LETTER Special Section on Artificial Reality and Telexistence

A Novel Test-Bed for Immersive and Interactive Broadcasting Production Using Augmented Reality and Haptics

Seungjun KIM[†a)], ***Student Member***, **Jongeun CHA**[†b)], **Jongphil KIM**[†], **Jeha RYU**[†c)], **Seongeun EOM**[†], **Nitaigour P. MAHALIK**[†], ***and*** **Byungha AHN**[†d)], ***Nonmembers***

SUMMARY In this paper, we demonstrate an immersive and interactive broadcasting production system with a new haptically enhanced multimedia broadcasting chain. The system adapts Augmented Reality (AR) techniques, which merges captured videos and virtual 3D media seamlessly through multimedia streaming technology, and haptic interaction technology in near real-time. In this system, viewers at the haptic multimedia client can interact with AR broadcasting production transmitted via communication network. We demonstrate two test applications, which show that the addition of AR- and haptic-interaction to the conventional audio-visual contents can improve immersiveness and interactivity of viewers with rich contents service.

key words: *broadcasting, augmented reality, haptics, interactive, viewer immersion*

1. Introduction

With the rapid growth of computing and telecommunication technology, development of recent digital multimedia systems appears to be highly dedicated to their functional enhancement in terms of users' immersion and interactions. In this context, much attention is newly given to the augmented reality (AR) technology because of its simple but excellent interactive display and tracking potential. It has been widely used in various applications within the industrial and scientific communities, edutainment, geographic visualization, and interactive interface design in public spaces [1].

Recently, the noteworthy research in the content of application of AR in broadcasting production has been carried out by BBC creative R&D groups [2], [3]. They showed three separate case studies on entertainment, education and information contents, and used them in the studio, classroom and home, respectively. Their main attention was to introduce prospective features of AR technology for future media production. Despite a short disappearance of virtual objects for the fast feature motions, AR technology shows excellence in interactive visualization while providing simple and user-friendly interface. The camera tracking system of virtual studios is robust and highly accurate; however, it is only effective in a special chroma-keying material space of a fully synthetic 3D stage. The people in the studio have to follow pre-scripted animation scenario of virtual contents in the empty space. In the meantime, Media Interaction Group of Philips Research in Eindhoven [4] reported recent efforts for the interactive television from its concepts and history to storytelling application. Enhanced future broadcasting network integrated with communication network technology can enable us to enjoy attractive bi-directional services by dynamically immersing the users into the broadcasting production.

In this paper, we demonstrate an immersive and interactive broadcasting system combining haptics with augmented reality based on haptically enhanced broadcasting chain, the concept of which had been suggested by authors [5], [6], [9]. Firstly, we have embedded AR broadcasting productions into our broadcasting chain that expands its operational domain into four phases; scene capture, AD (Auxiliary Data) editing, data transmission, and view & interaction. In the proposed chain, viewers may touch or manipulate objects in broadcasted contents, which can be simulated using haptic technology. One of the related researches was performed by O'Modhrain and Oakley who presented Touch TV by showing two potential scenarios: the creation of authored haptic effects for children's cartoon and the automatic capture of motion data to be streamed [10]. Also, J. Vallino and C. Brown demonstrated haptic-graphic, virtual-to-virtual haptic and virtual-to-real haptic interactions with relationships between multiple coordinate [11]. However, they did not demonstrate their work for the broadcasting chain.

In order to show feasibility and to find out technical issues, we have designed a prototype test-bed system, which is composed of 3D media database, AR module (called *AR process*) for broadcasting production and Haptic module (called *Haptic process*) for client interaction. We stream AR-related data in broadcasting production via communication network and demonstrate two test haptic interactive scenarios.

2. Test-Bed System Overview

A prototype test-bed system has been designed to test the haptically enhanced immersive multimedia broadcasting chain and relevant applications using augmented reality and haptics. As shown in Fig. 1, the test-bed system is implemented by configuring 3D media database, AR process,

Manuscript received March 11, 2005.
Manuscript revised June 20, 2005.

[†]The authors are with the Department of Mechatronics, Gwangju Institute of Science and Technology (GIST), 1 Oryoung-dong, Buk-gu, Gwangju, Korea.
a) E-mail: zizone@gist.ac.kr
b) E-mail: gaecha@gist.ac.kr
c) E-mail: ryu@gist.ac.kr
d) E-mail: bayhay@gist.ac.kr

DOI: 10.1093/ietisy/e89–d.1.106

streaming modules and haptic process according to the proposed broadcasting chain.

3D media database contains 3D object model information such as geometry, color, texture, depth, material and haptic (tactile & force or motion) data. These data are overlaid on the captured video, which thus makes haptic interaction possible at the client site. In order to retrieve relevant 3D objects data from database and to interact with them, object identifiers are used. There may be many different ways of data sending and interaction principles. For example, a haptic-related data (i.e. force and torque vector when touched to an object) may be produced and stored in the studio and be sent to the viewers for passive touching interaction. In the current test-bed, however, in order to give active haptic interaction, at the beginning of televising a broadcasting production, all the related 3D media are downloaded to client beforehand. Then the viewer can explore and manipulate the designated object at the client site. In this active touching method, haptic interaction data that is generated by collision detection and force computation procedure [12] is created and displayed in a standard desktop PC at the client site. Note in Fig. 1 that the haptic probe data is graphically rendered for visually showing haptic interaction point on the augmented reality scene.

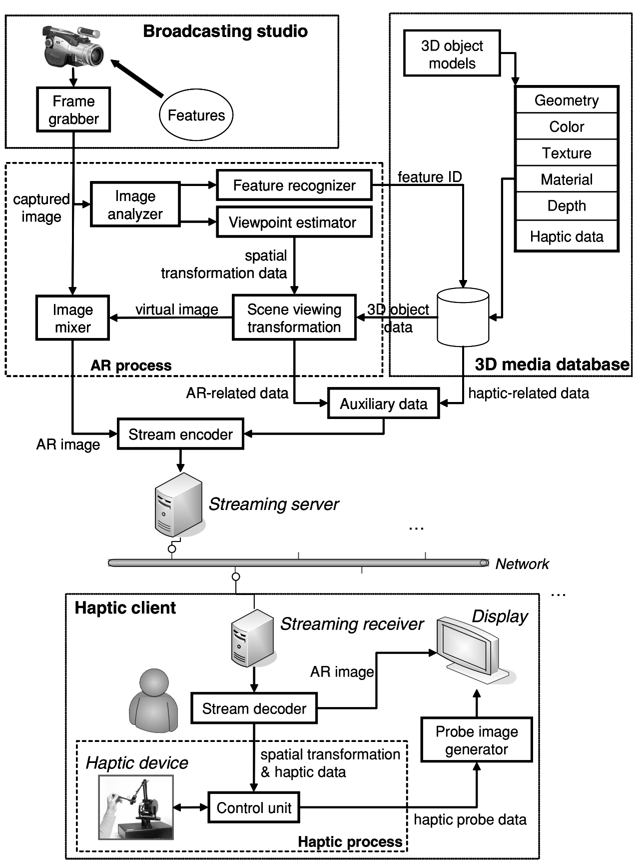


Fig. 1 Test-bed system architecture.

3. Data Streaming

In our test demonstration, TCP/IP network programming using the multimedia streaming technologies [13] are used. Two packet types, the first header byte for one is 0x01 ('0x' means hexadecimal value) and the other is 0x02, are assigned for transmission and reception, respectively. And then, key transmission data is specified as listed in Table 1. 3×4 transformation matrices for projection (P) and model viewing (T_{cm}) are transmitted as payloads, which are used to set up haptic probe navigation space. In addition, six haptic probe data (H) are streamed to represent the present position of the probe in the navigation space.

If haptic interaction time is given on the air, which means that AR process is also updating spatial transformation data, broadcasters in the studio guide viewers when and how they can touch or manipulate objects effectively. If the viewers want to interact with pre-downloaded 3D object models at any moment, the streaming server delivers the updated payload data continuously during the time.

At the same time, for future valid application, AR stream transmission over MPEG-2 TS has been considered [7]; however, we have hopefully excluded the technical issues of vision-touch incorporation in view of multimodal interaction because of the necessity of more intensive discussion as future works. To provide backward compatibility with the current digital television broadcasting system, AR-related data (e.g. pattern images, 3D computer-generated objects data, etc) including encoded video data should be delivered by using MPEG-2 TS (Transport Stream). Due to the limited bandwidth of broadcasting stream, however, all data for AR service are unable to be carried over MPEG-2 TS at a given time. Therefore, a two-stage transmission mechanism has been utilized; raw data for basic AR-service is conveyed over MPEG-2 TS and optional data, which is comparatively huge, for full AR-service is offered through Internet. This architecture requires additional internet provision of the broadcasting receiver.

In the transmission over MPEG-2 TS, our main focus is on two technical aspects: (1) PMT (Program Map Table) configuration for AR stream transmission and (2) AR-data encapsulation on the PES (Packetized Element Stream) level as PES packet data. Firstly, PMT should have (1) indication about which TS packet is carrying raw AR-data and (2) location information for optional AR-data which will

Table 1 Transmission data.

Packet type (send/receive)		Object	Payload
0x01	0x02	P	P_{00}, P_{01}, P_{02}, P_{03}, P_{10}, P_{11}, P_{12}, P_{13}, P_{20}, P_{21}, P_{22}, P_{23}
0x01	0x02	T_{cm}	T_{00}, T_{01}, T_{02}, T_{03}, T_{10}, T_{11}, T_{12}, T_{13}, T_{20}, T_{21}, T_{22}, T_{23}
0x02	0x01	H	x, y, z, θ_x, θ_y, θ_z

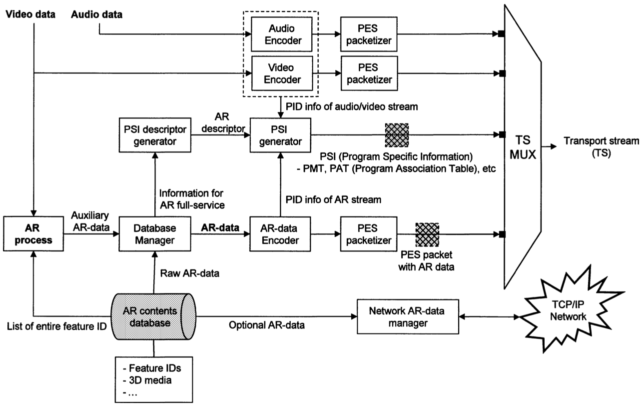


Fig. 2 Transmission architecture for AR service over MPEG-2 TS.

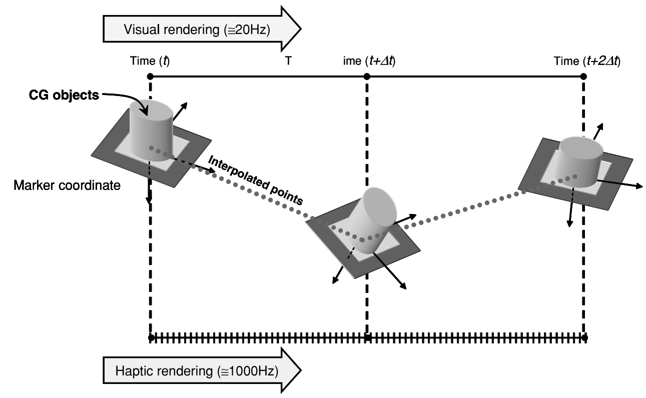


Fig. 3 The marker pose interpolation between the visual frame rate.

be transmitted through the Internet. Next, private_stream_1 (one of the private PES streams in MPEG-2 standard [14]) is selected to carry AR-data, which is useful to design integrated receiver/decoder chip sets because it has the same format as typical private streams (e.g. video or audio element stream). The proposed transmission architecture for AR service is presented in Fig. 2.

4. Consistency between Visual Augmentation and Haptic Interaction

In *AR process*, the captured video is analyzed to detect features which are used to determine the objects to be retrieved and derive their 3D spatial relationship between them and the camera in transformation matrix form. In this research, we have utilized ARToolkit [15] that can facilitate capturing of features of any kind of patterned boards. In ARToolkit, spatial transformation data are estimated in the frame capturing and processing rate, around 20 Hz. These data are very important for the seamless augmentation and visually consistent interaction. The image-based process causes high frequency errors making tremors in CG models' augmentation and in viewers' haptic interaction. It is almost unaware in the user's dynamic view state; however, it can disturb users' observation in their static view. For a single marker, the farther the marker is apart from the camera and the smaller the tilt angle is, the less the estimated accuracy. To estimate static registration errors, in this experiment, a fiducial marker (8×8 cm) is fixed on the wall perpendicular to the camera and a constant 70 cm apart is kept between the camera and marker. Although the marker is static in the scene, the translation ($x - y - z$ position) errors and the rotation (θ_{roll}-θ_{pitch}-θ_{yaw} angle) errors are time-varying. To eliminate the trembling errors [16], low-pass filtering has been applied by using threshold values: $x = 0.10$ cm, $y = 0.23$ cm, $z = 2.34$ cm, $\Delta\theta_{roll} = 0.80$, $\Delta\theta_{pitch} = 0.22$ and $\Delta\theta_{yaw} = 0.40$. If the variation is larger than the threshold, the motion of marker is regarded as in the dynamic state and the filtering pauses. In our demonstration, the trembling is almost completely eliminated so that haptic interaction becomes smoother in the static view state [8].

In addition, we should consider the fact that haptic is rendered more frequently than graphics. When the features move a little fast, their estimated moving distances can be discontinuous between frames. These conflicts are endurable in visual aspect, however, very crucial in haptic interaction aspect where haptical sensation is rendered in high resolution. Thus, as shown in Fig. 3, we have interpolated the marker pose between the visual frames. The results indicate that the interpolation is sufficient to give the apparent continuous force when the CG model is moving.

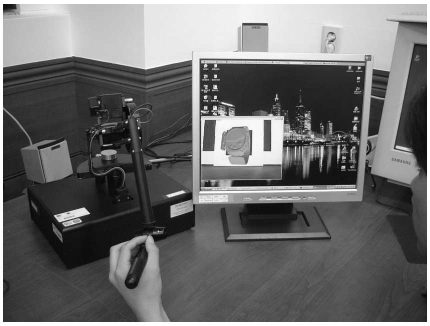
5. Demonstration

For the validation of the proposed framework, two test applications have been demonstrated; (1) an electric circuit for education aids and (2) a wrist-held MP3 player for a home shopping scenario as shown in Fig. 4.

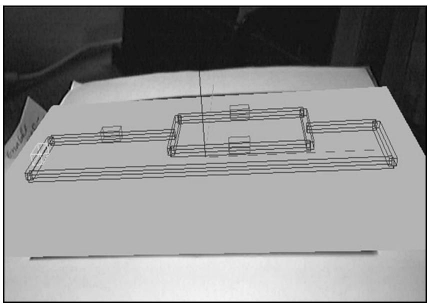
5.1 Test Applications

In the client site, the object-augmented scene is displayed and haptic probe is overlaid on it. The probe can navigate in the same 3D space of camera's viewpoint. The control unit processes the video media and the haptic data to control haptic device by using haptic rendering algorithm that is the main functions of *haptic process*. Viewers can feel the haptic effects synchronized with application scenarios, for example, passively by putting their hands on the vibrotactile display device. They also can actively touch, explore, and manipulate the 3D CG objects according to the preplanned path or the viewers' will. In our demonstration, we have used the 6-dof haptic devices, Phantom [17], which provides a very compelling sense of touching virtual objects especially when there is no conflict with the visual cues.

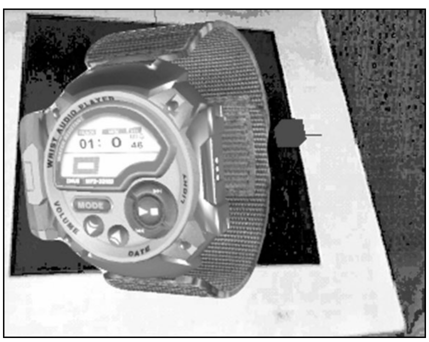
In the first 'electric circuit' application, we presented an engineering education program to show how the haptic interaction can be used for helping viewers' understanding of scientific experiments. A science teacher presents a feature board in front of camera, on which pre-designed CG circuit model is augmented, and starts its simulation. Then, viewers in the haptic client site touch graphical elec-



(a) Haptic client



(b) Electric circuit



(c) Wrist-held MP3 player

Fig. 4 Haptic interaction with test applications.

tric wires and feel the abstract current intensities and directions assigned to them. We note that it is very helpful for viewers' understanding of current flows in circuit. It can be extended to other scientific or engineering applications in which the conversion of physical or abstract data such as force, intensity, strength, and temperature to the physical sense is necessary.

In the second 'home shopping' scenario, a shopping host tries to advertise a product: wrist-held MP3 player. When the host puts a feature board in the camera range, viewers can watch the augmented 3D MP3 player model on it. Needless to bear the real product, the host can show all the aspects of the model by rotating the board while explaining the functions or features of the product. This board manipulation works more effectively when the product is too big and weighty to be handled. Then, the haptic interaction time is given. Using haptic device, viewers can try to push a functional button to know how it works and the corresponding operation or information is displayed. Also, they can actively contact the outline frame of the product.

5.2 Discussions

Major information in the broadcasted scene is conveyed through the live video image, and AR technique supplements reality rather than completely replacing it; AR can ensure the user's perception and interaction with real world unlike fully virtual environment. Thus, the users' interactivity by haptical & visual augmentation has shown the potentials of contributing to their immersiveness about the entire broadcasted information beyond in the case of the live video without any user-interactive object or the object divorced from real world. In this context, the term, *immersiveness*, has been conceived in this paper.

As noted in the test applications, AR production is capable of aiding viewer's understanding of the broadcasting contents as well as representing plenty of visual information. Enriched graphic effects (e.g. animation, object deformation, artificial effects, etc) in computing space are technically performed within specified boundaries of feature areas in the sequential broadcasted scenes. In present broadcasting, AR scene-like effects have occasionally appeared in sportscast or CFs (Commercial Films); In the case of soccer cast, for example, CG advertising fences or virtual guide lines near goal area are seamlessly aligned with the playground, virtual national flags are floating over the ground or the score is painted on the center of ground. The real objects look to be positioned on the ground. However, autonomous feature tracking is not incorporated in their production, so that all of the CG effects have been fully synthesized frame by frame. Or, it is necessary to prearrange cameras ensuring their recordable motion mechanism in the playground or studio. We have attempted to incorporate generic AR process in the broadcasting framework. The AR production will be created by service providers according to their purposes. And MCs or panellists of a program can use real physical panels or charts as the tracking references. Then, CG 2D/3D media is confidently augmented on them.

In addition, haptic interaction is implemented as an interactive multimedia service at client site. And the test scenarios have been initiated as feasible applications to validate our haptically-enhanced broadcasting chain. In several forums and exhibitions for interim result demonstration, we have confirmed the fact that viewers' interaction enhances their immersion in the broadcasting production. Most of participants showed great interest in experiencing artificial sense of touch on CG products. Considering the rapid growth of streaming technology and multimedia services via Internet, the users' interaction service such as Haptic on Demand (HOD) is expectable as a complementary service in future multimedia.

6. Conclusions

In this paper, we have presented a prototype test-bed system based on the newly proposed broadcasting chain using augmented reality and haptics. The system demonstrates future realistic multimedia broadcasting that includes sense-of-touch. Based on the proposed framework, two test applications are demonstrated to show that the haptic interaction can be incorporated in the exemplar broadcasting programs. Moreover, the combination of these technologies, AR and haptics, is expected to make producers take a new direction of future broadcasting production. In our present demonstration, sizes of test AR productions are insignificant, so that it was unnecessary to integrate the proposed two-stage transmission mechanism with the test-bed system. To represent viewer's immersive interaction in actual broadcasting environment, we are arranging high-quality 3D media in huge size and real-life experiments in the studio. After this, the specific evaluation of our broadcasting chain and the technical discussion in view of multimodal interaction will be intensively executed as future works.

Acknowledgements

This work is supported by the Ministry of Information and Communication (MIC) South Korea through the Realistic Broadcasting IT Research Center (RBRC) at Gwangju Institute of Science and Technology (GIST).

References

[1] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent advanced in augmented reality," IEEE Comput. Graph. Appl., vol.21, no.6, pp.34–47, Nov./Dec. 2001.
[2] V. Lalioti and A. Woolard, "Mixed reality production of the future," BBC R&D White Paper, published in the conference publication of the Int'l Broadcasting Convention (IBC 2003), Amsterdam, The Netherlands, pp.314–320, Sept. 2003.
[3] A. Woolard, V. Lalioti, N. Hedley, N. Carrigan, M. Hammond, and J. Julien, "Case studies in application of augmented reality in future media production," Proc. 2nd IEEE and ACM Int'l Symp. on Mixed and Augmented Reality (ISMAR 2003), pp.294–295, Oct. 2003.
[4] M. Bukowska, "Winky Dink half a century later. Interaction with broadcast content: concept development based on an interactive storytelling application for children," Technical Report, Media Interaction Group, Philips Research Laboratories, Eindhoven, Aug. 2001.
[5] J.E. Cha, J.H. Ryu, S.J. Kim, S.E. Eom, and B.H. Ahn, "Haptic interaction in realistic multimedia broadcasting," Proc. 5th Pacific Rim Conference on Multimedia (PCM 2004) on Advances in Multimedia Information Processing, Part III, pp.482–490, Nov./Dec. 2004.
[6] J.E. Cha, J.H. Ryu, S.J. Kim, and B.H. Ahn, "A haptically enhanced broadcasting system," Proc. 1st Joint EuroHaptics Conf. and Symp. on Haptic Interfaces for Virtual Environment and Teleoperator Systems (worldHAPTICS), Pisa, Italy, March 2005.
[7] S.D. Kim, J.I. Yun, S.J. Kim, and B.H. Ahn, "AR over MPEG-2 TS," Proc. 20th Int'l Technical Conf. on Circuits/System, Computers and Communications (ITC-CSCC2005), vol.1, pp.287–288, Jeju, Korea, July 2005.
[8] J.E. Cha, B.C. Lee, J.P. Kim, S.J. Kim, and J.H. Ryu, "Smooth haptic interaction in broadcasted augmented reality," Proc. 10th IFIP TC13 Int'l Conf. on Human-Computer Interaction (INTERACT 2005), LNCS 3585, pp.1046–1049, Rome, Italy, Sept. 2005.
[9] Y. Oktay, B.C. Lee, N.P. Mahalik, and J.H. Ryu "An electrotactile device for broadcasting the game applications," Int'l Conf. on Human Computer Interface, pp.388–394, Seoul, Feb. 2005.
[10] S. O'Modhrain and I. Oakley, "TouchTV: Adding feeling to broadcast media," Proc. European Conf. on Interactive Television: From Viewer to Actors, pp.41–47, 2003.
[11] J. Vallino and C. Brown, "Haptics in augmented reality," Proc. IEEE Int'l Conf. on Multimedia Computing and Systems (ICMCS), vol.1, pp.195–200, June 1999.
[12] J.P. Kim and J.H. Ryu, "Hardware based 2.5D haptic rendering algorithm using localized occupance map instance," Proc. 14th Int'l Conf. on Artificial Reality and Telexistence (ICAT2004), pp.132–137, Seoul, Nov. 2004.
[13] Windows Media Technology, http://www.microsoft.com/windows/windowsmedia.
[14] ISO/IEC 13818-1, "Information technology-Generic coding of moving pictures and associated audio information: system," Dec. 2000.
[15] ARToolKit, http://www.hitl.washington.edu/artoolkit.
[16] D.F. Abawi and J. Bienwald, "Accuracy in optimal tracking with fiducial markers: An accuracy function for ARToolKit," Proc. 3rd IEEE and ACM International Symp. on Mixed and Augmented Reality (ISMAR 2004), pp.260–261, Nov. 2004.
[17] Phantom, SensAble Technologies, http://www.sensable.com.