

Design and Implementation of Convergence Telepresence System Using Smart Device

Jong Ho Shin^{1*}, Ky Hyun Um¹, In Chul Kim², Chae Ho Cho² and Ill Chul Doo¹

¹*Dept. of LINC, Dongguk University, 30, Pildong-ro 1-gil, Jung-gu,
Seoul, 100-272, Korea*

²*M2Soft Co.,Ltd, Development Group, 7Fl Kolon Digital Tower II,
Seongdong-gu, Seoul, Korea*

{jhshin00, khum, mrdoo}@dongguk.edu, {mp3250, blackjoe}@m2soft.co.kr

Abstract

This paper presents a design and realization of 1080i Full-HD telepresence system by using commercialized mobile device and smart device. The existing telepresence system could establish Full-HD video for efficient transmission with exclusive H/W en-and de-coder and high-priced device. However, this technology realizes transmitter by using a camera of commonly-used smart phone for encoding Full-HD video, without expensive camera. Then, it realizes the video with decoder for Full-HD video by using low-price Android STB in order to output the video on large-sized screen. In addition, this technology puts various devices having totally different functions each other into one video call system by defining API (Application Program Interface) for video call, in a bid to secure efficient control structure from the combination between totally-different type of device. This paper describes general call processing system for video call, the way how to make it interwork through smart device and the way how to make it apply and realize on smart device.

Keywords: *Video Connectivity, Video Conferencing, High Definition Video, Telepresence*

1. Introduction

In this paper, a tele-presence system [1] has been established by further expanding the video communication service using with a set top box which utilizes mobile phones and proposes to show a method in realizing a system that is expanded to emphasize user experience by using the connectivity mechanism proposed in the previous paper [2].

Telepresence systems, which have been steadily developing during the past ten years, are video conference call solutions that can recognize the calling party's facial expressions and body movements. Currently, due to the rapid wide use of smart mobile devices and users roaming around, recent development trends put functions that can be easily applied to businesses and daily lives [3].

The telepresence system market has been mainly established by traditional major companies such as Polycom [7], Cisco [8] and LifeSize [9], and recently, other companies including Radvision [10], Teliris [11], Huawei [12] and Vidyo [13] have started building a new telepresence market with their unique advantageous features. Radvision has the advantage of providing a full solution comparable to that of other leading companies at a relatively low price. Teliris is an equipment and management services provider capable of

managing competitor's solutions within the same environment. Huawei has built various video end point and infra solution product lines. Vido has adopted a unique routing architecture to drastically reduce per port video conference call distribution costs and has the advantage of maximizing the quality of services according to network environments by uniquely utilizing H.264 SVC [15].

However, conference room based high-end telepresence solutions require investments up to several hundred thousand dollars, and since the system must be installed at two or more locations, impose a financial burden to a company. Although more companies are ambiguously defining video conference calls and telepresences on purpose to provide customers with HD video conference call solutions, the large capital expenditure needed to newly build telepresence systems have been an obstacle. Also, since most companies are not equipped with networks to support telepresences, additional costs to set up the network are considered obstacles, and since existing conventional telepresence equipment's interoperability is at a limited state, this fact also acts as an obstacle to the growth of the telepresence market [16].

Because of the high cost of the telepresence systems, recent trends are to introduce relatively inexpensive video conference systems to conference rooms, thereby replacing existing high cost telepresence systems [4].

Although there are recent clear movements from Skype, the industry leading video conference call system provider, to replace the current conference room-based telepresence market, due to expensive camera device equipments, which are the basis for providing HD level image services for telepresence systems, and complicated installation methods, Skype's system is not extensively in use in the conference room-based telepresence markets. In spite of the active linkage between telepresence products and the video conference call product lines, its resolution is somewhat inferior to that of existing telepresence systems.

Therefore, this paper proposes a solution to the telepresence system cost barrier which was to use high-end built-in encoder/decoder cameras and resolution wise, proposes a method to provide telepresence level full HD service to PCs, smart phones and televisions.

- We have replaced high-end cameras with user pre-owned smart phones that function as HD level encoder cameras, and proposed and developed a method of using the smart phone's hardware codec.
- By using its connectivity features, we have built a system where video calls are possible through existing Legacy TV by using the Android set top box which is available throughout the market, and propose its methods.
- In order to provide a seamless conference based service environment, the existing connectivity features have been extended, and while roaming, when the user makes a call from his/her mobile device, the screen is shared with the users in the conference room. Thus, the expansion method where mobile video conference calls are extended to the conference room-based telepresence is presented.

The paper will be organized as follows: First, the design on connectivity will be proposed. Then, the concept of extended connectivity where the mobile device is accessed and the screen base is transmitted to the telepresence system to provide a seamless conference based system environment [5], will be proposed. Thereafter, by actually using this, the method of building a telepresence system through a smart device will be proposed.

2. Video Connectivity Design and Implementation

Connectivity is a new protocol developed to overcome the limitations of conventional set top boxes when providing high resolution video calls via IPTV. Conventional IPTV set top boxes have been developed mainly for viewing VOD or other broadcasts, and are thought to provide an infinite number of channels where anyone can set up individual broadcasting stations [17].

However, most IPTV platform manufactures put priority on network investments where downloading is the focal point, and, therefore, for anyone to set up an individual broadcasting station, commercialization of IPTV STB took place partially to well decode Full-HD level video. Therefore, without changing all the STBs of users, it is impossible now to transit into interactive two-way TV channels, which is considered the basic TV application in recent Google TV and Apple TV.

At first, we found a method to implement full-duplex video communications in the IPTV platform without exchanging existing IPTV STBs. IPTV STBs were constructed based on the following limitations:

- IPTV STBs have very limited memory, and it is difficult to load large-scale library such as SIP.
- In order to save costs, IPTV STBs are comprised mainly of sigma chipsets, such as Broadcomm, which is optimized for decoders without an encoder.
- It is difficult to develop a device driver for connecting general USB cameras for IPTV STBs, and the driver cannot be applied to general purpose USB cameras

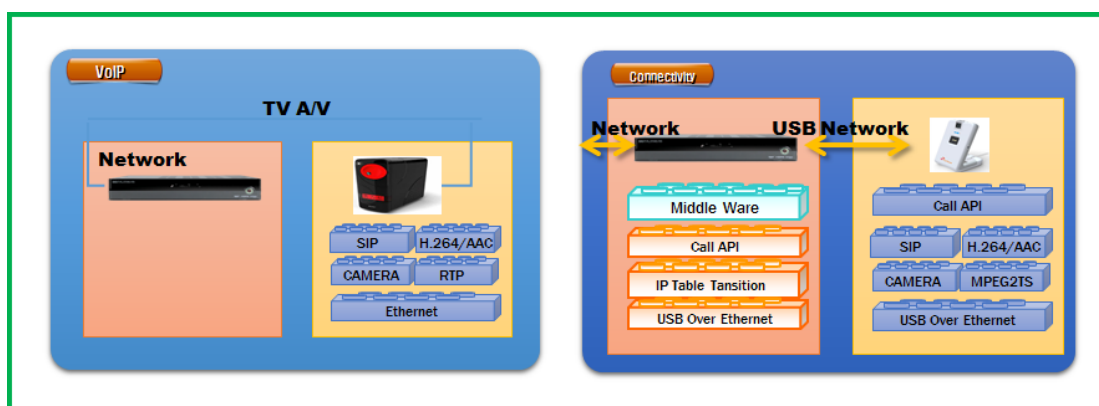


Figure 1-1. Terminal for video call Figure 1-2. Terminal for Connectivity

Due to this environment, the method for providing full-duplex video communications via IPTV can only be implemented in standalone type systems which independently process image processing terminals to the camera, H.264 encoder[6], network, SIP[18], etc.. As illustrated in Figure 1-1, one of these methods was developed by DMT [14] of Korea in the form of a completely hardware independent device where the system is built with no set top service and video conference call service, and where the image/voice signal of the set top box is entered into an independent device and then sent to the TV by mixing the image with the conference call within the hardware. http://www.dmtsai.com/product/MoIP_DMV1000.htm This method, however, simply plays the role of relaying physical audio/video signals, and though it is capable of making video conference calls, it cannot provide services connected to TV services.

Therefore, it was inevitable to find a basic method that provides video conference call services in connection with the TV menus, and, as described in Figure 1-2, a service model form combined with IPTV was developed.

- A 200 megapixels camera with a built-in H.264 encoder is required, and it must be able to process SIP signals.
- SIP and RTP/RTSP for transmitting images should be processed by using the STB's network rather than the network of the video conference call terminal.
- It should be possible to issue commands to the video conference call terminal by using the STB's menus.
- Call API is provided to the set top box to provide seamless services.

Through these concepts, as illustrated in Figure 2, Call APIs have been developed for the set top box, and it became possible by using the APIs to establish a video conference call with the terminal through changing the software at the set top box as an independent device even when the concept of image processing is not known.

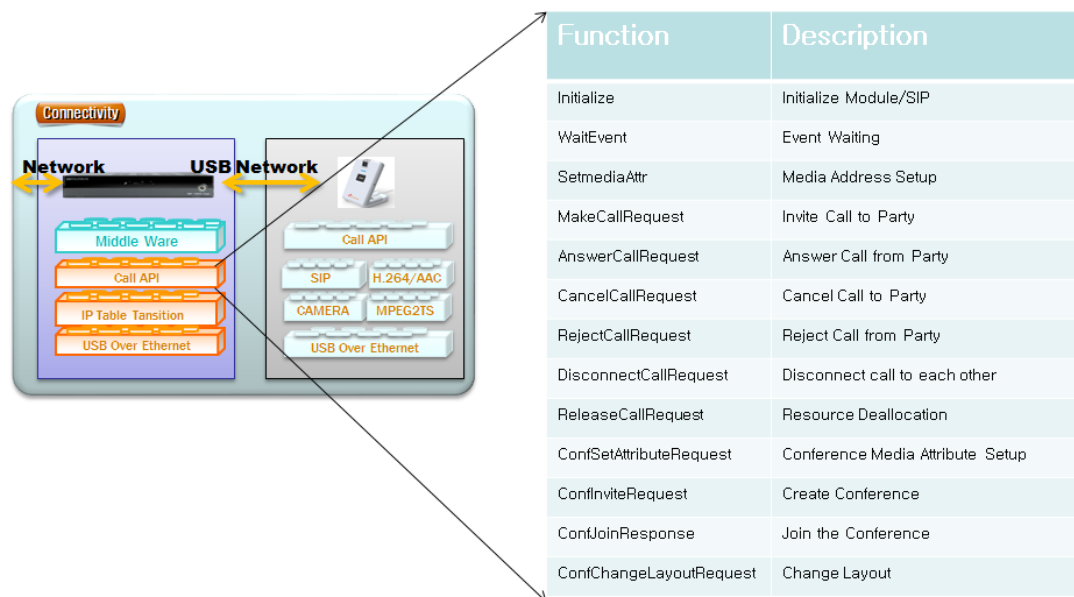


Figure 2. Connectivity Basic API

A video conference call module was provided through USB over Ethernet by providing minimum library to the STB as described above, and the logical connection between the STB and the module could be made through TCP/IP. This method of providing a module, however, also included problems for actual commercialization due to the cost of introducing physical modules, and so the researchers extended the characteristics of Connectivity by using a smart phone which has the characteristics of the module, especially H/W H.264 codec, a camera and a microphone, and provided the function in connection with the STB by using smart phones of most of the kinds.

However, the results of the prior study were limited to the method of providing the service through IP TV service providers and included difficult aspects in actually commercializing

the method. This paper describes a method in which the characteristics of Connectivity is extended so that common subscribers of networks as well as IP TV service providers can construct Tele-presence environments, and the method of using the system by installing an app without configuring the system by the user, and also the method implemented for providing a Connectivity environment seamlessly as needed after accessing the mobile terminal in order to maximize the mobility of the system which is the most salient features of mobile system.

3. System Architecture and Implementation

The system structure of Tele-Presence is generally comprised of a decoder for displaying the screen and an encoder for transmitting the screen and, in this paper, a portable Android device is used as a decoder and an iPhone or an Android smart phone as an encoder for implementing the system according to the paper. Figure 3 illustrates basic construction of the Tele-Presence system according to this paper.



Figure 3. Connectivity Tele-Presence Architecture

3.1. Decoder Android Device

For a decoder android device of the paper, MK802 II manufactured by Riko Magic was used and the corresponding decoder android device was connected to a TV through a HDMI port so that communication with an AP was performed through built-in WIFI module, and an app has been developed for performing Connectivity. The app runs at a service mode, and when the phone is called while the user is watching a video, working on a document or playing games, the phone wakes from the service mode and runs the program through the Call API s provided.

The decoder android device performs the task of decoding the video and audio transmitted from other party in real time by using OMX Interface [19] provided by Android.

Also, basic Primitives are provided so that the mobile encoder device in the same AP can be connected by broadcasting multicast messages on a regular basis for establishing seamless connection of Connectivity. Through the multicast message broadcast periodically, the mobile encoder device in the same AP automatically establishes connection with the decoder Android device.



Figure 4. Ongoing Video Communication

3.2. Mobile Encoder Device

The mobile encoder device performs the role of camera as illustrated in Figure 4, provides the environment of Preview of its own screen, SIP and WIFI, and the portion displayed in UI only shows whether the connection with the decoder Android device has been made. When the user brings his or her terminal to home, the app of the terminal automatically tries, in service mode, to connect to the decoder Android device which had been previously maintained connection and, when the connection has been made, the decoder Android device can provide the screen of other party through TV. If connection is not made between the decoder Android device and the mobile encoder device, the mobile encoder device only performs the role of a simple mobile device.

3.3. Mobile Device

Mobile Device performs its function at the same level of the mobile encoder device while the function is performed when connection with the decoder Android device is not made, and the difference with the mobile encoder device is that encoding and decoding are simultaneously performed in the terminal. If all the functions encoder and decoder have are ensured, the terminal performs its function in small screen.

3.4. Changing Tele-Presence from Mobile-Presence

When call is first received by the mobile device, the video conference call screen of other calling party can be shown on the TV through the connection with the decoder Android device, which is carried out by the following procedures.

1) The mobile device performs signal processing in WIFI or mobile environment, and the mobile device is in the state of continuing call with other party, and 2) the mobile device receives transmitted multicast message through decoder Android device on a regular basis in WIFI environment (when accesses WIFI while in the mobile environment, the existing connection with the network is maintained while accessing WIFI). 3) Received multicast message includes the address and port value available for the user to connect, and the mobile device implicitly tries to connect with the decoder Android device. When connection is made, the Tele-Presence button on the decoding screen of the mobile device become activated. 4) When the user pushes a corresponding button, the mobile device transmits Re-Invite signal to the other party with the address and port value which the decoder Android device has and which is available for decoding. 5) The moment Re-Invite signal is received, the sending port of other party is changed to decoder Android device, and 6) the other party transmits media to the corresponding address. 7) As soon as 200 OK is received on the Re-Invite signal, the decoder Android device changes to Tele-Presence mode and performs the function of a

decoder, and 8) the mobile device changes to a mobile encoder device and performs Preview mode. Figure 5 illustrates performing of this operation.



Figure 5. Mobile to Tele-Presence

5. Conclusion

In this theses, the implementation of Tele-Presence system for common retail market using Connectivity has been described. For this purpose, conventional high cost Tele-Presence system can be supplied at low price and, through the system of the paper, general customers will be able to construct conference room-based video conference call system by using a generally used Android player device and Android beam projector that can be used in video view or business share without having to purchasing high-priced Tele-Presence apparatus with no function of utilizing. Also, other than conventional Tele-Presence system which provides service in fixed form, the method proposed by the paper provides a scheme for On Going Video Sharing so that mobility is sufficiently ensured and the environment connected in the mobile environment can be moved to a conference room-based Tele-Presence system. By introducing the proposed system, it is expected to generalize various conference room-based Tele-Presence systems. The researchers will continue the research by introducing MCU-based multichannel Tele-Presence system in the future for the environment in which variety of Tele-Presence systems can be constructed.

Acknowledgements

This work is supported by Dept. of LINC, Dongguk University. (Project Title: Leaders in INdustry-university Cooperation.) The authors would like to thank Jeong-Hoon, Koh and Eu-Soo, Lee.

References

- [1] Zaka-ul-Haque, R. Siddiqui, R. Javaid and I. M. Hussain, "Tele-presence Based Patient Monitoring Robot", International Journal of Advanced Science and Technology, vol. 42, no. 5, (2012) May.
- [2] Y. H., Nam, H. J. Park, C. H. Cho and J. H. Park, "An Interactive IPTV System with Community Participation in Cloud Computing Environments", IEEE Systems Journal, (accepted for publication), (2013).
- [3] A. Jones, M. Lang, G. Fyffe, X. Yu, J. Busch, I. McDowall, M. Bolas and P. Debevec, "Achieving eye contact in a one-to-many 3d video teleconferencing system", ACM Transactions on Graphics, vol. 28, no. 3, (2009).
- [4] K. L. Kim and S. Y. Jeong, "An Exploratory Study for the Adoption and Use of Telepresence: Focusing on Supporting Trade Business Activity of SME", Journal of the Korea Academia-Industrial cooperation Society, vol. 11, Issue 9, (2010).
- [5] J. H. Choi, S. Y. Lee, I. K. Kim and Y. S. Woo, "Live-Video Service Using Multicast in Wireless Network", International Journal of Multimedia and Ubiquitous Engineering, vol. 7, no. 4, (2012).
- [6] S. Ghorbani and F. Zargari, "A Unified Architecture for Implementation of the Entire Transforms in the H.264/AVC Encoder", International Journal of Multimedia and Ubiquitous Engineering, vol. 8, no. 1, (2013).

- [7] Polycom, Telepresence, (2013), <http://www.polycom.com/products-services/hd-telepresence-video-conferencing.html>.
- [8] cisco, Telepresence, (2011), www.cisco.com/en/US/products/ps7060/index.html.
- [9] LifeSize, Telepresence, (2013), <http://www.lifesize.com/en/solutions/location/large-conference-room>.
- [10] Radvision, Telepresence, (2013), <http://www.radvision.co.kr/Products/Telepresence/>.
- [11] Teliris, Telepresence, (2013), <http://www.teliris.com/>.
- [12] Huawei, Telepresence, (2013), <http://enterprise.huawei.com/en/products/coll-communication/telepresence-videoconference/>.
- [13] Vido, Telepresence, (2013), <http://www.vidyo.com/>.
- [14] DMT, MoIP Device, (2013), http://www.dmts.com/product/MoIP_DMV1000.htm.
- [15] P. Karcher, M. Brown, H. Dewing and A. Smith, "The Forrester Wave : Room-Based Videoconferencing, Q3 2012", in FORRESTER, (2012) August.
- [16] E. J. Lee, S. J. Yeon and W. G. Ha, "Market trends and Current Status on Enterprise Telepresence", in Electronics and Telecommunications Trends, vol. 26, (2011) June.
- [17] I. K. Park, S. H. oh, S. K. Yoon and H. Y. Song, "An Implementatin of user-participated interactive IPTV service system", in IEEE International Symposium, (2008) April.
- [18] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261 (Proposed Standard), Internet Engineering Task Force, (2002) June, updated by RFCs 3265, 3853, 4320, 4916, 5393, 5621, <http://www.ietf.org/rfc/rfc3261.txt>.

Authors



Jong ho, Shin

He received a Master of Arts(M.A.) degree in digital broadcasting from Kyunghee University, Korea and an doctor of engineering in Information Technology from EARIST national University, Republic of the Philippines in 2007 and 2012, respectively. He was Director of Research at C-Channel IPTV Company from 2007 to 2009 and a General Manager at Korea Digital Convergence Association from 2009 to 2012. He is currently a Professor at Dongguk University LINC Project, Korea. His research interests include mass media convergence through Information Technology, mobile tele-presence system.



Ky Hyun, Um

He is now a professor at the department of Multimedia Engineering, Dongguk University. He received his B.S. degree in Applied Mathematics from Seoul National University in 1975, M.S. degree in Computer Science from KAIST in 1977, and Ph.D. degree in Computer Engineering from Seoul National University, Korea, in 1994. His research interests are multimedia application system design, serious game system, multimedia database system and its applications, and natural user interface design.



In Chul, Kim

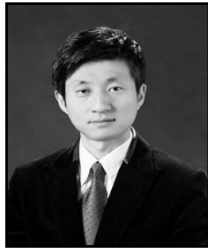
He received the B.S. degrees in information system engineering from Hanshin University, Korea in 2005. He was a Research Engineer at Ntelia Company from 2004 to 2007. He was a Senior Researcher at SN-Media Company from 2007 to 20011. He was also a Principal Researcher at M2Soft Company from 20011 to 2013. He is currently a Principal

Researcher at M2soft Company, Korea. His research interests include conference system, mobile tele-presence system.



Chae Ho, Cho

He received the B.S. and M.S. degrees in computer engineering from Ajou University, Korea in 1997 and 1999, respectively. He was a Research Engineer at Unisec Company from 1999 to 2002. He was a Principal Researcher at Point-I Company from 2002 to 2007. He was also a Chief Researcher at M2Soft Company from 2007 to 2013. He is currently a Chief Researcher at M2soft Company, Korea. His research interests include mobile video system, mobile tele-presence system.



Ill Chul, Doo

He received his M.E. degree in System Management from Sangmyung University, Korea, in 2000 and Ph.D. degree in Digital Culture & Contents from Hanyang University, Korea, in 2011

He was an adjunct professor in the Department of Digital Culture & Contents Hanyang University from 2009 to 2012. He is currently a Professor at Dongguk University LINC Project, Korea. His interests are mobile contents, and cultural industry, and cultural technology.

