

다특징점 정보 및 최적화 기반 비조정 카메라 영상으로부터 머리 움직임 추정 방법

Optimization Approach for Pose Determination of Human Head Using Multi Feature Points From an Uncalibrated Camera

송민규 · 김진영 · 나승유
MinGyu Song, Jin Young Kim, Seung You Na

전남대학교 전자공학과

요 약

머리의 자세 및 움직임 추적은 응시추적 및 시각운율 연구에서 필수적이다. 일반적으로 머리자세를 추정하는 방법은 보정된 카메라를 통해 추출된 얼굴의 특징점 정보를 이용한다. 그러나 실제 응용 분야에서는 보정되지 않은 카메라를 통한 머리 움직임을 추정해야 할 경우가 발생한다. 이에 따라 본 논문에서는 보정되지 않은 하나의 카메라를 이용, 단일특징점 정보를 이용한 머리 자세 추정 방법을 확장하여 최적화 기법을 도입한 다특징점 정보 기반 머리 자세 추정방법에 대하여 논하였다.

키워드 : 머리 자세 추정(Head Pose Determination), 다특징점(Multi Feature Points), 비조정 카메라(Uncalibrated Camera)

1. 서 론

머리의 자세 및 움직임 추정은 인간과 컴퓨터의 상호 작용 연구에 있어서 중요한 분야이다. 이 분야에 대한 연구 중 대다수가 보정된 단일 카메라를 사용하여 추출된 얼굴의 특징점 정보를 이용하여 머리의 자세 및 움직임을 추정하는 것이다[1]. 하지만 실제 응용 사항에서는 종종 보정되지 않은 카메라로부터 머리의 자세 및 움직임을 추정해야 하는 경우가 발생한다. 본 논문에서는 참고문헌[2]의 결과를 확장하여 최적화 기법을 도입한 다특징점 정보 기반 머리 자세 추정 방법에 대하여 논한다.

참고문헌[2]은 MPEG 7의 표준 얼굴모델에 기반한 단일특징점 정보 기반 머리 자세 추정방법에 대하여 논의하였다. 그러나 논문의 결과 Pitch, Yaw 그리고 Roll이 독립적으로 발생하는 경우는 좋은 결과를 얻었지만, 복합적으로 머리 움직임이 있는 경우에 대해서는 큰 어려움이 발생하였다. 본 논문은 이러한 단점을 극복하기 위하여 다특징점 정보를 이용, 연속된 영상에서 이전 영상의 기본 특징점 정보와 현재 영상의 측정된 특징점 정보들의 거리가 최소화되도록 최적화하는 방법에 기반하여, 카메라의 보정 없이 머리의 자세를 정확하게 추정하는 방법을 제안하였다.

본 논문에서 제시한 방법은 가상의 얼굴 모델을 통해 일차적으로 검증하였으며, 시간 운율(Visual Prosody) 분석을 위해 전남대와 콜로라도 대학에서 구축된 DB를 대상으로 검증하였다.

2. 머리 자세에 대한 정의

일반적으로 머리의 자세는 아래와 같이 Pitch, Yaw 그리고 Roll 3개의 이동 가능한 각도로 정의된다[2] :

1) Pitch 각도는 YZ평면상에서 Y축에서 Z축으로 시계

방향으로 회전하는 각도를 의미한다.

2) Yaw 각도는 XZ평면에서 X축에서 Z축으로 반시계 방향으로 회전하는 각도를 의미한다.

3) Roll 각도는 XY평면에서 X축에서 Y축으로 반시계 방향으로 회전하는 각도를 의미한다.

그림1은 머리 자세의 움직임 방향을 보여준다.

3. 머리의 자세 추정 방법

머리의 자세 추정은 각각의 Pitch, Yaw, Roll 각도가 알려진 영상의 특징점 정보와 새로운 영상의 특징점 정보를 비교하여 얻을 수 있다.

머리의 자세를 알 수 있는 이미지에서의 눈의 위치를 $U=(U_x, U_y)$ 라 가정하고 자세를 모르는 이미지에서의 눈의 위치를 $V=(V_x, V_y)$ 라 가정한다.

머리 회전의 중심점은 $C=(C_x, C_y, C_z)$ 이고 C 는 $(0 ES, -1290 ES, -270 ES)$ 로 초기화 한다. 여기서 ES 는 MPEG-7 표준에 따라 이미지의 두 눈 사이의 거리를 1024로 나눈 값이다. E_z 는 정면을 바라보는 영상에서 눈의 가운데 좌표의 깊이를 나타내며 1239 ES 이다.

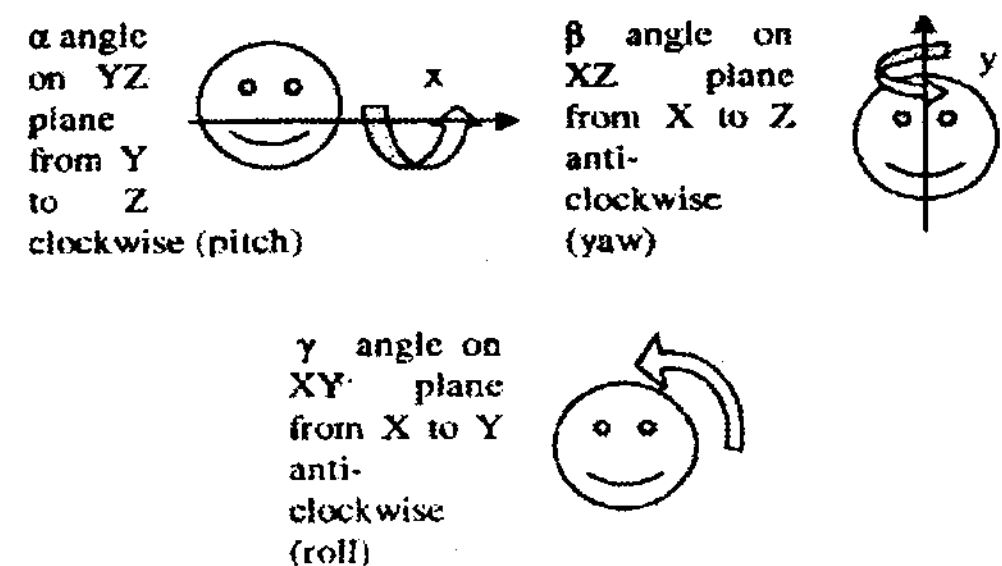


그림 1 머리 자세의 움직임 방향 (Pitch, Yaw, Roll)

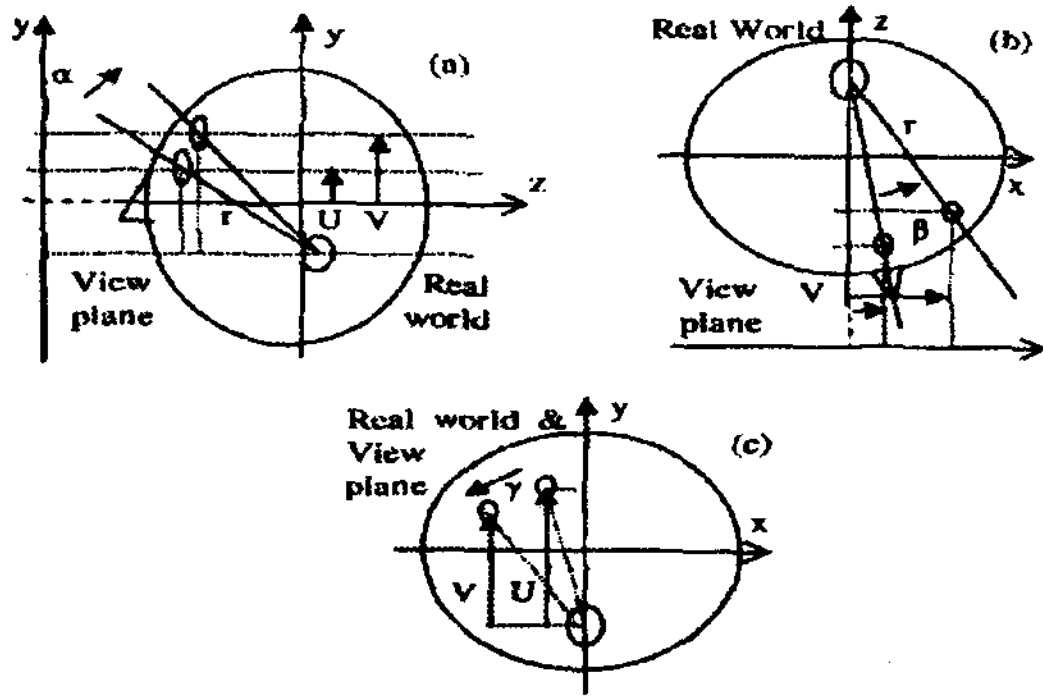


그림 2 머리 자세의 변화에 따른 눈의 위치 변화

단일 특징점 정보를 이용하여 머리 자세(pitch, yaw, roll)를 추정하는 경우 그림2의 눈의 위치 변화를 이용하여 다음 식 (1), (2), (3)으로 계산된다.

$$\alpha = \sin^{-1}\left(\frac{V_Y - C_Y}{r}\right) - \sin^{-1}\left(\frac{U_Y - C_Y}{r}\right) \quad (1)$$

$$\text{where } r = \sqrt{(U_Y - C_Y)^2 + (E_Z - C_Z)^2}$$

$$\beta = \sin^{-1}\left(\frac{V_X - C_X}{r}\right) - \sin^{-1}\left(\frac{U_X - C_X}{r}\right) \quad (2)$$

$$\text{where } r = \sqrt{(U_X - C_X)^2 + (E_Z - C_Z)^2}$$

$$\gamma = \tan^{-1}\left(\frac{V_Y}{-V_X}\right) - \tan^{-1}\left(\frac{U_Y}{-U_X}\right) \quad (3)$$

하지만 단일특징점 정보에 의해 Pitch, Yaw 그리고 Roll중에 둘 이상이 동시에 변화하는 경우 큰 에러가 발생한다. 그 문제를 해결하기 위해 다특징점 정보를 이용하였다. 머리의 자세를 알 수 있는 정면을 바라보고 있는 사람의 다특징점은 $P_i = (X_i, Y_i, Z_i)$ 이며 변화된 자세의 특징점들은 $\bar{P}_i = (\bar{X}_i, \bar{Y}_i, \bar{Z}_i)$ 이다. 측정된 특징점들은 다음(4)식에 의해 3D평면으로 매핑 된다.

$$P'_i = RP_i \quad (4)$$

(여기서 R은 2D평면을 3D평면으로 매핑 시켜주는 3D회전 행렬로 참고문헌[1]를 참조)

이 특징점들을 이용하여 머리의 기본 자세에서 움직임이 있는 경우, 기본 특징점과 측정된 특징점들과 거리가 최소화 하도록 최적화를 하는 방법을 통해 머리 자세의 움직임을 추정하였다.

본 논문에서는 하나의 카메라를 사용하기 때문에 \bar{Z}_i 는 측정할 수 없다. 따라서 X축과 Y축만을 고려하여 에러를 계산하며 값은 다음 식(5)와 같다.

$$E_i = \|P'_i - \bar{P}_i\|^2 \quad (5)$$

$$E_i = (\bar{X}_i - X_i)^2 + (\bar{Y}_i - Y_i)^2$$

$\sum_i E_i$ 를 최소화 하기위한 방법으로는 비선형 Least squares problem을 적용하였다.

4. 실험 및 결과

제안된 알고리즘을 검증하기위해 사용된 DB는 시각운율(Visual Prosody) 분석을 위해 전남대와 콜로라도 대학에서 구축한 총 남녀 총 22명에 대한 시청각DB[3]로써, 그림3은 DB의 한 장면이다.

그림3은 정면을 바라보고 있는 사진으로써 Pitch와 Yaw는 0라고 가정하고 P1, P2, P3, P4의 좌표를 이용하여 Roll값의 초기값을 결정하고 최적화 방법에 기반하여

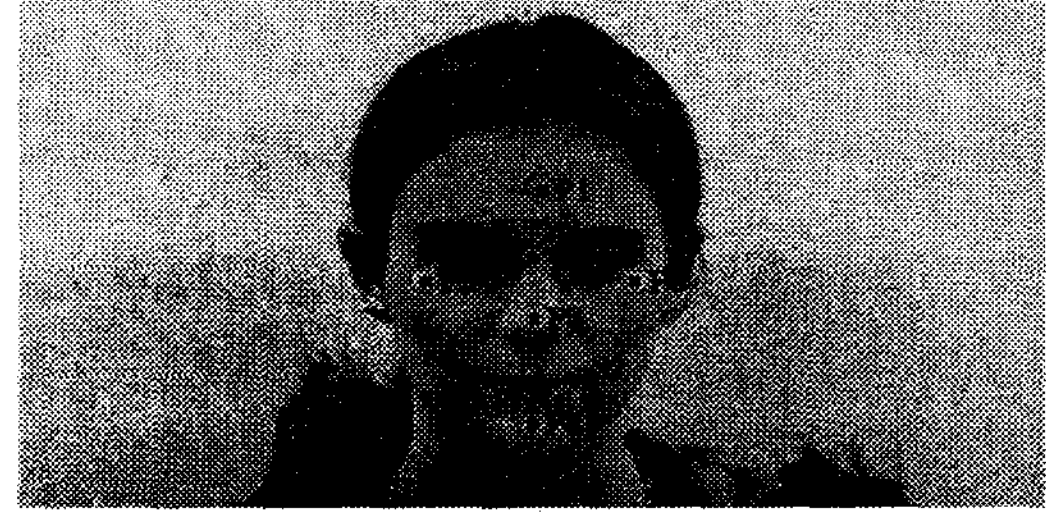


그림 3 실험에 사용된 DB의 한 장면

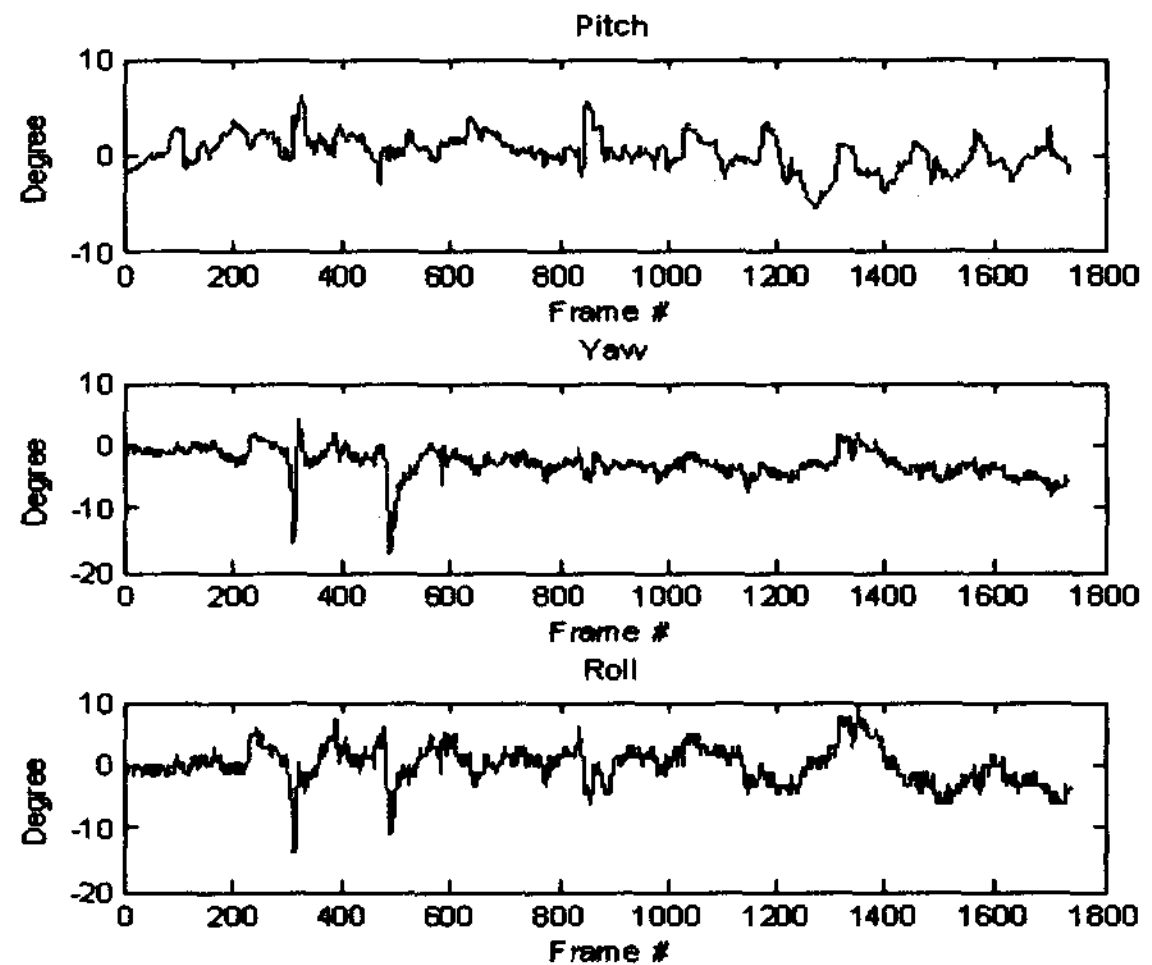


그림 4 머리자세의 움직임 추정 예(한 frame는 1/30sec)

머리 자세의 움직임을 추정하였다.

그림4는 55초 정도 되는 비디오에 제안된 방법을 적용하여 머리 자세의 움직임을 프레임별로 추정한 예이다. 실험 결과 보정되지 않은 단일 카메라를 사용하여 Pitch, Yaw, Roll의 변화가 복합적으로 존재하는 경우에도 큰 에러 없이 머리자세의 변화값을 추정하는 결과를 보였다.

5. 결론

본 논문에서는 연속되는 영상에서 변화하는 머리 자세의 움직임을 보정되지 않은 단일 카메라를 통해 추출한 다특징점 정보를 이용하여 추정하는 방법을 제안하였다. 실험결과 머리의 자세가 복합적으로 바뀌는 상황에서도 좋은 결과를 보였다. 앞으로 우리는 이 제안된 방법을 통해 얻은 머리 자세의 움직임을 이용하여 시청각운율을 분석하는 연구에 적용 할 것이다.

참 고 문 헌

[1] D. Dervinis, "Head Orientation Estimation using Characteristic Points of Face", internet.ktu.lt/lt/mokslas/zurnalai/elektr/z72/1392-1215-2006-11-07-61.pdf
 [2] B. Yip., W. Y. Siu and J.S. Jin, "Pose Determination of Human Head Using One Feature Point," 2004 IEEE International Conference on Multimedia and Expo(ICME), pp1182-1186
 [3] 송민규 외3인, "CU-CNU 시청각 데이터베이스 : 입술, 얼굴, 머리의 움직임에 관한 시청각 코퍼스", 2007 한국통신학회 춘계학술대회 논문집, pp95-98