Convergence Technologies by a Long-term Case Study on Telepresence Robot-assisted Learning

Mi-Suk Lim¹, Jeong-Hye Han²*
¹Adjunct Professor, Division of Convergence Computer & Media, Mokwon University
²Professor, Dept. of Computer Education, Cheongju National University of Education

Abstract The purpose of this paper is aimed to derive suggestions for convergence technology for effective management of distance education by analyzing a long-term case. The experiment was designed with notebook, smartphone or tablet based robot controlled by a remote instructor and a learner, who have experience of distance learning including robot assisted learning. The tablet based robot has the display system of feedback to speakers. During five months, three types of experiments were conducted randomly and a participant was interviewed thoroughly. The result, like the previous research, demonstrates that the task performance of the learner in telepresence robot-assisted learning was better than that in the notebook, and smartphone based. However, it is believed to be necessary to adjust the system for eye-contact and voice transmission for the remote instructor. The instructor required an additional sight by supplementing an extra camera and automatic direction control to source of sound.

Key Words : Video conferencing, Distance education, Telepresence, Robot assisted-learning, Convergence technology

요약 이 논문은 개인 대상으로 장기간의 원격영상 교육을 실시한 사례를 분석함으로써, 효과적인 운영을 위한 융합기술에 대한 제언을 도출하는 것을 목적으로 한다. 로봇보조학습을 포함한 원격영상교육 경험이 있는 교수학습자를 대상으로 텔레프리젠스를 노트북, 스마트폰 로봇, 그리고 음성인식 페드백이 제공되도록 제작한 페드형 로봇으로 실험 설계하여 5개 월간 원격수업 방식에 임의할당을 통해 진행하고, 과제 수행도를 측정함과 동시에 심층 인터뷰를 실시하였다. 학습자는 기존 연구와 마찬가지로 텔레프리젠스 로봇보조학습의 과제수행도가 가장 높은 것으로 나타났지만, 학습의 몰입도와 집중을 위하여 음성송출 페드백 외에 원격교수자와의 시선 맞춤이 필요하다고 생각했다. 원격 교수자는 원격교육의 영상입력 추가와 음원추적 자동 제어 기능을 요구했다.

주제어 : 원격영상교육, 원격교육, 텔레프리젠스, 로봇보조학습, 융합기술

*This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF=2015R1D1 A1A09060450).
*Corresponding Author : Jeong-Hye Han(hanjh@cje.ac.kr)

Received May 21, 2019 Revised June 28, 2019
Accepted July 20, 2019 Published July 28, 2019
1. Introduction

Distance learning based on temporal and spatial remoteness between instructor and learner utilizing electronic devices such as computers, pads, smartphones, tablets has deeply made an impact on our lives by taking an advantage of broadcasting applications such as Skype, Hangout, YouTube, Facebook live broadcasting, Face talk by Kakao, etc. In addition, the pursuit of cultivating 21st century competency as part of the OECD’s CERI in each country, it has been extended to education system with an access to the internet and advanced technological learning environments where global citizenship and virtual school are being promoted [1,2].

The practical implications for process innovation with intelligent robots and the effects of customer’s innovation for service robots by potential customers. are getting important[3,4]. The high–technological learning environment is beginning to drive the market for video conferencing and tele-operated (telepresence) robots. With the spread of smart phones and 5G network, video conference through smartphone is evolving into form of robots that are called telepresence robots. The prices of telepresence robot are getting cheaper and the quality of video is getting better. In fact, five major markets of telepresence robots: individual connections, teleconferences (enterprises), distance education, telemedicine, and remote monitoring. The market size is estimated to be upsurge to $ 136.9 million with 31,600–100,000 units in 2020 starting from 2016. Especially among the five markets, telepresence robots administers global training programs at low cost making it the second most key application. Main market opportunities are focused on Asia Pacific (China, Japan, Korea, India) and Europe (France, England, Germany), and it is anticipated that Asia–Pacific will be the largest market due to Chinese influence by 2022[5].

Although the effectiveness of the telepresence robot-assisted learning has been well established, most of the studies were researched in the classroom during a period of two months[6]. However, there is no case study until now on the finding of convergence technology for telepresence robot-assisted learning in long-term. Therefore, in this paper, we propose some convergence technologies to manage telepresence robot-assisted learning for individual using three electronic media groups which are (laptop, smartphone-based and pad-based robots) by a case study with in-depth interview conducted through long-term experiments.

2. Related Works

2.1 Video Conferencing

Since face-to-face learning, distance learning and hybrid learning did not portray a significant difference in learning motivation and attitude improvement of learners[7], distance learning is proposed for rural areas. In addition, collaborative learning with distant students through video conferencing is very effective in promoting learner awareness and motivation[1]. It has also been found that the creativity of students in elementary school was improved through collaborative video conferencing between South Korea and Australia[2]. The necessity of the collaborative video conferencing class for the creativity education and internationalization in the school is also suggested. However, learner’s attitude was important as non-verbal communication barriers between learner and the native speaker as an English instructor in class existed[8].

However, although the video conferencing classes by native speakers as English instructors in elementary schools are effective with enhanced interest and confidence in speaking and listening compared to face-to-face learning, it has been suggested that video English lessons in small group
activities below 15 pupils in one camera would be more effective[9]. In addition, elementary school students participating in one-to-one rather than in one-to-whole video conferencing class are expected to participate actively in learning since the individual learners are more likely to clearly recognize his/her activities and target[10].

In order to operate such a real-time video conferencing classroom efficiently, it is suggested that the stability of the system (solution of voice delay, ensuring of good sound quality, clear and bright screen, etc.) is a prerequisite and should be avoided during interaction[11]. Moreover, technological problems in communication such as low monitor resolution, and lack of eye contact with instructors during video conferencing classes results in lowered learning motivation amongst learners[12]. Since, in video conferencing the camera is located at the upper center or on sides of the display rather than the center of the display, instructors and learners can not make eye contacts. Therefore, their learning and teaching interaction can be reduced without eye-contact through the camera and computer display. Thus, some studies focused on development of gaze-fitting technology by the method of depth map generation and eye alignment using mid-view image synthesis[12]. Currently, platforms that provide individual video conferencing classes have guaranteed the stability of the network, but are yet to provide gaze-fitting technology. In this paper we would like to demonstrate the importance of eye contact between instructors and learners through long-term case study.

2.2 Robot-assisted Learning

Telepresence robot-assisted learning is more effective in contrast to computer-aided learning as the remote instructor interacts with robots to remotely navigate and actively interact with the learner in teaching and learning environment[6,13]. In addition, if the size of video face in the telepresence robot is larger, the authority of the instructor and attention of the learners increase[14,15].

Currently, although most telepresence robots use a tablet PC or a smartphone as their faces, they have not been able to solve the discrepancy of eye contacts. However, the robot using the robotic neck operated by instructor can look at the face of the learner. There is no case study as to whether or not such robotic neck looking at the learner’s face has the same expectation as the eye contact with the learner. Therefore, in this paper, it was conducted to see if a robot face alignment is meaningful to learners through deep interview of learners in telepresence robot-assisted learning.

3. Methodology

3.1 Experimental Design

The purpose of this paper is to find what technology is needed from an in-depth interview of instructors and learners regarding the task performance and interaction in distance education for over 5 months. It is necessary to quantitatively analyze whether there is a difference in long-term task performance among media types and what technology should be added to them for better interactions.

Fig. 1. The experimental scenes
Three media types (Laptop face talk, smartphone-based and pad-based robot) in the experimental design are shown in Fig. 1. After a month of pilot testing, a total of 37 sessions which incorporates 13 by the laptop, 11 by smartphone-based, and 13 by pad-based learning were randomly conducted during five months.

3.2 Methodology

The instructor and learner who participated in this experiment were selected after obtaining adequate experience (at least 30 hours of distance learning and robot-assisted learning) in order to minimize the novelty effect. Likewise, to eliminate the alternative communication between the instructor and learner through other channels, the instructor were selected from South Korea and the learner from USA, a 7th grade middle school student.

The subject of distance education was mathematics. As the learner was already familiar with the basic concepts, 5 minutes of the conceptual review was followed by exercises through combined problem solving. The learner in distance education has conceptual review and problem solving for about 10 minutes per day. And then a homework was given to him (about 10 pages). The instructor and the learner used the same textbooks each so that the distance learning can be done without sharing the contents of the textbook through additional video channel. When more information was required for solving the problem, they used the camera to show the textbook instead of their faces without using any other teaching and learning applications.

As soon as the learner completed the homework assignment, the remote instructor was called at the appointed time to proceed with the next class. Before the class, he was asked to report the time required to complete the assignment and the percentage of correct answers to the instructor the time required. If the learner failed to complete the assignment, the next class was not conducted for a maximum period of 5 days (120 hours). And then the instructor asked to him the completion rate (total number of solving problems/total number of problems) and proceeded the next class.

We utilized a telepresence robot KUBI made by Revolve Robotics Inc. with LED ears and t-shirts that respond to the learner’s sound that was made of the Arduino board (see Fig. 2). The LED control app based on Android 6.0 was developed by Java using Android Studio. Through this device, the learner could know whether his voice was transmitted through the robot’s microphone to the instructor.

![Fig. 2. The experimental robot made of KUBI and app](image_url)

This paper uses case study methodology to explore long-term interactions using three kinds media [16]. It is an empirical approach to deeply investigate and interview people who are directly involved why it is good or how it is uncomfortable in real-life phenomena. We interviewed the instructor and learner thoroughly and observed them during 5 months to provide some convergence technologies for effective management and utilization of telepresence robot.
4. Results

4.1 Performance analysis

The learning performance of three kinds of devices have stable and uniform distributions with different means so that we did not provide the time series. Table 1 shows the average completion rate and the correct answer percentage and time of assignments for each type of media for a total of 37 sessions. It shows that the learner by pad-type robot instructor has the highest average rate of completion 97.7, the highest percentage of correction 85 and the average 51 hours to complete the homework. The completion rate of smartphone based robot dropped with the lowest percentage, 84.4%, but the percentage of correction for laptop based learning has the lowest average 79.9%. These two are similar to each other in this case study. The smartphone based robot has remarkably lower scores of time, 32 hours, than the pad based, 51 hours. The instructor asked him about the reason. He said that in the case of smartphone, he started the homework immediately after the video conferencing. However in case of the robot, he did not do his homework right away since he had to disassemble the robot and put it away. Similarly he often surfed the Web or play games after video conferencing. It was obtained the standard deviation of the time for pad-based is very large (30.23), reflecting the additional time required for setting the robot. The standard deviation of the completion rate and the correct answer rate was 0.06~0.07, which is more stable than other smartphone-based and laptop.

Table 1. Averages & STDEV of the learner's performances

<table>
<thead>
<tr>
<th>Video Conference</th>
<th>completion rate</th>
<th>Correction rate</th>
<th>Time(h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laptop</td>
<td>89.7 (0.107)</td>
<td>79.9 (0.145)</td>
<td>52 (24.11)</td>
</tr>
<tr>
<td>Smartphone based robot</td>
<td>84.4 (0.11)</td>
<td>81.5 (0.092)</td>
<td>32 (14.21)</td>
</tr>
<tr>
<td>Pad-based robot</td>
<td>97.7 (0.06)</td>
<td>85 (0.077)</td>
<td>51 (30.23)</td>
</tr>
</tbody>
</table>

This long-term result is same as other studies on the authority of the robot's face size[14] and of robot's height[15]. Thus, the learner who has no novelty effect of robot-assisted learning in the long-term case had good completion rate and the average time through video conferencing robot with the bigger face.

4.2 Interview Analysis

Since, both the instructor and learner had sufficient previous experience of distance learning and robot-assisted learning, no unusual interaction was observed. After the classes of five months, intensive interviews were done regarding the inconveniences of telepresence robot-assisted learning. The followings are from the interviews.

**Instructor:** “When I was teaching through video, I think that I am able to grasp learner's reaction and understanding level. But there were many activities that need to be written in mathematics. So it was difficult to confirm the writing and the eye-contact on the screen. Since I were not sure of his question sometime I skipped or used the camera for the worksheet. In addition, I was difficult to read learner's facial expressions when the camera was focused on the worksheet if he understood the solution correctly. So I had to check with him several times to see if the solution was understandable. Whenever he prepared to learn, his parent helped him set the telepresence robot. I think robot-assisted learning seems more appropriate for the upper-grade learners. I had difficult time interacting by solving the problem when I controlled the angle of the robot neck to see what he was writing. However, I felt realistic when I controlled the telepresence robot. For example, when he stood up to get another reference book, I enjoyed his moving to the
other side in the room through the telepresence robot.”

**Learner:** "I felt that I was getting a great class with a robot instructor. But the most uncomfortable thing to me was that she did not look at my face or eyes. She was looking at her desk not me. My face on the screen is the same. I can not help eye-to-eye contact because the camera is at the center on the laptop and I should see the screen. I felt like she was not looking at me straightly but just book on my desk. So I did not look at her face on the display, and just listen to her looking at my workbook. And if she lean on the chair while she explained, her face looked smaller and far away. Then I felt like she did not concentrate to teach me well. She might think my face was small even if she looked at the laptop screen or robot. My mom asked her to sit closer to the camera to make her face look bigger. But if her face looked big sufficiently, she was too close to see the screen. I've done it myself, but when I put my face closer to the camera, the screen was getting too close, so I can not stare at the screen probably. I think that I feel like she is next to me when she gives the lesson she connected with the robot than that with the laptop. When I get up in class and go to the bookshelf, she can track me through the robot. So when I take classes with robots rather than with laptops, I get more nervous a little bit. But I think it’s faster to meet with her fact-to-face. Whenever setting the robot my mom bothered. There was sometime when a moving motor sounded a little noisy when I was solving problems silently. After the lesson, my mom said that robot could see our room like CCTV by hacking. Whenever my lesson ended, my mom always put the robot looking at the wall or removed its head immediately. Nevertheless, I like the robot assisted learning.

5. Conclusions

From the long-term case study, we were able to propose some technologies for efficient management of robot-assisted learning. Robot-assisted learning was still found to be most effective among the other type of learning in the long-term case study. Instructor felt the view too narrow through the robot camera and difficult to control it. She needed an autonomous robot that can track to the learner’s voice. And the distance between the camera and her face is close enough so that technologies can make her face look larger and eye-contact through the display screen should be needed. Those technologies were asked by the learner since his concentration was declined due to lack of proper eye contacts with the instructor and was further affected by the size of the face image on the screen. Now we propose convergence technologies to be installed for the effective operation of telepresence robot-assisted learning. First, for maintaining the aligned eye-to-eye contact between the learner and instructor, it is possible to consider providing dual lenses or a fisheye lens as a camera of the telepresence robot. Second, to align with the center of the display, an ultra-compact camera can be mounted above the screen with 3D printed frame allowing it to move freely over the center of the screen. This device has additional strength when the distance between the camera and the instructor is closer, the face size in the monitor can be larger like the ones in life-sized. This can be solved by attaching a telephoto magnifying lens on the camera and correcting by face tracking software. Third, the robot’s neck joint control can be provided with a fusion technology of learner’s face detection and voice tracking technique rather than manual. It can thereby reduce fatigue of the instructor caused by manual control. Furthermore, in order to prevent the
robot’s motor sound in learning, it can be considered convergence technique to synchronize the learners eye movement from the robot face to his textbook. Finally, it is necessary to provide an eye-patch for the robot camera to prevent hacking in privacy. In future research, four convergence technologies will be applied to find user’s satisfaction.

REFERENCES


임 미 숙(Mi-Suk Lim)  [정회원]

- 2019년 2월 : 목원대학교 IT공학과 (공학박사)
- 2019년 3월 ~ 현재 : 목원대학교 융합컴퓨터·미디어학부 겸임 교수
- 2017년 3월 ~ 현재 : 디지털컨텐츠 학회 이사

- 관심분야 : 정보보호, 안전, 로봇보조학습, SW교육
- E-Mail : misuk6389@mokwon.ac.kr

한 정 혜(Jeong-Hye Han)  [정회원]

- 1998년 2월 : 충북대학교 전자계산학 (이학박사)
- 2001년 3월 ~ 현재 : 청주교육대학교 컴퓨터교육과 교수
- 2011년 : 스탠포드 대학교 인간과학 기술융합연구소 방문학자

- 관심분야 : 로봇보조학습, 로봇융합교육, SW교육
- E-Mail : hanjh@cje.ac.kr