

## cGANs 기반 3D 포인트 클라우드 데이터의 실시간 전송 기법

### Real-time transmission of 3G point cloud data based on cGANs

Kwang-Seong Shin<sup>1</sup> · Seong-Yoon Shin<sup>2\*</sup>

<sup>1</sup>Associate professor, Department of Digital Contents Engineering, Wonkwang University, 54538 Korea

<sup>2\*</sup>Professor, Department of Computer Information Engineering, Kunsan National University, 54150, Korea

#### ABSTRACT

We present a method for transmitting 3D object information in real time in a telepresence system. Three-dimensional object information consists of a large amount of point cloud data, which requires high performance computing power and ultra-wideband network transmission environment to process and transmit such a large amount of data in real time. In this paper, multiple users can transmit object motion and facial expression information in real time even in small network bands by using GANs (Generative Adversarial Networks), a non-supervised learning machine learning algorithm, for real-time transmission of 3D point cloud data. In particular, we propose the creation of an object similar to the original using only the feature information of 3D objects using conditional GANs.

**Keywords** : GANs, 3D point cloud, conditional-GANs, hologram, tele presence

#### I. 서 론

텔레프레즌스(tele-presence)는 공간적으로 떨어져 있는 장소 또는 가상의 장소를 신체적으로 경험하는 것을 말하는데 고전적인 화상 전화나 화상 회의 시스템으로부터 발전되어 미국에 있는 사람과 영국에 있는 사람

이 한국의 특정 공간에 동시에 나타나 실시간으로 대화나 상호작용이 가능한 형태로 발전되어가고 있다. 컴퓨터와 대용량 통신회선을 이용해 신체적으로 가 있지 않은 다른 장소를 생생하게 경험한다는 점에서 원격현장감이라는 용어를 사용하기도 한다[1]. 즉 가상의 공간에 캡처된 3차원 실물객체의 움직임이 그대로 원격지에 전송되어 원격지에서 서로 상호작용이 가능하게 된다는 것이다. 이를 실현하기 위해서는 3D 포인트 클라우드 데이터로 표현되는 3차원 객체를 생성하고 이를 실시간으로 전송하기 위한 기술이 필요하다.

포인트 클라우드는 텔레프레즌스 분야에서 실제 업무 공간을 스캔하고 이를 통해 3D모델/맵을 컴파일(compile)하여 쓸 수 있게 해준다[2]. 일반적으로 depth camera를 장착한 키넥트 시스템(Kinect system)이나 3D 레이저스캐너를 통해 캡처할 수 있다. 3차원 객체를 표현하기 위해 3D 포인트 클라우드 데이터의 점들은 3차원 객체의 방향성 및 거리정보, 색 정보 등을 가지고 있고 시점에 따라 다른 정보를 제공해주기 때문에 정보의 양이 매우 크다. 따라서 이러한 매우 큰 데이터를 실시간으로 처리하고 전송하기 위해서는 고성능의 컴퓨팅 자원과 네트워크 자원을 필요로 한다. 현재까지 MicroSoft, CISCO등 글로벌 대기업을 중심으로 텔레프레즌스시스템이 연구 개발되고 있으나 방대한 용량의 3D 포인트 클라우드 데이터의 실시간 전송을 위한 기술개발은 아직도 많은 한계를 보이고 있다.

본 연구에서는 3D 포인트 클라우드 데이터의 실시간 전송을 위해 비지도학습 계열의 기계학습 알고리즘인 GANs(Generative Adversarial Networks)를 이용하여 적은 네트워크 대역에서도 다수의 사용자가 실시간으로 객체의 움직임 정보와 표정 정보를 전송할 수 있는 연구를 수행한다. GANs를 이용하면 진짜보다 진짜 같은 가짜 모델을 생성할 수 있는데 최근 중국의 CCTV 뉴스의 여성앵커는 GANs에 의해 생성된 가상의 인물이다. GANs의 응용 알고리즘 중 cGANs(conditional-GANs)를 이용하여 3차원 객체의 특징 정보와 스케레톤

Received 30 September 2019, Revised 30 September 2019, Accepted 10 October 2019

Open Access <http://doi.org/10.6109/jkiice.2019.23.11.1482>

print ISSN: 2234-4772 online ISSN: 2288-4165

데이터(skeleton data)만을 이용하여 원본과 유사한 객체를 생성하는 연구를 수행한다. 3차원으로 생성된 사람의 형상을 원격지로 전송하는 상황을 예로 들면 사람의 표정과 움직임의 정보 중 꼭 필요한 부분(움직임을 위한 스켈레톤 정보, 머리색, 피부색 등)만 전송한 후 cGANs 알고리즘을 이용하여 원하는 형태의 객체로 재현이 가능하게 된다. 이러한 경우 최소 조건만을 이용하여 임의의 재현된 3차원 객체의 정보는 결과적으로 원본 데이터의 구체적인 형상과 다르게 되지만 움직임에 대한 정보와 표정정보 등은 그대로 재현이 가능하고 일부 형태만 다르게 생성이 되는데, 유사도(level of Similarity)와 디테일의 정도(level of detail)에 따른 인간의 인지 변화도에 대한 반응관계를 survey를 통해 정량화할 수 있다.

## II. 관련연구

본 연구의 목표는 방대한 용량의 3D 포인트 클라우드 데이터의 실시간 전송을 위한 방법으로 cGANs를 이용한 특징기반 임의생성 시스템을 개발하는데 있다. 텔레프레즌스 시스템의 핵심 필요 기술인 3D 포인트 클라우드 데이터의 실시간 전송 기술 측면에서, depth camera가 장착된 키넥트 시스템(kinect system)을 통해 캡처된 3D 포인트 클라우드 데이터를 실시간으로 전송하기 위한 cGANs-인코더를 구현하고, 특징정보를 전송받은 모바일 단말에서 3차원 객체를 실시간으로 재현하기 위한 cGANs-디코더를 구현하며, 수신측에서 3차원 객체를 재현할 때 송신측에서 보낸 3차원 객체의 특징정보만을 입력받아 임의생성 됨으로 인해 발생하는 원본과의 차이에 대한 인지도의 정량화를 위해 유사도(level of Similarity)와 디테일의 정도(level of detail)에 따른 인간의 인지 변화도와 반응관계에 대해 연구한다.

3D 포인트 클라우드 형태로 생성된 방대한 양의 홀로그램 동영상 데이터를 네트워크를 통해 실시간으로 전송하기 위해서는 드라마틱한 압축 기술과 강력한 하드웨어 성능을 요구한다. 3D 포인트 클라우드 데이터의 용량을 줄이기 위해 포인트 클라우드 라이브러리를 사용하는 방법 등이 연구되었다[3]. 3D 콘텐츠를 실시간으로 전송하기 위한 방법으로 무손실 코딩기법이 제안되었고[4] MPEG-4를 이용한 홀로그램 콘텐츠의 압축 방법으로 h.264 기반 3차원 콘텐츠의 압축에 관련한 기

술 관련 연구가 다 방면에서 수행중이며 전송기술 관련 연구 또한 관심이 높은 영역이다[5]. 현재 MPEG을 중심으로 3차원 콘텐츠의 손실 압축 모델에 관한 연구가 진행되고 있다[6]. 가급적 원본의 형태 그대로 전송하는 것이 중요하겠지만 컴퓨팅 성능 및 네트워크 전송 이슈 등 제반 기술의 한계에 직면해 있고, 응용관점에서 본다면 원본 콘텐츠의 최소 정보만을 이용한 재생산된 콘텐츠의 표현은 가상휴면과의 상호작용과 co-presence의 영향 평가 측면에서 가치가 있다고 할 수 있다.

## III. cGANs기반 3D 전송 시스템

텔레프레즌스 서비스를 위해 방대한 양의 3D 포인트 클라우드 데이터로 이루어진 3차원 객체의 움직임을 실시간으로 모바일기기 등의 원격지에 전송하는 것은 네트워크 대역의 한계와 컴퓨팅자원의 한계로 인해 원본 그대로 지연 없이 보내는 것은 현재로서는 불가능하다. 따라서 데이터의 실시간 전송을 위해 원본데이터의 손실 압축 방법 등 다양한 방법이 연구될 수 있다. 본 연구에서는 생성된 포인트 클라우드의 움직임 정보와 얼굴 표정의 특징만을 추출하여 네트워크를 통해 전송하는 전송모형을 설계한다. 스케치 정보를 주면 그림을 완성시켜주는 PIX2PIX 알고리즘과 같은 GANs모델의 특징을 이용하여 최소정보만을 추출하여 전송한 후 수신측에서 복원시켜줄 수 있는 Conditional-GAN 모델을 설계/구현한다. 2014년 처음 GANs가 발표된 이후 Conditional-GAN, InfoGAN, f-GAN, Wasserstein GAN, DCGAN, BEGAN, Cycle GAN, DiscoGAN, EBGAN 등등 많은 응용모델들이 발표되었다. Conditional-GAN은 이미지 생성과정에서 컬러, 형태, 기타 특징들을 부가적으로 제어하여 원하는 형태의 이미지 생성이 가능하도록 해준다[7].

그림1은 모션캡처시스템(motion capture system)을 통해 캡처되는 3차원 객체 정보가 모바일 기기에 실시간으로 전송되는 시스템의 개요를 보여준다. 제안하는 텔레프레즌스 시스템을 위한 cGANs(conditional Generative Adversarial Networks) 기반 3D 포인트 클라우드 데이터의 실시간 전송 기법 연구를 통해 3차원 객체를 전송하게 되면 원본이미지와 복원된 이미지가 상이하게 표출된다. 상이하게 표출되는 정도는 설문조사를 통한 정성적 방법, physiological 정량적 방법에 의해 1단계부터

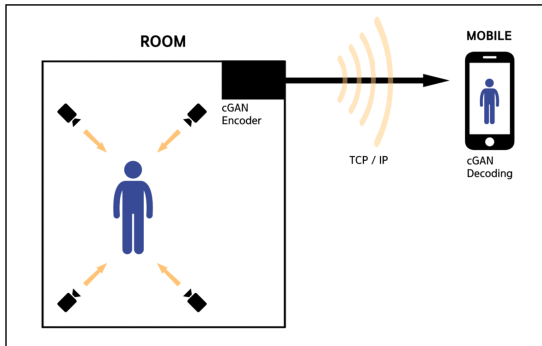


Fig. 1 Overview of creating and transmitting 3D point-cloud data

10단계 사이의 레벨로 구분할 수 있다. 이런 경우 두 가지 가설 설정이 가능한데 첫째, 상이함의 단계 별로 인지 수준이 1단계부터 10단계 사이에 uncanny valley가 존재한다는 가설과 둘째, 원본과 복원본이 서로 상이한 정도가 적어지는 특정 시점 이후부터는 saturation이 발생하여 레벨을 높여도 품질 대비 만족도는 변하지 않는다는 가설이다. 원본과 복원본의 상이함의 정도를 나타내는 요소로서 Visual Realism, co-presence, emotional response 등이 있는데 각 레벨을 구분하고 측정하기 위해 설문조사를 이용한 정량적인 방법과 정성적 방법에 따라 측정기준을 마련한다. 기 수행된 버추얼 휴먼(virtual human)과의 상호작용에 관한 연구에서 사용하였던 평가 시스템을 차용해서 복원데이터를 정량/정성적으로 평가하도록 한다[8].

#### IV. 결론

텔레프레즌스 기술은 방송, 교육, 의료 등 그 응용분야가 다양한 기술로써 미래의 유망한 기술 및 산업영역으로 기대를 모으고 있다. 본 연구는 어려운 제반 여건에서도 여러 가지 가능성을 제시한다. 예상되는 연구 결과는 관련 기술 수준과 상관없이 3D 콘텐츠 제작 및 전송 측면에서 다양한 활용이 가능하다. 연구결과의 활용성 측면에서 중요도를 판단한다면 일부 한정된 정보만으로 전체 정보를 복원할 수 있는 길이 열리기 때문에 개인의 프라이버시를 보호하는데 유리하고, 콘텐츠 생성 설비의 단순화와 대역폭에 덜 민감한 전송시스템 구축이 가능해져 비용절감 효과가 크다. 또한 가상 휴먼

기반의 텔레프레즌스 및 텔레컨퍼런스 모델에서 가상 객체를 통한 인간의 상호작용 연구가 가능하다. 본 연구의 결과는 기존의 교육, 문화/예술공연, 의료 분야 등 산업 분야에서 AR/VR의 시장 규모를 훨씬 넘는 수익이 창출될 것으로 기대한다.

#### ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT). (No. NRF- 2019R1G1A1087290)

#### REFERENCES

- [ 1 ] Computer Hope - telepresence [Internet]. Available: <https://www.computerhope.com/jargon/t/telepresence.htm>.
- [ 2 ] Medium Medium : new farmer - 3D points clouds for immersive real estate and telepresence experiences [Internet]. Available: <https://medium.com/new-farmer/3d-points-clouds-for-immersive-real-estate-and-telepresence-experiences-5cbdb03898b>.
- [ 3 ] R. B. Rusu, and S. Cousins, (2011, May). "Point cloud library (pcl)," In *2011 IEEE International Conference on Robotics and Automation* (pp. 1-4).
- [ 4 ] Y. H. Seo, H. J. Choi, and D. W. Kim, (2006). "Lossy coding technique for digital holographic signal," *Optical Engineering*, 45(6), 065802.
- [ 5 ] Peixeiro, J. P. Kim, C. Brites, J. Ascenso, and F. Pereira, (2018). "Holographic data coding: Benchmarking and extending hevc with adapted transforms," *IEEE Transactions on Multimedia*, 20(2), 282-297.
- [ 6 ] E. Darakis, and T. J. Naughton, (2009, May). "Compression of digital hologram sequences using MPEG-4," In *Holography Advances and Modern Trend, International Society for Optics and Photonics.s* (vol. 7358, pp. 735811).
- [ 7 ] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," *Berkeley AI Research (BAIR) Laboratory University of California, Berkeley*, 2016.
- [ 8 ] K. S. Shin, and D. S. Jo, "Exploring the Effects of the Virtual Human with Physicality on Co-presence and Emotional Response," *Journal of the Korea Society of Computer and Information*, 24.1 (2019): 67-71.