

As if the other person in the video conference is in the same space! Development of an immersive conference system with a sense of space

- Seongjun Kang of the School of Integrated Technology won the Best Paper Award at the International Conference of the Korea Internet Information Society



▲ (From left) Professor SeungJun Kim and master's student Seongjun Kang

GIST (Gwangju Institute of Science and Technology, President Kiseon Kim) graduate students developed a 'portal display' conference system that makes the other person participating in a non-face-to-face video conference feel as if they are in the same space through a monitor, and it won the best paper award at an international conference*.

* portal display: screen-based 3D stereoscopic conferencing system for immersive social telepresence)

A paper on the development of a 'portal display' conference system led by master's student Seongjun Kang (Advisor: SeungJun Kim) of the School of Integrated Technology and the verification of system effectiveness through user-based evaluation was hosted by the Korea Internet Information Society* at the 14th International Conference on Internet 'ICONI 2022'. Among a total of 122 papers, one paper was selected as the final paper (within the top 1%), and will be published in the SCI-level journal TIIS (KSII Transactions on Internet and Information Systems).

* Korea Internet Information Society: Established in 2000 to integrate computer-related groups, academia, and computer-related information technologies and contribute to technological development,

the Korea Internet Information Society has played a part in establishing the internet environment in Korea, which is called an internet powerhouse, as one of the leading Internet-related societies in Korea. It is one of the most influential academic societies in the domestic IT field with 2,700 members. In particular, it has been recognized in domestic and international academic circles by publishing the 'TIIS International Journal', an SCI-level English-language academic journal registered in SCIE.

Mr. Kang's thesis develops and verifies a system in which two different spaces are scanned and transmitted/received by a depth camera in real time, and spatial information is recombined to create a sense of depth, giving the impression that the two spaces are connected based on the monitor. It was carried out with support from the Korea Institute of Electronics Technology.



[Reference] An immersive non-face-to-face video system using a 'portal display' proposed by the research team. There is a parallax depending on the viewing direction, so you can feel the three-dimensional effect.

In a video conferencing system such as ZOOM, which is frequently used after the COVID-19 pandemic, spatial information between two different spaces is lacking, making it difficult for users to know who they are talking to and pointing to.

The research team proposes a method in which the face position is tracked in real time through image-based face recognition through a webcam, and the graphic space inside the monitor is linearly transformed* accordingly. They designed and developed a so-called 'portal display' where meetings and collaborations can take place in an environment where non-verbal information such as gestures and gaze information is not lost with just one webcam.

* linear transformation: Linear transformation of graphic space is a principle that can artificially create a stereoscopic view that generally occurs when a 3D object is viewed from various angles (stereo disparity, the difference in the position of an object in the two images according to the left and right visual fields).

The 'portal display' uses Lidar* cameras, which are mainly applied to autonomous vehicles, to scan remote users' RGB-D (Red, Green, Blue, Depth)* information in real time. It is a network formed through UDP communication* and works by transmitting and recombining RGB-D information in front of users in other spaces.

* Lidar: Light detection and ranging (Lidar) uses eye-safe laser beams to see the world in 3D and accurately display the irradiated environment on machines and computers.

* RGB-D: RGB-D usually refers to red, green, blue, and depth data captured by an RGB-D sensor. RGB-D images provide per-pixel depth information aligned with the corresponding image pixels.

* UDP communication: The User Datagram Protocol (UDP) is a communication protocol used on the Internet for particularly time-sensitive transmissions such as video playback or DNS lookups. Speeds up communication by not formally establishing a connection before data is transmitted.

Professor SeungJun Kim said, "As virtual reality technology develops, research to develop more immersive non-face-to-face video systems and conduct user-based evaluations is becoming increasingly important. We also plan to conduct research to enable higher quality interactions through deep learning-based 3D spatial composition models."



▲ Student Seongjun Kang, who won the Best Paper Award at the 2022 International Internet Conference, is giving a presentation.

Seongjun Kang, the first author, said, "Through this study, it is possible to conduct a remote conference with a sense of space even in a common situation using only one webcam. In the future, we plan to distribute free untact programs that are open to the public so that anyone can use the untact remote meeting platform with a sense of space."