

실사 영상을 사용한 얼굴 텍스처 생성

김동희[○], 윤종현, 박종승
인천대학교 컴퓨터공학과
{kdh86aa,jhyoon,jong}@incheon.ac.kr

Generating Face Textures using Real Images

Dong-Hee Kim[○], Jong-Hyun Yoon, Jong-Seung Park
Department of Computer Science and Engineering, University of Incheon

요 약

본 논문에서는 실사 영상을 사용하여 3차원 아바타의 얼굴 텍스처를 생성하는 기법을 제시한다. 먼저 UVW 맵을 기준으로 실사영상에서의 해당 영역을 수동으로 지정해 준다. 그 다음 지정된 영상 영역들을 사용하여 UVW 맵에 해당하는 텍스처 영상을 생성한다. 제안된 텍스처 생성 기법은 포토샵 등을 사용한 기존의 일반적인 방법에 비해서 수작업 시간을 단축할 수 있으며 실사 영상의 사실감을 높여 준다. 기존의 텍스처 생성 절차들은 3차원 모델의 메쉬 구조에 영향을 주었으나, 제안된 방법은 3차원 모델의 메쉬 구조를 수정하지 않고도 원하는 형태의 텍스처를 생성할 수 있는 장점이 있다. 실제 사람의 모습을 여러 각도에서 촬영하여 얻은 영상을 사용하여 주어진 UVW 맵에 적합한 얼굴 텍스처를 생성하였다. 생성된 텍스처를 사용하여 3차원 아바타를 렌더링한 결과 아바타 얼굴의 사실감이 증가되었음을 알 수 있었다.

1. 서론

애니메이션, 게임, 영화 등 엔터테인먼트 분야뿐만 아니라 교육에서도 3차원 아바타의 등장은 쉽게 찾아 볼 수 있다. 특히 최근 소프트웨어와 하드웨어의 발달로 실제 사람의 형상과 매우 유사하게 아바타를 만들 수 있게 되었다.

아바타의 생성을 위해서는 메쉬로 구성되는 3D 모델을 제작하여야 한다. 3D 모델링 저작도구로는 대표적으로 3D Max, Maya 등이 있다. 아바타의 사실감을 높이기 위해서는 메쉬의 구조뿐만 아니라 메쉬의 표면에 적용될 텍스처 역할이 중요하다. 사실적인 텍스처는 실사 영상들을 이용하여 만들 수 있다. 실사 영상으로부터 텍스처 맵을 제작하기 위해서 포토샵 등의 영상 편집 도구가 주로 사용된다. 텍스처 생성 과정으로 포토샵의 Liquify 기능을 사용하여 영상을 왜곡시키고, Blur 등의 기능을 사용하여 영상의 이음새를 부드럽게 한다. 그러나 포토샵을 사용한 텍스처 생성 방법은 수작업의 양이 많고 다중영상의 합성시의 이음새 적절히 제거해주지 못한다.

이러한 텍스처 생성시의 문제점을 해결하기 위해 실사 영상을 사용한 얼굴 텍스처 생성방법들이 연구되었다 [6][7][8]. 얼굴의 주요 특징을 표현하는 눈, 코, 입, 눈썹 등의 특징선을 지정하고 이들을 사진에 정합시키는 방법이 제안되었다[9]. 정면, 우측면, 좌측면에 대한 세장의 얼굴 사진을 사용하여 텍스처 맵을 생성하는 방법들이 연구되었다[1][2][3][4]. 서로 다른 여러 사진을 사용하는

방법들에는 사진의 합성 과정에서 이음새가 발생된다. 이음새로 인한 부자연스러운 텍스처 생성 문제를 해결하기 위해서 wavelet기반 병합 방법이 연구 되었다[1][2]. 5장의 영상으로부터 하나의 파노라믹 얼굴 영상을 생성하는 방법도 있다[5].

본 논문에서는 정면, 우측면, 좌측면의 세장의 실사 영상을 사용하여 주어진 UVW 맵에 대한 얼굴 텍스처를 생성하는 기법을 제시한다. 인접 영상들의 색 분포 정보를 사용하여 텍스처 생성 시에 발생하는 이음새를 제거한다.

본 논문의 구성은 다음과 같다. 2장에서는 일반적인 수작업의 텍스처 생성 절차에 대해서 설명한다. 3장에서는 실사 영상들로부터 텍스처 생성을 위한 제안된 기법을 설명한다. 4장에서는 다중영상들로부터 텍스처를 생성하는 실험 결과를 제시 한다. 5장에서 결론을 맺는다.

2. 실사영상으로부터 텍스처 맵의 생성

텍스처 맵은 흔히 포토샵 등의 영상 편집 프로그램을 사용하여 수작업으로 생성한다. 텍스처 맵의 생성 절차는 다음과 같다. 첫째, UVW 맵을 모델로부터 생성한다. UVW 맵은 3D Max 등과 같은 3D 모델 저작도구로부터 생성할 수 있다. 두번째, 텍스처 생성에 사용될 영상을 준비한다. 입력으로 사용될 영상은 생성되는 텍스처의 품질을 결정하므로 매우 중요하다. 영상의 획득 시의 조명에 따라 동일 물체도 다른 밝기로 나타난다. 따라서 여러 영상을 병합할 경우에 강한 이음새가 나타날 수 있

다. 세 번째, 입력영상들을 사용하여 UVW 맵에 대한 텍스처를 생성한다.

UVW 맵을 생성하기 위해서는 신체의 각 부위를 적당한 방법으로 펼치도록 한다. 펼치는 방법으로 Planar, Cylindrical, Spherical, Shrink, Box, Face, XYZ to UVW 등이 있다. 이 중에서 Planar와 Cylindrical이 흔히 사용된다. 신체의 각 부위에 대해서 일반적으로 Planar를 사용하여 펼친 UVW 맵을 적용하면 된다. 다만 얼굴 등과 같이 왜곡이 상대적으로 작아야 할 부위에 대해서는 Cylindrical을 사용하는 것이 더 적합하다. 그림 1과 같은 3차원 얼굴 모델의 메시 구조에 대해서 그림 2에서와 같이 펼치기를 위한 자른선 경계를 설정한다. Cylindrical로 펼친 UVW 맵이 그림 3에 있다.

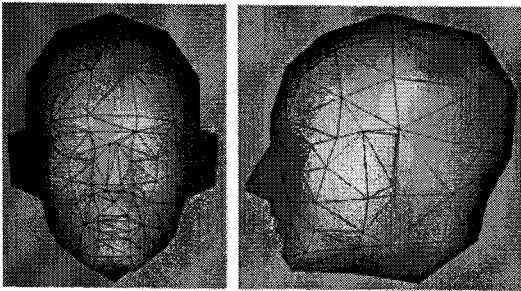


그림 1. 3차원 얼굴 모델의 메시 구조.

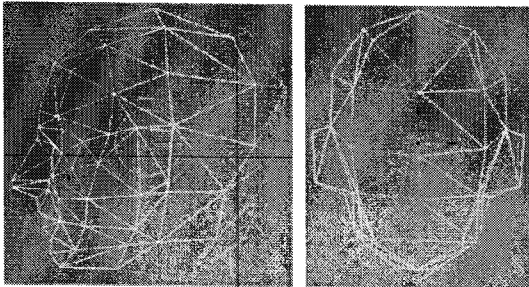


그림 2. 원통형으로 펼치기 위한 자른선 경계 설정.

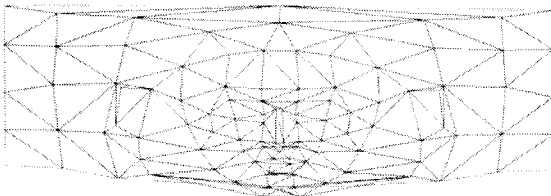


그림 3. 3차원 얼굴 메시 구조로부터 원통형으로 펼쳐서 생성한 UVW 맵.

UVW 맵이 주어지면 실제 얼굴 영상들을 사용하여 주어진 UVW 맵에 맞는 텍스처를 생성한다. 각 입력 영상에 대해서 UVW 맵 상에 겹치도록 표시한 후에 UVW 맵의 격자 구조에 맞도록 왜곡시킨다. 정면 영상에서 얼

굴의 중심선을 맞추는 후 눈, 코, 입의 위치를 맞춘다. 정면 영상에 대한 편집이 완료되면 측면 영상을 동일한 방법으로 UVW 맵 상에 겹치도록 표시하고 편집을 계속한다. 측면 영상에 경우에는 귀의 위치를 UVW 맵의 귀에 해당하는 위치에 맞추면 쉽게 정렬할 수 있다. 텍스처를 생성하는 모습이 그림 4에 있다. 주어진 모든 영상에 대한 편집이 완료된 후에도 UVW 맵에서 비어있는 공간이 있으면 이들 공간을 채워야 한다. 빈공간에 대해서는 입력 영상에서 적절한 부분의 영상패치를 따서 붙인다. 보통 머리카락 뒷부분의 영역에 빈 공간이 생길 수 있다. 그림 5는 빈 공간을 채워진 텍스처의 모습을 보여준다.

편집이 완료된 텍스처에는 각 영상 간의 경계선에 이음새가 생기게 된다. 이음새는 텍스처 품질을 크게 저하시키는 요인이므로 제거시켜야 한다. 이음새의 제거를 위해서 경계선 주변을 흐리게 함으로써 이음새가 선명히 보이지 않도록 만든다. 그림 6은 이음새를 흐리게 처리한 텍스처의 모습을 보여준다.

위에서 설명한 텍스처 생성 방법은 얼굴의 각 영역을 각각 왜곡시켜 주어야 하므로 편집을 위한 수작업의 양이 매우 많다. 또한 이음새 제거를 위한 별도의 수작업이 필요하다. 그림 7은 생성된 텍스처를 사용하여 렌더링 한 모습을 보여준다. 영상의 왜곡이 부정확하고 또한 이음새가 뚜렷이 보인다.

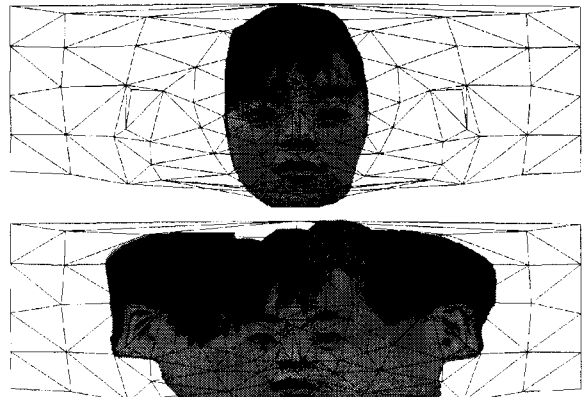


그림 4. 주어진 UVW 맵에 대하여 영상을 사용하여 텍스처를 수동으로 생성하는 모습.



그림 5. 빈 공간을 채워 완성한 텍스처의 모습.



그림 6. 이음새를 흐리게 처리한 텍스처의 모습.



그림 7. 그림 6의 텍스처를 3차원 모델에 맵핑하여 렌더링 한 모습.

최근 쉽게 텍스처를 생성할 수 있도록 하고 이음새가 제거 되도록 하는 기법들의 연구가 진행되었다. 한 예로 wavelet 기반 병합 기법[1,2]은 정면, 좌측면, 우측면의 세장의 영상의 크기와 밝기를 동일하게 조절한다. 그 다음 UVW 맵에 맞도록 각 영상을 변형시키고 변형된 영상들을 병합한다. 병합 과정에서 wavelet 변환을 사용하여 수직방향 성분을 제거시킨다. 이음새 부분의 수직 성분을 제거시키면 이음새가 흐려지는 효과가 있다. 이 기법은 영상의 변형 과정에서 특정 부분이 심하게 왜곡되는 현상이 발생할 수 있고 따라서 텍스처 일정 부분이 실제 모습과 다르게 될 수 있는 단점이 있다.

3. 영역 지정을 통한 텍스처 맵 생성 기법

본 절에서는 수작업의 양을 최소화하면서 인접 삼각형이나 인접 영상경계선에서 발생하는 이음새를 제거하는 텍스처 생성 기법을 제시한다. UVW 맵과 입력 영상들이 주어지면 이로부터 사실적인 텍스처를 생성하는 것이 목표이다.

텍스처 생성에 사용될 영상은 일반적으로 정면, 좌측면, 우측면의 세 개이다. 측면 영상이 하나만 있을 때에도 이를 대칭시켜 반대쪽 측면 영상으로 사용할 수 있으므로 두장의 영상으로부터도 텍스처를 생성할 수 있다. 입력 영상의 수가 많을수록 더 선명한 텍스처를 얻을 수 있다.

전체적인 텍스처 생성 절차는 다음과 같다.

단계 1. UVW 맵에서의 정점들에 대응되는 영상에서의 위치 지정.

단계 2. 각 영상의 칼라 분포 정규화.

단계 3. 텍스처 생성에 사용될 각 영상에서의 영역계산.

단계 4. 영상의 각 영역들을 블렌딩하여 텍스처 영상 생성.

첫 번째 단계에서는 UVW 맵과 각 영상간의 대응 관계를 지정한다. UVW 맵은 정점들과 삼각형의 에지들로 표현된다. UVW 맵에서의 영역과 대응시키기 위해서 삼각형의 정점들을 그 영역에 나타나는 영상에 지정한다. 정점에 대한 영상에서의 대응점을 지정할 때에 위치가 정확히 지정될수록 생성될 텍스처에서의 왜곡을 줄일 수 있다. 편리한 대응점 지정을 위해서 그래픽 인터페이스를 설계 및 구현하였다. UVW 맵은 정점 정보뿐만 아니라 삼각형 정보도 포함하고 있다. 영상에서는 정점들에 대한 대응점들의 지정만으로도 영상에서 대응되는 삼각형이 자동으로 생성된다. 또한 여러 입력 영상들을 돌려 보며 비교할 수 있도록 하여 UVW 맵에서 각 영역에 가장 적합한 입력 영상을 쉽게 선택할 수 있도록 하였다.

두 번째 단계에서는 각 영상의 칼라분포를 조절하여 모든 영상들의 칼라분포가 유사하도록 한다. 실사 영상 촬영시 촬영 시점에 따라서 빛의 분포가 일정치 않을 수 있고 카메라의 밝기 조정기능 등이 사용될 수 있으므로 영상들간의 밝기차가 크게 나타날 수 있다. 영상들간의 밝기 차는 텍스처를 생성할 때에 각 영상에서 사용되는 영역들의 경계선에 이음새가 심하게 나타나게 된다. 따라서 텍스처를 생성하기 전에 각 실사 영상의 칼라 분포를 정규화하는 것이 바람직하다. 입력영상이 세장인 경우, 실사 영상중 하나를 기준으로 선정하여 나머지 두개의 영상의 칼라분포를 선정된 기준영상과 비슷하게 만든다. 각 영상에 대해서 독자적으로 평활화를 수행하는 것이 아니라, 이웃 영상의 칼라분포를 영상의 칼라 분포로 변환한다.

세 번째 단계에서는 텍스처 블렌딩에 각 영상에서의 영역들을 계산한다. 각 영상에 대해서 대응점이 지정되어 있으면 해당하는 삼각형들의 집합을 결정한다. 삼각형 집합들로부터 각 영상에서 UVW 맵으로 맵핑할 영역을 계산한다. 각 영상의 맵핑 영역은 삼각형들의 합집합의 형태이므로 다각형으로 표현할 수 있고 오목(concave)의 형태 일수도 있고 볼록(convex)의 형태일 수도 있다. 영역의 경계선에서의 이음새를 제거하기 위해서는 맵핑 영역뿐만 아니라 일정한 거리의 이웃 영역까지 계산해 두어서 이음새 제거 시 이웃영역에서의 영상 정보를 사용할 수 있도록 준비해 둔다.

네 번째 단계에서는 영상의 각 영역들을 블렌딩하여 텍스처 영상을 생성한다. 텍스처 영상의 생성을 위하여 UVW 맵의 픽셀과 대응되는 입력 영상의 위치를 계산한다. UVW 맵에 존재하는 임의의 삼각형의 세 점을 $v_1 = (vx_1, vy_1), v_2 = (vx_2, vy_2), v_3 = (vx_3, vy_3)$ 이라 하고, 삼각형 안의 한 점을 $p(x, y)$ 라 한다. 이때 p 와 대응되는 입력영상에서의 점을 $p'(x', y')$ 이라고 하면 p' 의 계산은 식 (1)과 같이 한다.

$$p'(x', y') = v'_1 + t(v'_2 - v'_1) + u(v'_3 - v'_1) \quad (1)$$

여기서 v'_1, v'_2, v'_3 은 v_1, v_2, v_3 에 대응되는 입력 영상

에서 대응되는 삼각형을 이루는 점을 의미한다. u 와 t 는 UVW 맵의 삼각형인 v_1, v_2, v_3 를 사용하여 식 (2)와 식 (3)으로부터 계산된다.

$$u = \frac{(x - vx_1)(vy_2 - vy_1) - (vx_2 - vx_1)(y - vy_1)}{(vx_3 - vx_1)(vy_2 - vy_1) - (vx_2 - vx_1)(vy_3 - vy_1)} \quad (2)$$

$$t = \frac{(x - vx_1) - u*(vx_3 - vx_1)}{(vx_2 - vx_1)}$$

$$t = \frac{(vx_3 - vx_1)(y - vy_1) - (x - vx_1)(vy_3 - vy_1)}{(vx_3 - vx_1)(vy_2 - vy_1) - (vx_2 - vx_1)(vy_3 - vy_1)} \quad (3)$$

$$u = \frac{(x - vx_1) - t*(vx_2 - vx_1)}{(vx_3 - vx_1)}$$

위와 같은 과정을 거치면 이미지와 이미지상의 이음새가 부조화적인 텍스처가 생성된다. 사실감이 높은 텍스처 생성을 위해서는 이음새가 있는 부분을 제거하여야 한다. 한 텍스처를 구성하는 각 실사 영상 사이의 경계선 영역의 자연스러움을 위하여 블렌딩 과정을 거친다. 먼저, 이웃하는 두 실사 영상 사이의 경계선을 계산한다. 계산된 경계선을 기준으로 블렌딩을 위한 영역을 지정한다. 본 논문에서는 블렌딩을 위한 영역을 20 픽셀로 정의한다. 경계선이 생기는 블록 현상을 제거하기 위하여 픽셀의 주변 값들을 고려한다. 단순히 픽셀의 평균을 구하지 않고, 대상 픽셀과 인접한 픽셀일수록 가중치 적용을 통한 선형 보간을 통해 경계영역을 좀 더 자연스럽게 개선할 수 있다. 그림 8은 시스템의 전체적인 과정을 나타내고 있다.

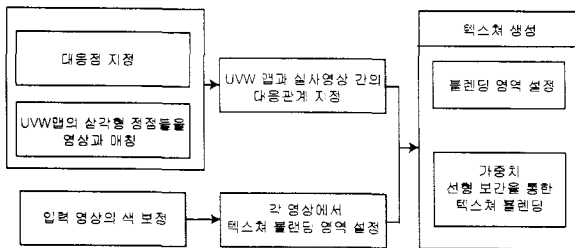


그림 8. 시스템의 전체적인 흐름도.

4. 실험 결과

3장에서 제시한 영역지정 기법을 사용하여 얼굴 영상으로부터 텍스처를 생성하였다. 그림 9와 그림 10은 각각 세장의 여자와 남자 얼굴 영상으로부터 삼각형 모양의 영역을 지정하여 텍스처 맵을 만드는 과정을 보여주고 있다. 구현된 사용자 인터페이스는 수작업의 양을 줄이기 위해서 최소한의 정점 지정만으로도 텍스처 생성이 가능하도록 구현하였다. 정점위치의 지정 시 UVW 맵의 정점에 대응되는 영상에서의 위치를 쉽게 알 수 있도록 톱플랫 텍스처를 함께 보여주도록 하였다.

정점의 지정이 완료되면 텍스처를 생성할 수 있다. 먼

저 서로 다른 영상간의 경계선을 찾고 경계선에 인접한 영역에 대해서는 색 혼합을 실시하여 이음새가 부드러워지도록 처리한다. 이음새 제거를 위해서 사용될 경계선 인접 영역이 그림 11에 있다. 인접 영역이 계산되면 인접영역을 블렌딩하여 이음새가 눈에 띄지 않도록 한다. 그림 12와 그림 13은 각각 남자와 여자의 얼굴 텍스처 생성 시 블렌딩을 적용하여 이음새를 부드럽게 한 텍스처와 그렇지 않은 텍스처를 보여주고 있다.

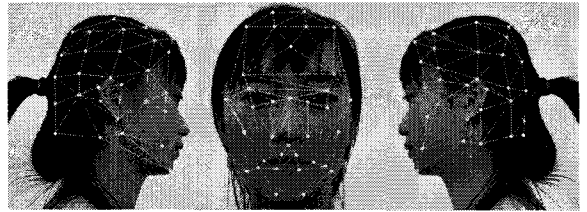
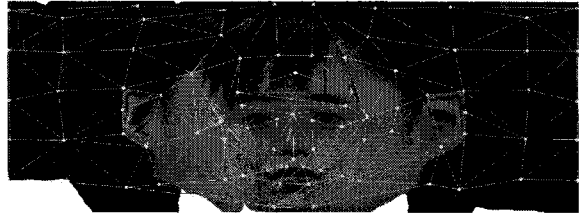


그림 9. 여자 얼굴 영상으로부터 텍스처의 생성 과정.



그림 10. 남자 얼굴 영상으로부터 텍스처의 생성 과정.

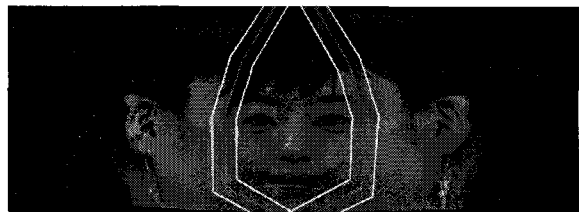


그림 11 이음새 제거를 위해서 사용할 경계선 인접 영역 계산.

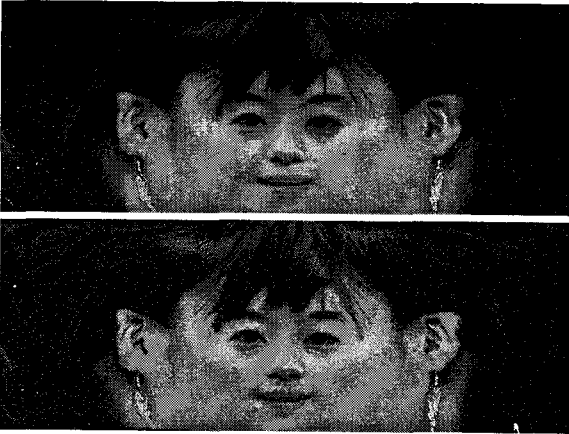


그림 12. 여자 얼굴 영상으로부터 블렌딩을 하지 않은 텍스처 맵(상)과 블렌딩을 적용시킨 텍스처 맵(하).

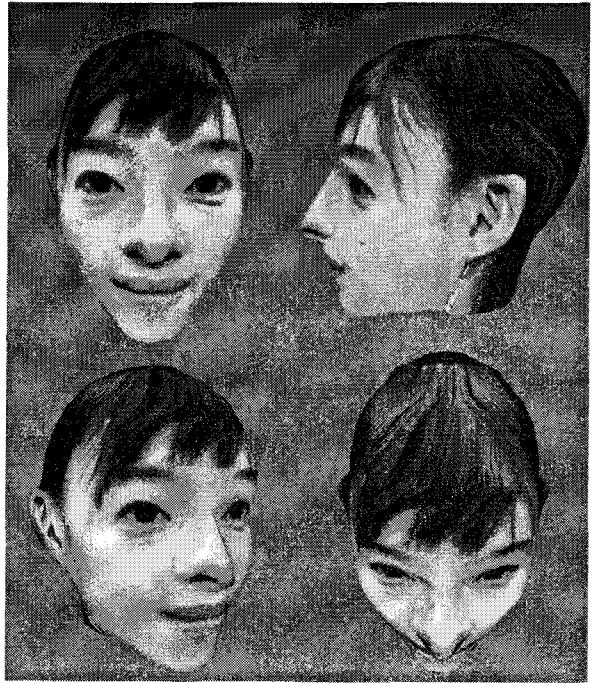


그림 14. 생성된 여자 텍스처 영상을 3D 모델에 맵핑하여 렌더링한 모습.



그림 13. 남자 얼굴 영상으로부터 블렌딩 하지 않은 텍스처 맵(상)과 블렌딩을 적용시킨 텍스처 맵(하).

생성된 텍스처 영상은 3차원 모델에 바로 맵핑하여 사용할 수 있다. 그림 14와 그림 15는 완성된 텍스처 맵을 가지고 3D 모델에 텍스처 맵핑하여 렌더링한 모습을 보여주고 있다. 본 논문에서 제시한 방법의 실험 결과는 실사를 거의 보존하면서 실사 합성으로 인한 이음새도 부드럽게 제거된 것을 볼 수 있다.

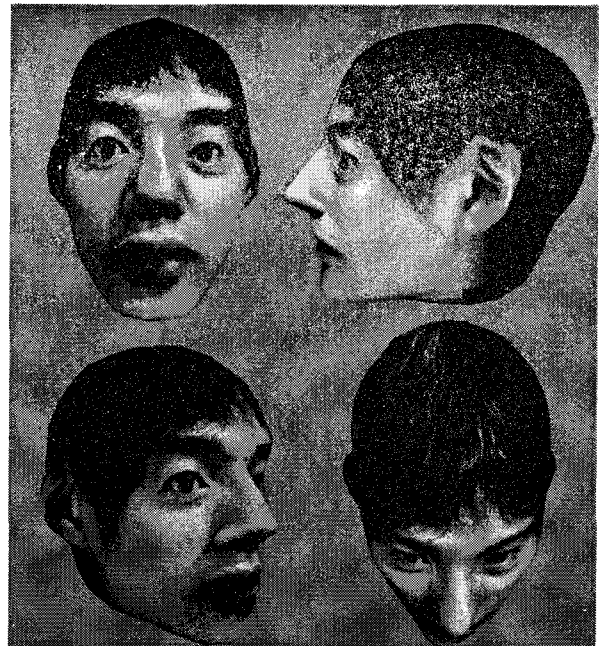


그림 15. 생성된 남자 텍스처 영상을 3D 모델에 맵핑하여 렌더링한 모습.

5. 결론

본 논문에서는 3차원 모델에 대한 UVW 맵을 생성하고 주어진 UVW 맵에 대한 얼굴 텍스처를 실사영상들로부터 생성하는 방법을 제시하였다. 기존의 수작업에 의존하는 방법보다 편집 시간을 크게 줄일 수 있었고 생성되는 텍스처의 사실감도 높일 수 있었다. 또한 텍스처를 생성할 때 이웃 영상의 경계영역에 인접한 영상 영역들을 블렌딩하여 경계영역에서 발생될 수 있는 부자연스러운 이음새가 뚜렷히 보이지 않도록 합성하는 기법을 제시하였다. 이음새는 텍스처 생성 시 이웃하는 두 영상 사이의 경계선 영역의 부자연스러움을 가중치 선형보간 방법을 사용한 블렌딩 방법을 사용함으로써 제거할 수 있었다. 텍스처 생성을 위한 각 실사 영상에는 칼라분포를 조절하여 모든 영상들의 칼라분포가 유사하도록 함으로써 촬영 시 환경적 요소에 의한 밝기 차이에 의한 문제점을 해결할 수 있었다.

제안된 방법으로 영상으로부터 생성한 얼굴 텍스처 파일은 3차원 모델의 해당 UVW 맵에 맞추어 제작되므로 추가적인 변환작업이 필요 없이 3차원 모델의 렌더링 시에 바로 사용될 수 있다. 또한 텍스처는 실 얼굴 영상의 사실감을 그대로 반영하므로 모델의 사실감을 높여준다. 아바타가 등장하는 가상환경이나 3차원 게임에서 아바타의 사실감 증대를 위해 효과적으로 활용될 수 있다.

향후 연구 과제로 주어진 UVW 맵에서 정점과 에지들을 자동 추출하여 메쉬 구조를 직접 생성하는 자동화 기법연구가 필요하다. 또한 두 인접 영상의 색보정 시 경계선 영역의 색 분포를 고려하기보다 객체의 영역을 구분하고 공통된 객체 영역에 대한 색 분포를 동일하게 변환하는 색보정 기법의 연구가 필요하다.

참고 문헌

- [1] S. Ferradal and J.C. Gomez, "TwinFaces: Seamless Textures for Rendering Head Models", IEEE 8th Workshop on Multimedia Signal Processing, pp. 176-181, 2006.
- [2] S. Ferradal and J. Gomez, "A graphical user interface for automatic facial texture mapping based on orthogonal photos", Proceedings of The 11th International Conference on Multi-Media Modeling(MMM2005), pp. 271-276, 2005.
- [3] Won-Sook Lee and Magnenat-Thalmann, N "Generating a population of animated faces from pictures", Proceedings of IEEE International Workshop on Modelling People, pp. 62-69 1999.
- [4] Won-Sook Lee, M. Escher, G. Sannier and N. Magnenat-Thalmann, "MPEG-4 compatible faces from orthogonal photos", Proceedings of Computer Animation, pp. 186 - 194, 1999.
- [5] Fan Yang, Michel Paindavoine, Hervé Abdi, "Fast Image Mosaicing for Panoramic Face Recognition", Journal of Multimedia, Vol. 1, No. 2, pp. 14-20, 2006.
- [6] W.-S. Lee and N. Magnenat-Thalmann, "Head modelling from pictures and morphing in 3D with image metamorphosis based on triangulation", Proceedings of The International Workshop on Modelling and Motion Capture Techniques for Virtual Environments (CAPTECH-98), pp. 254-267, 1998.
- [7] In Kye Park, Hui Zhang, Vladimir Vezhnevets, "Image-Based 3D Face Modeling System", EURASIP Journal on Applied Signal Processing, Vol. 13, pp. 2072-2090, 2005.
- [8] Alvin W.K.Soh, Zhang Yu, Edmond, C. Prakash, Tony K. Y. Chan and Eric Sung, "Texture mapping of 3D Human Face for Virtual Reality Environments", International Journal of Information Technology, Vol. 8, No. 2, pp.54-65, 2002.
- [9] 김항기, 김황수, "특징선 정합을 이용한 3차원 얼굴 모델링", 한국정보과학회 2000년 가을 학술발표논문집, 제27권, 제2호, pp. 505-507, 2000.
- [10] 황본우, 이성환, "적은 수의 특징점을 이용한 얼굴 영상 복원", 한국정보과학회, 2000년 봄 학술발표논문집, 제27권, 제1호(B), pp. 583-585, 2000.
- [11] 김종찬, 박경숙, 정선인, 허영남, 김용곤, "적은 수의 특징점을 이용한 얼굴 사진의 3차원 모델링 시스템", 한국정보과학회, 2001년도 봄 학술발표논문집, 제28권, 제1호(B), pp. 628-630, 2001.