous perception [62]. The goal of this paper is to address the challenging issue of remote haptic exploration (with vocal cues for colors on objects) and investigate the effectiveness of real-time haptic rendering in real-world for the remote exploration of individuals with VIs. For investigating a feasible solution, our current approach has been utilizing a low-cost equipment, which led us to start with a point-based haptic interaction and exploration. Cutaneous haptic feedback [63] and textural rendering [64] are our next goals to add on to this study, which we plan to pursue through collaboration.

7 CONCLUSION

In this paper, we have developed a framework for telepresence robotic system with haptic and auditory feedback for individuals with visual impairments. We have investigated the capability of this system in enabling the user to (i) teleoperate the robot to navigate in a remote 3D environment while feeling the 3D structure (with a haptic interface) and hearing the visual components (in a verbal description of color and distance) and (ii) take more closer observation of objects in display (i.e., collections in museums or art galleries) and feel the complex shapes and colors through the eyes of a telepresence robot. The results showed good indicators for a novel framework in achieving the goals, and we expect that our system will provide a basis for developing more realistic and high-performing telepresence robotic system with multi-modal feedback for individuals with visual impairments. We plan to study further by upgrading our system functionality with more sophisticated haptic sensations (e.g., textures and deformations in physical interactions) and contents adaptive multi-media handling.

ACKNOWLEDGMENTS

This research was supported in part by the US National Science Foundation under Grant No. 0940146. The authors would like to thank the Museum of Art at Seoul National University (MoA-SNU, http://www.snumoa.org/) and Seoul National University Museum (http://museum.snu.ac.kr/) in Korea for allowing access to their collections. The human subject experiments were performed under IRB Protocol H12125 and IRB Protocol BHS-1053. This work was supported in part by the NSF under Grant No. 0940146.

REFERENCES

[1] D. Pascolini and S. P. Mariotti, "Global estimates of visual impairment: 2010," *Brit. J. Ophthalmol. Online*, vol. 96, no. 5, pp. 614–618, 2011.

[2] U. R. Roentgen, G. J. Gelderblom, M. Soede, and L. P. de Witte, "Inventory of electronic mobility aids for persons with visual impairments: A literature review," *J. Vis. Impairment Blindness*, vol. 102, no. 11, pp. 702–724, 2008.

[3] E. L. Lamoureux, J. B. Hassell, and J. E. Keeffe, "The determinants of participation in activities of daily living in people with impaired vision," *Amer. J. Ophthalmol.*, vol. 137, no. 2, pp. 265–270, 2004.

[4] A. Amedi, N. Raz, H. Azulay, R. Malach, and E. Zohary, "Cortical activity during tactile exploration of objects in blind and sighted humans," *Restorative Neurol. Neurosci.*, vol. 28, no. 2, pp. 143–156, 2010.

[5] A. D. Angiulli, J. M. Kennedy, and M. A. Helle, "Blind children recognizing tactile pictures respond like sighted children given guidance in exploration," *Scandinavian J. Psychol.*, vol. 39, no. 3, pp. 187–190, 1998.

[6] E. D. Mynatt and G. Weber, "Nonvisual presentation of graphical user interfaces: Contrasting two approaches," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 1994, pp. 166–172.

[7] A. Afonso, A. Blum, B. F. Katz, P. Tarroux, G. Borst, and M. Denis, “Structural properties of spatial representations in blind people: Scanning images constructed from haptic exploration or from locomotion in a 3-d audio virtual environment,” *Memory Cogn.*, vol. 38, no. 5, pp. 591–604, 2010.

[8] R. Held, “Telepresence,” *J. Acoust. Soc. Amer.*, vol. 92, p. 2458, 1992.

[9] F. Michaud, P. Boissy, H. Corriveau, A. Grant, M. Lauria, D. Labonte, R. Cloutier, M. Roux, M. Royer, and D. Iannuzzi, “Telepresence robot for home care assistance,” in *Proc. AAAI Spring Symp.*, 2006, vol. 20006, pp. 50–55.

[10] D. Schloerb, “A quantitative measure of telepresence,” *Presence: Teleoperators Virtual Environ.*, vol. 4, no. 1, pp. 64–80, 1995.

[11] M. A. Hersh, *Assistive Technology for Visually Impaired and Blind People*. New York, NY, USA: Springer, 2007.

[12] California Dept. Social Services, Office of Services to the Blind, “Handbook of resources and services for persons who are blind or visually impaired,” 2012.

[13] West Virginia Dept. Education, Office of Special Programs, “Assistive technology for students who are blind or have low vision,” 2009.

[14] K. Salisbury, F. Conti, and F. Barbagli, “Haptic rendering: Introductory concepts,” *IEEE Comput. Graph. Appl.*, vol. 24, no. 2, pp. 24–32, Mar./Apr. 2004.

[15] M. Slater and S. Wilbur, “A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments,” *Presence: Teleoperators Virtual Environ.*, vol. 6, no. 6, pp. 603–616, 1997.

[16] E. Milios, B. Kapralos, A. Kopinska, and S. Stergiopoulos, “Sonification of range information for 3-d space perception,” *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 11, no. 4, pp. 416–421, Dec. 2003.

[17] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, “RGB-d mapping: Using Kinect-style depth cameras for dense 3d modeling of indoor environments,” *Int. J. Robot. Res.*, vol. 31, no. 5, pp. 647–663, 2012.

[18] C. H. Park and A. M. Howard, “Towards Real-time haptic exploration using a mobile robot as mediator,” in *Proc. IEEE Haptics Symp.*, 2010, pp. 289–292.

[19] C. H. Park and A. M. Howard, “Real world haptic exploration for telepresence of the visually impaired,” in *Proc. ACM/IEEE Int. Conf. Human Robot Interaction*, 2012, pp. 65–72.

[20] C. H. Park and A. M. Howard, “Real-time haptic rendering and haptic telepresence robotic system for the visually impaired,” in *Proc. World Haptics Conf.*, 2013, pp. 229–234.

[21] C. H. Park and A. M. Howard, “Robotics-based telepresence using multi-modal interaction for individuals with visual impairments,” *Int. J. Adaptive Control Signal Process.*, vol. 28, no. 12, pp. 1514–1532, 2014.

[22] I. Ulrich and J. Borenstein, “The GuideCane-applying mobile robot technologies to assist the visually impaired,” *IEEE Trans. Syst., Man Cybern., A, Syst. Humans*, vol. 31, no. 2, pp. 131–136, Aug. 2002.

[23] G. Jansson and L. Monaci, “Identification of real objects under conditions similar to those in haptic displays: Providing spatially distributed information at the contact areas is more important than increasing the number of areas,” *Virtual Reality*, vol. 9, no. 4, pp. 243–249, 2006.

[24] J. M. Loomis, R. G. Golledge, R. L. Klatzky, and J. R. Marston, “Assisting wayfinding in visually impaired travelers,” in *Applied Spatial Cognition: From Research to Cognitive Technology*. Mahwah, NJ, USA: Lawrence Erlbaum Associates, 2007, pp. 179–202.

[25] B. F. Katz, F. Dramas, G. Parseihian, O. Gutierrez, S. Kammoun, A. Brilhault, L. Brunet, M. Gallay, B. Oriola, M. Auvray, P. Truillet, M. Denis, S. Thorpe, and C. Jouffrais, “Navig: Guidance system for the visually impaired using virtual augmented reality,” *Technol. Disability*, vol. 24, no. 2, pp. 163–178, 2012.

[26] A. Tamjidi, C. Ye, and S. Hong, “An extended kalman filter for pose estimation of a portable navigation aid for the visually impaired,” in *Proc. IEEE Int. Symp. Robot. Sensors Environ.*, 2013, pp. 178–183.

[27] X. Qian and C. Ye, “3d object recognition by geometric context and Gaussian-mixture-model-based plane classification,” in *Proc. IEEE Int. Conf. Robot. Autom.*, 2014, pp. 3910–3915.

[28] V. Kulyukin, C. P. Gharpure, J. Nicholson, and S. Pavithran, “RFID in robot-assisted indoor navigation for the visually impaired,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2004, vol. 2, pp. 1979–1984.

[29] E.-S. Ryu and N. Jayant, “Home gateway for Three-screen tv using h.264 SVC and raptor FEC,” *IEEE Trans. Consum. Electron.*, vol. 57, no. 4, pp. 1652–1660, Nov. 2011.

[30] T. Haidegger, J. Sándor, and Z. Benyó, “Surgery in space: The future of robotic telesurgery,” *Surg. Endoscopy*, vol. 25, no. 3, pp. 681–690, 2011.

[31] J. Marescaux, J. Leroy, F. Rubino, M. Smith, M. Vix, M. Simone, and D. Mutter, “Transcontinental robot-assisted remote telesurgery: Feasibility and potential applications,” *Ann. Surg.*, vol. 235, no. 4, pp. 487–492, Apr. 2002.

[32] G. Ballantyne and F. Moll, “The da vinci telerobotic surgical system: The virtual operative field and telepresence surgery,” *Surg. Clinics North Amer.*, vol. 83, no. 6, pp. 1293–1304, Dec. 2003.

[33] J. Marescaux and F. Rubino, “The zeus robotic system: Experimental and clinical applications,” *Surg. Clinics North Amer.*, vol. 83, no. 6, pp. 1305–1315, Dec. 2003.

[34] D. Holt, A. Zaidi, J. Abramson, and R. Somogyi, “Telesurgery: Advances and trends,” *Univ. Toronto Med. J.*, vol. 82, no. 1, pp. 52–54, 2004.

[35] M. Lum, D. Friedman, G. Sankaranarayanan, H. King, K. Fodero, R. Leuschke, B. Hannaford, J. Rosen, and M. Sinanan, “The raven: Design and validation of a telesurgery system,” *Int. J. Robot. Res.*, vol. 28, no. 9, pp. 1183–1197, 2009.

[36] T. Stephens. (2012, Jan.). ‘Open-source’ robotic surgery platform going to top medical research labs [Online]. Available: http://news.ucsc.edu/2012/01/robotic-surgery.html

[37] H. King, B. Hannaford, K. Kwok, G. Yang, P. Griffiths, A. Okamura, I. Farkhatdinov, J. Ryu, G. Sankaranarayanan, V. Arikatla, K. Tadano, K. Kawashima, A. Peer, T. Schau, M. Buss, L. Miller, D. Glozman, J. Rosen, and T. Low, “Plugfest 2009: Global interoperability in telerobotics and telemedicine,” in *Proc. IEEE Int. Conf. Robot. Autom.*, 2010, pp. 1733–1738.

[38] C. Gunn, D. Stevenson, A. Krumm-Heller, S. Srivastava, P. Youngblood, L. Heinrichs, and P. Dev, “A remote interactive master class in surgery,” in *Proc. Australas. Comput.-Human Interaction Conf.*, 2004.

[39] M. Martini, R. Istepanian, M. Mazzotti, and N. Philip, “A Cross-layer approach for wireless medical video streaming in robotic teleultrasonography,” in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, pp. 3082–3085.

[40] G. Jansson, M. Bergamasco, and A. Frisoli, “A new option for the visually impaired to experience 3d art at museums: Manual exploration of virtual copies,” *Vis. Impairment Res.*, vol. 5, no. 1, pp. 1–12, 2003.

[41] A. Frisoli, G. Jansson, M. Bergamasco, and C. Loscos, “Evaluation of the pure-form haptic displays used for exploration of works of art at museums,” in *Proc. World haptics Conf.*, 2005, pp. 18–20.

[42] E. S. Axel and N. S. Levent, *Art Beyond Sight: A Resource Guide to Art, Creativity, and Visual Impairment*. New York, NY, USA: American Foundation for the Blind, 2003.

[43] S. J. Lederman and R. L. Klatzky, “Extracting object properties through haptic exploration,” *Acta Psychologica*, vol. 84, no. 1, pp. 29–40, 1993.

[44] O. Lahav and D. Mioduser, “Haptic-feedback support for cognitive mapping of unknown spaces by people who are blind,” *Int. J. Human-Comput. Stud.*, vol. 66, no. 1, pp. 23–35, 2008.

[45] W. Yu, R. Ramloll, and S. Brewster, “Haptic graphs for blind computer users,” in *Haptic Human-Computer Interaction*. New York, NY, USA: Springer, 2001, pp. 41–51.

[46] H. M. Kamel, P. Roth, and R. R. Sinha, “Graphics and users exploration via simple sonics (guess): Providing interrelational representation of objects in a non-visual environment,” in *Proc. 7th Int. Conf. Auditory Display*, 2001, pp. 261–266.

[47] W. S. Yeo, J. Berger, and Z. Lee, “Sonart: A framework for data sonification, visualization and networked multimedia applications,” in *Proc. Int. Comput. Music Conf.*, 2004, pp. 180–184.

[48] P. Parente and G. Bishop, “Bats: The blind audio tactile mapping system,” in *Proc. ACM Southeast Regional Conf.*, 2003, pp. 132–137.

[49] F. Candlin, “Dont touch! hands off! art, blindness and the conservation of expertise,” *Body Soc.*, vol. 10, no. 1, pp. 71–90, 2004.

[50] K. Hetherington, “The unsightly touching the parthenon frieze,” *Theory, Culture Soc.*, vol. 19, nos. 5/6, pp. 187–205, 2002.

[51] J. E. Colgate and G. G. Schenkel, “Passivity of a class of Sampled-data systems: Application to haptic interfaces,” *J. Robot. Syst.*, vol. 14, no. 1, pp. 37–47, 1997.

[52] J. J. Corso, J. Chhugani, and A. M. Okamura, “Interactive haptic rendering of deformable surfaces based on the medial axis transform,” in *Proc. Eurohaptics*, 2002, pp. 92–98.

[53] J. S. Mehling, J. E. Colgate, and M. A. Peshkin, "Increasing the impedance range of a haptic display by adding electrical damping," in *Proc. 1st Joint Eurohaptics Conf. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, 2005, pp. 257–262.

[54] C. H. Park, S. Remy, and A. M. Howard, "Visualize your robot with your eyes closed: A multi-modal interactive approach using environmental feedback," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2011, pp. 1368–1373.

[55] A. Hub, J. Diepstraten, and T. Ertl, "Design and development of an indoor navigation and object identification system for the blind," *ACM SIGACCESS Accessibility Comput.*, nos. 77/78, pp. 147–152, 2004.

[56] Y. Tian and S. Yuan, "Clothes matching for blind and color blind people," in *Computers Helping People with Special Needs*. New York, NY, USA: Springer, 2010, pp. 324–331.

[57] D. Shah, J. Ascenso, C. Brites, and F. Pereira, "Evaluating multi-view plus depth coding solutions for 3d video scenarios," in *Proc. 3DTV-Conf.: True Vis.-Capture, Transmiss. Display 3D Video*, 2012, pp. 1–4.

[58] D. De Silva, W. Fernando, H. Kodikaraarachchi, S. Worrall, and A. Kondoz, "A depth map Post-processing framework for 3d-tv systems based on compression artifact analysis," *IEEE J. Sel. Topics Signal Process.*, vol. PP, no. 99, p. 1, 2011.

[59] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. 6th Int. Conf. Comput. Vis.*, 1998, pp. 839–846.

[60] S. J. Lederman and R. L. Klatzky, "Haptic identification of common objects: Effects of constraining the manual exploration process," *Perception Psychophys.*, vol. 66, no. 4, pp. 618–628, 2004.

[61] A. M. Okamura, M. L. Turner, and M. R. Cutkosky, "Haptic exploration of objects with rolling and sliding," in *Proc. IEEE Int. Conf. Robot. Autom.*, 1997, vol. 3, pp. 2485–2490.

[62] K. J. Kuchenbecker, D. Ferguson, M. Kutzer, M. Moses, and A. M. Okamura, "The touch thimble: Providing fingertip contact feedback during point-force haptic interaction," in *Proc. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, 2008, pp. 239–246.

[63] E. P. Scilingo, M. Bianchi, G. Grioli, and A. Bicchi, "Rendering softness: Integration of kinesthetic and cutaneous information in a haptic device," *IEEE Trans. Haptics*, vol. 3, no. 2, pp. 109–118, Apr.–Jun. 2010.

[64] W. McMahan, J. M. Romano, A. M. Abdul Rahuman, and K. J. Kuchenbecker, "High frequency acceleration feedback significantly increases the realism of haptically rendered textured surfaces," in *Proc. Haptics Symp.*, 2010, pp. 141–148.

Chung Hyuk Park received the BS degree in electrical and computer engineering and the MS degree in electrical engineering and computer science from Seoul National University, Seoul, Korea. He received the PhD degree in electrical and computer engineering from the Georgia Institute of Technology with his thesis entitled "Robot-based Haptic Perception and Telepresence for the Visually Impaired." He is an assistant professor in electrical and computer engineering at the School of Engineering and Computing Sciences, New York Institute of Technology. His research interests include robotics and intelligent systems, assistive robotics, computer vision, haptic/multimodal systems, and human-robot interaction. Prior to joining NYIT, he was a postdoctoral research fellow at the Georgia Institute of Technology, where he was a member of the Human Automation Systems Laboratory (HumAnS Lab). From the outcome of his research and STEM outreach programs, more than 60 students with visual impairments have learned robot programming using multi-modal interfaces, which was recognized by the World Institute of Disability. He was also with LG Electronics Research Center for four years where he received the "LG Electronics Best R&D TDR Award." He is a member of the IEEE.

Eun-Seok Ryu received the BS, MS, and the PhD degree in computer science from Korea University in 1999, 2001, and 2008, respectively. He is an assistant professor in the Department of Computer Engineering, Gachon University, Seongnam, Korea. Prior to joining the University in 2015, he was a principal engineer at Samsung Electronics, Suwon, Korea, where he led a multimedia team. He was a staff engineer at InterDigital Labs, San Diego, CA, from January 2011 to February 2014, where he researched and contributed to next generation video coding standards such as HEVC and SHVC. From September 2008 to December 2010, he was a postdoctoral research fellow at GCATT in the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA. In 2008, he was a research professor in the Research Institute for Information and Communication Technology at Korea University, Seoul, Korea. His research interests are in the area of multimedia communications that includes video source coding and wireless mobile systems. He is a member of the IEEE.

Ayanna M. Howard is a professor at the Georgia Institute of Technology. Her area of research is centered around the concept of humanized intelligence, the process of embedding human cognitive capability into the control path of autonomous systems. This work, which addresses issues of autonomous control as well as aspects of interaction with humans and the surrounding environment, has resulted in more than 100 peer-reviewed publications in a number of projects from scientific rover navigation in glacier environments to assistive robots for the home. To date, her unique accomplishments have been documented in more than 12 featured articles including being named as one of the world's top young innovators of 2003 by the prestigious *MIT Technology Review* journal and in *TIME* Magazine's "Rise of the Machines" article in 2004. She received the IEEE Early Career Award in Robotics and Automation in 2005 and is a senior member of the IEEE.