

동영상과 3차원 얼굴 모델의 자동 정합

이 정 김창현
고려대학교 컴퓨터학과
{jay77, chkim1}@cgvr.korea.ac.kr

An Automatic Matching between Video Frames and 3D Facial Model

Jung Lee Chang-Hun Kim
Dept. of Computer Science & Engineering, Korea University

요 약

본 논문은 동영상 내의 얼굴을 특징인 얼굴로 자동 변환 및 정합하는 기술을 제안한다. 얼굴에 나타난 동작이나 표정은 높은 자유도로 인하여 기존에 사용되어온 2차원적이고 고정된 물체 위주의 동영상 정합 기술로는 자연스러운 결과물을 얻기가 어렵다. 본 논문에서는 입력 받은 정면 유사방향의 사진으로부터 3차원 얼굴 모델을 복원한다. 각 프레임에 등장한 얼굴의 3차원 방향을 추출하여 복원한 3차원 얼굴 모델에 적용한 후 대체할 얼굴 영역에 정합시킨다. 정합 과정 시 비디오 프레임 내의 조명효과와 얼굴색 등을 분석하고 3차원 얼굴 모델에 블렌딩하여 비디오 프레임과 자연스럽게 정합할 수 있도록 한다.

1. 서론

컴퓨터 시스템을 이용한 그래픽스 기술의 적용 범위가 영화나 애니메이션 등으로 넓어지면서 위험한 장면의 스텐트맨 얼굴을 주인공의 얼굴로 대체시키는 것과 같은 경우에 적용할 수 있도록 동영상 등장인물 얼굴의 특징인 얼굴로의 변환 기술이라는 새로운 분야가 등장하게 되었다.

하지만 고정된 물체를 대체하는 것과는 달리, 얼굴은 표정이나 여러 가지 동작의 자유도가 높다는 특징이 있기 때문에, 이미지 저작 도구를 이용한 2차원 변환기를 주로 사용하고 있는 기존의 기술들로는 자연스러운 얼굴 대체가 어렵다는 점이 지적되었다. 이와 같은 문제를 해결하기 위하여 대체시키고자 하는 인물의 3차원 얼굴 모델을 생성하고 변형한 후 비디오 프레임에 정합하는 기술이 필요하게 되었다.

본 논문은 동영상과 3차원 얼굴 모델의 자동 정합에 관한 것으로 각 비디오 프레임에 등장한 얼굴의 3차원 방향을 추출하여, 정면 유사방향의 사진으로부터 복원한 3차원 얼굴 모델에 추출된 3차원 얼굴 방향을 적용한다. 대체할 3차원 얼굴 모델이 원래의 얼굴과 비슷한

표정과 입 모양 등을 표현할 수 있도록 프레임에 등장한 얼굴을 분석하여 대체할 비디오 프레임에 등장한 얼굴과 유사한 표정과 입술 움직임을 생성시킨다. 또, 프레임 내에 나타난 조명효과를 분석하고 블렌딩과 같은 이미지 처리를 수행함으로써 특징인의 얼굴로 변화된 새로운 동영상을 얻게 된다.

2. 관련 연구

본 논문에서 사용한 3차원 얼굴 모델 복원 알고리즘은 사진으로부터 3차원 얼굴 모델을 복원하는 기존 연구[1, 2]가 완전한 정면사진의 입력을 요구하는 한계를 확장하여 정면 유사방향의 사진에서도 3차원 얼굴 모델의 복원이 가능하게 하였다.

본 논문에서 제안하는 시스템에서 사용한 3차원 얼굴 모델의 표정 생성 및 동작 제어 기술은 파라미터 기반의 얼굴 표정 생성 기법[3]을 토대로 대체할 3차원 얼굴 모델의 표정을 생성 및 변형하고 있다.

본 논문에서는 얼굴에 지정된 특징점으로부터 비디오 프레임에 등장한 얼굴의 3차원 방향을 추출한다. 기존 연구에서 3차원 방향 추출과정에서 부가적으로

입력 사진을 촬영한 카메라의 초점거리를 요구[4, 5]하던 제약조건을 단순화함으로써 3차원 얼굴 방향 추출에 소요되는 시간과 계산의 복잡도를 대폭 감소시켰다.

최종적으로 생성된 3차원 얼굴 모델을 각 비디오 프레임에 정합[6]함으로써 임의의 3차원 얼굴 모델이 비디오 프레임에 등장한 얼굴과 합성된 새로운 동영상을 생성하였다.

3. 제안 시스템의 특징

기존의 동영상 얼굴 대체 기술은 원본 동영상 이외에도 합성시키길 원하는 부분만을 따로 촬영한 동영상을 이용하여 교체시키는 방식이 주류를 이루고 있다. 하지만 이러한 방식은 어중으로 동영상을 촬영해야 하므로 비용이 많이 들고, 처음부터 얼굴의 합성을 고려하여 촬영한 동영상이 아니면 합성이 어렵다는 단점이 있다.

본 논문에서 제안하는 시스템은 정면 유사방향의 얼굴사진으로부터 3차원 얼굴 모델을 복원하고, 각 비디오 프레임 속에 나타난 얼굴정보(3차원 방향 및 표정 등)를 분석하여 앞서 복원된 3차원 얼굴 모델에 적용한다. 얼굴 정보에 따라 변형된 3차원 얼굴 모델을 비디오 프레임의 얼굴 영역에 정합한 후 조명환경, 얼굴색 등을 분석하여 블렌딩함으로써 최종적으로 합성된 동영상을 생성한다.

생성한 3차원 얼굴 모델을 프레임에 나타난 얼굴과 흡사하게 변형할 수 있도록 3차원 변형을 적용하므로, 기존의 2차원적인 얼굴 대체기술에 비해 사실적인 결과물을 얻을 수 있다.

4. 동영상과 3차원 모델 정합

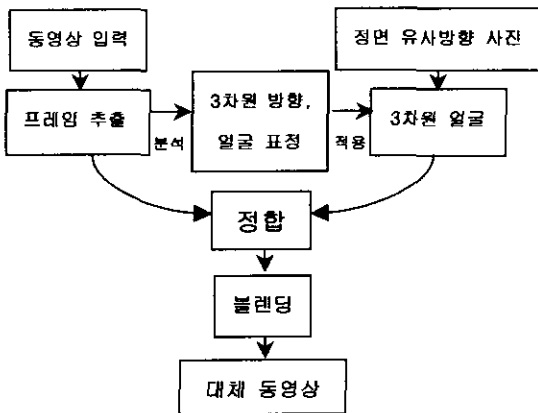


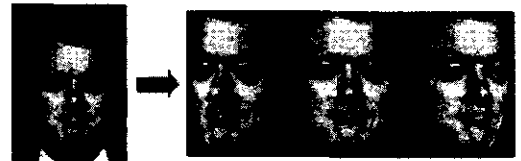
그림 1. 시스템 구성도

4.1 3차원 얼굴 모델 복원

동영상에 정합하게 될 특정 얼굴의 3차원 모델링 과정에서는 기존의 3차원 스캐너와 같은 고가의 장비사용은 복원결과에는 정확하지만 사용자가 원하는 모델의 즉

각적인 복원이 어렵고 비용이 많이 든다는 단점이 있다.

본 논문에서는 제안하는 이미지 기반 모델링 기법을 이용하면 완전 정면 사진(그림 2-a)뿐만 아니라 기존의 3차원 얼굴 복원 기술에서는 어려움이 있었던 정면 유사방향의 얼굴사진(그림 2-b)으로부터 3차원 얼굴 모델을 복원하는 것이 가능하다.



(a) 완전 정면 사진으로부터 복원한 예



(b) 정면 유사 방향의 사진으로부터 복원한 예
그림 2. 3차원 얼굴 모델 복원의 예

4.2 3차원 방향과 표정 분석 및 얼굴 정합

3차원 얼굴 모델을 생성한 후에는 합성할 동영상의 각 프레임을 동영상 전문 편집프로그램을 이용하여 추출해낸다. 그리고 3차원 얼굴 모델을 비디오 프레임의 원본 얼굴과 유사하게 정합하기 위해서 비디오 프레임에 등장한 얼굴에 몇 개의 특징점을 지정하여 3차원 얼굴 방향을 추출하고, 각 특징점의 위치관계로부터 각 얼굴 특징부위(눈, 눈썹, 입 등)의 이동량을 계산하여 3차원 얼굴 모델에 적용시켜 프레임 내의 얼굴과 유사한 방향과 표정을 짓도록 변형시킨다. 변형된 3차원 얼굴 모델을 해당 프레임에 정합시켜 특정한 얼굴로 합성된 프레임을 얻게 된다.

4.3 블렌딩 및 동영상 합성

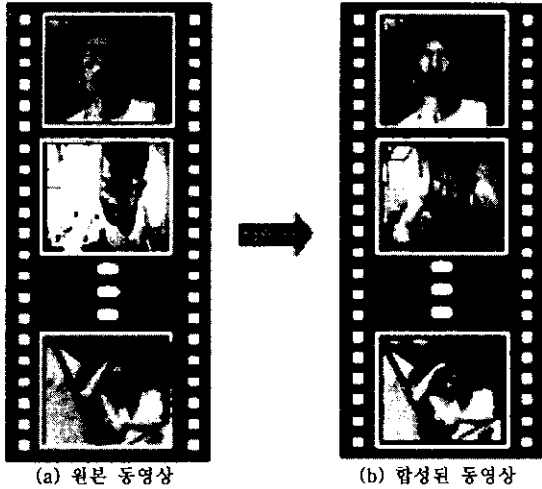
4.2절에서 프레임에 얼굴을 정합한 후(그림 3-a), 프레임 내의 조명환경과 얼굴색을 분석하여 블렌딩을 적용함으로써 최종적으로 생성될 프레임(그림 3-b)의 사실감을 높였다. 이 과정에서는 대체할 얼굴 각 부위의 색 분포를 조사하여 블렌딩 이미지를 내부적으로 생성하고 3차원 모델에 블렌딩을 적용하여 비디오 프레임 내의 조명환경에 일치시키고 있다.(그림 3)



(a) 블렌딩 적용 전 (b) 블렌딩 적용 후

그림 3. 블렌딩 적용 과정의 예

마지막으로 앞서의 과정들을 거쳐서 최종적으로 생성된 각 비디오 프레임들을 종합하여 특정 얼굴로 합성된 새로운 동영상상을 얻게 된다.(그림 4)

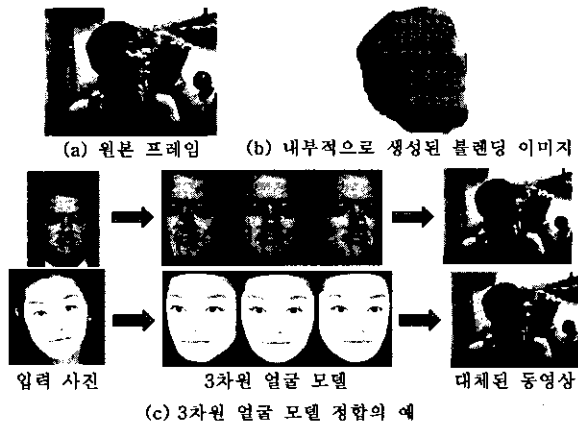


(a) 원본 동영상 (b) 합성된 동영상
그림 4. 특정 개인 얼굴로 합성된 동영상의 예

5. 실험 및 결과

본 논문에서 구현한 시스템은 펜티엄 III 550Mhz, 256MB RAM, OpenGL 가속 3D카드를 갖춘 PC에서 실험하였으며, OpenGL과 Borland C++ Builder 5를 이용하여 구현하였다.

본 논문에서 제안하는 시스템을 구성하는 과정별로 소요되는 시간은 3차원 모델링의 경우 특징점 지정 후 3차원 얼굴 모델 복원 과정은 상당히 짧은 시간(1초 미만)에 3차원 얼굴 모델의 복원이 가능하였으며, 복원된 3차원 얼굴 모델의 변형 및 비디오 프레임(320 x 240)과의 정합 과정은 평균적으로 초당 3~4 프레임 정도의 속도로 정합이 가능하였다.(그림 5)



(c) 3차원 얼굴 모델 정합의 예
그림 5. 생성된 3차원 모델과 정합된 최종 프레임 결과

6. 결론 및 향후 연구

본 논문에서는 3차원 얼굴 모델 복원 및 프레임 분석을 통해 동영상에 등장한 얼굴을 자동으로 대체하는 기법을 제안한다.

본 논문에서 제안한 시스템은 먼저 사용자가 입력한 정면 유사방향의 사진을 토대로 3차원 얼굴 모델을 복원하고 변형한다. 변형한 3차원 얼굴 모델을 동영상의 대체 얼굴 영역에 자동 정합함으로써 기존에 사용되어온 동영상 얼굴 대체 기법에 비하여 적은 시간과 노력으로 임의의 3차원 얼굴이 합성된 새로운 동영상의 생성이 가능하게 하였다. 동영상 얼굴 대체 과정에서는 프레임에 나타난 표정과 조명효과를 분석하여 3차원 얼굴에 적용함으로써 주변물체나 환경과의 조화가 유지된 사실적인 얼굴 대체가 가능하다.

본 논문에서 제안하는 시스템은 기존의 동영상 얼굴 대체 분야 뿐만 아니라 인터넷 디지털 영화와 같이 개개인의 주인공화를 요구하는 동영상의 제작 등에도 확장되어 사용될 수 있다.

향후 연구로는 프레임에 등장한 얼굴의 3차원 방향을 추출하는 과정에서 초점거리 복원 등을 통해 각 비디오 프레임에 나타난 원근효과를 반영함으로써 좀 더 정확한 3차원 방향을 추출할 것이다. 또, 블렌딩을 위한 프레임 내의 조명효과 분석과정에서 실제 조명의 위치추적 등을 통하여 좀 더 자연스러운 3차원 얼굴 모델의 정합이 가능하도록 할 것이다.

7. 참고 문헌

- [1] Won-Sook Lee, "Fast head modeling for animation," Journal Image and Vision Computing, Volume 18, Number 4, pp.355-364, Elsevier, 1 March, 2000.
- [2] Won-Sook Lee, "Generating Animatable 3D Virtual Humans from Photographs," Computer Graphics Forum, Volume 19, Issue 3, Eurographics 2000 Proc., pp.1-10, 2000
- [3] Parke, F. I., "Parameterized models for facial animation," IEEE Computer Graphics, Vol. 2, No. 9, pp. 61-68, 1982.
- [4] Thanarat Horprasert, "An Anthropometric Shape Model for Estimating Head Orientation," International Workshop on Visual Form, 1997.
- [5] Thanarat Horprasert, "Computing 3D Head Orientation from a Monocular Image Sequence," Automatic Face and Gesture Recognition, 1996.
- [6] Yongduck Seo, "Calibration-Free Augmented Reality in Perspective," IEEE Transactions on Visualization and Computer Graphics, Vol.6., No.4., 2000.