

5G 기반의 메타버스 모션 동기화 시스템의 설계 및 실험*

이 상 윤** · 이 대 식*** · 유 영 모**** · 유 현 수***** · 이상구*****

The Design and experiment of 5G-based metaverse motion synchronization system

Lee Sangyoon · Lee Daesik · You Youngmo · You Hyeonsoo · Lee Sangku

〈Abstract〉

In this paper, we design and experiment a 5G-based metaverse motion synchronization system with configuration of a mobile motion capture studio that has not been commercialized at home and abroad. As a result of the experiment, the average value of the latency test measurement using Wi-Fi is 0.134 seconds faster than the average latency test measurement value using the 5G network. Existing motion capture studios have spatial limitations as the motion capture range is limited to the Wi-Fi communication range. However, the 5G-based metaverse motion synchronization system configures a mobile motion capture studio so that motion performers can solve the spatial limitations by expanding the motion capture communication range indefinitely regardless of time and place. Therefore, it is possible to implement realistic metaverse contents by displaying a realistic and natural digital human because it is free from spatial constraints. The system which was tested in this paper can create a new business model by converging next-generation technologies that are receiving attention related to the digital virtual world, such as motion capture + 5G + digital human twin + metaverse. And it allows for research and develop a next-generation metaverse-based broadcasting solution at a recent time when the business value of digital human and metaverse technologies and functions has been proven and related sales are growing in earnest.

Key Words : 5G, Wi-Fi, Metaverse Contents, Digital Human, Motion Synchronization

I. 서론

산업발달 및 온라인 및 비대면 환경에 대한 요구가 증가되고 있으며, 가상현실과 실생활이 융합된 이른바 메타버스 콘텐츠에 대한 기대감이 확산되고 있다 [1]. 이러한 메타버스 콘텐츠는 다양한 콘텐츠를 비대면 환경으로 제공함으로써 시간 및 공간을 초월하여 많은 정보를 단시간에 제공할 수 있다는 면에서 큰 장점을 가지고 있다.

기존의 메타버스 콘텐츠는 주로 특정 콘텐츠를 대상으로 하고 사용자에게 일방적인 콘텐츠를 전달한다.

* 본 논문은 2023년도 과학기술정보통신부의 재원으로 정보통신평가원의 지원을 받아 수행된 연구임

(No.2022-0-00599, 5G 이동형 모션 트래킹 스튜디오 기반의 메타버스 방송 플랫폼 개발)

** (주) 트라이콤포텍 수석부장(주저자)

*** (주) 트라이콤포텍 연구소장(교신저자)

**** (주) 트라이콤포텍 책임연구원

***** (주) 트라이콤포텍 연구원

***** (주) 트라이콤포텍 대표이사

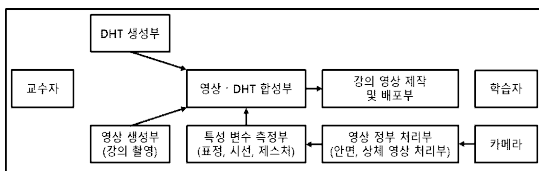
사용자의 다양한 참여가 가능하도록 가상현실 기술을 기반으로 디지털 휴먼을 생성하고, 이를 활용하는 기술로 사람의 역할을 대신하는 디지털 휴먼 기술에 있어서 어떻게 하면 인간과 실제로 비슷한 동작, 형태를 수행하는지에 대한 연구가 이루어지고 있다[3-6].

기존의 기술은 대상체와 친밀도 및 나이, 얼굴의 움직임, 바람 및 외부 힘을 고려하여 디지털 휴먼 영상을 생성함으로써 보다 정교하면서도 실감나는 디지털 휴먼 영상을 형성할 수 있도록 한다. 또한 대상체의 실시간 얼굴 형상 및 실시간 자세를 고려하여 정교하면서도 실감나는 디지털 휴먼 영상을 형성할 수 있다는 장점이 있으나, 고속의 데이터 전송 및 안정적인 데이터 전송이 이루어지지 않는 경우에 모션 동기화가 제대로 이루어지지 않으므로 메타버스 환경에서 디지털휴먼 영상이 어색하게 표현되는 문제가 있다[7, 8].

따라서 본 논문에서는 자연스러운 메타버스 콘텐츠를 구현하기 위한 디지털 휴먼의 모션이 동시에 이루어지는 경우 최적의 시스템 자원으로 각각의 모션을 동기화하는 5G 기반의 메타버스 모션 동기화 시스템을 제공하고자 한다.

II. 관련 연구

교감형 디지털 휴먼 튜터 생성 방법 및 시스템은 <그림 1>과 같다.



<그림 1> 교감형 디지털 휴먼 튜터 생성 방법 및 시스템[10]

<그림 1>은 교감형 디지털 휴먼 튜터 (DHT)의 생성 방법 및 시스템에 관한 것이다. 디지털 휴먼 튜터의 생성 방법은 강의용 영상 자료와 디지털 휴먼 튜터 모델을 생성하는 단계, 강의용 영상 자료와 DHT를 합성하여 강의용 영상 콘텐츠를 형성하는 단계, 영상 콘텐츠를 학습자에게 시현 하면서 학습자의 안면을 포함하는 인체 영상을 획득하는 단계, 인체 영상으로부터 학습자의 안면의 물리적 움직임에 대한 특징 정보를 추출하는 단계, 그리고 물리적 특징의 변화를 상기 DHT에 적용하여 학습자의 반응을 DHT의 외형에 모방시키는 단계를 포함한다.

디지털 휴먼 영상을 생성하고, 전자 장치에 저장하여 컴퓨터 관독 가능한 기록 매체에 저장된 프로그램의 순서는 <그림 1>과 같다.

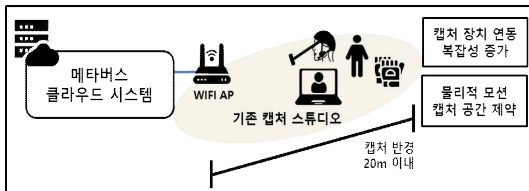


<그림 2> 디지털 휴먼 영상을 컴퓨터 관독 가능한 기록 매체에 저장된 프로그램[11]

<그림 2>에서 보면 디지털 휴먼 영상 형성을 위한 전자 장치는 대상체의 이미지를 획득하고, 대상체로부터 디지털 휴먼 영상의 얼굴 이미지 형성을 위한 제 1 특징점과 모발 이미지 형성을 위한 제 2 특징점을 추출하고, 대상체와 디지털 휴먼 영상과의 친밀도

및 대상체의 나이를 고려하여 제 1 특징점에 제 1 가중치를 부여하고, 대상체의 움직임, 바람 및 대상체 외부에서 작용하는 힘을 고려하여 제 2 특징점에 제 2 가중치를 부여하고, 대상체의 실시간 얼굴 형상 및 실시간 자세를 고려하여 제 1 특징점 및 제 2 특징점에 제3 가중치를 부여하고, 제 1과 3 가중치에 반영된 제 1 특징점과 제 2와 3 가중치에 반영된 제 2 특징점으로 대상체와 일치도가 가장 높은 대역 모델을 선택한다. 대역 모델을 3차원으로 스캐닝하고, 다양한 각도에서의 이미지를 촬영으로. 제 1과 3차원 데이터를 형성한다. 제 1 영상은 제 1과 3차원 데이터를 이용하고, 제 2와 3차원 데이터는 설정된 표정 템플릿을 웹 카메라로 촬영한다. 제 2 영상은 제 2와 3차원 데이터를 이용하여 대상체의 표정을 형성하고, 3차원 애니메이션 포인트를 이용하여 제 1과 2 영상을 합성하여 디지털 휴먼 영상을 형성한다.

Wi-Fi[10]를 사용한 기존의 모션 캡처 스튜디오 시스템 구조는 <그림 3>과 같다.



<그림 3> 기존의 모션 캡처 스튜디오 시스템 구조

<그림 3>에서 보면 디 모션 캡처 장치 연동은 상용화된 모든 얼굴 및 몸 모션 캡처 장치와 모션 캡처 스튜디오는 동일한 Wi-Fi 망에서 연동되어야 하고 다른 Wi-Fi망으로 전환 시 초기 설정에 대한 복잡성이 높아 작업 효율성을 저하시킨다. 또한 얼굴 및 몸 모션 캡처 장치는 Wi-Fi 및 유선 연결을 기반으로 모션 캡처 스튜디오와 연동됨으로 인한 공간제약이 발생시킨다.

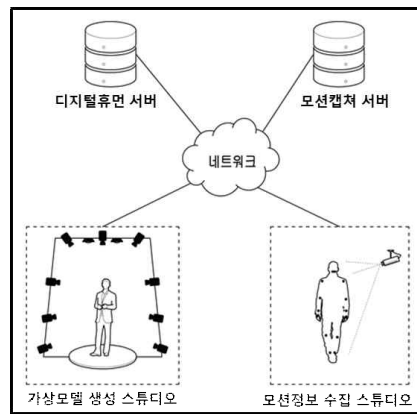
<그림 1>, <그림 2>, <그림 3>에서 보면 기존의 기술은 대상체와의 친밀도 및 나이, 얼굴의 움직임, 바

람 및 외부 힘을 고려하여 디지털 휴먼 영상을 생성함으로써 보다 정교하면서도 실감나는 디지털 휴먼 영상을 형성할 수 있도록 한다. 또한 대상체의 실시간 얼굴 형상과 실시간 자세로 정교하고 실감나는 디지털 휴먼 영상을 형성할 수 있는 장점이 있다.

그러나 고속의 데이터 전송 및 안정적인 데이터 전송이 이루어지지 않는 경우에 모션 동기화가 제대로 이루어지지 않으므로 메타버스 환경에서 디지털휴먼 영상이 어색하게 표현되는 문제가 있다. 특히 실시간 얼굴 형상 및 자세를 고려한 가중치를 특징점에 부여할 뿐, 얼굴이나 몸의 움직임이 동시에 이루어지는 경우 각각의 모션정보의 수집 및 전달 시간에 따른 차이를 전혀 고려하지 않으므로 실제 구현되는 대상체의 현실감이 떨어지는 문제가 있다. 따라서 아직 구현되지 않은 5G로 이동형 모션캡처 스튜디오 시스템을 설계하여 5G 기반의 메타버스 모션 동기화 시스템의 설계하고 지연시간을 실험한다.

III. 5G 기반의 메타버스 모션동기화 시스템

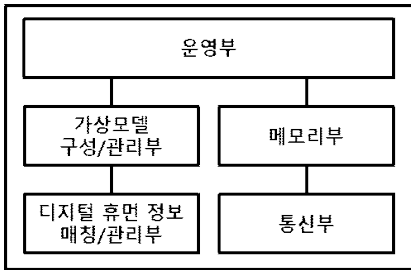
5G 기반의 메타버스 모션동기화 시스템의 구성도는 <그림 4>와 같다.



<그림 4> 5G 기반의 메타버스 모션동기화 시스템 구성도

<그림 4>에서 보면 디지털휴먼 서버 및 디지털휴먼 서버와 5G 네트워크 연결된 모션캡처 서버를 포함한다. 디지털휴먼 서버는 미리 설정된 가상모델 구성 방법에 기초하여 가상모델을 구성하며 모션캡처 서버로부터 디지털휴먼 모션정보를 전달받는 기능 및 디지털휴먼 모션정보를 미리 설정된 모션동기화 방법에 기초하여 가상모델과 매칭하여 디지털휴먼 정보를 생성하는 기능을 수행한다.

디지털휴먼 서버의 구성도는 <그림 5>와 같다.

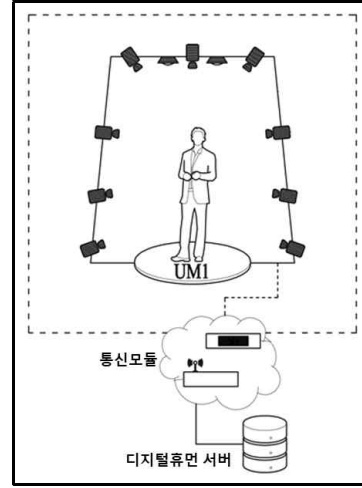


<그림 5> 디지털휴먼 서버의 구성도

<그림 5>에서 보면 메모리부는 가상모델 생성 스튜디오의 사용자 이미지 정보, 모션정보 수집 스튜디오의 모션정보, 가상모델 구성 및 디지털휴먼 정보 생성 과정에서 발생된 데이터를 저장한다. 통신부는 통신모듈을 포함하고, 가상모델의 구성 및 관리가 이루어지도록 하는 가상모델 구성/관리부, 디지털휴먼 정보의 생성, 동기화 및 관리가 이뤄지도록 하는 디지털휴먼 정보 동기화/관리부 및 서버의 전반적인 운영이 이뤄지도록 하는 운영부를 포함한다.

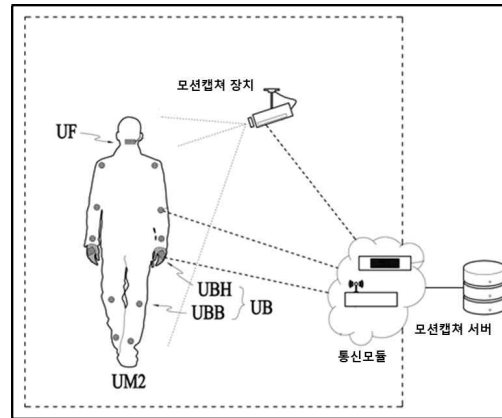
디지털휴먼 서버와 가상모델 생성 스튜디오를 연결하는 네트워크는 구성도는 <그림 6>과 같다.

<그림 6>에서 보면 전신 캡처 장치는 복수 개의 카메라와 조명으로 구성되며, 복수 개의 카메라 및 조명을 제어하고 사용자 이미지 정보를 획득하여 네트워크를 통해 디지털휴먼 서버로 전달하는 기능을 수행하는 통신모듈로 구성된다.



<그림 6> 디지털휴먼 서버와 네트워크 연결 구성도

모션캡처 서버는 모션정보 수집 스튜디오의 모션캡처 장치를 통해 사용자 모델(UM)의 모션정보를 수집할 수 있다. 모션캡처 장치의 구성 및 모션캡처 서버와의 네트워크 연결을 나타내는 구성도는 <그림 7>과 같다.



<그림 7> 모션캡처 서버와 네트워크 연결 구성도

<그림 7>에서 보면 모션정보 수집 스튜디오는 모션캡처 장치를 포함하며 모션캡처 장치를 제어하고 사용자 모델의 얼굴의 움직임을 나타내는 페이스 모

선정보(UF)와 사용자 모델의 손이나 손을 제외한 몸의 움직임을 나타내는 바디 모션정보(UBB, UBH, UB) 중 어느 하나의 정보를 획득하여 네트워크를 통해 모션캡처 서버로 전달하는 기능을 수행하는 통신 모듈을 포함한다. 통신모듈은 5G 네트워크를 거쳐 캡처 서버로 전달된다.

모션캡처 장치는 카메라 형태로 이루어져서 사용자 모델의 신체 전체를 촬영하면서 바디와 얼굴의 움직임을 각각 구분하여 저장, 인식할 수 있도록 구성한다.

바디 모션정보(UBB)를 수집하는 모션캡처 장치는 16개 이상의 마그네틱 센서가 사용자 모델의 관절에 해당하는 위치에 구비되어 각각의 센서가 사용자 모델의 바디모션을 캡처하도록 구성된 스마트 슈트로 이루어진다.

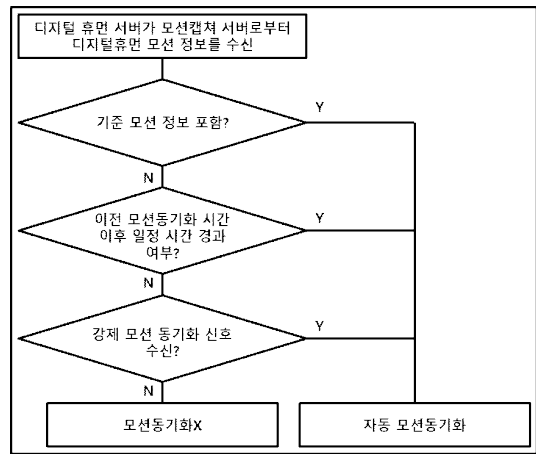
손의 움직임에 따른 바디 모션정보(UBH)를 수집하는 모션캡처 장치는 24개 이상의 마그네틱 센서가 사용자 모델의 손가락 관절에 해당하는 위치에 구비되어 각각의 센서가 사용자 모델의 손의 움직임을 나타내는 바디 모션을 캡처하도록 구성된 스마트 글로브로 이루어진다.

모션캡처 서버는 이러한 특징점의 변화를 통해 사용자 모델의 페이스 모션정보(UF)를 수집할 수 있다. 또한 마이크 및 녹음기능이 구비되어 사용자 모델의 음성정보(UV)를 수집하여 모션캡처 서버로 전달한다. 사용자 모델의 음성정보는 사용자 모델의 음색이나 톤 또는 말투 정보를 포함하고, 특정 단어로 구성되어 메타버스 모션동기화를 위한 기준 모션정보를 구성한다.

모션캡처 서버는 모션캡처 장치와 네트워크 연결되어 사용자 모델의 움직임 정보로서 얼굴의 움직임을 나타내는 페이스 모션정보와 바디 및 손의 움직임을 나타내는 바디 모션정보와 음성정보를 포함하는 디지털휴먼 모션정보를 수집하고, 이를 네트워크를 통해 디지털휴먼 서버로 전달한다.

네트워크는 기존의 저속 네트워크망으로 구성되면 디지털휴먼 정보의 데이터와 가상현실 콘텐츠 정보를 전달하는 과정에서 데이터 손실이나 시간 지연이 발생 가능성이 높고, 대용량의 데이터 전송이 이루어질 수 있도록 한다는 점에서 5G 기반의 네트워크로 구성한다.

5G 기반의 메타버스 모션동기화 방법의 순서도는 <그림 8>과 같다.



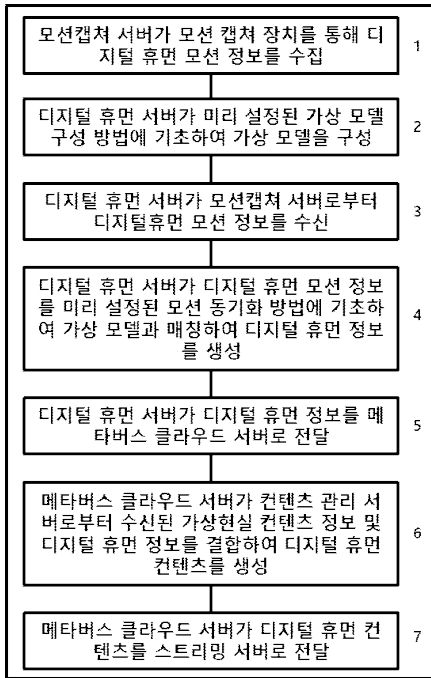
<그림8> 메타버스 모션동기화 방법의 순서도

<그림 8>에서 보면 디지털휴먼 서버가 모션캡처 서버로부터 디지털휴먼 모션정보를 수신하는 경우 우선 기준 모션정보가 포함되어 있는지 여부를 판단하고, 기준 모션정보가 포함된 경우 자동으로 모션동기화가 이루어지도록 하고 그렇지 않은 경우에는 이전 모션동기화 시간 이후 일정 시간이 경과하였는지 여부를 판단한다. 만일 이전 모션동기화 시간 이후 일정 시간이 경과하였다면 자동으로 모션동기화가 이루어지도록 하고 그렇지 않은 경우에는 강제 모션동기화 신호가 수신되었는지를 판단한다.

강제 모션동기화 신호는 강제적으로 모션동기화가 이루어져야 하는 상황에서 해당 신호가 전달되는 경우 자동으로 모션동기화가 이루어지도록 함으로써

디지털휴먼 정보가 자연스럽게 구성한다. 예를 들면 디지털휴먼 모션정보를 전달하는 네트워크의 상황이 변경되는 경우에 모션캡처 서버로부터 자동으로 강제 모션동기화 신호가 디지털휴먼 서버로 전달되도록 구성한다. 디지털휴먼 서버는 강제 모션동기화 신호가 수신되는 경우 제1 시간 및 제2 시간의 시간차이를 반영하여 자동으로 모션동기화가 이루어지도록 하고 그렇지 않은 경우에는 모션동기화를 자동으로 수행하지 않도록 구성한다.

5G 기반의 메타버스 모션동기화 방법에 따라 생성되는 디지털휴먼 콘텐츠의 생성에 관한 순서도는 <그림 9>와 같다.



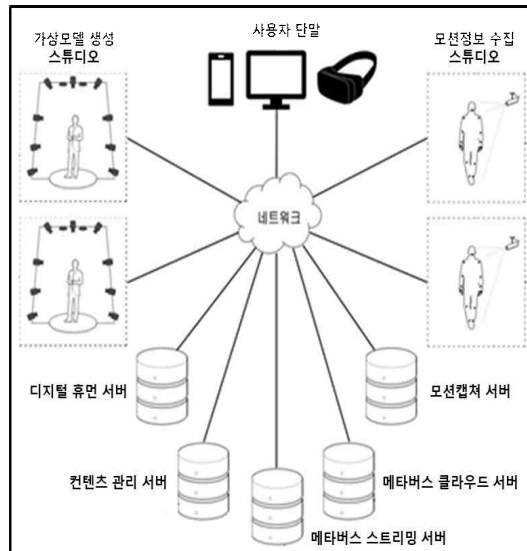
<그림 9> 디지털휴먼 콘텐츠의 생성에 관한 순서도

<그림 9>에서 보면 5G 기반의 디지털휴먼 콘텐츠의 생성 방법은 모션캡처 서버가 디지털휴먼 모션정보를 수집하는 단계 1, 디지털휴먼 서버가 미리 설정

된 가상모델 구성방법에 기초하여 가상모델을 구성하는 단계 2, 디지털휴먼 서버가 모션캡처 서버로부터 디지털휴먼 모션정보를 수신하는 단계 3, 디지털휴먼 서버가 디지털휴먼 모션정보를 가상모델과 매칭하여 디지털휴먼 정보를 생성하는 단계 4를 포함한다.

다음으로 클라우드 서버에 디지털휴먼 정보를 전달하는 단계 5와 디지털휴먼 콘텐츠를 생성하기 위한 메타버스 클라우드 서버가 콘텐츠관리 서버로부터 수신된 디지털휴먼 정보와 가상현실 콘텐츠 정보를 결합하여 디지털휴먼 콘텐츠를 생성하는 단계 6을 포함한다. 메타버스 클라우드 서버가 디지털휴먼 콘텐츠를 메타버스 스트리밍 서버로 전달 단계 7, 사용자 단말에서 네트워크를 통해 디지털휴먼 콘텐츠의 실시간 스트리밍이 표시된다.

5G 기반의 메타버스 모션동기화 시스템의 구성도는 <그림 10>과 같다.



<그림 10> 5G 기반의 메타버스 모션동기화 시스템의 구성도

<그림 10>에서 보면 디지털휴먼 서버는 복수의 가

상모델 생성 스튜디오를 통해 각각 가상모델을 구성하며, 모션캡처 서버는 복수의 모션정보 수집 스튜디오를 통해 각각 사용자별 모션정보를 수집한다.

<그림 9>의 6 단계(디지털휴먼 콘텐츠를 생성하는 단계)에서는 4 단계(디지털휴먼 정보를 생성하는 단계)에서 생성된 디지털휴먼 정보와 콘텐츠관리 서버로부터 수신된 가상현실 콘텐츠 정보의 결합시 메타버스 모션동기화 방법과 유사하게 디지털휴먼 정보와 가상현실 콘텐츠 정보의 동기화가 이루어진다. 즉 데이터의 수신 시간의 차이에 따른 모션동기화 설정, 기준 모션정보의 포함여부나 모션동기화 시간이나 강제 모션동기화 신호 수신에 따른 자동 모션동기화 및 네트워크 환경 변화에 따른 미리 설정된 가중치가 반영된 사전 모션동기화 등의 방법이 디지털휴먼 콘텐츠 생성과정에서 이용될 수 있으며, 더욱 현장감 있는 실시간 디지털휴먼 콘텐츠 스트리밍이 이루어진다.

메타버스 클라우드 서버는 콘텐츠관리 서버로부터 수신되는 가상현실 콘텐츠 정보를 콘텐츠관리 서버에서 송신하는 시간과 메타버스 클라우드 서버에서 수신하는 시간 차이를 통해 가상현실 콘텐츠 정보에 대한 지연시간 정보를 획득한다. 다음으로 디지털휴먼 정보를 디지털휴먼 서버에서 송신하는 시간과 메타버스 클라우드 서버에서 수신하는 시간 차이를 통해 디지털휴먼 정보에 대한 지연시간 정보를 획득한다.

메타버스 클라우드 서버는 가상현실 콘텐츠 정보에 대한 지연시간 정보와 디지털휴먼 정보에 대한 지연시간 정보를 반영하여 가상현실 콘텐츠 정보와 디지털휴먼 정보의 결합시 모션동기화가 이루어지는 디지털휴먼 콘텐츠를 생성한다.

메타버스 스트리밍 서버는 복수의 사용자별 디지털휴먼 모션정보가 반영된 복수의 사용자별 디지털휴먼 콘텐츠를 메타버스 클라우드 서버로부터 수신하고 이를 사용자 단말에 각각 전달할 때 복수의 사용자별 디지털휴먼 콘텐츠가 하나의 메타버스 공간

상에 표시되는 과정에서 서로간의 자연스러운 동작 이루어질 수 있도록 사용자별 모션동기화를 수행할 수 있다. 복수의 사용자별 디지털휴먼 콘텐츠의 스트리밍에서의 사용자별 모션동기화는 각각의 사용자별 디지털휴먼 콘텐츠를 메타버스 클라우드 서버에서 송신한 시간과 메타버스 스트리밍 서버에서 이를 수신한 시간의 차이에 기초하여 이루어진다.

메타버스 모션동기화는 네트워크를 통한 정보전달 과정에서 오는 시간 차이를 고려하고, 디지털휴먼 모션정보가 수집되고 가상모델의 생성, 가상모델과 디지털휴먼 모션정보가 결합하여 디지털휴먼 정보를 생성하는 과정에서, 데이터 전송과 텍스처링 및 리깅 등의 후처리가 필요하므로 복수개의 모션이 동시에 이루어지는 경우 이를 동기화하여 정확한 매칭이 이루어진다.

미리 설정된 모션동기화 방법은 모션캡처 서버가 모션캡처 장치로부터 디지털휴먼 모션정보를 수집한 제1 시간 및 디지털휴먼 서버가 디지털휴먼 모션정보를 전달받은 제2 시간의 차이에 기초하여 동기화가 이루어진다. 디지털휴먼 정보를 생성함에 있어서 5G 네트워크를 이용하는 경우 대용량의 데이터를 전송하더라도 전송시간이 매우 짧으므로 디지털휴먼 모션정보가 빠르게 가상모델과 매칭될 수 있다. 그러나 사용자 모델의 이미지나 센서의 수가 많아 디지털휴먼 정보를 구성하는 디지털휴먼 모션정보의 데이터 크기가 매우 크거나 5G 네트워크에 문제가 발생하여 원활한 데이터 전송이 이루어지지 않는 경우 또는 5G 네트워크를 대체하여 사용 가능한 Wifi 등의 무선 네트워크의 속도가 늦거나 시간당 데이터 처리량 또는 전송 용량이 매우 적은 경우에는 디지털휴먼 모션정보의 전달이 지연됨에 따라 가상모델과 디지털휴먼 모션정보의 매칭시 현실감이 떨어지는 디지털휴먼 정보의 구성이 이루어진다.

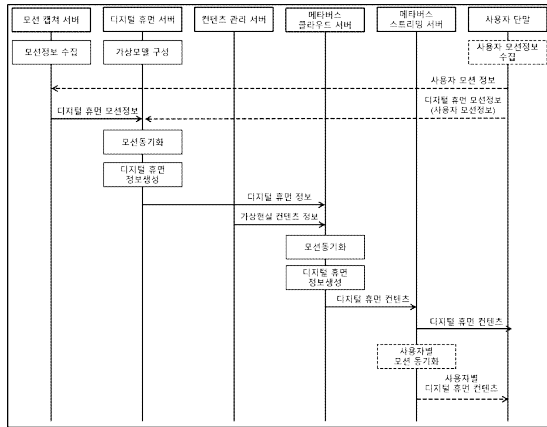
따라서 디지털휴먼 모션정보가 모션캡처 서버에서 수집되는 시간(제1 시간)과 디지털휴먼 모션정보가

디지털휴먼 서버에서 수집되는 시간(제2 시간)의 차이를 측정하고 이러한 시간 차이를 가상모델과 매칭시 고려하여 매칭함으로써 끊김을 최소화하는 자연스러운 디지털휴먼 정보가 생성된다.

디지털휴먼 서버는 모션캡처 서버로부터 디지털휴먼 모션정보와 함께 모션캡처 서버에서 수집 및 전송하는 디지털휴먼 모션정보의 수집 및 전송시간 정보를 전달받을 수 있으며, 하나의 모션 다음 모션 사이나 새로운 모션의 시작시에 해당 시간 차이를 반영하여 가상모델과 매칭되도록 함으로써 정확한 메타버스 모션동기화가 이루어진다.

IV. 5G 기반의 디지털 휴먼 콘텐츠 스트리밍 구성도

5G 기반의 메타버스 모션 동기화 시스템에서 이루어지는 디지털 휴먼 콘텐츠 스트리밍 구성도는 <그림 11>과 같다.



<그림 11> 디지털 휴먼 콘텐츠 스트리밍 구성도

<그림 11>에서 보면 메타버스 메타버스 동기화 시스템은 사용자 단말로부터 각각의 사용자의 디지털

휴먼 모션정보를 수집하고, 디지털휴먼 서버로 직접 전달하거나 사용자 단말로부터 수집된 각각의 사용자의 디지털휴먼 모션정보를 모션캡처 서버를 거쳐 디지털휴먼 서버로 전달하도록 구성된다.

사용자 단말은 각각의 사용자 모션정보를 수집할 수 있도록 하는 모션캡처 장치를 구비한다. 사용자 단말에 구비된 모션캡처 장치는 카메라 및 각종 센서이다. 모션캡처 장치를 통해 미리 설정된 사용자 그룹별로 사용자의 디지털휴먼 모션정보를 제공할 수 있으며, 표시수단을 통해 모션동기화가 이루어진 디지털휴먼 콘텐츠를 사용자별로 스트리밍한다.

메타버스 클라우드 서버는 콘텐츠관리 서버로부터 수신된 가상현실 콘텐츠 정보와 사용자 단말로부터 수신된 사용자별 모션정보와 결합하여 생성된 사용자별 디지털휴먼 정보를 서로 결합하여 사용자별 디지털휴먼 콘텐츠를 생성하고, 사용자별 디지털휴먼 콘텐츠를 메타버스 스트리밍 서버로 전달한다. 또한 메타버스 클라우드 서버는 가상현실 콘텐츠 정보와 사용자별 디지털휴먼 정보를 결합하는 과정에서 모션동기화를 통해 자연스러운 사용자별 디지털휴먼 콘텐츠를 생성한다. 사용자별 디지털휴먼 콘텐츠는 사용자별 디지털휴먼 정보에 가상현실 콘텐츠 정보가 결합된다. 예를들면 하나의 상황(모델 선발대회, 콘서트, 전시회, 교육 등)에 하나의 가상현실 콘텐츠 정보를 수신하고, 동일한 가상현실 콘텐츠 정보에 각각의 사용자별 디지털휴먼 정보를 위치 및 시간만 달리하여 배치하는 형태로 구성한다. 또한 각각의 가상현실 콘텐츠 정보가 각각의 사용자별 디지털휴먼 정보에 맞춤형으로 별도로 구성되어 각각 결합하도록 함으로써 사용자별로 독창적인 디지털휴먼 콘텐츠가 생성되도록 구성한다.

미리 설정된 사용자 그룹은 가상모델과 매칭하여 사용자별 디지털휴먼 정보를 생성하는 제1 사용자 모션정보를 제공하는 제1 사용자 그룹과 제1 사용자 그룹에서 제공되는 제1 사용자 모션정보의 변경이 가능

하도록 하는 제2 사용자 모션정보를 제공하는 제2 사용자 그룹을 포함한다.

디지털휴먼 서버는 제2 사용자 모션정보를 미리 설정된 범위 내에서 가상모델과 매칭하여 사용자별 디지털휴먼 정보를 생성한다.

제2 사용자 모션정보는 모션캡처 장치를 통해 수집된 제2 사용자 그룹에 포함된 제2 사용자의 움직임 정보이며 얼굴의 움직임을 나타내는 페이스모션정보와 바디 및 손의 움직임을 나타내는 바디모션정보 중 적어도 어느 하나를 포함한다.

사용자 그룹은 가상현실 내에서 다양한 종류 및 형태의 콘텐츠를 직접 또는 간접으로 경험하도록 하는 다양성을 제공하기 위해 등급 또는 단계의 형태로 구분한다. 가상현실 콘텐츠 정보는 모델 선발대회, 교육, 전시, 공연 등이다. 예를 들면 공연의 경우 공연장에서 직접 공연하는 공연자와 해당 공연장에서 공연을 관람하는 관람자를 필요로 하고, 모델 선발대회의 경우 직접 워킹 및 각종 퍼포먼스를 제공하는 다수의 모델과 사회자 및 관람자를 필요로 한다.

제1 사용자 그룹과 제2 사용자 그룹은 각각 공연자와 관람자이고, 각각 공연장에서 행하는 동작과 허용되는 움직임의 범위 등이 다르므로 가상현실 콘텐츠 정보와 결합하여 디지털휴먼 콘텐츠를 생성하기 위한 디지털휴먼 정보를 구성하는 모션정보에도 제한한다.

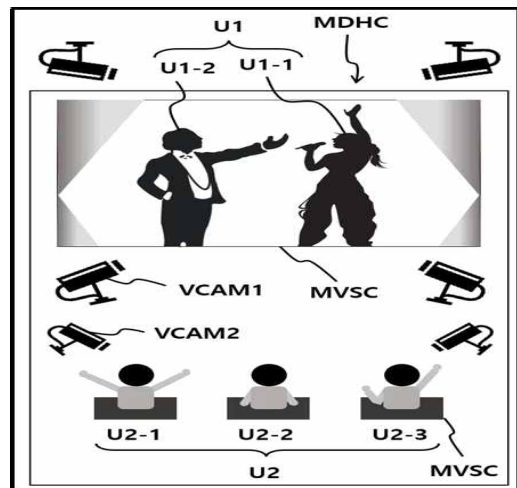
제1 사용자 그룹에 해당하는 사용자는 실제 공연을 행하는 사람이므로 사용자의 모든 모션정보를 그대로 제공하여 전술한 가상모델과 매칭하여 사용자별 디지털휴먼 정보를 생성한다. 이를 위해 모션정보 수집 스튜디오는 제1 사용자 그룹의 사용자가 직접 방문하여 사용자별 디지털휴먼 모션정보를 수집하고, 수집된 제1 사용자 그룹의 사용자별 디지털휴먼 모션정보를 모션캡처 서버를 거쳐 디지털휴먼 서버로 전달하도록 구성될 수도 있다.

제2 사용자 모션정보는 바디 움직임 중에서 상반

신에 대한 바디모션정보만을 포함하도록 구성한다. 제2 사용자 그룹에 해당하는 사용자는 공연이나 대회를 관람하는 사람이므로, 무대 아래에서 무대 위로의 이동을 방지하거나 일정한 크기 이상의 소리를 내지 못하도록 제한하는 것이 필요하며, 사용자의 일부 모션정보만을 제공하여 전술한 가상모델과 매칭하여 사용자별 디지털휴먼 정보를 생성한다.

예를 들면 제2 사용자 그룹에 해당하는 사용자가 제공하는 제2 사용자 모션정보는 박수를 치거나, 공연자를 응원하는 플래카드를 흔드는 행동이다. 이 때 사용자별 디지털휴먼 콘텐츠의 스트리밍이 실시간으로 이루어지는 라이브 공연의 경우 이러한 행동이 제1 사용자 그룹의 사용자에게 보여질 수 있도록 함으로써 제1 사용자 그룹에서 제공되는 제1 사용자 모션정보의 변경이 가능할 수 있으며, 공연자와 관람자간의 다양한 형태의 상호 교감을 이끌어낼 수 있도록 모션동기화를 통한 자연스러운 움직임이 이루어지도록 한다.

5G 기반의 메타버스 모션동기화 시스템에서 이뤄지는 디지털휴먼 콘텐츠 스트리밍의 모식도는 <그림 12>와 같다.



<그림 12> 디지털휴먼 콘텐츠 스트리밍의 모식도

<그림 12>와 같이 공연자(제1 사용자 그룹의 사용자)의 모습(디지털휴먼 정보, U1) 및 공연 무대(가상 현실 콘텐츠 정보, MVSC, Metaverse Virtual Space Contents)가 사용자별 디지털휴먼 콘텐츠(MDHC, Metaverse Digital Human Contents)로 스트리밍되도록 구성한다.

제2 사용자 그룹의 사용자 단말에 제공되는 사용자별 디지털휴먼 콘텐츠는 제2 사용자 그룹의 사용자가 공연 무대가 아닌 관람석에 앉아 있는 제2 사용자 그룹의 다른 사용자의 모습(디지털휴먼 정보, U2)일 수도 있다.

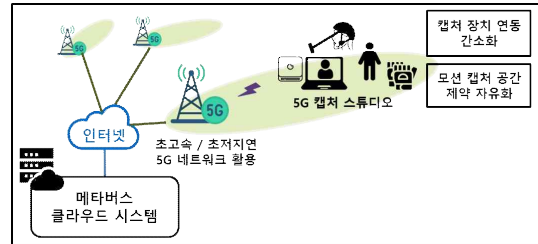
예를 들면 제1 사용자 그룹의 사용자 단말은 VR 장치 또는 2차원이나 3차원 표시수단을 포함하고, 공연자의 시점에서 바라보는 공연 무대나 관람자(제2 사용자 그룹의 사용자)의 모습이 사용자별 디지털휴먼 콘텐츠로 스트리밍되도록 구성한다. 또한 제1 사용자 그룹의 사용자 단말에서 제1 사용자의 선택에 따라 공연자인 자신의 모습이 어떻게 보여지는지를 확인할 수 있도록 가상카메라의 각각의 위치에서 바라보는 이미지/영상 정보를 선택하여 스트리밍하도록 구성한다.

모션동기화 시스템은 모션동기화에 기초하여 디지털휴먼 정보와 가상현실 콘텐츠 정보를 결합하되 디지털휴먼 정보와 가상현실 콘텐츠 정보의 자연스러운 결합이 이루어지도록 하는 모션동기화가 이루어진 디지털휴먼 콘텐츠를 생성한다. 따라서 5G 기반의 메타버스 모션 동기화 시스템의 설계 및 지연시간을 실험하였다.

V. 실험

본 논문에서는 국내외 상용화된 모션 캡처 장치 및 스튜디오 중에서 아직 5G로 설계되지 않은 시스템을 <그림 13>과 같이 설계하여 Wi-Fi를 사용한 기존의

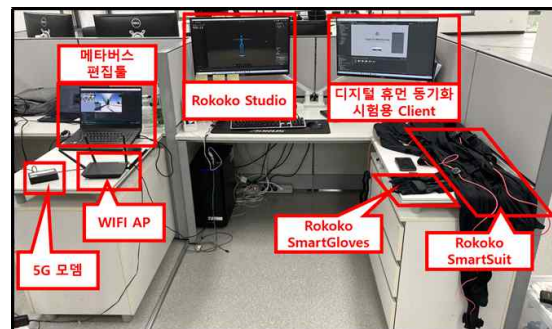
모션 캡처 스튜디오 시스템 구조인 <그림 3>과 지연 시간을 실험하였다.



<그림 13> 5G망 모션 캡처 스튜디오 시스템 구조

본 논문에서 시험 구성도는 모션 캡처 장비 (Rokoko SmartSuit, Rokoko Remote)들은 WIFI AP를 통해 모션 캡처 스튜디오와 연결되고, 모션 캡처 스튜디오의 클라이언트는 5G 모뎀을 통해 메타버스 클라우드 시스템에 접속되고 디지털 휴먼 모션 데이터를 전달한다. 메타버스 편집 툴은 메타버스 클라우드 시스템에 접속되며 메타버스 클라우드 시스템으로부터 모션 데이터를 전달받아 디지털 휴먼 모션과 동기화한다.

전체 시험 환경 구성도는 <그림 14>와 같다.



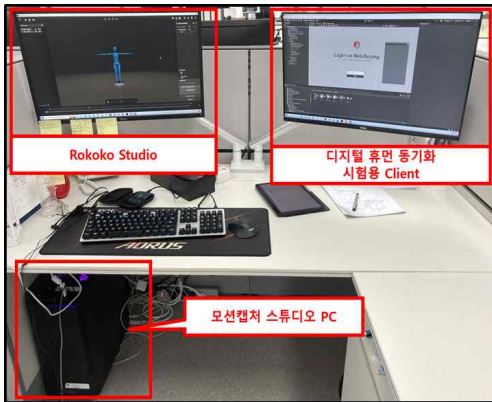
<그림 14> 전체 시험 환경 구성도

모션 캡처 장비는 <그림 15>와 같다.



<그림 15> 모션 캡처 장비

모션캡처 스튜디오는 <그림 16>과 같다.



<그림 16> 모션캡처 스튜디오

메타버스 편집 툴은 <그림 17>과 같다.



<그림 17> 메타버스 편집 툴

시험 장비 구성은 <표 1>과 같다.

<표 1> 시험 장비 구성

장비명	규격	수량	비고
Rokoko SmartSuit	마그네틱 센서 16개 WIFI 지원	1	몸 모션 캡처 모션 데이터 전송
Rokoko SmartGloves	마그네틱 센서 24개 WIFI 지원	1	손 모션 캡처 모션 데이터 전송
iPhone 13 Pro-Rokoko Remote	Apple A15 SoC, RAM 6GB, SL센서 (TrueDepth)	1	얼굴 모션 캡처
모션캡처스튜디오	Intel(R) Core(TM) i7-10700K, RAM 64GB, RTX 2070	1	RokokoStudio 시험용 Client
5G 모뎀	HE-900S, 5G NSA	1	모션데이터 전송용, 인터넷망 연결

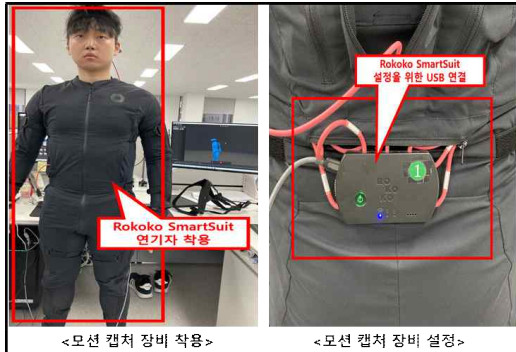
본 논문에서 시험 목적은 <표 2>와 같다.

<표 2> 시험 목적

시험 목적	요구 기능	시험 설명	판정 기준
연기자는 몸 동작 특정 모션 및 손 동작 특정 모션을 취하고 측정자는 연기자, 메타버스 편집 툴의 디지털 휴먼 트윈을 슬로우모션 비디오 촬영하고 녹화 영상을 플레이하여 동작 모션 동기화가 이루어지는지, 지연시간 발생 여부를 측정한다	1. 몸 모션 캡처 및 모션 데이터 전송 2. 손 모션 캡처 및 모션 데이터 전송 3. 디지털 휴먼 몸/손 리깅 동기화 4. 5G 모뎀 연동 데이터 전송	1. 동작 모션 동기화가 이루어지는지 확인한다. 2. 지연시간 발생여부를 측정한다.	1. 연기자는 몸 동작 특정 모션 및 손 동작 특정 모션을 취하고 측정자는 연기자, 모션 캡처 스튜디오의 더미 아바타를 슬로우모션 비디오 촬영 후 지연시간을 확인한다. 2. 측정 결과값은 5회 측정 후 평균값으로 한다.

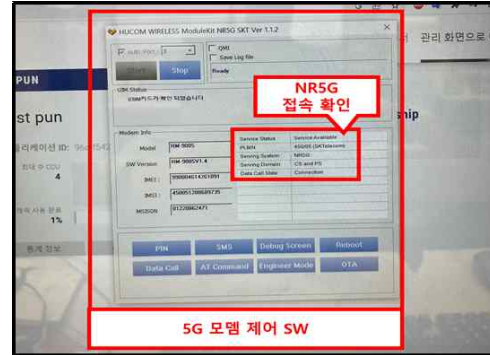
본 논문에서 시험 절차는 다음과 같다.

- (1) <그림 13>과 같이 시험환경을 구성한다.
- (2) 모션캡처 장치들을 WIFI-AP 무선망을 통해 모션캡처 스튜디오와 연동한다.

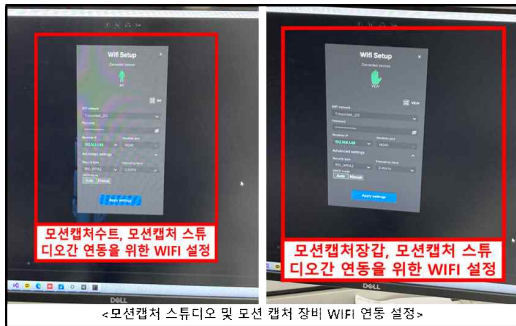


<모션 캡처 장비 착용>

<모션 캡처 장비 설정>



5G 모뎀 제어 SW



모션캡처스튜디오, 모션캡처 스튜디오간 연동을 위한 WIFI 설정

모션캡처장갑, 모션캡처 스튜디오간 연동을 위한 WIFI 설정

<모션캡처 스튜디오 및 모션 캡처 장비 WIFI 연동 설정>

(5) 연기자는 몸 동작 특정 모션 및 손 동작 특정 모션을 취하고 측정자는 연기자, 모션 캡처 스튜디오의 더미 아바타를 슬로우모션 비디오 촬영하고 녹화 영상을 플레이하여 동작 모션 동기화가 이루어지는 지, 지연시간 발생 여부를 측정한다.

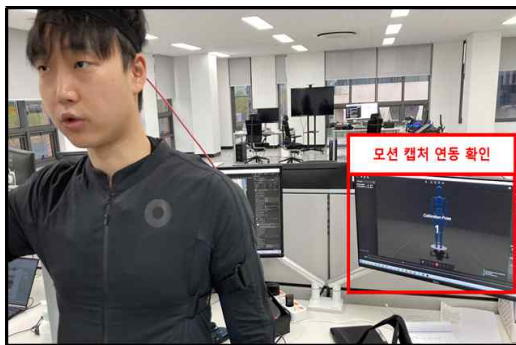
(3) 클라이언트를 실행하여 디지털 휴먼과 몸/손리깅 동기화를 확인한다.



Rokoko SmartSuit Rokoko Gloves 몸, 손 모션 캡처

메타버스 편집을 클라이언트 디지털 휴먼 모션 동기화 확인

지연시간 측정



모션 캡처 연동 확인

(4) 5G 모뎀을 통해 NR5G망에 접속 되어 있는지 확인한다.

<그림 3>와 같이 Wi-Fi를 사용한 기존의 모션 캡처 스튜디오 시스템을 5회 지연시간 시험 측정 평균값은 <표 3>과 같고, 본 논문에서 설계한 5G망을 사용한 모션 캡처 스튜디오 시스템을 5회 지연시간 시험 측정 평균값은 <표 4>와 같다.

<표 3> Wi-Fi를 사용한 지연시간 평균값
시간 단위 : 초(sec)

시험 횟수	시작 시각	종료 시각	콘텐츠 지연시간
1	02:08.54	02:08.85	0.31
2	02:49.23	02:49.50	0.27
3	03:39.65	03:39.89	0.24
4	04:11.58	04:11.83	0.25
5	04:42.47	04:42.76	0.29
평균 지연시간			0.272

<4> 5G망을 사용한 지연시간 평균값
시간 단위 : 초(sec)

시험 횟수	시작 시각	종료 시각	콘텐츠 지연시간
1	02:50.03	02:50.38	0.35
2	03:25.33	03:25.78	0.45
3	03:50.32	03:50.69	0.37
4	04:28.74	04:29.14	0.40
5	05:12.25	05:12.71	0.46
평균 지연시간			0.406

<표 3>과 <표 4>에서 보면 Wi-Fi를 사용한 지연시간 시험 측정 평균값이 5G망을 사용한 지연시간 시험 측정 평균값보다 0.134초 더 빠르다.

그러나 <그림 3>의 Wi-Fi를 사용한 기존의 시스템은 모션 캡처 장치 연동 설정 복잡성에 있어서 상용화된 모든 얼굴 및 몸 모션 캡처 장치와 모션 캡처 스튜디오는 동일한 Wi-Fi 망에서 연동되어야 하고 다른 Wi-Fi 망으로 전환 시 초기 설정에 대한 복잡성이 높아 작업 효율성을 저하시킨다. 모션 캡처 공간 제약에 있어서 얼굴 및 몸 모션 캡처 장치는 Wi-Fi 및 유선 연결을 기반으로 모션 캡처 스튜디오와 연동됨으로 인한 공간제약이 발생한다. 기능적 한계에 있어서 실시간 전송에 있어서 고품질 얼굴 영상을 스트리밍하는 과정에서 간헐적으로 품질저하(Jitter, Latency 증가)가 발생할 수 있다.

<그림 13>과 같이 국내의 상용화되지 않은 5G 망

을 사용한 시스템은 모션 캡처 장치 연동 간소화에 있어서 모션 캡처 장치 및 스튜디오에서 한번 설정된 설정값(메타버스 디지털 휴먼 URL 세션)으로 메타버스 클라우드 시스템과 연동할 수 있으며, 메타데이터 API를 통해 얼굴, 몸 모션데이터 및 사용자 음성을 초고속/초저지연을 특징으로 하는 5G 네트워크를 통해 실시간 데이터를 안정적으로 전송할 수 있다. 모션 캡처 공간 자유화에 있어서 전국 5G망을 활용하여 이동형 스튜디오를 구성하여 시간 장소에 구애 없이 디지털 휴먼 트윈과 인터랙션 연동 기능을 구현할 수 있다.

모션 캡처 장치 연동은 상용화된 모든 얼굴 및 몸 모션 캡처 장치와 모션 캡처 스튜디오는 동일한 Wi-Fi 망에서 연동되어야 하고 다른 Wi-Fi망으로 전환 시 초기 설정에 대한 복잡성이 높아 작업 효율성을 저하시킨다. 또한 얼굴 및 몸 모션 캡처 장치는 Wi-Fi 및 유선 연결을 기반으로 모션 캡처 스튜디오와 연동됨으로 인한 공간제약이 발생시킨다.

따라서 5G망 모션 캡처 스튜디오 시스템은 이동형 모션캡처 스튜디오를 구성하여 모션 연기자는 시간 장소에 구애 없이 실시간으로 모션 데이터를 메타버스 플랫폼에 전달할 수 있다.

VI. 결론

기존의 메타버스 모션 동기화 시스템은 WIFI 망에서 연동되어야 하고 다른 Wi-Fi망으로 전환 시 초기 설정에 대한 복잡성이 높아 작업 효율성을 저하시킨다. 또한 얼굴 및 몸 모션 캡처 장치는 Wi-Fi 및 유선 연결을 기반으로 모션 캡처 스튜디오와 연동됨으로 인한 공간제약이 발생시킨다.

본 논문에서는 국내외 상용화된 모션 캡처 장치 및 스튜디오 중에서 아직 5G로 설계되지 않은 시스템을 이동형 모션캡처 스튜디오를 구성하여 모션 연기자

는 시간 장소에 구애 없이 현장감있고 자연스러운 디지털휴면을 표시함으로써 실감있는 메타버스 콘텐츠를 구현할 수 있다. 또한 모션의 종류 및 무선 네트워크의 종류에 따른 데이터 전달 시간의 차이로 인해 디지털휴면의 모션이 어색하게 구현되는 현상을 최소화할 수 있는 장점이 있다.

본 논문을 통해 모션 캡처 + 5G + 디지털 휴먼 트윈 + 메타버스 등 디지털 가상세계 관련 주목받는 차세대 기술들을 융복합하여 새로운 신사업 모델을 창출할 수 있고, 근래 디지털 휴먼, 메타버스 기술 및 기능에 대한 사업성이 증명되고 관련 매출이 본격적으로 성장하는 시점에서 차세대 메타버스 기반 방송 솔루션을 연구 개발할 수 있다.

참고문헌

- [1] 김우상 · 나건, “몰입형 가상현실의 아바타 디자인 특성에 관한 연구,” 한국디자인문화학회지, 제24권, 제1호, 2018, pp.91-104.
- [2] 류소현 · 김수영 · 김윤상, “메타버스 방송 콘텐츠에 관한 연구: 사례 분석과 제언,” 한국디지털콘텐츠학회논문지, 제24권, 제6호, 2023, pp.12-25.
- [3] 김창식 · 이윤희 · 안현철, “메타버스에 관한 연구: 뉴스 빅데이터 서비스 활용과 사례 연구를 중심으로,” 디지털산업정보학회 논문지, 제17권, 제2호, 2021, pp.85~101.
- [4] 김태경 · 정성민, “메타버스 보안 모델 연구,” 디지털산업정보학회 논문지, 제17권, 제4호, 2021, pp.95~102.
- [5] 송원철 · 정동훈, “메타버스 해석과 합리적 개념화,” 정보화정책, 제28권, 제3호, 2021, pp.3-22.
- [6] 정재현 · 김건하, “메타버스 특성과 가치 인식 서비스 이용의도에 미치는 영향,” 서비스마케팅학회 학술대회 발표논문집, 제2022권, 제6호, 2022, pp.18-19.
- [7] 신동기 · 이영민 · 이영순, “3G/4G/5G/Wi-Fi(이중대역)용 광대역 모노폴 슬롯 안테나 설계,” 한국인터넷방송통신학회 논문지, 제22권, 제1호, 2022, pp.127-134.
- [8] 이하은 · 김이길 · 김경태, “VR환경에서 디지털휴먼 얼굴 표현이 현실감에 미치는 요인 분석,” 한국공간디자인학회논문집, 제16권, 제1호, 2021, pp.383-391.
- [9] 권선희, “디지털 휴먼 이미지 활용에 관한 연구,” 한국휴먼이미지디자인, 제4권, 제2호, 2022, pp.16-34.
- [10] 황민철, “교감형 디지털 휴먼 튜터 생성 방법 및 시스템,” 대한민국 특허청, 2월, 2021.
- [11] 오병기, “디지털 휴먼 영상 형성을 위한 전자 장치 및 방법과, 그를 수행하도록 컴퓨터 판독 가능한 기록 매체에 저장된 프로그램,” 대한민국 특허청, 6월, 2021.

■ 저자소개 ■



이상윤
(Lee Sangyoon)

1999년 월~현재
트라이콤텍 엔지니어 총괄
수석부장
1996년 월~1999년 월
텔넷코리아 연구원
1995년 2월
충북대학교 통계학과(이학사)
2015년 8월
한국방송통신대학교 정보과학과
이학석사)

관심분야 : 무선/이동 전송, 차세대 이동통신,
메타버스 소프트웨어 및
어플리케이션, AR/VR
어플리케이션
E-mail : sangyoon@tricomtek.com



이 대 식
(Lee Daesik)

2011년 4월 ~ 현재
트라이콤텍(주) 연구소장
1995년 2월 가톨릭관동대학교 전자계산공학과
(공학사)
1999년 8월 가톨릭관동대학교 전자계산공학과
(공학석사)
2004년 2월 가톨릭관동대학교 전자계산공학과
(공학박사)

관심분야 : 시스템소프트웨어, 멀티미디어,
유비쿼터스 통신
E-mail : daesik@tricomtek.com



유 영 모
(You Youngmo)

2009년 7월 ~ 현재
㈜트라이콤텍 책임연구원
2010년 2월 경동대학교 멀티미디어통신학부
(공학사)

관심분야 : 무선/이동 전송, 차세대 이동통신,
AR/VR 어플리케이션
E-mail : youngmo@tricomtek.com



유 현 수
(You Hyeonsoo)

2021년 7월 ~ 현재
㈜트라이콤텍 연구원
2021년 8월 한국폴리텍대학교
디지털콘텐츠과(전문학사)

관심분야 : 메타버스 플랫폼 구축,
VR/AR콘텐츠 제작, 빅데이터 AI,
게임제작 및 상업화, App제작 및
상업화, Tool제작 및 상업화
E-mail : hyunsu@tricomtek.com



이 상 구
(Lee Sangku)

1998년 11월 ~ 현재
트라이콤텍(주) 대표이사
2020년 ~ 현재
VRARA(세계VARA협회) 한국 대표
1985년 2월 한양대학교 전자공학과 (공학사)

관심분야 : 이동통신시스템(5G, OTA, 인증),
AR/VR/XR 솔루션
E-mail : sangku@tricomtek.com

논문접수일 : 2023년 7월 26일
수정접수일 : 2023년 8월 21일
게재확정일 : 2023년 9월 1일