

픽셀 수와 정점들 간의 현격한 차이를 보완하는 Morphable 모델 Morphable model to interpolate the difference between the number of pixels and the number of vertices

고방현, 홍태화*, 이종원, 문현준, 김용국, 문승빈
세종대학교, 삼성전자 정보통신총괄 통신연구소*

Ko Bang-Hyun, Hong Tae-Hwa*, Lee Jong-Weon, Moon Hyeon-Joon, Kim Yong-Guk, Moon Seung-Bin
Sejong Univ., Samsung Electronic Telecommunication Network Business*

요약

본 논문은 GeometriX장비를 통해 획득된 데이터를 가공하여 Morphable 모델을 생성한다. 기존의 MPI[3]에서 만들어진 Morphable 모델은 Cyberware[10]장비로 획득된 가공하지 않은(raw) 데이터를 사용하였기 때문에 픽셀의 수와 정점의 수가 같아 원칙적으로 대용량의 데이터를 가진다. 이를 해결하고자 모델 생성에 참조되는 픽셀의 수가 정점의 수 보다 월등히 많은 모델을 사용한 Morphable 모델을 생성한다.

Abstract

This paper presents Morphable Model Construction which is based GeometriX tool processing various data array such as texture space and geometry(shape space) in order to reduce calculation cost due to rapid advancement of face recognition speed. It introduces efficiently vertex and pixel reduction for raw data, which is based GeometriX tool using stereo scan.

I. 서론

영상 정보 중에서 이미지상의 얼굴(face)은 육안으로 매우 쉽게 식별할 수 있다. 하지만, 카메라를 사용하여 얼굴 식별 및 인식하는 비전시스템은 여전히 사람과 비교하여 부정확하고 시간과 비용이 많이 소모된다. 초기에 2차원 이미지로 얼굴 인식을 하였으나 지금은 3차원 기반의 Morphable 모델[3]을 사용한 단계로 발전했다.

하기 위해서 우선 이미지들의 두 눈의 각 중점과 입의 중점을 기록하여 목록(파일 Geye.txt)을 만든다.



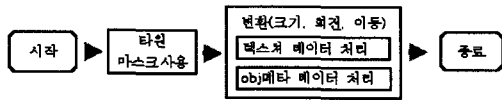
▶▶ 그림 1. GeometriX로 획득된 비가공데이터

II. 전처리

GeometriX장비로 스캔하여 두 개의 파일을 생성한다. 첫 번째는 얼굴 굴곡(face geometry)과 텍스처(texture) 참조정보를 가진 obj 파일이다. 두 번째는 jpg 확장자를 가지는 텍스처 파일이다. 가공하지 않은 데이터는 Morphable 모델 생성에 불필요한 데이터를 가졌기 때문에 데이터 처리에 대용량이 필요하다. 전처리를 통하여 필요한 부분만 사용하기 위해서 먼저 이미지들의 두 눈의 각 중점과 입의 중점을 기록하여 목록(파일 Geye.txt)을 만든다.

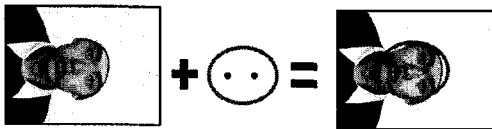
GeometriX장비로 스캔한 가공하지 않은 데이터는 Morphable 모델 생성에 불필요한 데이터를 가져서 자료 처리에 대용량을 요구한다. 전처리를 통하여 필요한 부분만 사용

또한, 여러 사람은 각각 자세와 중심위치가 조금씩 틀리다. Morphable 모델은 최종적으로 이들을 하나로 합쳐서 동일 위상의 정점 간에 평균을 구한다. 그래서 인물들은 확대, 축소, 이동 그리고 회전등의 변환을 통하여 같은 크기와 시선을 가질 필요가 있다. 작업과정 중에 obj파일에서 정점데이터와 텍스처 참조데이터 그리고 정점 토폴로지 내용도 병행하여 변경한다.



▶▶ 그림 2. 가공하지 않은 데이터를 전처리

데이터로 획득된 이미지파일(jpg 형식)에서 두 눈과 입술의 좌표를 구한다. 이때 왼쪽 눈, 오른쪽 눈 그리고 입술은 중심점 위치를 잡는다.



▶▶ 그림 3. 마스크와 변환을 사용한 전처리

- A. 누워있는 원본 이미지를 읽는다.
- B. 두 눈 사이 좌표와 입술 중앙 점을 타원의 두 초점으로 둔다.
- C. 가상의 타원을 그리고 내부의 각 픽셀을 읽는다.
- D. 각 픽셀을 축소 이동 회전을 적용하여 세운다.
- E. C'D까지의 명령을 타원 내부의 모든 픽셀에 적용한다
- F. A'E까지의 명령을 Geye.txt파일에 있는 인물 목록 모두에 적용한다.



▶▶ 그림 4. 전처리된 이미지

전 처리된 텍스처 파일에서 참조되지 않는 매쉬와 정점을 새로 만들어진 obj 파일에서 제외한다.

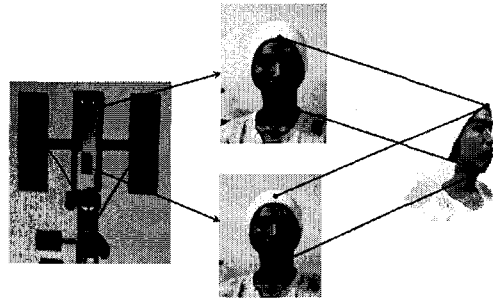
III. Pixel-to-Vertex Map (PVM)

3.1. GeometriX에 적합한 계산량 축소 방법 필요

테스트에 사용된 가공하지 않은 데이터에서 전처리 적용 후 필요한 부분만을 추출한 데이터양은 정점 약 4000개, 이미지(텍스처)픽셀 120X160개이다. 전체 데이터를 대표할 수 있는 일부 데이터만을 사용하여 계산량을 줄이면서 원본과 거의 유사한 정확도를 가지는 다변량 데이터 처리 방법인 PCA(Principal Component Analysis)를 사용한다. PCA는 얼굴 굴곡 데이터를 처리하는 셰이프 함수와 픽셀을 처리하는 텍스처 함수에 각각 적용한다[6].

3.2. PCA에 부적합한 GeometriX 데이터

GeometriX[4]로 획득된 가공하지 않은 데이터는 스테레오 카메라를 사용하여 두 장의 이미지를 얻는다. 이를 바탕으로 얼굴 굴곡을 구성하여 정점의 위치 정보를 가지는 셰이프(shape)데이터를 생성한다. 이때 전체 픽셀 중 해당 정점에 대응하는 이미지상의 값(텍스처가 될 RGB 값)은 일부 픽셀만 사용된다. 즉 정점 수는 픽셀보다 현저히 작아서 PCA[2]를 적용할 수 없다.



▶▶ 그림 5. 입체 3차원 인물은 GeometriX로 찍은 사진 두 장으로부터 복원된다.

3.3. Vertex Average By PixelMap

위의 문제를 해결하기 위하여 전체 정점을 참조하는 픽셀에 연결한 픽셀맵(pixelmap)을 만든다. 모든 얼굴(face)데이터에 대하여 픽셀맵을 만들며 결과적으로 얼굴마다 상이한 픽셀맵이 생성되어 픽셀간의 위치와 숫자가 틀리다. 상이한 픽셀맵들을 더해서는 평균모델을 생성할 수 없다. 트레인셋(train set)에서 특정한 하나를 기준(reference)모델로 정하고 이에 해당하는 픽셀맵과 동일하도록 나머지 트레인셋들을 타겟으로 정한다. 기준의 픽셀맵과 동일하게 타겟의 픽셀맵을 맞춘다. 기준 픽셀맵의 동일 위치에 타겟 픽셀맵의 픽셀이 없을 경우 보간을 통해 생성한다[5].



▶▶ 그림 6. 마스크의 사용 예

3.4. Face Space based on PCA

레이저 스캔을 통해 생성된 M+1개의 3D 얼굴을 확보하여 전처리를 하였다. 그 중에서 하나를 기준 얼굴로 두고 나머지 M 개를 타겟 얼굴로 둔다. 타겟 얼굴들을 각각 기준 레이저 스캔과 일치시킨다[7].

각 얼굴은 N_V 개 공간상에서 정점들과 N_V 개의 픽셀 값을

가지고 있다. 정점과 픽셀은 각각 쌍으로 있으며 이들을 순서에 따라 기준과 일치 시킨다. 타겟에 속하는 모든 얼굴들을 한결같이 다룬다. 3차원 공간에서 모든 얼굴의 정점 모음을 세이프 매트릭스 S라 두고 이에 대응하는 모든 얼굴 텍스처 2차원 평면에서 사용되는 픽셀 모음을 텍스처 매트릭스 T라 둔다.

$$S = \begin{pmatrix} x_1 x_2 \dots x_{N_v} \\ y_1 y_2 \dots y_{N_v} \\ z_1 z_2 \dots z_{N_v} \end{pmatrix}, \quad T = \begin{pmatrix} r_1 r_2 \dots r_{N_v} \\ g_1 g_2 \dots g_{N_v} \\ b_1 b_2 \dots b_{N_v} \end{pmatrix} \quad (1)$$

선형 얼굴 공간을 만들 때, M개 각 얼굴의 세이프 S_i와 텍스처 T_i의 선형 조합을 사용하여 새로운 얼굴을 만들 수 있다.

$$S = \sum_{i=1}^M a_i S_i, \quad T = \sum_{i=1}^M \beta_i T_i \quad (2)$$

식 (2)은 세이프와 텍스처에서 균등한 분포를 갖는다. 이러한 분포는 a_i 또는 β_i >> 1일 경우 얼굴의 형태를 가질 수 없기 때문에 세이프 공간과 텍스처 공간이 Gaussian 확률 분포를 가진다고 가정한다. PCA는 공분산 매트릭스를 대각화하는 것(de-correlate)과 같이 공간 변환을 하는 통계적인 방법을 제공한다. PCA는 세이프와 텍스처가 서로 간에 미치는 상관성(correlation)을 분리 적용한다. PCA를 세이프와 텍스처에 직접 적용하기 위해서 세이프의 평균 S를 개개인의 얼굴에서 빼고 이들을 각각 나열하여 결과를 차곡 차곡 스택(stack)하여 데이터 매트릭스 A를 만든다. A에 Singular Value Decomposition[7]을 적용하여 생성된 공분산 매트릭스 C에서 기저벡터를 구한다.

$$\bar{S} = \frac{1}{M} \sum_{i=1}^M S_i, \quad a_i = \text{vec}(S_i - \bar{S}), \quad A = (a_1, a_2, \dots, a_M) = U \Lambda V^T$$

$$C = \frac{1}{M} A A^T = \frac{1}{M} U \Lambda^2 U^T \quad (3)$$

vec(S)는 S의 칼럼(column)을 스택 한다. 직교 매트릭스 U의 칼럼 M개는 공분산 매트릭스 C의 기저벡터고, σ_i² = λ_i²/M는 U의 고유치(λ_i는 내림차순으로 정렬된 대각화된 매트릭스 Λ의 요소)들이다. S_i = U U⁽³⁾에 의하여 U의 i칼럼을 U_i라 두고, 3×N_v 매트릭스가 되는 주요소(principal component) i를 둔다. a_m⁽ⁿ⁾ 표시는 m×1 벡터를 n×(m/n) 매트릭스로 폴드(fold)한 것이다.

100명의 한국인 남녀들의 두 부분 세이프와 텍스처 각각에 PCA를 적용한다.

$$S = \bar{S} + \sum_{i=1}^M a_i S_i, \quad T = \bar{T} + \sum_{i=1}^M \beta_i T_i \quad (4)$$

식(4)가 각각 텍스처와 세이프 각각에 적용된 선형조합으로 Morphable 모델을 생성한다.



▶▶ 그림. 5 생성된 Morphable 모델

IV. 모델 정밀도 간략화

만들어진 모델은 이마와 뺨 부분을 중점으로 불필요한 데이터를 가지고 있다. 간선과 양쪽 두 정점을 한 쌍으로 하는 페어(pair)를 제거하고 새로운 정점으로 대체하는 압축(contraction)을 반복한다. 이때 새로운 정점의 위치는 주위 평면과 제곱 거리 합을 벡터적으로 해석한 QEM(Quadric Error Metricx)[8][9]를 사용한다.

4.1. QEM사용한 새로운 정점 생성

페어를 (v₁, v₂)로 정의한다. 페어가 모서리 또는 PBR₁₋₂PBR₂ (ℓ는 문턱 값 인수)이면 삭제 대상이다.

각 정점에서 만나는 삼각형 면의 집합 해는 각 정점이다. 각 정점에서 면 집합을 결합하면 해당 면에 대한 거리 제곱의 합으로 정점 오차(error)를 정의할 수 있다.

$$\Delta(v) = \Delta([v_x v_y v_z 1]^T) = \sum_{f \in \text{faces}(v)} (d^T v)^2 \quad (5)$$

(P=[abcd]^T는 a²+b²+c²=1인 식 ax+by+cz+d=0) 정점에서 면 집합은 해당 정점을 충족하는 삼각형의 면으로 초기화된다. 위 식의 오차는 아래의 이차 식으로 계산한다.

$$K_p = \Delta^T = \begin{bmatrix} a^2 & ab & ac & ad \\ ab & b^2 & bc & bd \\ ac & bc & c^2 & cd \\ ad & bd & cd & d^2 \end{bmatrix} \quad (6)$$

$$\begin{aligned}
 \Delta(v) &= \sum_{p \in \text{planes}(v)} (v^T p)(p^T v) \\
 &= \sum_{p \in \text{planes}(v)} v^T (pp^T) v \\
 &= v^T \left(\sum_{p \in \text{planes}(v)} K_p \right) v
 \end{aligned} \tag{7}$$

오차 이차 K_p 는 공간상의 특정 정점과 평면 P 와의 거리 제곱을 찾으며 K_p 를 모두 더하면 Q 매트릭스가 된다. 반복 압축하는 동안 정점이 없어질 경우 오차 정도를 알려줄 4×4 매트릭스 Q 를 각 정점에 할당하고 오차 정점 $v = [v_x, v_y, v_z]^T$ 를 2차 식 $\Delta(v) = v^T Q v$ 로 정의한다. $\Delta(v) = \epsilon$ 는 이차 서페이스(surface)를 가지며 모든 정점 집합에서 Q 에 관한 오차는 ϵ 를 가진다. 압축할 pair $(v_1, v_2) \rightarrow \bar{v}$ 에서 \bar{v} 에 대한 오차를 근사하는 새로운 매트릭스 \bar{Q} 를 유도해야 한다. 단순 가산 공식 $\bar{Q} = Q_1 + Q_2$ 사용하기 위해 선택했다. 압축 $(v_1, v_2) \rightarrow \bar{v}$ 는 $\Delta(\bar{v})$ 를 최소화하는 \bar{v} 위치를 찾는다. 에러 함수 Δ 가 이차이며 최소 해법은 선형으로 $\partial \Delta / \partial x = \partial \Delta / \partial y = \partial \Delta / \partial z = 0$ 을 풀어 \bar{v} 를 찾는다.

$$\begin{bmatrix} q_{11} & q_{12} & q_{13} & q_{14} \\ q_{12} & q_{22} & q_{23} & q_{24} \\ q_{13} & q_{23} & q_{33} & q_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \bar{v} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{8}$$

4.2. Algorithm Summary

인접 그래프 구조를 사용한 모델링을 한다. 모든 정점, 간선 그리고 면은 분명하게 표현되어 함께 연결된다. 유효한 페어 셋(set)을 추적하기 위해 각 정점은 자신의 인접 정점 목록을 유지한다. 알고리즘은 간단히 다음과 같이 요약할 수 있다.

1. 모든 초기 정점에서 Q 매트릭스를 구한다.
2. 모든 유효한 페어를 선택한다.
3. 각각 유효한 pair (v_1, v_2) 에 대한 최적의 contraction 타겟 \bar{v} 를 구한다. 이 타겟정점 $\bar{v}^T(Q_1 + Q_2)\bar{v}$ 의 에러는 그 페어 압축 비용이 된다.
4. 최소 비용을 기준으로 하는 힙자료구조의 원소에 모든 페어를 넣는다.
5. 반복적으로 힙으로부터 비용이 최소 되는 pair (v_1, v_2) 를 삭제하면서 v_1 에 관한 모든 유효한 페어의 비용을 갱신한다.



▶▶ 그림 6. ▶▶ 그림 7. ▶▶ 그림 8.
Mesh = 2345 Mesh = 1689 Mesh = 1343
정규화된 모델의 메쉬를 줄인 결과.

V. 결론 및 향후 발전계획

변화가 심한 환경 중 한 가지 요인인 조명에 대한 효과가 없다. 모델의 사용처를 늘리기 위해서는 조명과 피부의 상태를 모델링할 필요가 있다. 또한, 생성된 모델에서 데이터를 제거하는 것만이 얼굴인식에 좋은 결과를 가져오는 것은 아니다. 많은 시간과 처리 비용을 허용하더라도 더욱 정확한 인식 결과를 얻고자 한다면 Curved PN Triangles[11]를 사용하여 정점을 증가시켜 보는 것도 하나의 방법이 될 수 있다. 더욱 세분화된 mesh를 가지고 정교한 인식이 가능하기 때문이다.

■ 참고 문헌 ■

- [1] W. ZHAO, R. CHELLAPPA, P. J. PHILLIPS, AND A. ROSENFELD, "Face Recognition: A Literature Survey"
- [2] Lindsay I Smith, "A tutorial on Principal Components Analysis", February 26, 2002.
- [3] Volker Blanz, Thomas Vetter, "A Morphable Model For The Synthesis Of 3D Faces"
- [4] www.geometrix.com
- [5] Taehwa Hong, Hagbae Kim, Hyeonjoon Moon, Yongguk Kim, Jongweon Lee and Seungbin Moon "Face Representation Method Using Pixel-to-Vertex Map (PVM) for 3D Model Based Face Recognition"
- [6] <http://sasmania.com/zb41/contents.php?id=227&PHPSESSID>
- [7] Volker Blanz, Thomas Vetter, "Chapter 10. Morphable Models of Faces"
- [8] Michael Garland Paul S. Heckbert, "Surface Simplification Using Quadric Error Metrics"
- [9] Michael Garl and Paul S. Heckbert, "Simplifying Surfaces with Color and Texture using Quadric Error Metrics"
- [10] <http://www.cyberware.com/>
- [11] Alex Vlachos, J'org Peters, Chas Boyd, Jason L. Mitchell, "Curved PN Triangles"

Session III

제 3 논문발표장(평생교육원 204호 강의실)

Ⅲ-A e-Sports&Leisure콘텐츠-1
>>> 좌장 김 흥 설(배재대)

Ⅲ-B e-Sports&Leisure콘텐츠-2
>>> 좌장 송 강 영(동서대)
