

단안 기반 하에서 시간에 따라 변화하는 비강체 물체의 평균 모양 및 움직임 복원에 관한 연구

김태원* · 허남호* · 김진웅* · 이수인* · 홍기상**

1. 서 론

1.1 비강체 복원 연구 소개

비강체 복원(Nonrigid recovery)이란 3차원 공간상에서 시간에 따라 변화하는 비강체의 3차원 움직임 (Motion)과 모양(Shape)을 동시에 복원하는 것이다. 지난 수 십년 동안 컴퓨터 비전(Computer vision) 분야의 수 많은 연구자들이 단안 (Monocular) 또는 다안(Multiple viewpoints) 기반 비강체 복원 연구를 수행해 왔다[13,17,18,20, 25,26]. 그 중에서 FEM(Finite Element Method) 기반 방법[7,8,16,17,20]은 비강체표면(Surface)을탄성(Elastic property)[2]을 가지도록 모델링하고, 시간적으로 변화하는 모양을 진동모드(Vibration modes) 개념을 이용하여 이를 방정식으로 표현한다. 이렇게 표현된 방정식은 FEM 기법을 이용해서 풀 수 있게 된다. 또한 광학흐름(Optical flow)과 같은 시각 정보(Visual cues)를 방정식에 포함하여 모델링함으로써 비강체의 모양뿐만 아니라 움직임까지도 복원할 수 있다[9]. 그러나 FEM 기반 방법과 같은 모델 기반 방식들(Model-based approaches)은 물체 표면의 탄성 계수나 제한된 수의 진동모드 등의 특수한 파라미터에 의존하기

때문에 적용하고자 하는 응용 분야에 맞게 파라미터를 재설정해야 하는 어려움을 가지고 있다. 이러한 모델 기반 방식들에 대한 더 자세한 내용을 알고 싶은 독자들은 [1]을 참조하기 바란다. 최근에 Blanz et al.[3]은 조밀(Dense)하면서 정확한 비강체 복원을 위해 모퍼블 모델(Morphable model)에 기반한 방법을 제안했다. 이 방법은 초기에 3차원 스캐너를 이용해서 획득한 3차원 얼굴 모델들로부터 PCA(Principal Component Analysis) 기법을 이용하여 텍스처(Texture)까지 모델링할 수 있는 모퍼블 모델을 생성한다. 이렇게 생성된 모퍼블 모델을 이용하면 사람의 얼굴을 찍은 단일 영상을 입력으로 받아 모퍼블 모델로부터 그에 해당하는 3차원 얼굴을 정밀하게 복원할 수 있다. 그러나 이 방식은 초기에 예제 얼굴 모델들 간의 조밀한 대응점(Dense correspondence) 정보가 필수적이어서 일반적인 활용 측면에서 보았을 때 그 응용성이 떨어지는 한계를 가지게 된다. Vedula et al.[25,26]은 복수의 카메라(Multiple cameras)에 기반하여 광학흐름 정보를 계산하고 광학흐름과 비강체 모양/움직임 사이의 관계식을 설정함으로써 비강체 복원을 할 수 있는 방법을 제안하였고, 이를 실제 비디오에 적용하여 얻어진 비강체 복원 결과를 보여줬다.

최근에 Bregler et al.[6]은 어파인 카메라(Affine

* 한국전자통신연구원 전파방송연구단 방송시스템연구그룹
 ** 포항공과대학교 전자전기공학과

camera) 가정 하에서 추적된 특징점들(Tracked feature points)로 구성된 측정행렬(Measurementmatrix)에 분해기법(Factorization method)을 적용하여 비강체의 움직임과 시간적으로 변화하는 모양을 동시에 얻을 수 있는 방법을 제안하였다. 이 방법은 기존의 강체(Rigid objects)에 적용된 분해기법[22]을 비강체로 확장한 방법으로 비강체의 모양 변화를 다수의 기저모양(Basis shape)로 모델링하는 방식을 이용한다[4,5]. 이러한 비강체 분해기법에 기반하여 Toressani et al. [24,23]은 측정행렬의 랭크제한조건(rank constraints)을 이용하여 비강체의 특징점을 자동적으로 정확하게 추적할 수 있는 방법을 제안하였다. 이 방법의 특징은 특징점 추적 시에 특징점이 물체에 의해 가려질 경우(Occlusion problem)에도 강건하게 추적할 수 있는 장점을 가지고 있다. 마지막으로, Bue et al.[4,5]은 스테레오 카메라(Stereo camera)를 이용하여 추적된 특징점들로 구성된 측정행렬을 분해하여 비강체의 움직임과 시간적으로 변화하는 모양을 동시에 정확하게 복원할 수 있는 방법을 제안하였다. 그러나 이 방법은 두 개의 비디오에서 추적된 특징점들 간의 동기화(Synchronization)가 선수조건(Prerequisite)으로 필요로 하는 문제가 있다.

1.2 모양과 움직임 사이의 모호성(Shape and motion ambiguity)

일반적으로 단안 기반 하에서는 사전정보 없이 비강체 움직임과 시간적으로 변화하는 모양을 동시에 복원할 시에는 모호성(Ambiguity) 문제가 발생한다(그림 1 참조). 그림 1에서는 시간적으로 변화하는 비강체가 시간 $t=0$ 에서 4개의 검은 점으로 표시되어있다. 그림에서 I_0 , I_1 그리고 I_1' 은 각각 시간 $t=0$ 와 $t=1$ 에서의 영상을 나타내고,

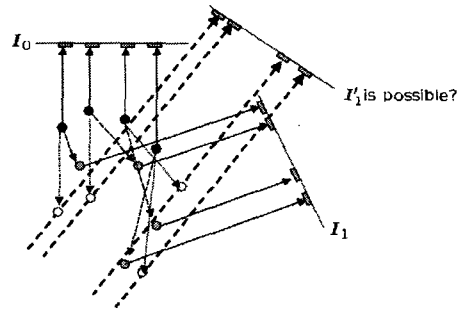


그림 1. 단안 기반 카메라(Monocular affine camera) 하에서의 움직임과 모양의 모호성 도시

화살표는 오소그래픽 카메라 사영(Orthographic camera projection)[22]을 나타낸다. 그림에서 볼 수 있듯이 4개의 검은 점은 점으로 이루어진 비강체는 시간 $t=1$ 에서 각각 4개의 흰 점과 4개의 회색 점 위치로 움직이는 것을 볼 수 있다. 이렇게 움직인 4개의 흰 점과 4개의 회색 점들은 시간 $t=1$ 에서 해당 카메라에 각각 사영 되었을 경우 영상 I_1 과 영상 I_1' 가 서로 같은 영상을 생성하게 된다. 도시된 그림을 통해서 유추할 수 있는 것은 단안 기반 카메라로부터 비강체를 복원할 시에는 같은 영상일 경우에도 모양과 움직임이 서로 다른 비강체 복원 결과를 얻을 있다는 것이다. 이러한 현상은 비강체의 움직임과 변화하는 모양이 서로 상관 관계를 갖고 있기 때문에 발생하는 문제에 해당한다[14,15]. 즉, 비강체의 움직임(Motion)에 의해서 나타나는 3차원 점의 위치 변화는 비강체의 모양 변화만으로도 표현할 수 있기 때문에 나타나는 현상으로 기존의 비강체 분해 방식[6]에서는 이러한 모호성을 고려하지 않아 결과적으로 최종 복원된 비강체 모양과 움직임이 실제와는 다른 복원 결과를 얻을 수 있다. 본 논문에서는 이러한 모호성을 제거하기 위해 비강체 움직임을 비강체의 평균 모양 움직임으로 정의함으로써 단안 방식이 항상 지니고 있는 모호성을 없앨 수 있는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 제 2 장에서는 기존의 강체와 비강체 분해기법을 살펴보고 기존 방법에서 고려하지 않는 비강체 움직임과 모양 사이에 존재하는 모호성 문제를 보인다. 제 3 장에서는 이러한 모호성을 제거하기 위한 DoN(Degree of Nonrigidity) 개념을 도입하고, 제 4 장에서는 DoN 개념에 이용하여 ICRF(Iterative Certainty Reweighted Factorization) 방법을 제안하며, 실제 비디오에 적용하기 위한 Robust ICRF 방법을 제안한다. 제 5 장에서는 제안된 방법과 기존 방법을 적용하여 얻어진 실험 결과와 그 비교분석 자료를 제시한다. 마지막으로 제 6 장에 결론을 맺고자 한다.

2. 강체/비강체 분해기법(Rigid/Nonrigid factorization methods): 개관

2.1 강체 분해기법

(Rigid factorization method)

추적된 특징점으로 이루어진 $2F \times P$ 측정행렬은 다음과 같이 정의된다[22]:

$$W = \begin{bmatrix} x_{11} & \dots & x_{1P} \\ x_{21} & \dots & x_{2P} \\ \vdots & & \vdots \\ x_{F1} & \dots & x_{FP} \end{bmatrix}, \text{ or } W = \begin{bmatrix} w_1^T \\ w_2^T \\ \vdots \\ w_F^T \end{bmatrix} \quad (1)$$

여기서 $x_{ij} = (x_{ij}, y_{ij})$ 는 총 F 개의 프레임으로 이루어진 비디오에서 추적된 i 번째 프레임 영상의 j 번째 특징점을 나타낸다. 이때 측정행렬 W 가 강체의 특징점으로 이루어져 있을 경우 측정행렬에 분해기법[22]을 적용하여 아래와 같이 움직임과 강체 모양으로 각각 분해할 수 있다.

$$W \approx M_{2F \times 3} S_{3 \times P} = \hat{M} G G^{-1} \hat{S}$$

$\hat{M} = [R_1 \dots R_P]^T$, $\hat{S} = [X_1 \dots X_P]$ 이다. 여기서 R_i 는 2×3 회전 행렬을, X_j 는 3×1 3차원 점을, G 는

비특이(Nonsingular) 행렬을 나타낸다. 특히, 비유클리디언(Non-Euclidean) 움직임과 모양에 해당하는 행렬 M 과 S 는 행렬 G 에 의하여 유클리디언(Euclidean) 움직임과 모양으로(\hat{M} 과 \hat{S})로 변환될 수 있다(자세한 내용은 [22] 참조).

2.2 비강체 분해기법(Nonrigid factorization method)

서론에서 언급 했듯이 Bregler et. al.[6]은 기존의 강체 분해기법을 확장하여 비강체에 대해서도 분해기법을 적용할 수 있는 새로운 방법을 제안하였다. 이때 측정행렬은 강체의 경우와 같이 비강체로부터 추적된 특징점들로 이루어진다. 그들은 이렇게 구성된 측정행렬을 K 개의 기저모양(S_k)을 이용하여 근사하고 최종적으로 랭크제한조건을 가하여 측정행렬을 분해함으로써 비강체의 움직임과 모양을 동시에 얻을 수 있는 방법을 처음으로 제안했다. 측정행렬은 아래와 같이 분해된다.

$$W \approx \hat{W} = \begin{bmatrix} l_{11}R_1 & l_{12}R_1 & \dots & l_{1K}R_1 \\ l_{21}R_2 & l_{22}R_2 & \dots & l_{2K}R_2 \\ \vdots & & & \vdots \\ l_{F1}R_F & l_{F2}R_F & \dots & l_{FK}R_F \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_K \end{bmatrix} \quad (2)$$

$$= \hat{M} \hat{S},$$

여기서 l_{ij} 는 기저모양에 붙는 계수이다(자세한 분해 과정은 [6]을 참조). 그러나 비강체 분해기법은 서론에서도 언급 했듯이 비강체 움직임과 시간적으로 변화하는 모양 사이에 존재하는 모호성 문제를 고려하지 않았기 때문에 잘못된 복원 결과를 만들어 낼 수 있다. 실제로 비강체 움직임을 보이는 얼굴을 촬영한 비디오(그림 6 참조)에 식 (2)를 적용하여 얻어진 복원 결과(시간 $t = 0$)를 그림 2에 표시했다. 그림 2에서 볼 수 있듯이 최종 복원된 얼굴 모양을 살펴보면 입 주변이 돌출되어 있는 것을 볼 수 있다. 이러한 결과는 제 1.2절에서 설명한 비강체의 움직임과 모양 사이에 존재하는

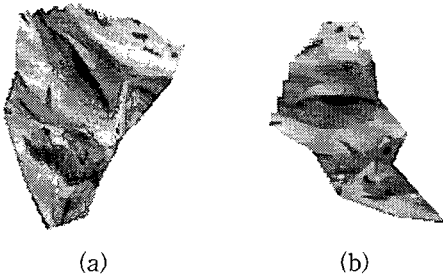


그림 2. 비강체 분해기법을 적용해 복원된 일정한 시점에서의 3차원 얼굴 모양: (a) Upper view, (b) Side view

모호성을 고려하지 않기 때문에 발생하게 된다. 즉, 움직임을 나타내는 행렬 \hat{M} 이 모양(\hat{S})에 해당하는 정보를 표현할 수 있기 때문에 발생하는 현상으로 다음 장에서 이러한 문제점을 제거할 수 있는 실마리를 제공하는 DoN 개념에 대해 설명한다.

3. DoN(Degree of Nonrigidity)

DoN 개념을 도입하기 위해서 우선 시간적으로 모양을 변화하면서 움직임을 보이는 비강체를 가정하자(예들 들면 사람의 얼굴). 말을 하면서 움직이고 있는 사람의 얼굴 모양을 살펴보면 얼굴의 일부분은(예들 들면 이마, 코 등) 강체와 같이 거의 모양 변화가 없이 움직임을 주로 보이는 부분(Near-rigid surface)이 존재하면서, 다른 부분(예들 들면, 입 주변)은 움직임뿐만 아니라 모양 변화가 많이 발생하는 부분(Non-rigid surface)도 존재한다. 즉, 얼굴 경우와 같이 비강체의 모양 변화가 평균 모양에서 거의 변화하지 않는 부분과 상당히 많이 변화하는 부분이 동시에 존재하는 비강체를 가정한다. 그리고 시간 $t = i$ 에서 P 개의 특징점으로 이루어진 비강체 모양을 $X_i = [X_{i1}, \dots, X_{iP}]$ 로 표시하면 비강체의 평균 모양, $\bar{X} = [\bar{X}_1, \dots, \bar{X}_P]$ 은 다음과 같이 구할 수 있다.

$$\bar{X} = \frac{1}{F} \sum_{i=1}^F X_i,$$

여기서 F 는 총 프레임 개수를 나타낸다. 이러한 평균 모양을 이용하여 아래와 같이 DoN을 정의한다.

$$\text{DoN} \equiv \sum_{i=1}^F (X_{ij} - \bar{X}_j)(X_{ij} - \bar{X}_j)^T \quad (3)$$

DoN의 2차원 사영(2D projection of DoN)

식 (3)에서 DoN를 3차원 공간 상에서 정의하였다. 아래와 같이 재투영 오차(Reprojection 오차)를 정의하도록 하자.

$$e_{ij} = x_{ij} - \bar{x}_{ij} \quad (4)$$

여기서 $x_{ij} = R_i X_{ij}$, $\bar{x}_{ij} = R_i \bar{X}_j$. 식 (4)의 재투영 오차를 이용하여 DoN을 2차원 영상으로 사영시킬 경우 아래와 같이 표현될 수 있다.

$$\begin{aligned} \sum_i^F e_{ij} e_{ij}^T &= \sum_i^F R_i (X_{ij} - \bar{X}_j)(X_{ij} - \bar{X}_j)^T R_i^T \\ &\equiv C_j \end{aligned} \quad (5)$$

여기서 C_j 는 j 번째 특징점에 해당하는 2×2 분산행렬(Covariance matrix)을 나타낸다. 식 (5)를 살펴보면 P 개의 점들 중 평균 모양에서 크게 변화하지 않는 부분에 해당하는 점들은 작은 분산을 갖게 되고 평균 모양에서 크게 변화하는 부분에 해당하는 점들은 큰 분산을 갖게 됨을 알 수 있다. 그림 3에 제 2 장에서 소개된 강체 분해기법을 적용하여 얻어진 얼굴 모양을 평균 모양으로 이용하여서 얻어진 특징점의 분산행렬을 타원으로 표시했다.

그림 3 (b)에서 표시된 특징점들의 분산행렬을 살펴보면 평균 모양이 제대로 구해졌을 경우에 작은 분산을 갖어야 하는 이마와 코 부분(Near-rigid surface)에서 상당히 큰 분산을 갖는 특징점들이 존재함을 알 수 있다. 그 이유는 강체 분해기

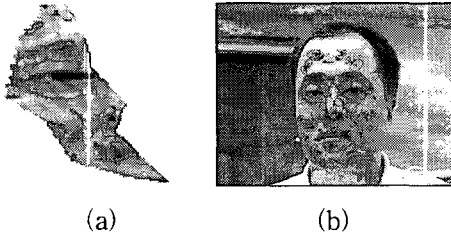


그림 3. (a) 강체 분해기법을 적용하여 얻어진 얼굴 모양, (b) (a)에서 얻어진 모양을 평균 모양으로 이용하여 계산된 분산행렬 표시

법에 의해서 구해진 얼굴 모양이(그림 3 (a) 참조) 실제 얼굴의 평균 모양과 많은 차이를 갖고 있기 때문이다. 다음 장에서는 이러한 분산행렬을 이용하여 비강체의 평균 모양과 그에 해당하는 움직임 정보를 복원할 수 있는 방법을 제안한다.

4. 제안된 방법

그림 3 (a)에서 알 수 있듯이 비강체의 특징점으로 구성된 측정행렬에 강체 분해기법을 적용하여 얻어진 모양은 일반적으로 평균 모양에 해당하지 않는다. 본 논문에서는 DoN에 기반하여 기존의 강체 분해기법을 확장하여 비강체의 특징점으로 구성된 측정행렬로부터 직접 평균 모양을 분해할 수 있는 방법을 제안한다. 기본 아이디어는 Irani et al.[12]이 제안한 특징점의 불확실성(Uncertainty)을 고려한 확실성 가중치 분해기법(Certainty weighed factorization method)을 평균 모양을 계산하는데 응용하는 것이다. 확실성 기반 분해기법을 요약하여 설명하면 확실성이 높은 특징점에 더 많은 가중치를 두고 분해하는 기법으로 실제 비디오의 경우 추적된 특징점들에 잡음이 많이 존재하게 되는데 이때에도 강건하게 강체의 모양과 움직임 정보를 계산해 낼 수 있는 분해기법에 해당한다. 이러한 개념은 비강체의 평균 모양 복원에도 사용될 수 있다. 즉 비강체의

평균 모양으로부터 크게 변화하는 부분에 해당하는 특징점들에는 강체 분해기법 적용 시에 낮은 확실성을 주고, 평균 모양으로부터 크게 변화하지 않는 부분에 해당하는 특징점들에는 높은 확실성을 줌으로써 최종 분해된 모양이 평균 모양에 가깝도록 하는 것이다. 여기서 확실성은 제 3 장에서 도입된 분산행렬과 일맥상통하는 개념으로 확실성은 분산행렬의 역(Inverse)에 해당한다고 할 수 있다.

4.1 ICRF(Iterative Certainty Re-weighted Factorization)

그러나 Irani et al.이 제안한 확실성 가중치 분해기법은 각 특징점들의 확실성이 미리 주어져 있다는 가정을 한다. 즉, 비강체 복원에 적용할 경우 비강체의 평균 모양이 미리 주어져 있어야 분산행렬(또는 확실성)을 계산할 수 있음을 의미한다. 아래에 평균 모양과 그에 해당하는 움직임 정보뿐만 아니라 분산행렬도 동시에 계산할 수 있는 알고리즘을 단계별로 설명한다.

- (1) 비강체 특징점으로 구성된 측정행렬 W 에 강체 분해기법을 적용하여 초기 움직임($M^{(0)}$)과 모양($S^{(0)}$)으로 분해한다.
- (2) (1)에서 구해진 $M^{(0)}$ 과 $S^{(0)}$ 를 식 (5)를 이용하여 각 특징점들의 초기 분산행렬($C_j^{(0)}$)과 재투영 오차($W-M^{(0)}S^{(0)}$)를 계산한다.
- (3) 아래의 과정을 반복한다.
 - 분산행렬 $C_j^{(i-1)}$ 의 역을 확실성 가중치 Irani et al.이 제안한 확실성 가중치 분해기법을 적용하여 W 로부터 움직임($M^{(i)}$)과 모양($S^{(i)}$)을 분리해 낸다.
 - 현재 분해된 움직임($M^{(i)}$)과 모양($S^{(i)}$) 정보를 이용하여 분산행렬 $C_j^{(i)}$ 과 재투영 오차 ($W-M^{(i)}S^{(i)}$) 다시 계산한다.
 - 재투영 오차가 더 증가하지 않을 경우에

반복을 멈추고, 그 반대의 경우 계속 현재 과정을 반복한다.

(4) 마지막으로, 분해된 움직임(M)과 모양(S)을 유클리디언 값으로 변환시킨다.

위에 단계별로 정리된 알고리즘을 인공 데이터 및 실제 비디오에 적용하였을 경우에 보통 몇 번의 반복 과정 후에 수렴하는 결과를 보였다(제 5장의 실험 결과 참조).

4.2 Robust ICRF

그러나 4.1에서 제안된 ICRF 방법을 실제 비디오에 적용할 경우에는 추적된 특징점들에 아웃라이어(Outliers)가 발생하여 잘못된 수렴을 하는 문제점이 있을 수 있다. 보통의 경우 아웃라이어들은 평균 모양에서 많이 변화하는 특징점들에 존재하게 되는데, 이러한 아웃라이어 특징점들의 영향을 없애기 위해 ICRF에 강건성을 부여하여 실제 적용 시에 정확한 결과를 얻을 수 있도록 해야 한다. 이를 위해서 본 논문에서는 M-estimator를 도입한다. $d_j \equiv \det \sqrt{C_j}$ 을 사용하여 분산 행렬을 아래와 같이 재정의 할 수 있다.

$$\sqrt{C_j} = \begin{cases} \sqrt{C_j} & d_j \leq \sigma \\ \sqrt{\sigma} \sqrt{C_j} / d_j^2 & \sigma < d_j \leq 3\sigma \\ 0 & d_j > 3\sigma \end{cases}$$

여기서 표준분산 σ 는 일반적으로 d_j 의 중간값(Median) 값으로 설정된다.

$$\sigma = \frac{med d_j}{0.6745}$$

5. 실험 결과

제안된 방법을 테스트하기 위해서 본 논문에서는 인공 데이터(Synthetic data)와 실제 비디오(Real video)를 가지고 실험한 결과를 제시한다.

우선 실험에 사용된 인공 데이터를 생성한 절차에 대해 설명하기로 한다.

(1) 시간적으로 변화하는 비강체 모양

- 처음에 총 30개의 점을 지름 320의 반구(Half sphere) 내에서 무작위(Random)로 생성
- 각 점들의 좌표 값(x, y, z)을 시간에 따라서 $[-L, L]$ 구간에서 생성
- 각 점들에 대해서 L 은 두 가지로 경우로 나누어서 결정

- Case 1(평균 모양으로부터 변화가 크지 않는 점): L 을 $[1 \sim 20]$ 구간에서 선택
- Case 2(평균 모양으로부터 변화가 큰 점): L 을 $[100 \sim 200]$ 구간에서 선택

(2) 아래와 같이 총 100개의 프레임에 해당하는 비강체 움직임 생성

- 비강체 전이(Translation)는 무작위 값으로 생성하여 설정
- 비강체의 회전(Rotation)은 무작위로 생성된 두 개의 쿼터니언 벡터(Quaternion vector)를 이용하여 보간 기법(Interpolation technique)을 적용하여 생성

(3) 각 점들을 재투영 해서 얻어진 영상 좌표에 표준분산 $\sigma = 1.0$ (Pixel)에 해당하는 Gaussian 잡음 첨가

위의 절차에 따라서 총 9개 set에 해당하는 인공 데이터를 생성하였고, 각각의 set은 Case 1과 Case 2에 해당하는 점들의 비율을 0.1에서 0.9까지 변화하면서 각 set에 대하여 총 100개씩의 데이터를 생성했다. 이렇게 생성된 데이터에 강제 분해기법(Rigid factorization) 비강체 분해기법(Nonrigid factorization) 그리고 제안된 ICRF(Iterative Certainty Re-weighted Factorization)

방법을 적용한 결과를 그림 4에 표시한다. 그림 4에서는 3가지 방법을 적용하여 계산된 움직임/평균 모양을 참 값에 해당하는 움직임/평균 모양과 비교하여 그 오차를 표시했다(비강체 분해기법의 경우 계산된 시간적으로 변화하는 모양의 평균을 취하여 평균모양을 계산했다). 그림에서 알 수 있듯이 제안된 ICRF 방법이 가장 정확한 결과를 보임을 알 수 있다. 그림 5에는 회전 움직임 값의 비교 결과를 회전각(roll/pitch/yaw) 별로 나누어서 보여준다. 회전 움직임의 경우도 제안된 ICRF 방법이 더 정확한 결과를 보임을 알 수 있다. 그림 5에 표시된 결과를 보면 제안된 ICRF 방법을 이용하면 좀 더 정확할 복원 결과를 얻을 수 있을 뿐만 아니라 계산된 값이 다른 두 방식에 비해서 부드럽게 변화하는 곡선을 이루고 있음을 알 수 있다.

인공 데이터 이외에도 제안된 방법을 실제 비디오(그림 6 참조)에 적용하여 테스트를 수행했다. 그림 6에 총 200 프레임으로 이루어진 비디오의 예시를 위해 몇 개의 예제 프레임은 보여준다. 실제 비디오에서 추적된 30개의 특징점들 중에는 아웃라이어 존재하기 때문에 ICRF 방법 대신 ro-

bust ICRF 적용하여 테스트를 수행했다. 그림 7에 robust ICRF를 적용했을 경우에 iteration에 따라 변화하는 분산행렬을 도시했다. 8번째 iteration에서 최종 수렴을 하였고, 최종 수렴된 분산행렬을 살펴보면 입 주변에서 가장 큰 분산을 가지고, 이마와 코 부분에서 작은 분산을 갖는 것을 확인할 수 있다. 이를 통해서 최종 복원된 얼굴 평균 모양과 움직임이 합리적으로 복원되었음을 예상할 수 있다. 그림 8에 최종 수렴된 얼굴의 평균

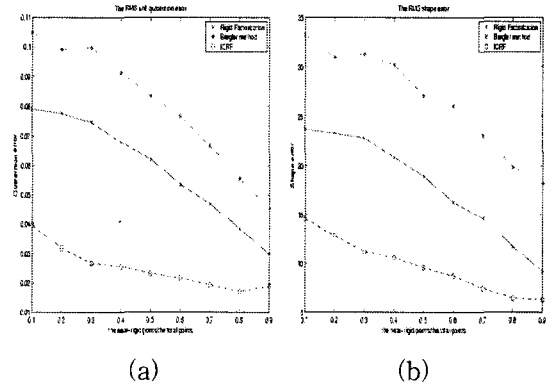


그림 4. X축은 Case 1과 Case 2에 해당하는 점들의 비율을 표시하고, Y축은 (a) Quaternion error (b) RMS(Root Mean Square) shape error

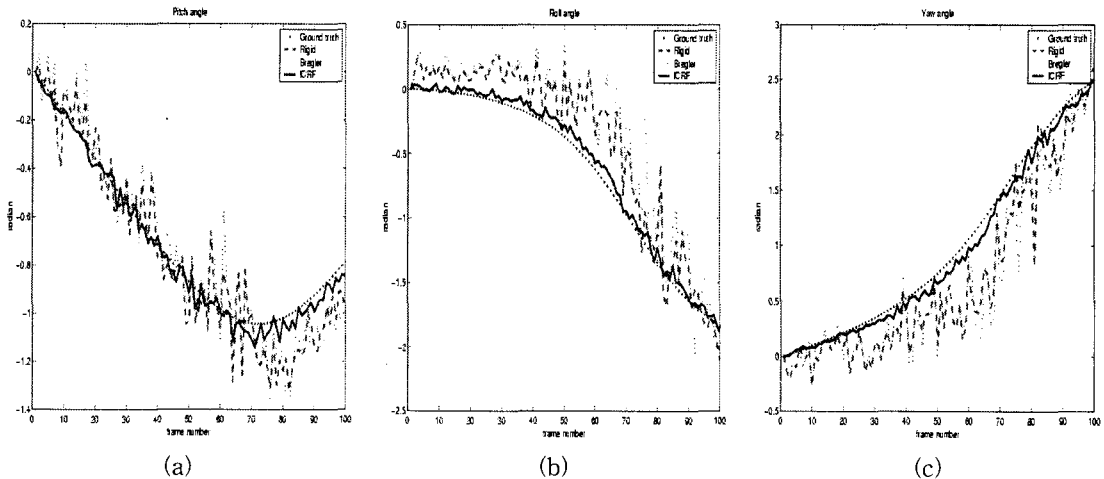


그림 5. X축은 프레임 번호, Y축은 radian angle: (a) Roll angle (b) Pitch angle (c) Yaw angle

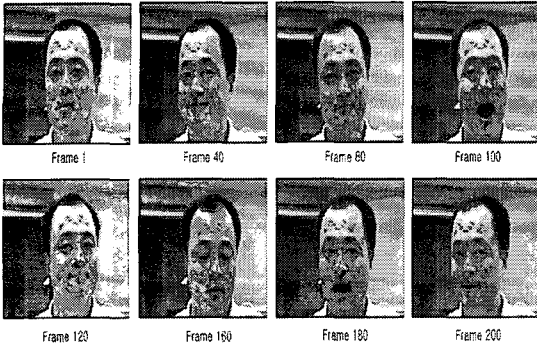


그림 6. 총 200 프레임으로 구성된 비디오에의 예제 프레임 도시 (KLT Tracker(21)를 이용하여 총 30개의 특징점 추적)

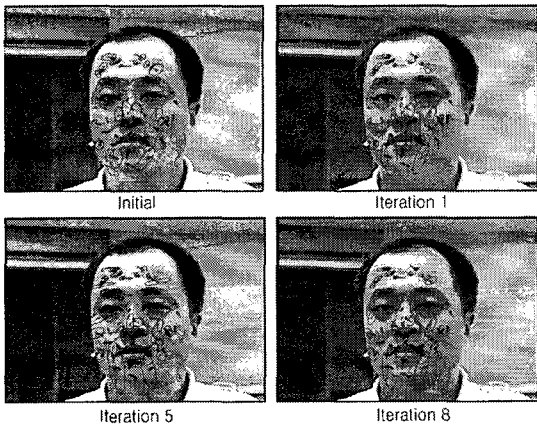


그림 7. Robust ICRF 방법을 적용하였을 경우 매 iteration 마다 변화하는 분산행렬 도시

모양을 보여주고 그림 9에는 강체 분해기법, 비강체 분해기법 그리고 제안된 robust ICRF 방법을 적용하여 얻어진 움직임 정보를 비교하여 보여준다. 그림 9에서 알 수 있듯이 강체/비강체 분해기법을 통해서 얻어진 움직임 정보는 거의 비슷한 결과를 보이지만 robust ICRF 방법을 적용해 얻어진 결과는 실제 얼굴의 움직임과 유사한 결과를 보인다(그림 10 참조). 그림 10에 세 가지 방법을 이용해서 얻어진 얼굴의 움직임 정보를 실제 영상 프레임에 좌표축의 움직임으로 표시했다. 그림 10에서 알 수 있듯이 robust ICRF 방법을 적용해서 얻어진 복원 결과가 실제 얼굴의 움직임에 가장 가까운 것을 확인할 수 있다.

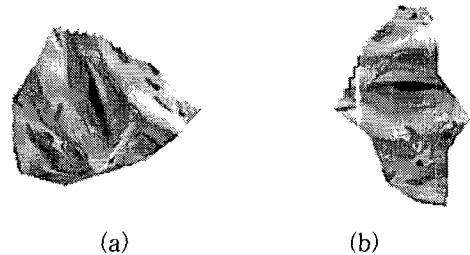


그림 8. Robust ICRF 방법을 적용하여 얻어진 얼굴의 평균모양 (a) Upper view (b) Side view

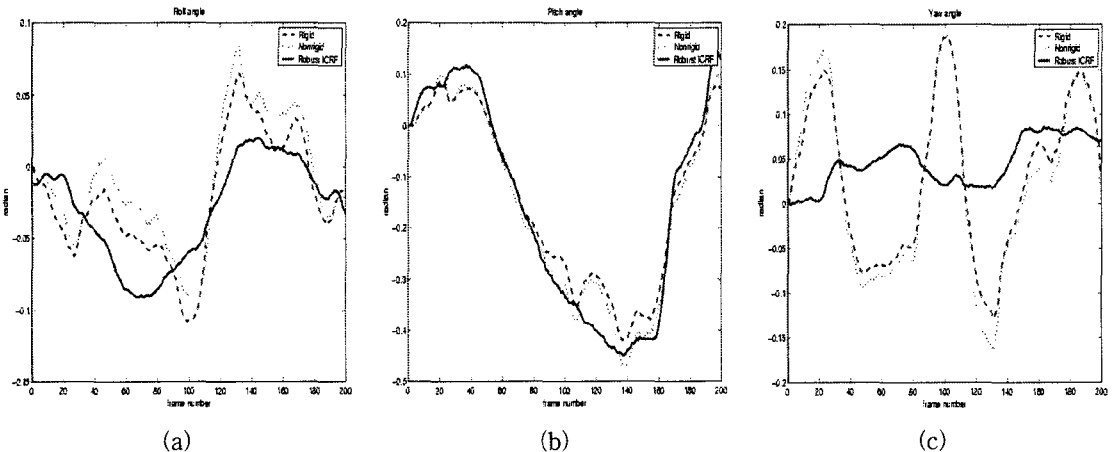


그림 9. Robust ICRF 방법을 적용하였을 경우 움직임(회전) (a) Roll angle (b) Pitch angle (c) Yaw angle

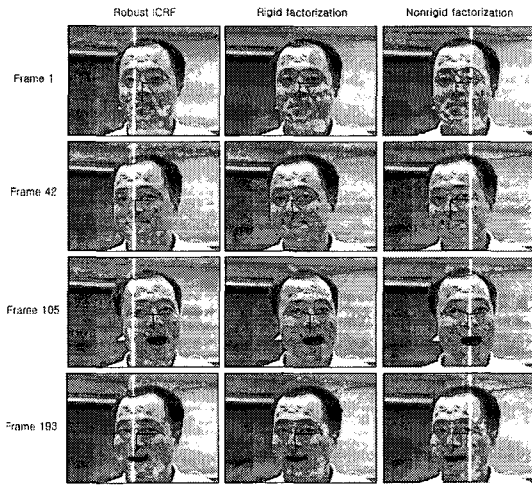


그림 10. 3 가지 방법에 의해서 구해진 얼굴의 움직임 (회전)을 좌표축의 움직임으로 도기한 결과 (Z축은 첫 번째 프레임서 평면에 수직 방향이 되게 위치)

6. 결론

본 논문에서는 단안 기반 하에서 시간적으로 변화하는 비강체의 평균 모양과 그에 해당하는 움직임을 복원할 수 있는 방법을 제안하였다. 일반적으로 어떠한 제약 조건 없이 모양이 변화하는 비강체 복원(모양 및 움직임)은 사전 정보 없이는 불가능하다. 그러나 비강체 중에서 일정 영역이 모양 변화를 크게 하지 않는 부분이 있는 경우에는 비강체의 평균 모양 및 그에 해당하는 움직임을 복원 시에 중요한 정보로 사용될 수 있음을 본 논문에서 이론적/실험적으로 제시했다. 이를 위해서 DoN 개념을 도입하고 이를 바탕으로 ICRF 방법을 제안하였다. 제안된 방법을 사용하면 단안 기반 하에서 비강체 평균 모양과 그에 해당하는 움직임을 복원할 수 있다. 구해진 평균 모양과 그에 해당하는 움직임 정보는 기존의 비강체 복원 방법[19,24]에 필요한 초기 복원 값으로 사용될 수 있어 좀 더 정확한 비강체 복원 결과를 얻을 수 있을 것으로 예상된다. 또한, 향후 연구로

본 논문에서 제안된 방법을 확장하여 비강체의 시간적으로 변화하는 모양까지 복원 할 수 있는 방법에 대한 연구를 수행할 예정에 있다.

참고 문헌

- [1] J. Aggarwal, W. Liao Q. Cai, and B. Sabata, "Nonrigid motion analysis: Articulated and elastic motion," *Computer Vision and Image Understanding*, Vol. 70, No. 2, pp. 142-156, 1998
- [2] R. Bajcsy, "Multiresolution elastic matching," *Computer Vision, Graphics and Image Processing*, pp. 46:1-21, 1989.
- [3] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3d faces," In *SIGGRAPH'99*, 1999.
- [4] M. Bran, "Morphable 3d models from video," In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 456-463, 2001.
- [5] M. Brand and R. Bhotika, "Flexible flow for 3d nonrigid tracking and shape recovery," In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 315-322, 2001.
- [6] C. Bregler, A. Hertzmann and H~iermann, "Recovering non-rigid 3d shape from image streams," *CVPR00*, pp. 690-696, 2000.
- [7] A. Del Bue and L. Agapito, "Non-rigid 3d shape recovery using stereo factorization," In *Asian Conference on Computer Vision*, 2004.
- [8] C. Davatzikos, D. Shen, A. Mohamed and S.K. Kyriacou, "A framework for predictive modeling of anatomical deformations," *IEEE Trans. on Med. Imaging*, Vol. 20, No. 8, pp. 836-843, August 2001.
- [9] D. Decarlo and D. Metaxas, "The integration of optical flow and deformable models with applications to human face shape and motion estimation," In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 231-238, 1996.

- [10] P. Huber, "Robust Statistics," John Wiley and Sons, 1981.
- [11] M. Irani, "Multi-frame optical flow estimation using subspace constraints," In Proc. Int. Conf. on Computer Vision, pp. 626-633, 1999.
- [12] M. Irani and P. Anandan, "Factorization with uncertainty," In Proc. European Conf. on Computer Vision, 2000.
- [13] M. Kass, A. Witkin and D. Terzopoulos, "nakes: Active contour models," Int. Journal of Computer Vision, Vol. 1 No. 4, pp. 312-331, 1987.
- [14] T. Kim and K.S. Hong, "On the reconstruction of approximate motion and average shape of deforming objects using a monocular view: application to the nonrigid human face," Asian Conf. on Computer Vision, 2004.
- [15] T. Kim and K.S. Hong, "Estimating approximate average shape and motion of deforming objects with a monocular view," International Journal of Pattern Recognition and Artificial Intelligence, Vol. 19, No. 4, pp. 585-601, 2005.
- [16] J. Martin, A. Pentland and R. Kikinis, "Shape analysis of brain structures using physical and experimental modes," Technical Report TR-276, MIT, 1994.
- [17] C. Nastar and N. Ayache, "Classification of nonrigid motion in 3d images using physics-based vibration analysis," In Proc. IEEE Workshop on Biomedical Image Analysis, 1994.
- [18] A.P. Pentland and B. Horowitz, "Recovery of nonrigid motion and structure," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 13 No. 7, pp. 730-742, July 1991.
- [19] H. Aanaes and F. Kahl, "Estimation of deformable structure and motion," In Workshop on Vision and Modelling of Dynamic Scenes, ECCV02, Copenhagen, Denmark, 2002.
- [20] S. Sclaroff and A. Pentland, "Physically-based combinations of views: Representing rigid and nonrigid motion," In Proc. IEEE Workshop on Nonrigid and Articulate Motion, 1994.
- [21] J. Shi and C. Tomasi, "Good features to track," In Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 593-600, 1994.
- [22] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: A factorization method," Int. Journal of Computer Vision, Vol. 9, No. 2, pp. 137-154, Nov 1992.
- [23] L. Torresani and C. Bregler, "Space-time tracking," In Proc. European Conf. on Computer Vision}, 2002.
- [24] L. Torresani, D.B. Yang, E.J. Alexander and C. Bregler, "Tracking and modeling non-rigid objects with rank constraints," In Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 493-500, 2001.
- [25] S. Vedula, S. Baker, P. Rander, R. Collins, and T. Kanade, "scene flow," In Proc. Int. Conf. on Computer Vision, pp. 722-729, 1999.
- [26] S. Vedula, S. Baker, S. Seitz and T. Kanade, "Shape and motion carving in 6d," In Proc. IEEE Conf. Computer Vision and Pattern Recognition, June 2000.



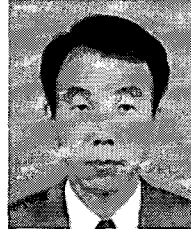
김 태 원

- 1996년 포항공과대학교 전자전기공학파(공학사)
- 1998년 포항공과대학교 전자전기공학파(공학석사)
- 2005년 포항공과대학교 전자전기공학파(공학박사)
- 2005년~현재 한국전자통신연구원 선임연구원
- 관심분야 : 컴퓨터비전, 3DTV, 증강현실, 패턴인식



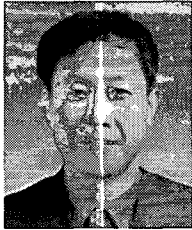
허 남 호

- 1992년 포항공과대학교 전자전기공학과(공학사)
- 1994년 포항공과대학교 전자전기공학과(공학석사)
- 2000년 포항공과대학교 전자전기공학과(공학박사)
- 2000년~현재 한국전자통신연구원 선임연구원
3DTV시스템연구팀 팀장
- 관심분야 : 3DTV, 휴먼 팩터, 모터 제어



이 수 인

- 1985년 경북대학교 전자공학과(공학사)
- 1989년 경북대학교 전자공학과(공학석사)
- 1996년 경북대학교 전자공학과(공학박사)
- 1990년~현재 한국전자통신연구원 책임연구원, 방송시스템연구그룹 그룹장
- 관심분야 : 지상파 DTV/DMB/DCATV 방송시스템, 3DTV 전송 시스템



김 진 응

- 1981년 서울대학교 전자공학과(공학사)
- 1983년 서울대학교 전자공학과(공학석사)
- 1993년 미국 Texas A&M University(공학박사)
- 1983년~현재 한국전자통신연구원 책임연구원
- 1983년~1984년 스웨덴 LM Ericsson사 방문연구원
- 2004년~2005년 미국 Columbia University 방문연구원
- 관심분야 : 신호처리, 대화형 방송시스템, 3DTV, 3D DMB 시스템



홍 기 상

- 1977년 서울대학교 전자공학과(공학사)
- 1979년 KAIST 전기전자공학과(공학석사)
- 1984년 KAIST 전기전자공학과(공학박사)
- 1984년~1986년 한국에너지기술연구원 선임연구원
- 1988년~1989년 CMU, Visiting professor
- 1986년~현재 포항공과대학교 전자전기공학과 교수
- 관심분야 : 3차원 물체인식, 컴퓨터비전, 영상처리