

다해상도 3D 얼굴 모델의 압축

박동희* · 이종석* · 이영식** · 배철수*

*관동대학교, **경동대학교

Multiresolution 3D Facial Model Compression

Dong-hee Park* · Jong-seok Lee* · Young-sik Lee** · Cheol-soo Bae*

*Kwandong University, **Kyungdong University

E-mail : easthee1@chollian.net

요 약

본 논문에서는 효율적인 압축 기법과 멀티미디어를 위한 다 해상도 3D 얼굴 모델 전송, 그리고 저비트율 응용에 대해 제안하고자 한다. 일반적으로 얼굴 모델은 3D 레이저 디지털화에 의해서 얻어지게 되고 애니메이션, 비디오 게임, 비디오 회의와 같은 응용 범위에 따라 여러 해상도로 재양자화 되어진다. 3D 디지털화된 얼굴 모델을 정합하고 재양자화 하기 위해서 2D 템플릿을 변형함으로써 압축 모델을 얻을 수 있다. 현재까지의 연구에서 다섯 가지 해상도로 계층적 2D 얼굴 와이어프레임 템플릿을 만들었다. 변형 과정에서 2D 템플릿은 얼굴 특징점과 제안된 PCAT(piecewise chainlet affine transformation)에 의해 바뀌게 된다. 재양자화된 후 3D 디지털화된 모델은 인지하지 못할 정도로 손실이 줄어들게 된다. 더욱이, 본 논문에서 제안한 계층적 데이터 구조를 갖는 다 해상도 얼굴 모델은 통신망에서 점진적으로 알려지고 사용되어질 것이다.

ABSTRACT

In this paper, we proposed an approach to efficiently compress and transmit multiresolution 3D facial models for multimedia and very low bit rate applications. A personal facial model is obtained by a 3D laser digitizer, and further re-quantized at several resolutions according to different scope of applications, such as animation, video game, and video conference. By deforming 2D templates to match and re-quantize a 3D digitized facial model, we obtain its compressed model. In the present study, we create hierarchical 2D facial wireframe templates are adapted according to facial feature points and the proposed piecewise chainlet affined transformation(PACT) method. The 3D digitized model after requantization are reduced significantly without perceptual loss. Moreover the proposed multiresolution facial models possessed of hierarchial data structure are apt to be progressively transmitted and displayed across internet.

키워드

PCAT, OESL, ITRI

1. 서 론

MPEG-4 표준 SNHC(Synthetic/Natural Hybrid Coding)는 자연 영상, 합성된 영상을 포함한 가상 객체가 각각 압축되고, 코딩되고, 전송되어야 한다고 제안하고 있다[1]. 멀티미디어 통신과 저비트율 응용에 있어서 사람의 얼굴과 신체모델의 합성객체가 가장 보편적이다. 본 논문에서는 얼굴 모델을 위한 효율적인 압축과 전송, 표현방법에 대하여 제안하였다.

멀티미디어 환경에서 다양한 응용 분야를 조사하는 것은 예를 들어 컴퓨터 애니메이션에서 얼굴 모델은 자체 충실도를 보존하는 고 해상도로 나타낼 수 있고, 사람과 컴퓨터 환경, 비디오 회의에서 얼굴 모델은 낮은 충실도를 갖는 중간 해상도로 나타낼 수 있다. 더욱이 실시간과 양방향 환경, 즉, 인터넷/비디오게임과 가상 현실(Virtual Reality)과 같은 환경에서 얼굴 모델은 가장 낮은

해상도로 나타낼 수 있고, 받아들여질 수 있다. 이런 응용분야에서 얼굴 모델의 데이터는 계산량을 줄이기 위해서 확실히 감소함으로써 다양한 여러 응용 분야에 이용하기 위해서 다중 해상도 얼굴 모델 압축에 관한 접근법을 제안하고자 한다.

본 논문에서는 얼굴 모델[2]로부터 데이터 영역을 획득하기 위해서 3D 레이저 디지털라이징 시스템을 이용한다. 디지털화 된 모델은 보통 수 천개의 영역들에서 고찰한 데이터의 정점들로 이루어져 있다. 대부분의 정점들은 중복되어 있으며 감소시켜야 한다[3]. 예를 들어 얼굴모델에서 이마와 뺨의 평평한 표면은 매우 많이 샘플링 되어진다. 이런 중복된 정점들을 줄이기 위한 연구는 표면의 반복적 재분할[6,7], 다 해상도 웨이블릿 분석[8,9], 강제적 Delaunay 피라미드[10], 계층적 삼각분할 표면[11], 그리고 삼각 십진 알고리즘[12]을 통해서 이루어지게 된다. 이러한 접근법들은 압축율을 낮추었지만 저 해상도에서 인식이 떨어지게 되었고, 얼굴 특징들, 즉 눈, 코, 입 등은 손상을 입게 되고 또 얼굴 특징으로써 인식 못할 수도 있다. 얼굴 골격과 근육을 고찰할 때 [13], 다섯 가지 해상도로 계층적 2D 얼굴 와이어프레임 템플릿을 구성하고, 각 템플릿마다 고품질을 갖는 얼굴 특징들을 관리한다. 2D 템플릿은 3D 디지털화된 얼굴 모델과 매칭하기 위해서 변형하게 되고 변형된 템플릿은 이 모델을 양자화하기 위해서 사용된다.

2D 얼굴 템플릿은 템플릿이 변형하는 동안 영역 윤곽선의 뒤틀림 부분을 보호하기 위해 제안된 PCAT방법에 의해 변형되며 이 방법은 영역에서 얼굴 템플릿을 분할한다. 각 영역 안의 정점들은 각 영역의 윤곽선에 의해 제한되고 적용되어지며 다른 영역의 변형에 의해서는 영향을 받지 않는다. 예를 들어 템플릿의 코 부분이 변형되어질 때 입 부분의 영역은 영향을 받지 않게 된다.

저 대역폭 채널을 통해 얼굴 모델 데이터를 효율적으로 전송하기 위해서, "단계의 세분화" 그리고 "향상된 디스플레이와 전송"[8] 특성을 지닌 계층적 데이터 구조를 사용하게 된다. 즉, 얼굴 모델의 데이터는 저, 중, 고 해상도 정점에 의해서 연속적으로 전송된다. 수신기에서, 복잡한 얼굴 모델은 우선 재구성되고 추가적으로 들어오는 고해상도의 데이터에 의해 점진적으로 향상되어진다. 각각의 건본 얼굴에 대한 다섯 가지 해상도의 압축률과 재양자화된 얼굴 모델의 투영도를 논의하였다. 본 연구에서 제안한 방법을 10회에서 1440회까지 압축률을 실험하였다.

II. 다해상도 얼굴 모델 압축 방법

다 해상도 얼굴 모델 압축 절차는 다음과 같다. 우선 계층적 2D 템플릿을 구성하고 특징점 추출, 디지털화된 모델의 윤곽선 추출, 자동 윤곽선 고정, 템플릿 변형 과정, 그리고 디지털화된

3D 모델로부터 변형된 각각의 정점의 깊이를 찾게 된다.

2.1 다 해상도 2D 템플릿

그림1(a)의 "남자" 얼굴은 2D 템플릿을 구성하기 위해서 사용한다. 다른 분야에 적용하기 위해서 이 얼굴 모델로부터 다섯 가지 해상도 즉, 고, 중, 저, 중저, 극저(그림2(a-e))로 템플릿을 구성한다. 2D 템플릿의 와이어프레임 구성은 그림1(b,c)와 같이 얼굴의 국부 특성에 대한 표면 분석을 함으로써 구성된다.

템플릿을 구성하는 동안, 중요한 얼굴 특징들과 부드러운 얼굴 표면 모두를 고려하게 된다. 고, 중, 저 해상도 템플릿에서 중요한 얼굴 특징 즉, 눈, 코, 입은 그림 2(a-c)에서와 같이 고해상도로 유지해야 한다. 이마, 안구, 볼, 턱과 같은 부드러운 얼굴 표면은 그림2(a-e)와 같이 계층적인 망에 의해서 점진적으로 다운-샘플링된다. 즉, 저 해상도 템플릿에서 다각형은 여러 다각형을 합치고 고해상도 템플릿의 정점들을 제거함으로써 구성되어진다.

중저 템플릿은 중요한 얼굴 특징들과 다각형을 더 줄이게 된다. 극저 템플릿은 단지 복잡한 얼굴을 표현하기 위해서 필수 정점들과 다각형을 유지한다.

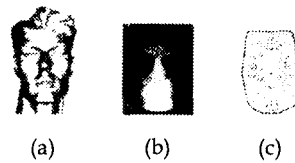


그림1. (a) 3D 디지털화된 모델 (b) 영역 이미지 (c) 윤곽선

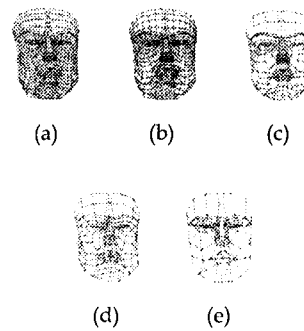


그림2. 얼굴 템플릿

(a) 고해상도 (b) 중간 해상도 (c) 저해상도 (d) 중저해상도 (e) 극저해상도

2.2 얼굴모델 디지털화와 특징점 추출

제안한 방법을 시험하기 위해서 그림4와 같은 모델을 사용한다. 얼굴은 OESL/ITRI[2]을 이용하

여 3D 디지털타이저의 플랫폼에서 0.1mm해상도로 스캔하였다. 디지털 데이터는 그림3에서와 같이 얼굴 영역을 가로지르는 다중 선의 형태로 재구성된다. 특징점들의 2D 좌표열과 디지털화된 얼굴 윤곽은 그림4와 같이 디지털화된 모델로부터 추출한다. 특징점들은 머리와 얼굴로 제한된다 [14,15].



그림3



그림4

그림 3. 디지털화된 다중선

그림 4. 얼굴 특징점과 추출된 윤곽선

2.3 템플릿 정합에 의한 얼굴모델 재양자화

특징점과 얼굴 윤곽에 의한 다 해상도 템플릿들은 3D 디지털 모델의 2D 투영을 정합하기 위해 변형된다. 인종, 나이, 성과 같은 요인들은 개인간 상당한 차이를 나타낸다[14]. 다양한 얼굴 모델을 수용하기 위해서 정교하게 템플릿을 변형하는 것은 어려운 과제이다.

본 논문에서 템플릿은 얼굴 윤곽과 국부적인 얼굴 특징점에 의해서 변형된다.

2.3.1 자동 얼굴 윤곽 고정

얼굴 윤곽 C_M 은 디지털 모델로부터 추출된다. C_T 와 비교해서 이산 템플릿 윤곽 C_M 은 연속적이라고 간주될 수 있다. C_M 과 C_T 는 다음과 같다.

$$C_M = r_M(\theta_j), \quad j=1, \dots, m \quad (1)$$

$$C_T = r_T(\theta_i), \quad i=1, \dots, n \quad (2)$$

여기서 $r_M(\theta_j)$ 는 각 함수이고 θ_j 는 윤곽의 j 번째 점의 각이다. 그리고 $m \gg n$ 이다.

C_T 와 C_M 간의 자동 윤곽고정은 각이 θ_i 에 가장 가까운 점을 찾기 위해서 C_M 에 있는 점들을 조사함으로써 이루어지게 된다.

$$C'_M = r_M(\theta_j \approx \theta_i), \quad j=1, \dots, n \quad (3)$$

여기서 C'_M 은 고정된 윤곽이고 그 결과는 그림5와 같다.

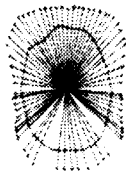


그림 5. 자동 윤곽 고정

2.3.2 PCAT에 의한 자동 템플릿 변형

본 논문에서는 국부 템플릿 변형을 수행하기 위해서 PCAT 방법을 제안한다. PCAT는 각각의 특징점을 결정하기 위해 제공된 얼굴 모델의 국부 윤곽 곡선을 유지하는 성능을 가지고 있고, 그리고 변형 과정을 수행하는 동안 PCAT는 다른 얼굴 영역의 왜곡없이 국부적으로 템플릿을 조작할 수 있다. 템플릿 변형은 윤곽영역을 변형하는 것과 각 영역속의 정점을 조작하는 두 단계로 이루어진다.

우선 국부 얼굴 영역, 즉 이마, 눈, 코, 입, 뺨, 턱, 안구(그림6) 등에서 2D 템플릿을 분해한다. 각각의 윤곽영역은 여러 개의 불연속 체인으로 구성되며, 이들은 그림7처럼 템플릿의 체인 정점들로 이루어져 있다. 그림8과 같이 윤곽에 각 체인 두 끝 정점은 그림10처럼 PCAT 방법에 의해 다른 체인의 정점을 조작하기 위한 제어 점으로 쓰이게 된다.



그림6



그림7

그림 6. 얼굴 분할 그림 7. 체인렛 구조



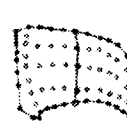
(a)



(b)

그림8. 영역 윤곽 변형

(a) 이마 템플릿의 윤곽 (b) 변형된 윤곽



(a)



(b)

그림 9. 각 영역에서 정점의 변형된 템플릿 (a) 이마 템플릿 (b) 변형된 템플릿

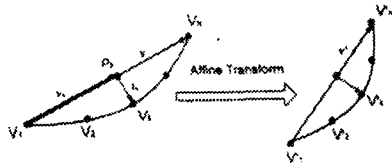


그림 10. PCAT 방법

다음 단계로 각 영역의 와이어프레임은 국부 체인으로 분해된다. 각 체인의 양끝은 윤곽영역에 존재하고 그림9와 같이 제어 점으로 쓰이게 된다. 그리고 이 체인들의 정점을 조작하기 위해서 PCAT 기법을 사용하게 된다. 이 PCAT 기법은 다음과 같다.

- I) 템플릿의 체인 C는 n개 $V_1, V_2, \dots, V_k, \dots, V_n$ 의 체인 정점으로 구성된다. 정점의 끝 V_1 과 V_n 은 그림10(a)와 같이 제어 점으로 쓰인다. 변형된 템플릿의 응답 체인 C'에서 그림10(b)와 같이 제어 점으로 쓰이는 두 끝 정점 V'_1 과 V'_n 을 갖게 된다.
- II) C'에서 각 임의의 정점 $V'_k, k=2\dots n-1$ 에 대한 좌표를 얻을 수 있다. 과정은 다음과 같다.

체인 C 또는 C'에서 벡터 v나 v'는 다음과 같이 정의한다.

$$v = V_n - V_1 \quad (4) \quad v' = V'_n - V'_1 \quad (5)$$

벡터 v_k 는 다음과 같다.

$$v_k = P_k - V_1 \quad (6)$$

여기서 P_k 는 그림10에서 v위에 있는 V_k 투영점이다. 그리고 벡터 t_k 는 다음과 같다.

$$t_k = V_k - P_k \quad (7)$$

이러한 관계식에서 다음과 같은 스�কে링 인수를 갖게 된다.

$$s_1 = |v'|/|v| \quad (8) \quad s_2 = |v_k|/|v| \quad (9)$$

그리고 요구되는 정점은 식10과 같다.

$$V'_k = V'_1 + s_2 v_k + s_1 t_k \quad (10)$$

2.3.3 3D 디지털 모델에 변형 템플릿을 정합

전 단계에서는 3D 디지털모델의 2D 투영으로 쓰이는 변형 템플릿을 얻었다. 디지털 모델에 변형 템플릿을 정합하기 위해서는 그 자체의 깊이를 구해야 완전한 3D 얼굴모델을 재구성할 수 있다. 변형 템플릿의 각 정점 $V(x,y)$ 에서 그림11에 나타나 있는 4개의 정점 $V_1(x_1,y_1,z_1), V_2(x_2,y_2,z_2),$

$V_3(x_3,y_3,z_3), V_4(x_4,y_4,z_4)$ 를 구하기 위해 디지털 모델의 점 (x, y) 에 가장 가까운 2D 좌표를 갖는 정점을 조사한다. V의 깊이는 다음과 같은 양선형 변환에 의해서 계산된다.

먼저 x와 y 방향에 있는 w_1 과 w_2 를 계산한다. 그러면 각각 아래 식과 같다.

$$w_1 = |y - y_1| / |y_3 - y_1| \quad (11)$$

$$w_2 = \frac{|x - x_1 - (x_3 - x_1)w_1|}{|x_2 - x_1 - (x_3 - x_2 - x_3 + x_1)w_1|} \quad (12)$$

V에 대한 깊이 z는 식13과 같다.

$$z = z_1 + (z_3 - z_1)w_1 + (z_2 - z_1)w_2 + (z_4 - z_2 - z_3 + z_1)w_1 w_2 \quad (13)$$

2.4 계층적 데이터 구조

다 해상도 얼굴 모델을 효율적으로 전송하고 품질을 향상시키기 위해서 계층적 데이터 블록 구조를 재배치한다. 저, 고 해상도에서 얼굴 특징 영역의 정점, 다른 얼굴 영역의 저, 중, 고 해상도 와이어프레임. 각각의 개체는 인코더와 디코더에서 똑같은 구조를 갖게 되고, 단지 정점들만이 전송된다. 채널을 통해서 비트스트림을 연속적으로 전송하고 수신기에서 복잡한 얼굴모델은 우선 데이터를 가지고 구성되고 고해상도의 정점을 추가함으로써 다듬어지게 된다. 그리고 다 해상도 특징을 갖는 계층적 얼굴 모델은 애니메이션이나 가상현실, 비디오게임과 같은 다양한 목적에서 다른 해상도로 바꾸기가 편리하다.

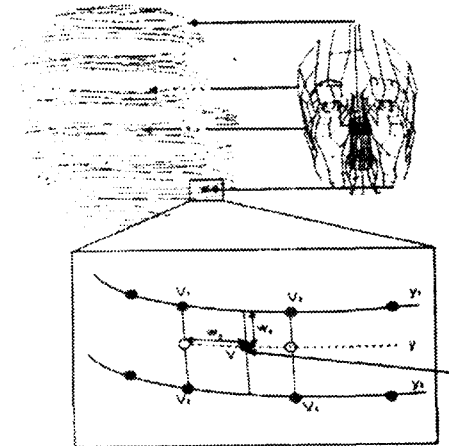


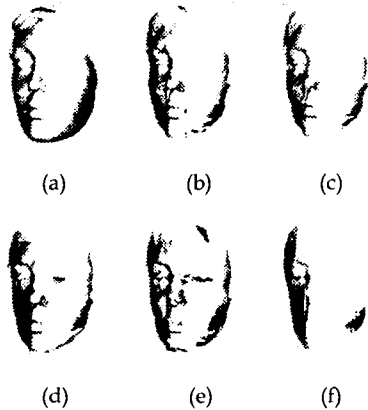
그림 11. 3D모델에 2D템플릿 정합

III. 실험 결과

본 논문에서는 여러 개의 정면 얼굴 석고상을 가지고 테스트하였다. 압축률은 표1과 같다. 실제

디지털 모델을 가지고 투시 모델을 비교하였다. 그림13과 같이 고해상도에서 얼굴모델의 압축 버전을 나타내는 것은 실제 모델과 같게 인식되어지며, 저해상도에서 핵심 얼굴 특징들은 보호된다. 극저해상도에서 113개의 정점만이 얼굴 특징들이 덜 손상입은 완전한 얼굴 형태를 보호하는데 필요하다.

소 정점만으로도 손상을 입지 않은 얼굴형태를 얻을 수 있었다. 템플릿은 제안된 PCAT 방법에 의해서 변형되며, 이 방법에서 각 체인의 국부 곡선은 모델의 회전, 이동 등의 변수임을 알 수가 있었다 또한, 다양한 응용분야에서 다른 해상도로 쉽게 변환될 수 있으며, 점진적으로 이용되어질 수 있을 것이다.



(a) 3D모델의 원영상 (b) 고해상도 (c)저해상도 (d) 저해상도 (e) 중저해상도 (f) 극저해상도
그림 13. 실험결과

Ratio	Compressed Model					
	High	Middle	Low	Very	Ultra	
3D Models	904	690	436	264	113	
A	9247	10.2	13.4	21.2	35.0	81.8
B	11842	13.1	17.2	27.2	44.9	104.8
C	988	10.1	13.2	20.8	34.4	80.4
D	9520	10.5	13.8	21.8	36.1	84.2
E	162814	180.1	236.0	373.4	616.7	1440.8
F	137659	152.3	199.5	315.7	521.4	1218.2

표1. 테스트 모델의 압축률

V. 결 론

본 논문은 멀티미디어를 위한 다 해상도 3D 얼굴 모델 압축과 압축 모델을 효율적으로 묘사하기 위해서 계층적 데이터 구조를 제안한다. 3D 디지털 얼굴 모델들은 많은 중복 샘플들로 구성되어 있는데, 디지털 3D 모델의 양자화에서 다 해상도 템플릿을 적용함으로써 애니메이션이나 가상현실, 비디오게임 등과 같은 다양한 분야에서 다른 해상도로 바꾸기가 편리하게 됨을 알 수가 있었으며, 고해상도에서의 얼굴모델은 실제모델과 같게 인식되어지며 극저해상도에서는 113개의 최

참고문헌

- [1] "MPEG-4 Overview (Tokyo Version), Coding of moving pictures and audio, ISO/IEC JTC1/SC29/WG11 N2196, Mar. 1998.
- [2] B.T. Chen, W.S.Lou, C.C.Chen, and H.C.Lin, "3D Digitizer: Method and Analysis," in *Proc. CVGIP, Taiwan*, pp. 406-418, 1997.
- [3] M.Deering, "Geometry Compression," in *Proc. Computer Graphics (SIGGRAPH)*, pp. 13-20, 1995.
- [4] T. Sakguchi, M. Ueno, S. Morishima, and H. Harashima, "Analysis and Synthesis of Facial Expression Using High-Definition Wire Frame Model," in *Proc. IEEE Int'l. Workshop on Robot and Human Communication*, pp. 194-199, 1993.
- [5] Y. Lee, D. Terzopoulos, and K. Waters, "Realistic Modeling for Facial Animation," in *Proc. Computer Graphics(SIGGRAPH)*, pp. 55-62, 1995.
- [6] H. T. Tanaka and F. Kishino, "Adaptive Mesh Generation for Surface Reconstruction," in *Proc. IEEE Computer Vision and Pattern Recognition*, pp. 88-94, 1993.
- [7] M. Vasilescu and D. Terzopoulos, "Adaptive Meshes and Shells: Irregular Triangulation, Discontinuities, and Hierarchical subdivision," in *Proc. IEEE Computer Vision and Pattern Recognition*, pp. 829-832, 1992.
- [8] M. Eck, T.D. DeRose, T. Duchamp, H. Hoppe, M. Lounsbery, and W. Stuetale, "Multiresolution Analysis for Arbitrary Meshes," in *Proc. Computer Graphics.(SIGGRAPH)*, pp. 173-182, 1995.
- [9] M. Lounsbery, T.D. DeRose, and J. Warren, "Multiresolution Analysis for Surface of Arbitrary Topological Type," *ACM Transactions on Graphics*, vol. 16, pp. 34-73, Jan. 1997.
- [10] A. Voigtmann, L. Becker, and K. Hinrichs, "A Hierarchical Model for Multiresolution Surface Reconstruction," *CVGIP: Graphic Model and Image Processing*, vol. 59, pp. 333-348, Sep. 1997.

- [11] L.D. Floriani and E. Puppo, "Hierarchical Triangulation for Multiresolution Surface Description," *ACM Transactions on Graphics*, vol. 14, pp. 363-411, Oct. 1995.
- [12] W.J. Schroeder, J.A. Zarge, and W.E. Lorensen, "Decimation of Triangle Meshes," *Computer Graphics*, vol. 26, pp. 65-70, July, 1992.
- [13] A.J. Fridlund, *Human Facial Expression*, Academic Press, 1994.
- [14] L.G. Farkas, *Anthropometry of the head and Face*, 2nd ed., Raven Press, 1994.
- [15] C.J. Kuo, R.S. Huang, and T.G. Lin, "Synthesizing Lateral Face from Frontal Facial Image Using Anthropometric Estimation," in *Proc. ICIP*, vol. 1, pp. 133-136, 1997.
- [16] RC. Gonzalez and R.E. Woods, "Representation and Description," in *Digital Image Processing*, Addison-Wesley, 1993.