Real-time Video Streaming with Multi-Camera for a Telepresence Wheelchair

Van Kha Ly Ha[1], Tuan Nghia Nguyen[2], Hung T. Nguyen[3]

[1, 2, 3] Centre for Health Technologies, Faculty of Engineering and Information Technology,
University of Technology Sydney, New South Wales, Australia.
E-mail: [1]VanKha.L.Ha@student.uts.edu.au, [2]Tuannghia.nguyen@uts.edu.au, [3]Hung.Nguyen@uts.edu.au

Abstract— **This paper presents a new approach for telepresence wheelchairs equipped with multiple cameras. The aim of this system is to provide effective assistance for the elderly and people with disabilities. The work explores the integration of the Internet of Things, such as multimedia, wireless Internet communication, and automation control techniques into a powered wheelchair system. In particular, multiple videos are streamed in real-time from an array of cameras mounted on the wheelchair, allowing wide visualization surrounding the wheelchair. By using video communication and interaction, remote users can assist to navigate a wheelchair via the Internet through wireless connections in a distant location. The experimental results show that video streaming can achieve high-quality video with the streaming rate up to 30 frames per second (fps) in real-time. The average round-trip time is under 27 milliseconds (ms). The results confirmed the effectiveness of the proposed system for tele-monitoring and remote control to achieve safer navigation tasks for wheelchair users.**

Keywords-multi-camera; Internet of Things; telepresence; wheelchair; assistive technology; healthcare.

I. INTRODUCTION

In recent years, there has been an increasing population of people with disabilities and the elderly. It is challenging for these people moving around in daily life at home or in a hospital. Smart wheelchairs which are integrated with new technologies, such as so-called telepresence wheelchairs, are in high demand. A telepresence wheelchair is a system that allows a person to control the wheelchair without actually being there. The telepresence wheelchair has attracted increasing attention from academia, industry, society, and so on, but technical challenges remain to improve the telepresence wheelchairs [1].

Recently, extensive research in the field of telepresence has developed telepresence robots for many purposes. The efficacy of these systems has been demonstrated in meeting and tele-education robots [2, 3], and references therein. Meanwhile, there has been limited work for healthcare, such as for remote visiting, home care of the elderly [4]. Most of these papers mentioned to the robotic hardware designs and applications. In contrast to all the approaches as mentioned earlier, the research in [5] presented a brief method with open protocol framework for telepresence robots. However, no such investigation has been applied for a telepresence wheelchair to assist wheelchair users for healthcare purposes.

It is worth noting that the vast majority of existing telepresence robots offer one camera for video communication. Therefore, a single view is available only at any one time. Currently, spherical cameras have been introduced for a 360-degree view. These have been previously explored in academic research projects for telepresence with a full view [1]. It has been proven that the solution for telepresence systems with a spherical camera is feasible. However, there are a few limitations. These existing cameras are high cost and typically designed for capturing images and videos without audio. Moreover, these devices increase the power consumption of the system and they do not support video or voice over Internet protocol for a long distance communication. In order to address these drawbacks, an ideal alternative solution would be to use multi-camera for a telepresence wheelchair to have a wider view around the wheelchair at the same time. This work allows a remote assistant user to predict the real-world target and process the action in real-time for safety purposes.

Motivated by the above observations, our approach involves various key design objectives, namely real-time multi-video communication and remote interaction ability for a wheelchair from a long distance. It is considerably challenging due to the mobility and the remote controllability requirements of the wheelchair. The investigation of a telepresence wheelchair with multiple cameras is explored for the first time to create solutions for improving health care.

The main objective of this paper is to propose a novel telepresence wheelchair with multiple cameras. More specifically, we have developed and implemented appropriate techniques for a telepresence wheelchair system based on a communication framework. Notably, we introduce the multi-video streaming of real world surrounding a wheelchair to remote location in real-time. Moreover, we highlight an interesting remote control methodology which allows a user to be able to control the wheelchair from a long distance. Then, the video streaming performance, the round-trip time, and the remote control tasks are also assessed to evaluate the feasibility and efficiency of the designed system.

The rest of this paper is structured as follows. In Section II, we introduce the designed system and the proposed methodology in detail. Section III presents experiments and results of video streaming and remote controlling. Finally, conclusions and future works are presented in Section IV.

II. DESIGNED SYSTEM AND PROPOSED APPROACH

A. Design of the Telepresence Wheelchair

Our goal is to develop a telepresence wheelchair which is capable of providing video conferences and interaction from an unknown starting location. The telepresence wheelchair platform that we have developed is shown in Fig. 1. This designed system is expected to be useful in remote monitoring and collaborative control to provide assistance and enhance safety mobility for people with disabilities.

For the purpose of the ability to use voice, video, and remote interaction as a function of telepresence, the system consists of many components. Our powered wheelchair is equipped with a Mac mini computer, multimedia components, control interfaces (joystick and a multifunction data acquisition National Instruments device NI USB-6008), and dynamic control systems (DX power module, DX master remote) [6, 7]. These devices are connected in order to process, control and transmit video around the wheelchair while it is moving to a remote location over the wireless network in real-time.

Moreover, to have a wider view of an unknown environment surrounding the wheelchair from a remote site, along with a number of accessories for wheelchair platform, a 7-port USB extension hub is used to extend connections for the vision system. Four cameras are also mounted in four directions at an angle of 90 degrees in front of the wheelchair to capture images surrounding the wheelchair, as can be observed from Fig. 1. There is one touchscreen display which is used for two-way video communication.

In order to handle the movement of the wheelchair, the National Instrument device (NI) USB-6008 with a USB interface is connected to the minicomputer at the back of the wheelchair. The NI USB-6008 has 8 analog inputs and 2 analog outputs which allow the system to transform command tasks to control the wheelchair.
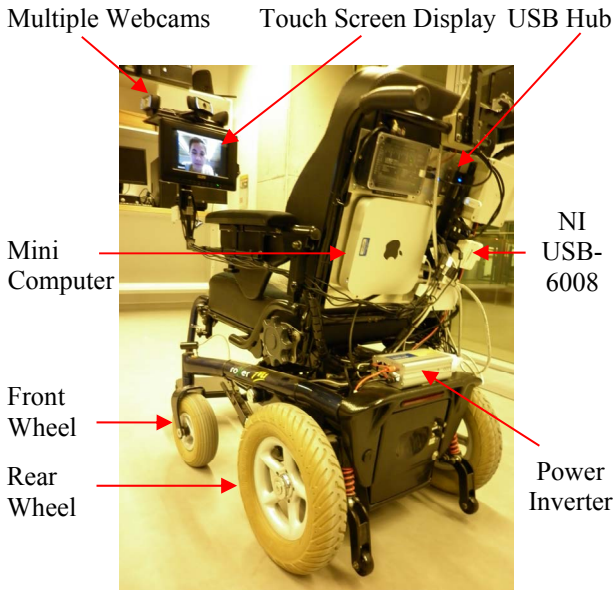

Figure 1.   The telepresence wheelchair system description.

B. Multi-Video Acquisition and Processing

We have developed an application module for multi-video acquisition, and processing plugged in Skype framework, a well-known application [8]. This module will access, perform video processing, and display multimedia streams from ubiquitous cameras. Multi-cam is a free and open source, and can be developed from Microsoft DirectShow [9]. This technique is based on the Component Object Model (COM) technology. COM objects in DirectShow are called filters. A typical DirectShow application comprises three kinds of filters: source, transform, and render. They are connected to process a media stream from source to render by a filter graph manager. The filters need to be synchronized in order to keep media samples in synchronization with media streaming process.

To be more specific, source filters support multimedia stream acquisition from digital cameras. Once the media stream has been captured, DirectShow filters can be used to transform it. Transform filters have been written to resize video images. Media streams can also be multiplexed together, taking two or more streams and making them one synchronized video. Transform filters act on the acquired multimedia stream and alter it in some way. After the transform filters have been implemented, the final one task is to render the media stream to the display, speakers, or a device [10].

In this work, the process of multi-cam streaming is described in Fig. 2. The four cameras namely cam#1, 2, 3, 4 are connected to a computer. Live video frames are sent from these cameras to the central processing unit (CPU) of the computer to process. Cameras are actual hardware devices that communicate with Skype through CPU. In the Windows operating system, a filter can read and write requests to the Skype. In fact, it multiplexes the video source from physical cameras. It passes video data from one or more physical cameras to Skype by appropriately transforming the data. Skype will stream incoming frames from a live video source with a reference frame.
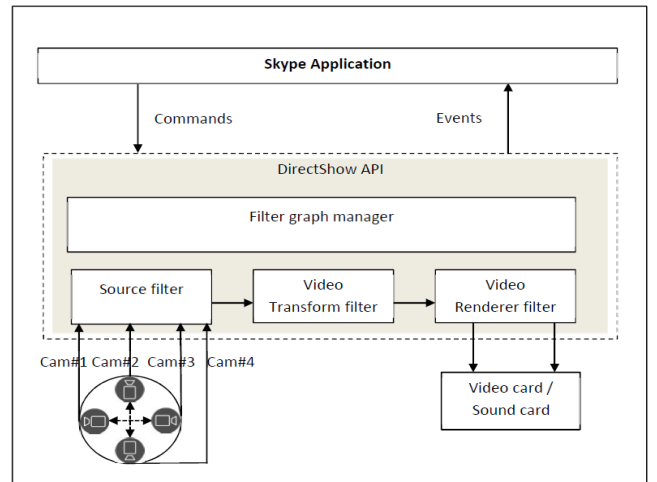

Figure 2.   The diagram for a stream of multi-video data from multi-cam.

C. Telepresence Wheelchair Control

By observing multiple views around the wheelchair from the multi-video data, a remote user can send an appropriate control command to the telepresence wheelchair. To be able to control the telepresence wheelchair moving in different directions, the telepresence wheelchair system in Fig. 1 is designed to have two front freewheels and two rear driving wheels. The two rear wheels are connected and controlled by direct current motors. A hardware connection diagram of the wheelchair control components is shown in Fig. 3. It consists of an add-on DX power component and a virtual joystick module. Instead of using the joystick, the computer generates appropriate control signals following the calculations of assistive navigation software which is integrated into Skype as a third-party application programming interface (API).

To control the movements of the telepresence wheelchair, five states, namely go forward, turn left, turn right, go backward and stop, have been designed. When the remote user sends commands, the API of the wheelchair will receive them and then transforms them into appropriate moving states. These digital signals from the computer will be converted into the analog forms and transferred to the virtual joystick module via the analog output pins of the NI USB – 6008 to control the wheelchair. In practice, as the digital values are obtained, they are converted into the appropriate voltage values and then fed into the virtual joystick module. Define that u_1 and u_2 are the input voltages, v and ω stand for the output velocity and the angular output velocity of the wheelchair, respectively. The range of the speed v of the vehicle is 0 to 1m/s, then when $v = 0$, the signal value for input voltage u_1 sent to the AO_0 of the module is expected to be 2.5V and if $v =$1m/s, u_1 should be 3.77V. The relationship between u_1 and v can be determined by

$$u_1 = 2.5 + 1.27v \text{ (V)} \quad (1)$$

Similarly, the signal u_2 is fed into the pin AO_1 of the module can be calculated based on the value of the rotational velocity or turning speed ω in radian per second by

$$u_2 = 2.5 + 1.3\frac{2\omega}{\pi} \text{ (V)} \quad (2)$$

According to (1), the wheelchair velocity v can be expressed as

$$v = \frac{1}{1.27}(u_1 - 2.5) \quad (m/s) \quad (3)$$

It follows from (2) that the turning speed ω is given by

$$\omega = \frac{\pi}{2.6}(u_2 - 2.5) \quad (rad/s) \quad (4)$$

In our proposed approach, the speed signal is approximately 2.5V±1.27V. The 2.5V value is defined as the stop signal. The forward signal with the speed of 1m/s is 3.77V, and the backward signal is 1.24V. The rotational signal range is about 2.5V±1.38V. The right turn signal is greater than 2.5V and the values less than 2.5V is represented as the left turn state of the wheelchair.
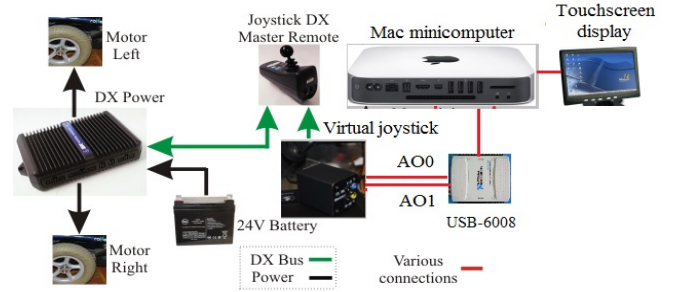

Figure 3. The connection of the wheelchair control components.

D. System Evaluation

From the technology perspective, real-time communication is the technical challenge for such a telepresence system. The data information must be transferred wirelessly in real-time to guarantee the system executes commands in real-time to avoid risk and make it safer with desired actions. In order to evaluate the performance of a telepresence wheelchair system, the data transmission between the remote user and the wheelchair must be considered in the wireless network. To ensure the system performs well and suit for real-time communication, the system demands some standard quality of video streaming and response time. The network processing delay would affect to the interactivity of the data communication. There are many existing methods and techniques approached to evaluate a real-time system in the telecommunication and computer network society. The criteria to assess the system performance in existing reports shown that frame rate of video streaming and round-trip time are the key parameters taken into account to evaluate the real-time system [11, 12].

Frame rate or frame frequency refers to the number of individual frames that an imaging device displays in each second, as measured by frame per second (FPS). The maximum frame rate allowable for each of the cameras depends on the resolution of the images and the bandwidth (BW) of the network. The frame rate can be roughly approximated using the following general formula from (5) and (6) (assuming all cameras are at the same resolution). An estimate for the number of packets per frame (PPS) can be determined according to the following:

$$PPS = \frac{Image\ size \times Bytes\ per\ Pixel}{Bytes\ per\ Packet}. \quad (5)$$

$$FPS = \frac{BandWidth}{\frac{(Pixel\ per\ Frame \times Bytes\ Per\ Frame)}{Number\ of\ Camera}}. \quad (6)$$

Round-trip time (RTT) is defined as the time required for a packet to transmit from the source to the destination and back [11]. The round-trip time calculation in a wireless network was presented in [13]. The RTT calculation is as follow.

$$RTT_e = (1-\alpha) \times RTT_e + \alpha \times RTT_m \quad (7)$$

RTT_e: Estimated Round-trip time.
RTT_m: Round-trip time measurement for a sample segment.
α: Constant weighting factor and recommend value is $\alpha = 1/8$.
By substituting into (7), we obtain (8).

$$RTT_e = 0.875 \times RTT_e + 0.125 \times RTT_m. \quad (8)$$

III. EXPERIMENTS AND RESULTS

A. Multi-Video Streaming for a Telepresence Wheelchair

Different from the existing telepresence robots designed with one camera based on Skype framework [14], [15]. Here four cameras and the telepresence technology were applied to the wheelchair for safety health care purpose while navigating the wheelchair. We developed a third-party application to stream multi-video in order to achieve a wider visualization at the remote site. To that end, four Logitech C930 webcams were connected to a computer via a USB extension hub.

The experiments were carried out at various locations where a remote user was far away from a wheelchair. We carried out experiments in five scenarios. Fig. 4 illustrates the experiment positions. The detail information of the scenario is listed in Table I. A remote user was in room A and the wheelchair was in different rooms which were labelled room A, B, C, D, and the balcony E, respectively. We conducted with ten trials for each scenario to stream multi-video with a resolution of 640x480 pixels each camera from a wheelchair to a remote client using the Internet through wireless connections based on Skype framework. Before doing the tests, the wheelchair battery was fully charged, and the wheelchair speed was set to 1m/s to ensure that the wireless connection could be maintained. Surprisingly, by obtaining the multi-view video, at the remote site a remote user can hear and observe different angles of view around the wheelchair and feels as he or she is present at the wheelchair as illustrated in Fig. 7.

To assess the performance of video streaming, a set of data was measured then collected and analysed to evaluate the quality of video streaming rate. In addition, an analysis of network data was recorded in detail, including the system information and the round-trip time. In this work, RTT is a measure of the consecutive transceiver data packets between the remote user and the wheelchair.

TABLE I. THE TEST POSITIONS OF THE EXPERIMENTS

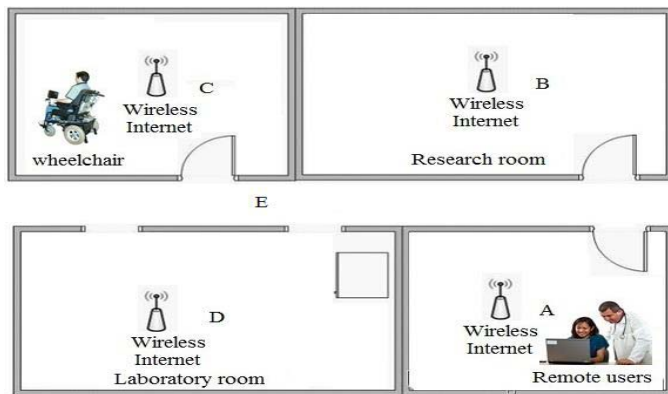| Experiments	Remote Position	Wheelchair Position	Distance (m)
Scenario 1	A	A	10.00±5.00
Scenario 2	A	B	30.00±5.00
Scenario 3	A	C	60.00±5.00
Scenario 4	A	D	40.00±5.00
Scenario 5	A	E	20.00±1.00

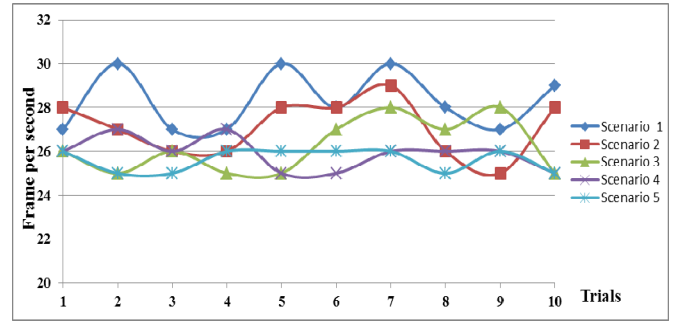Figure 4. System experiments at the University of Technology Sydney.

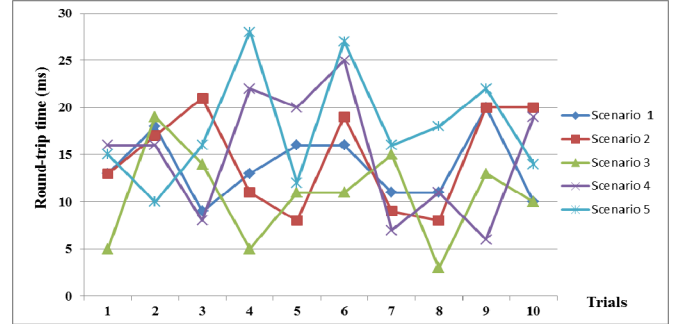Figure 5. The average streaming rate of ten trials in the experiments.

Figure 6. The average round-trip time of ten trials in the experiments.

The results of the experiments are described in Figs. 5 and 6. It can be observed from Fig. 5 that the streaming rate in all scenarios fluctuates considerably between 25 and 30 fps. The scenario 1 peaks the highest average streaming rate up to 30 fps. On the other hand, the scenario 5 has the lowest average streaming rate of 25 fps. It is not surprising that the results of the scenario 5 are lower than the others due to the weak signal at the balcony areas where the wireless coverage is limited.

However, in comparison with [11], it is worth noting that the standard video frame rate is 25 fps (Europe) or 30 fps (North America). Thus, the overall streaming rate performance of the system has reached the standard of real-time video communication. Furthermore, the system has higher performance than the previous similar work [16] which presented an approach of the multi-view video stream with the frame rate was 15 fps in wireless streaming.

Regarding the round-trip time, according to Telecommunication standardization sector of the International Telecommunication Union (ITU-T) recommendations for the quality of services and transmission performance [12], it is worth noting that the RTT of less than 400 milliseconds (ms) is considered as real-time multimedia communications. The results in Fig. 6 illustrate the average round-trip time of ten trials in of each scenario in the experiments. It can be observed from Fig. 6 that the RTT fluctuates from 3 to 27 ms. Obviously, it is an excellent streaming quality under the RTT threshold value [12].

In further comparison with findings in the recent previous telepresence robot study [15], the authors presented that the telepresence robot with less than 40 milliseconds delay was considered good performances. The results have confirmed the efficacy of our approach for a real-time telepresence wheelchair system.
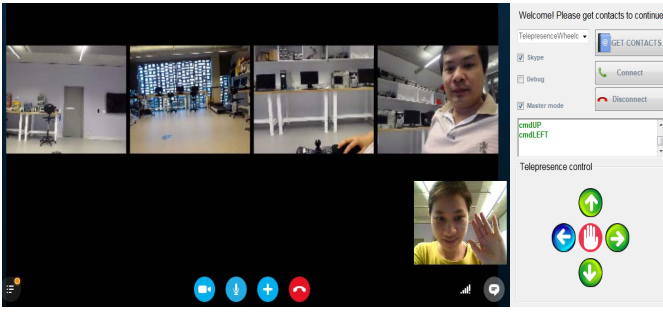
Figure 7.   Multi-view video obtained from a telepresence wheelchair and the control interface at the remote site.

B.   Telepresence Wheelchair with Remote Assistance

In order to fulfill the requirements of the remote assistance, we have developed an application with a very user-friendly interface. This application is able to connect with the NI USB 6008 for control processing. From the remote location, the user can establish a connection to the wheelchair by clicking a connect button and starting a muti-video conference. Then, the user can control the wheelchair with the function of telepresence from five buttons (go forward, turn left, turn right, go backward and stop) as shown in Fig 7. The explanation of the five buttons is listed in Table II. The experiments were carried out at five scenarios with ten trials for each position.

TABLE II.   COMMAND BUTTONS AVAILABLE DURING EXPERIMENTS.

Button	Command	Explanation
	Stop	Wheelchair halts
	Go Forward	Wheelchair moves directly forward
	Go Backward	Wheelchair moves directly backward
	Turn Left	Wheelchair keeps turning left until next command task
	Turn Right	Wheelchair keeps turning right until next command task

TABLE III.   THE AVERAGE ACCURACY OF COMMAND TASKS.

Experiments	Upload BW [kBps]	Download BW [kBps]	Number of command task	Accuracy (%)
Scenario 1	132	70	50	100
Scenario 2	156	80	50	100
Scenario 3	148	9	50	100
Scenario 4	135	40	50	100
Scenario 5	78	3	50	97
Average				99.4

The results from Table III show that the wheelchair can be controlled remotely in different types of bandwidth connections. The wheelchair can successfully navigate at various distances with the average accuracy of 99.4 %. During the experiments, it can be observed that there is a correlation between the fluctuations of wireless signal strength, the streaming rate, round-trip time and accuracy of command tasks due to the quality of the wireless signal. It may be developed by using more wireless access points and wireless repeaters to guarantee the signal strength. Nevertheless, these overall results indicate that the proposed method is promising for full view display approach and would be more effective for telepresence wheelchair systems in real-time.

IV.   CONCLUSION

In this paper, we have presented a novel telepresence wheelchair system using multi-camera which is able to improve the quality of life for the elderly and people with disabilities with the Internet of Things. The experimental results obtained from the Internet-enabled wheelchair in real-world environments demonstrated the effectiveness of the proposed techniques and the feasibility the designed system. The results indicate that the video is streamed smoothly, and the wheelchair is efficiently controlled over the Internet using wireless networks in real-time. Future work will consider transferring the wheelchair user's heart rate, blood pressure, and body temperature for advanced monitoring in assistive healthcare applications.

REFERENCES

[1]   V. K. L. Ha, T. N. Nguyen, and H. T. Nguyen, "Real-time transmission of panoramic images for a telepresence wheelchair," 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015, pp. 3565-3568.

[2]   K. Oh-Hun, K. Seong-Yong, K. Young-Geun, and K. Dong-Soo, "Telepresence robot system for English tutoring," in IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO), 2010, pp. 152-155.

[3]   S. C. Herring, "Telepresence Robot for Academics," presented at the ASIST, Montreal, Quebec, Canada, 2013.

[4]   J. Gonzalez-Jimenez, C. Galindo, and J. R. Ruiz-Sarmiento, "Technical improvements of the Giraff telepresence robot based on users' evaluation," in IEEE Conference on RO-MAN, 2012, pp. 827-832.

[5]   P. S. a. H. Khoswanto, "Open protocol framework for telepresence robot " ARPN Journal of Engineering and Applied Sciences vol. 9, p. 4, 12 Dec 2014.

[6]   N. Tuan Nghia, S. Su, and H. T. Nguyen, "Neural Network Based Diagonal Decoupling Control of Powered Wheelchair Systems," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 22, pp. 371-378, 2014.

[7]   J. S. Nguyen, S. W. Su, and H. T. Nguyen, "Experimental study on a smart wheelchair system using a combination of stereoscopic and spherical vision," in 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013, pp. 4597-4600.

[8]   M. Skype (Accessed: 2015, 10 Oct). Available: http://www.skype.com

[9]   J. MacCormick, "Video chat with multiple cameras," presented at the Proceedings of the 2013 conference on Computer supported cooperative work companion, San Antonio, Texas, USA, 2013.

[10]   M. Pesce, "Programming Microsoft DirectShow for digital video and television", Microsoft Press, 2003.

[11]   ETSITR102479, "Review of available material on QoS requirements of Multimedia Services," Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN), 2006.

[12]   I.-T. R. G.1010, "Transmission systems and media, digital systems and networks," in Quality of service and performance, ed. Geneva, Switzerland, 2001.

[13]   P. D. R. S. Sergio Esteban Gabal´an Villar, Jos´e Luis Sandoval Alaguna, "RTT calculation in wireless home network," presented at the Introduction To Computer Networking, 2013.

[14]   P. Corke, K. Findlater, and E. Murphy, "Skype : a communications framework for robotics" in Proceedings of the 2012 Australasian Conference on Robotics and Automation, D. Carnegie and W. Browne, Eds., ed. Wellington, New Zealand: Australian Robotics & Automation Association, 2012.

[15]   M. Denojean-Mairet, Q. Tan, F. Pivot, and M. Ally, "A Ubiquitous Computing Platform - Affordable Telepresence Robot Design and Applications," in 2014 IEEE 17th International Conference on Computational Science and Engineering (CSE), 2014, pp. 793-798.

[16]   S. Kodera, T. Fujihashi, S. Saruwatari, and T. Watanabe, "Multi-view video streaming with mobile cameras," in 2014 IEEE Global Communications Conference, 2014, pp. 1412-1417.