Network 가상환경을 위한 단순화된 2.5D 비디오 아바타 생성*

이원우, 우운택
광주과학기술원 U-VR 연구실
{wlee, wwoo}@kjist.ac.kr

Simplified 2.5D Video Avatar Generation for Networked Virtual Environment

Wonwoo Lee, Woontack Woo
KJIST U-VR Lab.

요 약

본 논문에서는 네트워크를 통해 영상기반 2.5D 비디오 아바타의 효율적인 전송 및 증강을 위한 실시간 단순화 알고리즘을 제안한다. 영상 기반 2.5D 비디오 아바타는 사용자의 모습을 실시간으로 반영할 수 있는 장점을 가지고 있으나 모델을 구성하는 데이터의 양이 많기 때문에 네트워크를 통해 전송, 증강하는 경우 지연이 발생할 수 있다. 제안된 단순화 알고리즘은 카메라를 통해 생성된 2.5D 비디오 아바타 모델을 실시간으로 단순화하고, 입력된 영상으로부터 얻어진 텍스쳐를 적용하여 단순화된 모델의 사실감을 높인다. 모델을 구성하는 데이터의 양을 줄임으로써 제한된 네트워크 대역폭 내에서 효과적인 전송및 증강이 가능하다.

Keyword : video avatar, augmentation, mesh simplification, background subtraction

1. 서론

가상현실 기술의 발전과 함께 이에 대한 관심이 증가하면서 단순히 가상의 객체를 보는 것 보다는 가상공간에서의 네비게이션 및 가상객체와의 상호작용에 대한 관심이 증가하고 있다. 가상공간을 네비게이션하고 가상객체와 상호작용하기 위해서는 사용자를 대신하는 아바타(avatar)가 필요하다. 아바타는 사용자의 상태를 반영하는 특성 때문에 가상환경을 이용한 응용 시스템의 주요 구성요소로 부각되고 있으며 사용자를 가상 환경에 통합하려는 연구가 엔터테인먼트, 산업, 통신 분야에서 진행되고 있다. 또한 원격지에 위치하는 사용자의 비디오 아바타를 가상공간에 증강함으로써 상호작용 및 협업 환경을 구성하기 위한 연구도 진행되고 있다.

EVL 의 Rajan 은 사용자의 머리를 3D 모델로 생성하고 이를 네트워크를 통해 원격지의 가상공간에 증강하는 시스템을 구성하였다 [1]. Tokyo 대학의 Tesuro 등은 2.5D 비디오 아바타를 이용한 원격지 간의 협업 환경을 구축하였으며, video avatar 를 plane model, depth model, voxel model, face model 등으로 분류하고, 응용 분야에 따라 서로 다른 아바타 생성 방법을 제안하였다 [2][3]. 또한, 특별한 장비를 사용하지 않고 자연스런 배경으로부터 비디오 아바타를 생성해내는 방법에 대한 연구도 진행 중이다 [4]. 그 외에도 원격지의 비디오 아바타

* 본 연구는 광주과학기술원과 실감방송 연구센터를 통한 정보통신부 ITRC 사업의 지원에 의한 것임

들을 가상 공간에 증강하고 문화 컨텐츠를 공유할 수 있도록 하는 연구도 이루어지고 있다 [5].

영상기반 2.5D 비디오 아바타는 모델을 구성하는 데이터의 양이 많기 때문에 네트워크를 통하여 전송, 증강하는 경우 넓은 대역폭을 필요로 하며, 대역폭이 충분치 않을 경우, 아바타 전송 및 증강에 지연이 발생할 수 있다. 이러한 전송 및 증강의 지연은 아바타가 사용자의 현재 상태를 반영하지 못하도록 하며, 사용자의 몰입감을 떨어뜨릴 수 있다. 따라서 2.5D 비디오 아바타 모델을 단순화하여 전송할 수 있는 방법이 요구된다.

본 논문에서는 영상 기반 2.5D 비디오 아바타의 실시간 단순화 알고리즘을 제안한다. 모델의 단순화를 통해 전송할 정보의 양을 줄임으로써 네트워크를 통해 효율적인 전송 및 증강을 할 수 있다. 제안된 비디오 아바타 단순화 방법은 세 단계로 이루어져 있다. 첫 번째 단계에서는 자연스러운 배경으로부터 사용자를 분리하고 비디오 아바타 생성에 사용될 점 데이터를 얻는다. 두 번째 단계에서는 단순화 알고리즘을 통해 첫 번째 단계에서 얻은 점의 수를 줄이고 각 점들 사이의 연결 정보를 생성한다. 세 번째 단계에서는 텍스쳐 매핑을 통해 단순화된 모델의 사실감을 높인다. 그림 1은 단순화된 비디오 아바타 생성 과정을 나타낸 블럭도이다. 제안된 단순화 알고리즘을 통해 네트워크를 통해 전송할 데이터의 양을 줄일 수 있으며, 아바타 모델의 단순화 정도를 조절함으로써 적응적인 전송이 가능하므로 네트워크의 제한된 대역폭 내에서 효과적인 전송 및 증강이 가능하다.

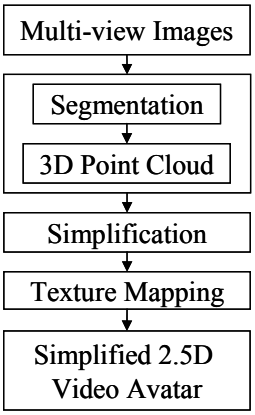


그림 1. 단순화된 2.5D 비디오 아바타 생성 과정

본 논문의 구성은 다음과 같다. 2 장에서는 제안된 비디오 아바타 단순화 알고리즘을 각 단계별로 설명한다. 3 장에서는 실험 결과를, 4 장에서는 결론 및 추후 과제에 대해서 언급한다.

2. 단순화된 2.5D 비디오 아바타 생성

2-1. 데이터 획득

자연스러운 배경에서 비디오 아바타를 구성할 점 데이터를 얻기 위해 본 논문에서는 RGB 컬러 모델과 정규화된 RGB 컬러 모델에 기반한 배경 분리 알고리즘을 사용한다 [6]. 멀티뷰 카메라로부터 입력된 영상으로부터 사용자를 분리하고, 양안차 맵을 이용해 분리된 사용자 영상의 각 픽셀에 대한 3 차원 점 정보를 얻는다. 이를 통해 사용자를 구성하는 3 차원 point cloud 를 얻는다.

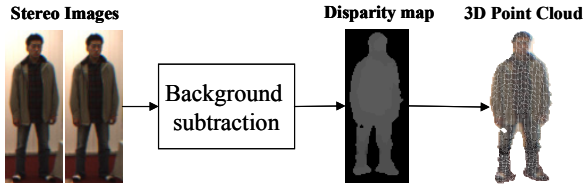


그림 2. 점 데이터의 획득

2-2. 2.5D 비디오 아바타 단순화

네트워크를 통해 실시간으로 비디오 아바타를 전송, 증강하기 위해서는 모델링 과정 뿐만 아니라 메쉬 단순화 과정 역시 실시간으로 행해져야 하며, 자동화된 방법이 필요하다. 따라서 본 논문에서는 실시간 단순화를 위해 다른 알고리즘에 비해 속도가 빠른 Vertex Clustering 방법을 사용한다 [7].

제안된 단순화 알고리즘은 다음과 같다. 먼저 데이터 획득 과정에서 얻은 point cloud 의 좌표의 최소값과 최대값에 근거하여 의 바운딩 박스를 생성한다. 그림 3과 같이 생성된 바운딩 박스를 일정한 크기를 갖는 작은 셀들로 다시 세분하고, 임의의 한 셀 내부에 분포하는 점들을 하나의 대표점으로 매핑하여 모델을 구성하는 점들의 수를 줄인다. 바운딩 박스를 작은 셀로 세분화하는 과정에서, 셀의 가로 및 세로 방향 갯수를 결정하기 위해 황금 비율을 사용한다. 단순화 이후에도 비디오 아바타의 형태를 유지하기 위해서 대표점의

결정은 셀 내부에 위치한 점들의 median 값을 사용한다. 임의의 셀 내부에 분포하는 점들은 서로 간의 깊이 값 차이가 크지 않기 때문에 vertex clustering 으로 생성된 대표점이 셀 내부에 있는 점들의 깊이값을 유지할 수 있다.

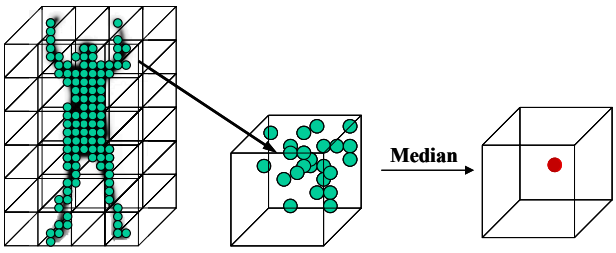


그림 3. Vertex clustering 을 통한 대표점의 생성

비디오 아바타를 구성하는 점의 수를 줄인 후에는 대표점들 사이의 연결 정보를 생성하기 위해 각 셀 사이의 인접성을 판단한다. 대표점이 존재하는 유효한 셀들을 일정한 규칙에 따라 연결함으로써 대표점들에 대해 삼각화를 수행한다. 이 과정을 통해 단순화된 메쉬 모델을 생성할 수 있다. 또한 셀의 갯수를 동적으로 조절함으로써 메쉬 모델의 단순화 정도를 조절하는 것이 가능하다.

2-3. 텍스쳐 매핑

텍스쳐 매핑은 컬러 정보가 결여되어 있는 메쉬 모델에 컬러 정보를 부여함으로써 보다 실감나는 아바타를 생성할 수 있도록 한다. 본 논문에서는 실시간으로 입력되는 영상으로부터 텍스쳐를 동적으로 생성하고 이를 비디오 아바타에 적용한다. 그림 4 는 텍스쳐 매핑 과정을 개념적으로 나타낸 것이다.

네트워크를 통해 전송할 데이터의 양을 최소화하기 위해 텍스쳐는 사용자를 둘러싸는 사각형의 영역으로부터 생성한다. 카메라로부터 얻은 2 차원 이미지와 사용자를 배경으로부터 분리한 후 얻는 깊이 맵을 통해 입력된 영상의 화소와 3 차원 점의 대응 관계를 알 수 있다. 이로부터 3 차원 공간상의 각 점에 대한 텍스쳐 좌표를 계산한다. 임의의 한 셀 내부에 존재하는 점 들의 텍스쳐 좌표로부터 median 값을 취하고 이를 각 셀의 대표점에 대한 텍스쳐 좌표로 사용한다. 실시간으로 입력되는 영상으로부터 매 프레임 마다 텍스쳐를 생성하여 적용하므로 사용자의 현재 상태를 아바타에 지연 없이 반영하고, 이를 통해 사용자의 몰입감을 높일 수 있다.

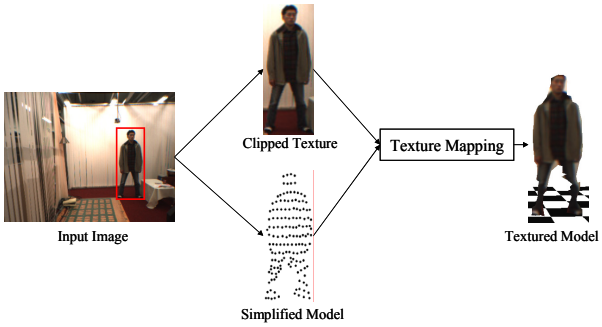


그림 4. 텍스쳐 매핑 과정

3. 실험 결과

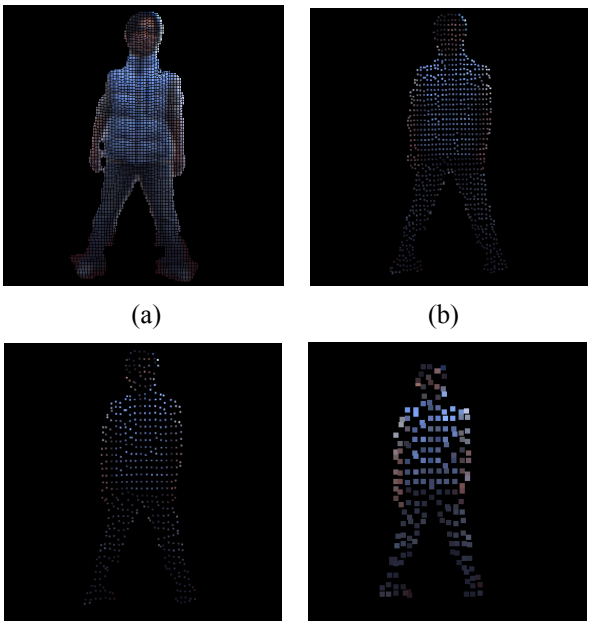
제안된 2.5D 비디오 아바타 단순화 알고리즘을 검증하기 위하여 세 단계의 단순화된 모델을 생성하고, 각각 Case 1, Case 2, Case 3 라 하였다. 이를 단순화 알고리즘 적용 전의 2.5D 비디오 아바타 모델과 비교하였다.

표 1은 각 단계별로 생성된 비디오 아바타의 데이터 양을 비교한 것이다. 셀의 갯수에 따라서 점과 면의 수가 달라짐을 알 수 있다.

표 1. 단순화 정도에 따른 데이터 양의 변화

	Cells (가로×세로)	Faces	Points
단순화 이전	•	12653	6652
Case 1	35×56	1557	884
Case 2	25×40	810	478
Case 3	15×24	290	190

그림 5는 단순화 알고리즘을 적용 후 줄어든 point cloud 를 단순화 정도에 따라 나타낸 것이다. 그림 5(a)는 단순화 알고리즘을 적용하기 전의 point cloud 이며, 그림 5(b), 그림 5(c), 그림 5(d)는 단순화의 결과로 생성된 새로운 point cloud 이다. 각 셀의 point cloud 로부터 median 값을 취함으로써 점의 수가 줄어들었지만 전체적인 형태를 유지하고 있는 것을 볼 수 있다.

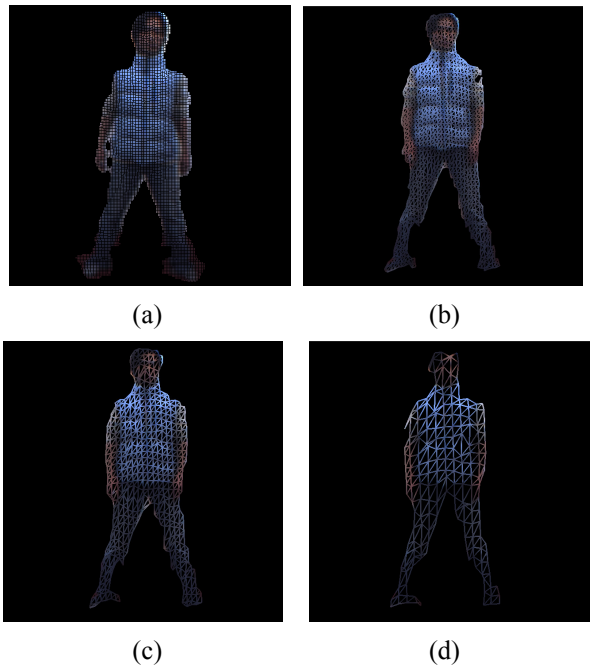


(a) (b)

(c) (d)

그림 5. 단순화된 point cloud (a)단순화 알고리즘 적용전 (b)Case 1 (c)Case 2 (d)Case 3

그림 6은 단순화된 점들로부터 생성된 메쉬 모델을 각 단계별로 보인 것이다. 그림 6(a)에서 보이듯이, 단순화 이전에는 메쉬가 매우 조밀하다. 그러나 그림 6(b), 그림 6(c), 그리고 그림 6(d) 에서 보이듯이 단순화 이후에는 적은 수의 삼각형들 만으로 메쉬가 구성된 것을 볼 수 있다.

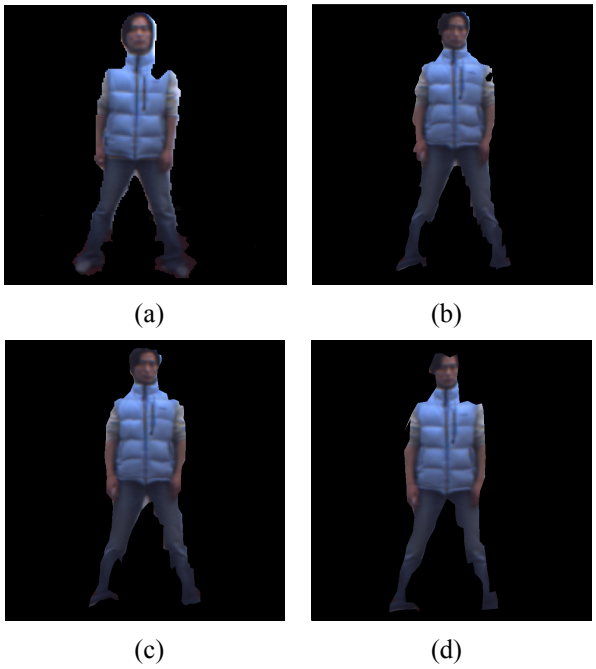


(a) (b)

(c) (d)

그림 6. 단순화된 메쉬 모델 (a)단순화 알고리즘 적용전 (b)Case 1 (c)Case 2 (d)Case 3

그림 7은 단순화된 모델에 텍스쳐를 적용한 결과를 보인 것이다. 단순화된 메쉬 모델에 텍스쳐를 적용함으로써 단순화된 모델의 사실성을 높인 것을 볼 수 있다. 그림 7(a)는 단순화 알고리즘을 적용하기 전의 모델이며 그림 7(b), 그림 7(c), 그림 7(d)는 각 단계별로 텍스쳐를 적용한 모델이다.



(a) (b)

(c) (d)

그림 7. 텍스쳐를 적용한 비디오 아바타 (a)단순화 알고리즘 적용 전 (b)Case 1 (c)Case 2 (d)Case 3

4. 결론 및 추후과제

본 논문에서는 네크워크를 통해 효율적으로 전송하여 증강할 수 있는 단순화된 2.5D 비디오 아바타 생성 방법을 제안하였다. 제안된 알고리즘은 자연스런 장면으로부터 사용자를 분리하여 2.5D 비디오 아바타를 생성할 점 데이터를 획득하고, 단순화 과정을 수행한다. 단순화된 비디오 아바타 모델에 텍스쳐를 적용하여 사실감을 높인다. 비디오 아바타 모델의 단순화를 통해 데이터의 양을 줄임으로써 제한된 네트워크 대역폭 내에서 효율적인 전송을 할 수 있다. 현재는 단순화된 모델을 전송하기 때문에 비디오 아바타를 전송받아 렌더링 하는 쪽에서는 단순화된 모델만을 증강할 수 있다. 향후 메쉬 모델이 계층적인 데이터 구조를 갖도록 하여 비디오 아바타 모델을 전송 받아 증

강하는 경우에도 계층적인 증강이 가능하도록 하는 연구를 진행할 예정이다.

참고문헌

[1] V. Rajan, S. Subramanian, D. Keenan, A. Johnson, D. Sandin, T. DeFanti, “A Realistic Video Avatar System for Networked Virtual Environments”, IPT2002, 2002.

[2] T. Ogi, T. Yamada, K. Tamagawa, M. Kano, M. Hirose, Immersive Telecommunication Using Stereo Video Avatar, IEEE VR2001, pp.45-51, 2001.

[3] T. Ogi, T. Yamada, Y. Kurita1, Y. Hattori1, M. Hirose. Usage of Video Avatar Technology for Immersive Communication, Proc. First International Workshop on Language Understanding and Agents for Real World Interaction, pp. 24-31, July (2003).

[4] Y.Suh, D.Hong, and W.Woo, 2.5D Video Avatar Augmentation for VRPhoto, ICAT02, pp. 182 -183, Dec. 3-6, 2002.

[5] Y.Suh, D.Hong, and W.Woo, 2.5D Video Avatar for Networked VRPhoto System, HCII03, pp. 533-537, 2003

[6] D.Hong, W.Woo, Background subtraction for a vision-based interface, PCM03 CD Proceeding: 1B3.3, 2003

[7] Rossignac, J., Borrel, P. (1993) Multi-resolution 3D approximation for rendering complex scenes, 2nd Conf. on Geom. Model. in Comp. Graph., 453-465