

3 차원 정보를 활용한 물체의 텍스처 교체 기법

김주현, 박종일¹

한양대학교

jhkim@mr.hanyang.ac.kr, jipark@hanyang.ac.kr¹

Texture replacement technique using 3D information

Joohyeon Kim, Jong-Il Park¹

Hanyang University

요 약

AR 기술과 장비가 발전하며 다양한 분야에 이를 접목한 콘텐츠들이 생겨나고 있다. 가상광고는 기존의 광고 방법으로는 적용될 수 없는 영역에도 광고가 가능함으로써 광고분야에 새로운 패러다임으로 떠오르고 있다. 본 논문에서는 AR 환경의 가상 광고에 활용할 수 있는 기법을 제안함으로써 관련 기술의 상용화에 기여하고자 한다. 제안하는 기법은 RGB-D 카메라를 사용하여 물체의 3 차원 정보를 복원하고, 선택된 영역의 텍스처를 교체하는 기법과 모델 기반의 카메라 추적기술을 활용하여 실시간으로 물체의 텍스처가 교체될 수 있음을 보일 것이다.

1. 서론

오늘날 AR 기술은 사용자에게 새로운 경험을 제공하는 핵심 기술로 발전하고 있다. 최근에 출시된 포켓몬 고는 AR 기술을 활용하여 많은 인기를 끌 수 있었다. AR 기술이 다른 기술과 차별화 될 수 있는 것은 현실 세계와 가상 세계의 결합으로 사용자들에게 실감나는 콘텐츠를 제공할 수 있다는 것이다. 하지만 사용자가 영상에서 이질감을 느끼지 않도록 하기 위해서는 다양한 기술적인 문제들이 존재하기 때문에 AR 기술을 활용한 상용화 어플리케이션을 만드는 것은 기술적으로 매우 도전적인 문제이다.

최근에 광고 시장은 다양한 방법으로 발전해 왔는데, 그 중에서도 컴퓨터 그래픽 기술을 활용한 가상 광고는 2000 년초에 스포츠경기에 등장하여 새로운 패러다임을 제시하였다. 기존의 가상 광고 기법은 광고 시간과 광고 삽입 위치에 제약이 있는데, 이는 시청자들이 가상 광고로 인한 불편함을 느끼지 않도록 하기 위함이다. 예를 들어 사람이 있는 영역에는 가상 광고를 할 수 없다는 제약이 있는데, 이러한 제약은 이미 촬영된 영상의 광고 영역에 가상광고를 할 경우에는 해당사항이 없다. 하지만 광고 영역에 가상광고를 삽입하는 자동화 기술이 없기 때문에 영상의 프레임마다 사람이 작업을 해야 하는 불편함이 있었다. Kim [1]은 비디오 클립의 환경에서 간판 영역을 교체할 수 있는 기법을 제안함으로써 가상 광고 삽입 작업에 활용될 수 있는 실용적인 방법을 제안한 바 있다. 다가오는 미래에 AR 기술이 활성화되면 가상광고는 방송 분야 뿐만 아니라 AR 기술 기반의 콘텐츠에도 접목이 가능할 것으로 예상되기 때문에 이를 위한 핵심 기술에 대한 연구가 필수적이다.

본 논문에서는 AR 기반의 가상광고에 적용이 가능한 기법을 제안하고자 한다. AR 환경에서의 가상광고는 어떤 물체의 텍스처를 교체하는 방법으로 이루어질 가능성이 크다. 이를 위해서는 카메라 영상으로 획득한 객체를 정확하게 추정할 수

있어야 하며, 교체하고자 하는 텍스처로 선택된 영역에 합성되어야 하는 기술이 필요하다. 제안하는 기법은 RGB-D 카메라를 이용하여 물체의 3 차원 정보를 복원하고, 일부 영역의 텍스처를 교체한 모델링 정보를 생성한다. 그리고 Seo [2]가 제안한 모델기반 카메라 추적 기법으로 RGB 카메라의 입력영상에서 실시간으로 교체된 텍스처를 증강함으로써 제안하는 방법이 AR 기술을 활용한 가상광고에 적용될 수 있는 기법임을 보일 것이다.

2. 3 차원 정보를 활용한 물체의 모델링

텍스처를 교체하고자 하는 물체를 모델링하는 것은 가장 먼저 수행해야 하며 필수적인 과정이다. 본 논문에서는 RGB-D 카메라로부터 획득한 물체의 포인트 클라우드(Point cloud)들을 Iterative closest point (ICP) 알고리즘을 활용하여 모델링하였고 CGAL [3]에서 제공되는 메쉬(mesh) 생성 알고리즘으로 모델을 메쉬화 할 수 있었다.

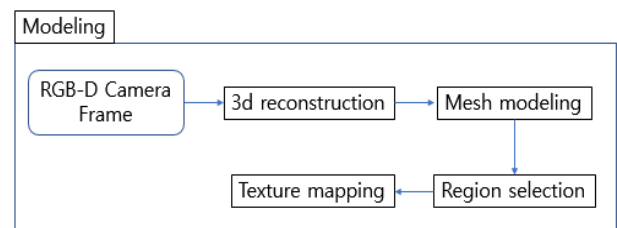


Figure 1. 모델링 순서도

모델을 메쉬화 하였으면, 사용자는 텍스처를 교체하기 위한 영역을 선택하고, 360 도 이루어진 파노라마 영상으로 텍스처 맵핑을 수행한다. 모델링에 관한 순서도는 그림 1에서 볼 수 있다.

¹ 교신저자

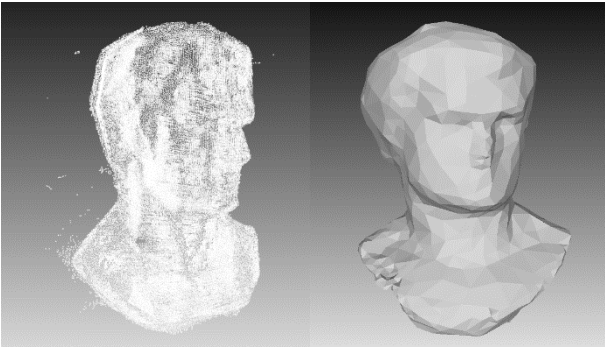


Figure 2. 모델링 결과
(좌: 포인트 클라우드로 복원 우: 메쉬화된 모델)

ICP 알고리즘은 입력 데이터와 목표 데이터(p_i, q_i)가 있고, 모델의 법선 벡터 n_i 을 알고 있을 때 입력과 모델에 정합하기 위하여 아래 식 (1)과 같이 정의되는 오차를 최소화하기 위한 회전행렬 R 과 이동 벡터 t 를 계산하는 기법이다[4].

$$e = \sum_i ((R \cdot p_i + t - q_i) \cdot n_i)^2 \quad (1)$$

객체의 3 차원 정보를 복원하기 위하여 카메라는 고정하고, 물체를 한바퀴 회전한 영상을 사용하였으며, 영상의 ROI(region of interest) 영역의 포인트 클라우드를 ICP 알고리즘으로 정합하여 모델링할 수 있었다.

교체영역의 선택방법은 사용자가 원하는 영역을 마우스로 선택하여 이루어진다. 우리는 이미지에서 선택된 2 차원 점의 좌표에서 모델방향으로의 광선 벡터를 만들고 좌표계의 변환을 통하여 월드좌표계의 광선 벡터를 생성하였다. 그리고 광선벡터와 메쉬와의 교점을 확인하여 사용자가 선택한 메쉬를 정확하게 찾을 수 있었다. 일반적으로 3 차원 모델에는 전방영역과 후방영역이 있기 때문에 보통 두개의 메쉬가 선택되는데, 카메라 시점과 가까운 교점의 메쉬가 보이는 영역이므로 이를 선택하여 영역을 지정할 수 있었다.

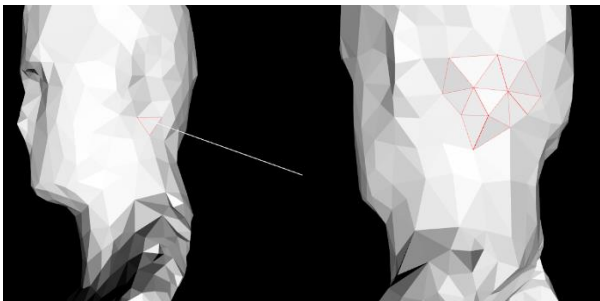


Figure 3. 교체영역 선택
(좌: 카메라 시점에서 모델방향으로의 광선벡터 생성
우: 선택된 메쉬 영역)

텍스처 교체는 360 도 파노라마 영상의 텍스처를 사용하였다. 이는 3 차원 모델에 2 차원 영상의 텍스처로 합성할 경우 왜곡이 되는 현상이 발생하기 때문이다. 360 도 파노라마 영상을 3 차원 모델로 투영시키기는 방법은 교체하고자 하는 영역의 좌표를 구좌표계로 변환하고, 이에 대응되는 파노라마 영상의 위치를 계산하면 맵핑 좌표를 찾을 수 있다.

3. 모델기반 카메라 추적 기법

물체에 대한 모델링이 완료되었으면 영상에서 카메라의 포즈 따라 모델의 텍스처 부분이 교체된 것을 증강시키는 부분이 필요하며 이에 관한 순서도는 그림 4 와 같다.

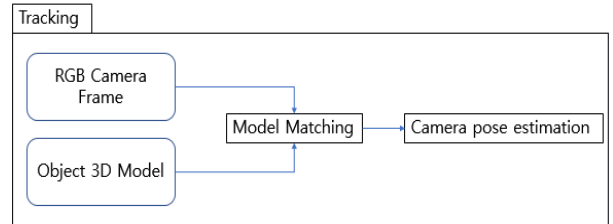


Figure 4. 추적 순서도

물체를 추적하기 위해서는 카메라의 입력영상에 존재하는 물체의 외형 정보와 모델링 정보를 정합하기 위한 과정과 이를 기반으로 카메라의 포즈를 추정하는 과정으로 이루어져 있다.

3.1 모델의 정합

모델의 정합은 영상에서 검출된 에지와 물체의 모델로부터 검출된 에지 사이의 대응 관계를 알아내는 것이다. 가장 먼저 객체 모델의 에지를 추출하기 위해서 영상으로 투영한 뒤에 모델의 법선 벡터를 활용한 은면 제거를 하여 카메라 시점에서 보이는 가장 외곽선의 에지를 추출한다. 그리고 영상에 있는 모델의 에지는 캐니 에지 검출기를 사용하여 추출한다. 영상에서 검출된 에지와 모델의 에지를 정합하기 위해서는 각 선의 법선 방향으로 가장 가까운 에지를 대응점으로 정의하여 카메라 포즈를 추정할 때 사용한다.

3.2 카메라 포즈 추정

카메라의 포즈 추정은 이전 프레임에서의 카메라 포즈를 갱신함으로써 진행되며 식 (2)와 같이 표현된다.

$$E_t = E_{t-1} \Delta E \quad (2)$$

이 때, ΔE 는 동차 좌표계의 3 차원 강체 변환의 표현인 리 그룹(Lie Group) SE(3)으로 표현된다[5]. 움직임은 G_i 의 가중치 합에 의하여 생성되며, 이는 아래 식 (3)과 같이 x, y, z 축을 따라 회전과 변환을 나타내는 4x4 행렬이 된다[6].

$$\Delta E = \sum_{i=0}^5 a_i G_i \quad (3)$$

G_i 는 동차 좌표계 상의 x, y, z 방향에 대한 평행 이동 및 회전 변환의 기저 행렬이며, a_i 는 각 변환에 대한 변화량을 나타낸다. 대응점간의 거리를 최소화 하는 비용함수는 식 (4)와 같다.

$$\hat{a} = \operatorname{argmin} \left\{ \sum_j^N e \left\| d^j - \sum_{i=0}^5 a_i \left(n_{\perp}^j \cdot l_i^j \right) \right\|^2 \right\} \quad (4)$$

여기서 d는 이전 카메라 포즈 E_{t-1} 로 투영시킨 모델의 선과 이에 대응되는 영상의 에지간의 거리이며, n_{\perp} 은 투영된 선의 법선 벡터, N_e 는 대응점의 개수이다. l_i 는 투영행렬 P로 투영된 선의 작은 카메라 움직임에 의한 변화량 벡터이다.

$$l_i = \left[\left(\frac{u}{w} \right)' \left(\frac{v}{w} \right)' \right]^T = \left[\frac{u'}{w} - \frac{uv'}{w^2} \quad \frac{v'}{w} - \frac{vw'}{w^2} \right]^T \quad (5)$$

$$\left[u' v' w' \right]^T = PG_i [X \ Y \ Z \ 1]^T \quad (6)$$

최소자승법으로 계산된 a로부터 카메라 움직임 ΔE 가 계산되고 현재 프레임의 카메라 위치 및 자세 E_t 가 갱신된다. 카메라 위치 및 자세 E_t 의 갱신은 반복적으로 수행되는데, 재투영 오차가 일정 값 이하로 작아지거나 미리 정해진 횟수를 넘게 되면 카메라의 위치 및 자세 E_t 의 갱신이 종료된다.

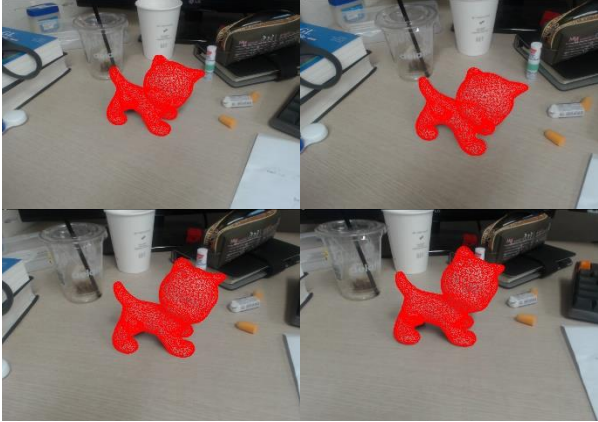


Figure 5. 모델 기반 카메라 추적 영상

4. 실험 및 결과

4.1 구현 및 실험 환경

모델링을 위하여 사용된 RGB-D 카메라는 Kinect2 를 사용하였으며, 객체 추적 및 증강을 위해서 사용한 RGB 카메라는 Logitech c920 을 사용하였다. 기본적인 구현 및 실험환경은 데스크탑에서 이루어졌으며, Intel 사의 i7-4790K 4.00GHz Quad Core CPU, 32GB RAM, NVIDIA GeForce GTX970 의 조건하에 수행되었다. 포인트 클라우드를 처리하기 위해서는 PCL library 에서 제공되는 기본 함수들을 사용하였고, 모델의 메쉬 생성은 CGAL 에서 제공하는 알고리즘을 적용하였다. 그리고 3 차원 모델링 인터페이스 및 그래픽 렌더링은 OpenGL2.0 을 이용하였다.

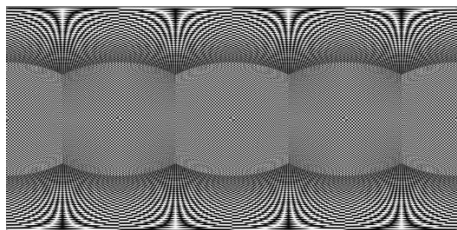


Figure 6. 실험에 사용된 텍스처 영상

실험에 사용된 텍스처는 360 파노라마 형식으로 생성된 체스보드 영상을 사용하였으며 2448×1224 크기를 지니고 있다.

4.2 실험 결과

실험은 두 모델에 대하여 수행되었다. 사용자로부터 교체 영역을 선택받고, 이를 RGB 카메라로 객체를 추적하여 텍스처를 보여주는 것은 비교적 자연스럽게 수행되는 것을 확인할 수 있었다. 하지만 조명환경을 고려하지 않았기 때문에 텍스처의 색이 눈에 띄게 구별되는 것을 볼 수 있었으므로 향후에는 이를 개선하기 위한 연구가 필요하다.

추적은 실시간으로 수행이 가능하지만 객체 모델에 따라 영향을 받는 것을 확인할 수 있었다. 예를 들어 별다른 특징이 없는 모델보다 비교적 돌출된 영역이 많이 존재하는 모델이 잘 추적되는 것을 확인할 수 있었다. 그리고 모델의 메쉬 개수에 따라 계산량이 많아지기 때문에 수행시간에 영향을 주는데, 실시간 처리를 위해서는 특징을 잘 보여줄 수 있는 적절한 개수의 포인트 클라우드를 정확하게 생성하는 알고리즘이 필요하다는 것을 느낄 수 있었다.

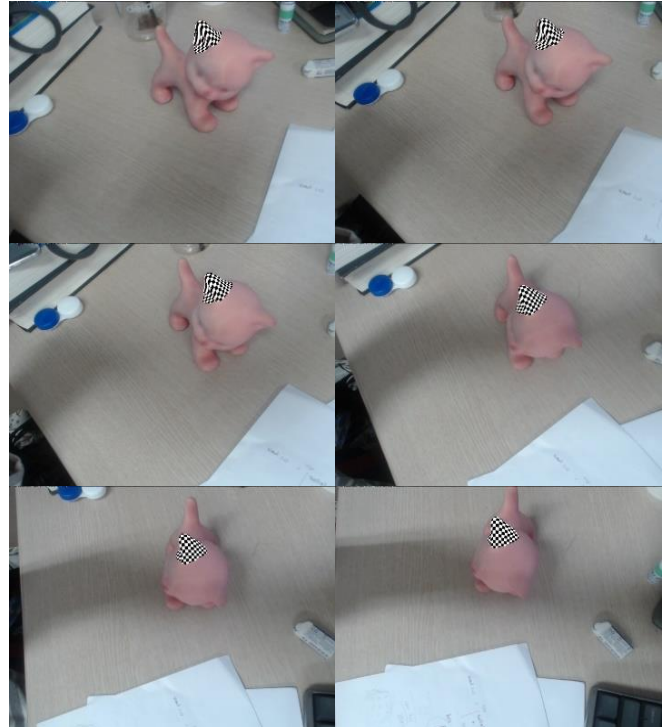




Figure 7. 텍스처 교체 및 추적 결과

5. 결론

본 논문에서는 AR 기반의 가상광고에 활용 가능한 물체의 텍스처 교체 기법을 제안하였다. 실험 결과, 제안하는 방법은 실시간으로 카메라를 추적하여 지정된 영역에 텍스처를 정확하게 증강시키는 것을 확인할 수 있었다. 다만, 3차원 모델에 텍스처를 합성하기 위하여 360 파노라마 사진을 사용해야 한다는 제약이 있으며, 향후에는 조명에 따른 색 보정 작업에 대한 연구를 하여 자연스러운 증강이 될 수 있도록 개선할 계획이다.

감사의 글

"본 연구는 미래창조과학부 및 정보통신기술진흥센터의 대학 ICT 연구센터육성 지원사업의 연구결과로 수행되었음" (IITP-2017-2012-0-00628)

참고문헌

- [1] 김주현, 박종일 "비디오 클립에서의 간판영역 교체 기법", *방송공학회논문지*, vol.22, no. 3, pp339-349, May, 2017
- [2] SEO, Byung-Kuk; PARK, Jong-Il; PARK, Hanhoon. Camera tracking using partially modeled 3-D objects with scene textures. In: *VR Innovation (ISVRI), 2011 IEEE International Symposium on*. IEEE, p. 293-298. 2011.
- [3] <http://www.cgal.org/>
- [4] Chen, Y., & Medioni, G., "Object modeling by registration of Multiple range images", *IEEE International Conference on Robotics and Automation*, 135. 1991.
- [5] T. Drummond and R. Cipolla, "Application of lie algebras to visual servoing," *International Journal of Computer Vision*, vol. 37, no. 1, pp. 21-41, 2000.
- [6] Martin A. Fischler & Robert C. Bolles. "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". *Comm. of the ACM*. 24 (6): 381-395. June, 1981.