

워터셰드 변환을 이용한 영상기반의 3D 얼굴 모델링

*신현실, 이상은, 장원달, 윤태수, 양황규
*동서대학교 인터넷공학부
e-mail : Spinwind75@hotmail.com

3D Face Modeling based on Image Using Watershed Transform

Hyun-Shil Shin*, Sang-Eun Lee, Won-Dal Jang, Tae-Soo Yun, Hwang-Kyu Yang
*Division of Internet Engineering, Dong-Seo University

요 약

본 논문에서는 얼굴 영상으로부터 워터셰드 변환을 이용하여 3차원 얼굴 모델을 구성하는 방법을 제안한다. 워터셰드 변환으로 분할된 각각의 영역으로부터 얼굴의 특징점들을 추출하고 MPEG-4에서 정의해놓은 FDP(Facial Definition Parameter)를 기반으로 얼굴 메쉬모델을 생성한다. 워터셰드 변환시 발생하는 영역 기반의 과분할 결과에서 얻어지는 정확한 정보와 MPEG-4의 FDP를 기반으로한 Candide Model을 이용함으로써 매우 간편하게 3D 얼굴 모델을 생성할 수 있고 영상 압축 및 전송에 매우 효율적으로 이용될 수 있다.

1. 서론

3차원의 형태로 사물의 움직임을 나타내는 3D 기술은 현실감 넘치는 최첨단 표현기술로 대두되면서 미래 핵심 산업인 영화, 게임, 디지털 콘텐츠, 인터넷 등의 산업에 미치는 효과와 부가가치가 매우 크다.

최근 영화나 애니메이션, 게임 등에서는 3D 레이저 스캔이나 모션캡처 장치를 이용하여 3D 캐릭터와 애니메이션 데이터들을 제작하는 것이 일반적인 방법으로 대두되고 있다. 이러한 장비나 기술의 한 분류로 영상을 기반으로한 3D 데이터 추출 및 활용에 대한 연구도 매우 활발히 이루어 지고 있다.

영상을 기반으로한 3D 모델링은 여러 가지 방법에 의해 제안되어져 왔으나 위치 정보와 특징점을 찾는 데 정확한 정보를 구하기가 힘들었다. 또한 전반적인 일들을 수작업을 통한 보정 또는 특정 메쉬모델의 정보로 유사 데이터를 찾아 메쉬모델을 생성함으로써 효율성이 매우 떨어졌다.

본 논문에서는 정확한 위치정보와 특징점을 찾기 위해 워터셰드 변환을 이용하였다[2]. 워터셰드 변환은

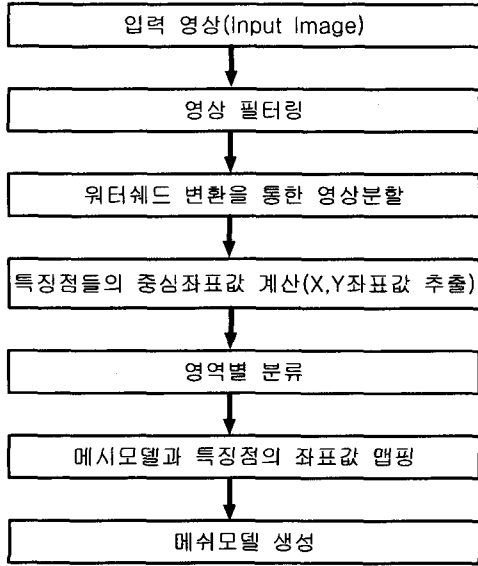
비교적 영상을 정확하게 분할하는데 영상 내부의 잡음으로 인한 다수의 국부적 최소점들 때문에 과분할 현상이 발생한다. 이러한 영상의 과분할은 얼굴의 특징영역들을 보다 정확하게 분할 할 수 있기 때문에 얼굴의 특징 점의 위치 정보를 정확하게 찾을 수 있다.

또한 MPEG-4의 FDP를 기반으로한 Candide-3 Model을 이용함으로써 보다 효율적이고 간편하게 3D 얼굴 모델을 생성할 수 있다[1]. MPEG-4를 기반으로 하고 있는 Candide-3 Model은 데이터 압축 및 전송에 매우 효율적으로 이용될 수 있다.

분할되어진 영역들의 위치 정보를 추출하기 위해서는 분할 영역에서의 중심좌표 값을 계산하였다.

이렇게 추출되어진 점들을 MPEG-4의 FDP를 기반으로한 Candide-3 Model과 매핑시켜 입력영상의 3차원 얼굴 모델을 생성한다.

제안되어진 방법의 전체 개요도는 그림 1에서 나타내고 있다.

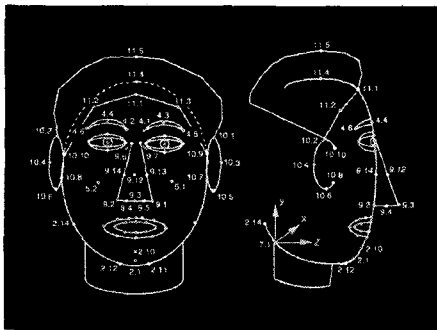


[그림 1] 전체 개요도

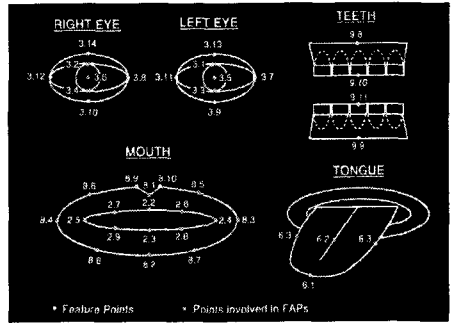
2. 얼굴모델

그림 2,3에서는 MPEG-4에서 얼굴을 움직이는 것으로 얼굴의 표정과 눈, 입의 움직임을 재현하는데 얼굴 모델을 구성하는 파라미터를 FDP(Facial Definition Parameter)라 부른다. FDP의 각 특징 파라미터를 움직이기 위하여 전송하는 움직임 파라미터를 FAP(Facial Animation Parameter)라 부른다[4]. MPEG-4에서는 FAP를 모두 MPEG-4의 파라미터로 정의하고 있으며 이들을 관련된 항목으로 통합한 10개의 그룹으로 분류하고 각 그룹 단위로 FAP를 변경할 수 있도록 되어있다.

그림 4에서 보여지는 메시 모델은 MPEG-4를 기반으로 만들어진 Canded-3 Model 이다[5].



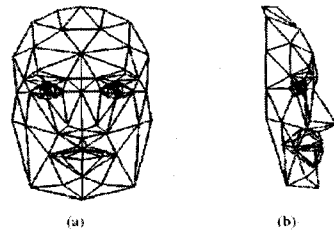
[그림 2] MPEG-4에서 정의하고 있는 각 얼굴 영역의 특징점(정면, 측면)



[그림 3] MPEG-4에서 정의하고 있는 각 얼굴 영역의 특징점(눈, 입, 이, 혀)

	그룹명	그룹의 FAP의 수
1	visemes and expressions	2
2	jaw, chin, inner lowerlip, cornerlips, midlip	16
3	eyeballs, pupils, eyelids	12
4	eyebrow	8
5	cheeks	4
6	tongue	5
7	head rotation	3
8	outer lip positions	10
9	nose	4
10	ears	4

[표 1] MPEG-4에서 정의하는 FAP의 특징점



[그림 4] Candide Model (a)정면 (b)측면

3. 워터셰드 변환을 이용한 얼굴 분할 및 특징추출
3.1 영상 필터링

전통적인 워터셰드 분할은 잡음이나 양자화 에러에 의해 많은 지역적인 최소값을 생산하는데 이것이 워터셰드 알고리즘을 수행할 때 영역을 과분할 시키므로 일반적으로 모폴로지 연산을 사용한다. 모폴로지

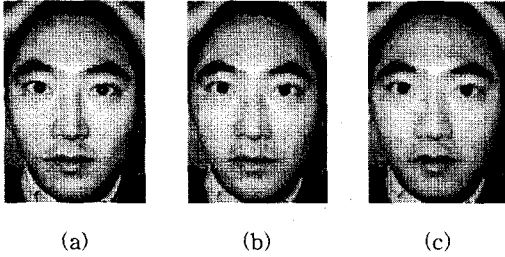
연산 중에서 팽창(dilation, 식 1)연산과 침식(erosion, 식 2)연산을 조합한 효과적인 모폴로지 연산인 제거(Opening, 식3)와 채움(Closing, 식3) 연산을 이용하여 영상 속에 있는 노이즈와 양자화 에러를 걸러낸다. 모폴로지 연산을 이용한 노이즈 제거이후 노이즈 제거에 효과적이고 널리 사용되는 Median 필터링을 통해 다시 한번 제거한다.

$$A \oplus B = \{x(B)_x \cap A \neq \emptyset\} \quad (1)$$

$$A \ominus B = \{x(B)_x \subseteq A\} \quad (2)$$

$$A \circ B = (A \ominus B) \oplus B \quad (3)$$

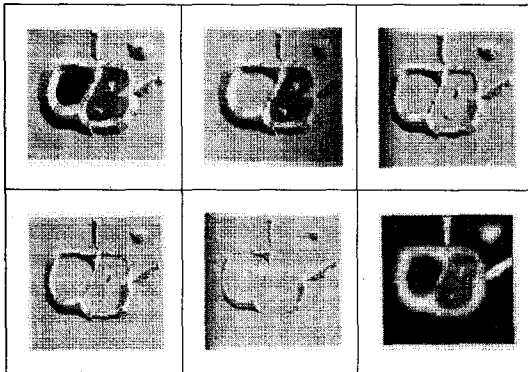
$$A \bullet B = (A \oplus B) \ominus B \quad (4)$$



[그림 5] (a)Gray영상 (b)Morphological Filtering (c)Median Filtering

3.2 워터셰드 변환

워터셰드 변환 알고리즘은 영상 처리 분야에서 인접한 영역을 효과적으로 군집화 하는 과정을 통해 영상을 분할하는 목적으로 사용되는 기법이다. 일반적으로 워터셰드 알고리즘의 기본적인 구성은 Vincent와 Soille에 의한 방법과 Meyer에 의한 방법으로 대표된다. 그림 6은 워터셰드 변환의 방법을 순차적으로 보여 준다.



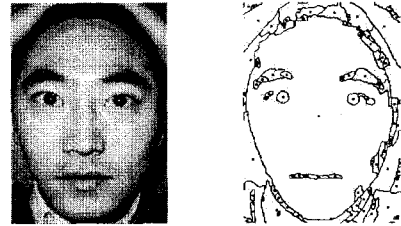
[그림 6] 워터셰드 변환 방법

영상의 명도를 하나의 위상학적 고도로 생각함으로써 위상학적 표면에 물을 떨어뜨리면 물은 낮은 고도(Minima Altitude)에서부터 흘러 들어가 높은 고도(Maxima Altitude)까지 물이 차 들어가는 원리를 이용하였다. 가장 낮은 높이에서부터 수면 높이를 점차로 증가시

키면서, 고도를 낮히는 과정으로 서로 다른 담수 영역(Catchment Basin)이 서로 접하게 될 경우 가상의 댐을 쌓아서 물이 찬 부분을 서로 분리하고, 인접한 영역의 유사성을 조사하여 유사한 영역을 서로 합치는 과정이다. 워터셰드 알고리즘에 의한 영상 분할은 영역 내부에 존재하는 임의의 점들로 구성된 국부적인 최소점들로부터 범람(Flooding)을 시작하고 담수 영역을 점진적으로 포함시켜 국부적인 최소점들을 그 영역의 경계까지 확장시킴으로써 영역 경계인 워터셰드를 추출하는 방법이다[3].

3.3 특징 점들의 중심 좌표값 추출

워터셰드 변환을 통해 분할되어진 집수지역(basin)들에 대한 정보는 다음과 같다. 집수지역 사이의 워터셰드 라인의 에지 정보와 인접한 영역에 대해서 화소의 개수, 밝기정보를 저장하고 있다. 이를 이용하여 각 영역의 중심값들을 계산하고 이를 얼굴의 특징점이라 한다. 계산되어진 중심값들은 3D 얼굴 모델링의 맵핑데이터로 이용되어진다.

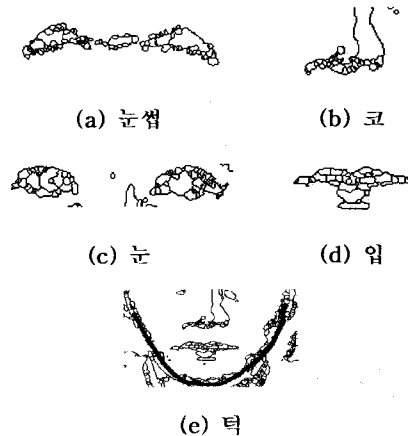


(a) 원 영상 (b) 워터셰드 변환 후 중심좌표점

[그림 7] 워터셰드 변환후 영역 중심점

4. 영역별 분류

그림 8에서는 워터셰드 변환을 통해 분할되어진 각 영역들을 얼굴영역의 특징을 나타내는 특징점으로 놓고 정규화된 모델의 그룹들 별로 구분하여 나눈다.



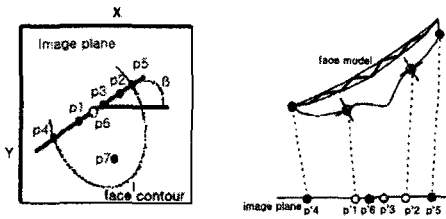
[그림 8] 분류된 각 얼굴 영역들

5. 매쉬모델과 특징점의 좌표값 맵핑

분류되어진 각 영역들의 중심점을 MPEG-4에서 정의하고 있는 그룹별로 얼굴 모델 좌표점과 맵핑시킨다. 맵핑시키기 위한 방법으로 2단계의 과정을 수행한다[6]. 첫 번째로 분할되어진 얼굴 영상의 위치와 크기, 눈과 입의 중심점을 정의되어진 매쉬 모델과 맵핑시킨다. 두 번째로 각각의 특징점을 나타내는 그룹별로 세부적인 위치와 크기를 조정하여 각 영역별로 맵핑시킨다.

5.1 분할 영상의 위치와 크기 조정

분할 영상의 위치와 크기를 조정하기 위해 얼굴 모델을 y축과 z축에 의해 회전시킨 다음 3차원 공간상의 크기와 위치를 조정한다.



[그림 9] 분할된 얼굴 특징점과 매쉬모델의 맵핑

y축 회전에서 얼굴 모델과 분할 영상의 회전 각도를 맞추고 z축 회전으로 특징점으로 분할되어진 각 영역 중 눈의 중심과 입의 중심, 턱 외곽의 특징점으로 얼굴 모델과의 위치와 크기를 조정한다(그림9). 식 5와 식 6에서 분할 영상과 모델의 y축에 대한 각각의 회전 값을 구할 수 있다.

$$r_{real} = \frac{|P_3 - P_6|}{|P_4 - P_5|} \quad (5)$$

$$r_{model} = \frac{|P_3 - P_6|}{|P_4 - P_5|} \quad (6)$$

식 7,8,9에 의해 각각 x축, y축, z축으로 얼굴 모델의 크기를 조절한다.

$$S_x = \frac{|P_1 - P_2|}{|P_1 - P_2|} \quad (7)$$

$$S_y = \frac{|P_3 - P_7|}{|P_3 - P_7|} \quad (8)$$

$$S_z = \frac{1}{2} (S_x + S_y) \quad (9)$$

5.2 세부 그룹별 영역의 중심점 맵핑

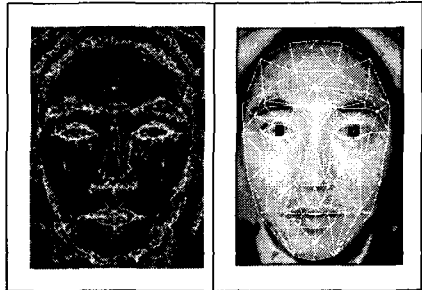
얼굴 모델의 위치와 크기를 조절한 다음 얼굴 모델의 각 그룹별 특징점들을 그림 10의 방법으로 분할 영역의 중심점들과 맵핑시킨다.



[그림 10] 입 영역의 중심 값 맵핑

6. 실험 결과 및 고찰

본 논문에서 제안하고 있는 방법은 PentiumIII CPU 시스템에서 Visual C++와 OpenGL Library를 이용해 구현하였다. 워터셰드 변환을 이용한 얼굴 영역의 특징 추출은 영역 기반의 영상 분할이라는 점에서 특징 영역의 중심점을 쉽게 계산 할 수 있었다. Candide 모델은 OpenGL의 GLM 클래스를 이용하여 Object data 파일로 저장하였다. 얼굴모델의 에지와 점들은 각각 컨트롤 할 수 있고 영상의 얼굴 모델로 변형되어진 정보를 Object data 파일로 재 저장할 수 있다. 또한 특징 영역의 점들과 비교적 쉽게 매핑을 할 수 있었다. 그림 11은 워터셰드 변환에서 찾아진 영역의 중심점으로 얼굴 모델을 조정후 원 영상과 매핑시킨 결과이다.



[그림 11] 결과 영상

7. 결론

본 논문에서는 각 얼굴 영역 특징점들을 워터셰드 변환의 과분할 결과를 이용하여 보다 정확하게 찾아내었다. 또한 MPEG-4에서 정의하고 있는 각 특징점들을 적용한 Candide 모델을 사용함으로써 영상 압축과 전송의 효율성을 개선하여 3차원 얼굴 모델을 생성하였다. 향후 MPEG-4에서의 FAP정보를 이용한 자동 애니메이션의 연구로 확장하고자 한다.

8. 참고 문헌

[1] L. Zhang, "Automatic adaptation of a face model using action units for semantic coding of videophone sequences", IEEE Tran.on Circuits Syst. Video Tech, 8(6), October 1998, pp. 781-795.
 texture-based estimation texture and optical flow-based estimation
 [2] W.D. Jang, T.S. Yun, and H.K. Yang, "Image Segmentation Based on Watershed Transform by Image Reconstruction," Proc. of KIPS Fall Conference, vol. 9, No. 2, pp. 637-640, Nov, 2002
 [3] 최영규, 이상범, "An Efficient Character Image Enhancement and Region Segmentation Using Watershed Transformation", 한국정보처리학회 논문지B, vol.9-B NO.04, pp. 0481 ~ 0490 2002. 08.
 [4] 미키 스케이치, " 다양한 영상,음성을 자유자재로 부호화하는 MPEC-4의 세계 ", 대영사, pp 100-101,1999.
 [5] http://www.bk.isy.liu.se/candide
 [6] Markus Kampuann, "Automatic 3-D Face Model Adaptation for Model-Based Coding of Videophone Sequences", IEEE Tran. on Circuits Syst. Video Tech, vol. 12, NO.3, March 2002