

반복적 확장 칼만 필터를 이용한 얼굴의 3차원 움직임량 추정

박 강 령[†] · 김 재 회^{††}

요 약

컴퓨터 시각 인식 방법을 이용하여 얼굴의 3차원 움직임량을 추정하고자 하는 연구는 가상 현실 환경에서 얼굴 움직임에 의한 3차원 그래픽 화면 조정, 시뮬레이터에서의 훈련자 얼굴 움직임에 의한 화면 조정 및 모니터상의 시선 위치 파악 등을 위해 필수적으로 요구되는 기술로서 최근 활발히 연구되고 있다. 기존에 얼굴의 3차원 움직임량을 추정하고자 하는 연구들은 대부분 확장 칼만 필터(extended kalman filter)를 이용하였으나, 이러한 방법은 필터의 초기값을 정확하게 설정해야하는 제약 요소를 갖고 있으며, 또한 얼굴의 회전 방향 변화시 이에 대처하지 못하는 경우 역시 종종 발생한다. 본 논문에서는 이러한 문제점을 해결하기 위하여 확장 칼만 필터의 변형 형태인 반복적 확장 칼만 필터를 이용하여 얼굴의 3차원 움직임량을 추정하였다. 반복적 확장 칼만 필터에서는 확장 칼만 필터에서 계산되어 나오는 추정 오차 공분산 행렬 값이 미리 정해진 임계치보다 커지는 경우, 현재 얼굴의 움직임량을 제대로 추정하지 못하고 있는 것으로 판단하여 칼만 필터에서 사용하는 회전 및 이동 속도, 그리고 회전 및 이동 가속도를 변형함으로써 얼굴의 움직임량을 정확하게 추정할 수 있도록 하는 방법이다. 실험 결과 반복적 확장 칼만 필터를 사용하였을 경우에 얼굴의 급격한 회전 방향 변화에도 얼굴의 3차원 움직임량을 정확하게 추정할 수 있음을 알 수 있었다.

3-D Facial Motion Estimation Using Iterative Extended Kalman Filter

Kang-Ryoung Park[†] · Jai-Hie Kim^{††}

ABSTRACT

Accurate 3-D facial motion estimation using computer vision is required for 3-D view control in desktop VR system or simulator and gaze detection on a monitor, etc and it has recently been researched vigorously. Most previous researches use a EKF (Extended Kalman Filter) for 3-D facial motion estimation, but it has restriction of choosing accurate initial filter parameters. In addition, the EKF cannot estimate 3-D facial motion when a user changes his face direction abruptly. To overcome such problems, we use IEKF (Iterative Extended Kalman Filter) transformed from EKF. IEKF is that when the estimated error covariance value calculated from IEKF exceeds in a predetermined threshold, it adjusts the initial parameters (rotational and translational velocities or accelerations) adaptively so that the 3-D facial motion can be estimated exactly. As experimental results, IEKF can even estimate 3-D facial motion accurately in case of abrupt facial directional change.

1. 서 론

얼굴의 3차원 움직임량을 정확하게 추적하는 연구들은 컴퓨터 비전과 휴먼 컴퓨터 인터페이스 등에서 많은 응용 범위를 가지고 최근 활발히 연구가 진행되고 있다. 그 대표적인 예로는 화상회의 시스템에서 사용자의 얼굴 움직임에 따라 원격 카메라를 자동으로 포커싱(focusing)하는 시스템, 그리고 공장 자동화 시스템에서 관리자의 양손을 입력 수단으로 사용하지 못하는 상황에서 사용자의 얼굴 움직임에 의해 프로세스 제어를 하는 시스템, 가상 현실 환경에서의 얼굴 움직임에 의한 3차원 그래픽 화면 조정, 시뮬레이터에서의 훈련자 얼굴 움직임에 의한 화면 조정, 모니터상의 시

선 위치 파악 및 model-based image coding 등이 있다[1]. 기존에 사용된 얼굴의 3차원 움직임 추정 알고리즘으로는 신경망(Neural Network), 광류 분석법(optical flow methods) 및 확장 칼만 필터(Extended Kalman Filter)등이 있는데, 이 가운데 신경망과 광류분석법은 처리시간이 오래 걸리고 움직임 추정 에러 크다는 단점 때문에 확장 칼만 필터를 많이 이용하고 있다[2-7, 11, 13]. 그러나 이러한 확장 칼만 필터 방법 역시 필터의 초기값에 따라 얼굴 움직임을 제대로 추정하지 못하고 발산하는 경우가 종종 발생하며, 또한 얼굴의 회전 방향 변화시 이에 대처하지 못하는 경우가 종종 발생한다. 본 논문에서는 이러한 문제점을 해결하기 위하여 확장 칼만 필터의 변형 형태인 반복적 확장 칼만 필터를 이용하여 얼굴의 3차원 움직임량을 추정하였다.

[†] 정 회 원 : LG 전자 기술원 선임연구원

^{††} 정 회 원 : 연세대학교 전기·컴퓨터공학과 교수

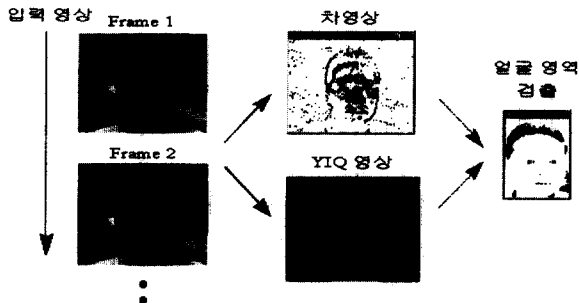
논문접수 : 2000년 7월 18일, 심사완료 : 2001년 1월 26일

2. 얼굴 특징점 추출

얼굴의 3차원 움직임량을 추정하기 위하여, 본 논문에서는 2차원 카메라 영상에서 관측된 얼굴 특징점(양 눈, 코, 입)의 위치 변화도를 이용하였다. 입력 영상에서 얼굴 특징점의 위치를 추출하기 위하여 이 논문에서는 먼저 얼굴 영역을 검출한 후 추출된 얼굴 영역내의 제한된 범위 내에서 양 눈과 코 및 입의 양 끝점을 추출하였다.

2.1 얼굴 영역의 추출

이 논문에서는 (그림 1)과 같이 시간적으로 연속된 두 영상간의 차영상 정보와 칼라 정보를 이용하여 얼굴 영역을 검출한다. 얼굴의 살색 영상 처리부에서는 입력된 얼굴의 살색 칼라 정보에 대한 RGB신호를 YIQ model로 변환함으로써 얼굴의 살색 정보에 민감한 I성분 구간(110~150)을 바탕으로 얼굴 영역을 검출한다[1].

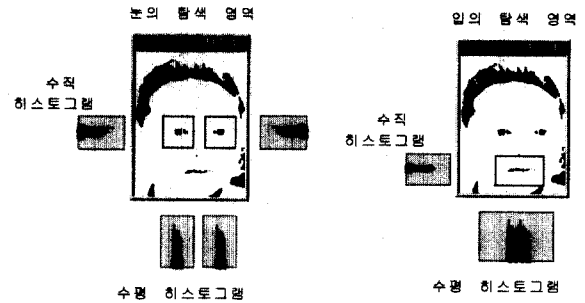


(그림 1) 차영상과 살색 정보를 이용한 얼굴 영역 검출

2.2 눈동자, 코 및 입의 양 끝점 추출 및 움직임 추적

얼굴 영역을 추출한 후, 추출된 얼굴 영역에서 조도의 변화에 따른 영향을 줄이기 위하여 히스토그램 평활화(histogram equalization) 및 이진화(binairization) 과정을 통해 이진 영상으로 변환하였다. 얼굴 영역에 대한 이진 영상이 구해진 후, (그림 2-a)처럼 얼굴 내 눈의 위치에 대한 사전지식을 이용하여 이진 영상의 제한된 범위 내에서 수직, 수평 히스토그램의 피크 위치(peak position)를 파악함으로써 눈의 위치를 정확하게 추출할 수 있었다. 양 눈이 검출된 후 (그림 2-b)처럼 입의 위치에 대한 존재 가능 범위를 설정한 후, 이 영역에 대한 이진화 및 수직 방향 히스토그램으로 입선의 수직 위치를 먼저 추출하였다.

추출된 입선의 수직 위치로부터 입의 양 끝점을 추출하기 위해 입선에 대한 수평 히스토그램을 구하여 임계치(수평히스토그램의 평균값)이상으로 갑작스럽게 변화되는 지점을 입의 양 끝점으로 결정하였다. 콧구멍 역시 눈동자와 같은 방법으로 추출하였으며, 두 콧구멍의 근접도가 큰 관계로 양 콧구멍을 포함한 영역에 대한 수평 히스토그램을 통해 두 개의 피크 위치를 추출함으로써 양 콧구멍의 수평

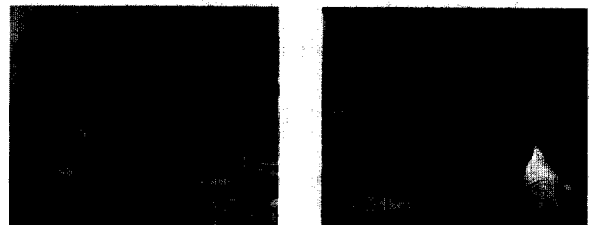


(a) 눈 영역 검출

(b) 입 영역 검출

(그림 2) 눈 및 입 영역 검출

위치를 파악하였다. 눈 위치 추출시 안경 사용자의 경우, 안경알에 의해 반사되어 눈동자 흑화소가 많이 소실될 수 있다. 이러한 문제점을 해결하기 위하여 눈 영역의 이진화 시 p-tile method를 사용하여 흑화소의 개수를 일정하게 유지함으로써 흑화소 소실 문제를 해결하였다. (그림 3)은 본 논문의 방법을 사용하여 검출된 사용자의 얼굴 영역 및 눈, 코, 입의 양 끝 위치를 나타낸 것이다. 초기 영상에서의 얼굴 특징점 추출 방법과는 달리 이후 연속 영상에서는 매 영상마다 얼굴 영역을 추출하지 않고 이전에 추출된 특징점 부근을 탐색하는 방법을 이용하여 특징점의 위치를 추적하였다. 이때 이전 몇 장의 영상에서 특징점의 움직임 정보를 바탕으로 현재 특징점의 위치를 예측하는 알고리즘을 사용하였다. 예측된 특징점 부근에 정해진 크기의 탐색 영역을 설정하여 이 탐색 영역내의 히스토그램 분석을 통해 특징점들을 추적하였다. 이때 전체 영상 크기는 320×240 pixel이고, 이중 얼굴 영역의 평균 크기는 144×208 pixel, 눈 영역의 평균 크기는 8×4 pixel, 콧구멍의 평균 크기는 4×4 pixel, 그리고 입의 끝점의 평균 크기는 3×3 pixel이다. 여기서 외부 조명은 급격히 변하지 않는다고 가정하였다. 실험 결과, 얼굴 특징점의 실제 위치(사람이 눈으로 보고 직접 추출한 특징점의 위치)와 본 논문의 방법에 의해 추출된 특징점 위치 사이의 최소 자승 에러(RMS error)는 <표 1>과 같다.



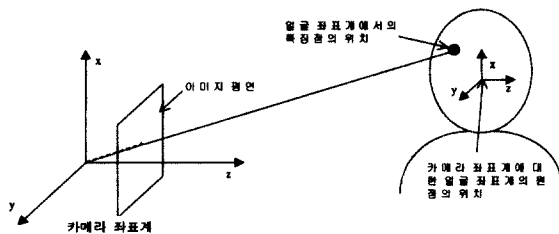
(그림 3) 추출된 얼굴 영역 및 눈, 코, 입 특징점 위치

<표 1> 얼굴 특징점의 추출 정확도

특징점	눈	코	입
최소 자승 에러 (pixel)	2	5	3

3. 카메라 및 얼굴 좌표계

추출된 얼굴 특징점의 움직임 정보를 이용하여 본 논문에서는 얼굴의 3차원 움직임량을 추정한다. 이때, 본 논문에서 카메라 좌표계에 대한 얼굴 움직임을 이동(translation)과 회전(rotation)으로 분리하여 얼굴의 이동은 카메라 좌표계 원점에 대한 얼굴 좌표계 원점의 이동으로, 얼굴의 회전은 얼굴 축을 기준으로 한 회전으로 정의하였다. 이처럼 얼굴의 움직임을 정의하기 위해 3차원상의 얼굴 특징점을 카메라 중심 좌표계와 얼굴 중심 좌표계 두 부분으로 나누어 나타냈다. 이렇게 얼굴의 특징점을 두 부분으로 나타냄으로써 얼굴의 회전을 카메라를 기준으로 본 것이 아니라 얼굴 축을 기준으로 하여 일관성 있는 변수로 모델링할 수 있는 장점이 있다. (그림 4)는 본 논문에서 사용하는 3차원 얼굴 좌표계와 카메라 좌표계를 나타낸 것이다. 이때 본 논문에서는 (그림 4)와 같이 초기에 카메라 좌표계에 대해 얼굴 좌표계의 y축의 원점은 서로 일치한 것으로 가정했으며, 동시에 x축 역시 초기에만 서로 평행한 것으로 가정하였다. 이를 위하여 초기에 사용자가 수동으로 카메라 좌표계에 대해 얼굴 좌표계의 x, y축을 맞추어 주도록 하였다. 여기서 얼굴 좌표계를 별도로 설정한 이유는 얼굴의 이동만 존재하는 경우 이를 카메라 좌표계에서 표현할 수 있지만, 얼굴의 회전의 경우 얼굴 좌표계에서 설정된 얼굴 축(x, y, z축)을 기준으로 이루어지므로 카메라 좌표계에서는 표현하기 어렵기 때문이다.



(그림 4) 카메라 좌표계 및 얼굴 좌표계

3.1 카메라 투영 모델

(그림 4)에서 처럼 3차원 상에서의 얼굴 특징점은 카메라의 2차원 이미지 평면에 식 (1)과 같이 원근 투영 변환(perspective projection transform)에 의해 맺히게 된다.

$$X = f \cdot \frac{x}{z} \quad Y = f \cdot \frac{y}{z} \quad (1)$$

x, y, z는 카메라 좌표계를 기준으로 한 3차원 상에서의 얼굴의 특징점의 좌표이고, X와 Y는 이 특징점이 카메라의 2차원 이미지 평면에 투영된 좌표이다. 이때 초기 z은 3차원 움직임 추적 센서인 polhemus sensor를 이용하여 측정된 값을 사용하였다. polhemus sensor는 전자기장 센싱 방법에 의해 삼차원 이동량(translation) 및 회전량(rotation)을 정확하게 측정할 수 있는 장비이다(평균 0.08 cm의 이동량

측정 에러 및 0.15°의 회전량 측정 에러)[17]. f는 카메라의 초점거리이며, 본 논문에서는 카메라의 spec 상에 나와 있는 값을 참조하였으며(f=8mm), 이 값의 정확도를 측정하기 위하여 렌즈 캘리 브레이션을 별도로 수행하여 정확한 카메라 초점 거리를 구하였다(f=7.8mm) [18]. 이 때 <표 1>과 같이 2차원 카메라 좌표계에서 관측된 얼굴 특징점의 위치 추출 자체에는 에러가 포함되어 있으므로, 위의 식은 다음과 같이 모델링할 수 있다.

$$X = f \cdot \frac{x}{z} + n_x \quad Y = f \cdot \frac{y}{z} + n_y \quad (2)$$

n_x 와 n_y 는 이미지 평면에서의 추출 에러로 본 논문에서는 평균이 0인 독립적 가우스 분포(independent zero mean Gaussian distribution)로 모델링한다.

3.2 얼굴의 3차원 움직임 모델

전술한 바와 같이 본 논문에서는 얼굴의 움직임을 이동과 회전에 분리하기 위해 3차원 상에서 얼굴 특징점의 좌표를 카메라 중심 좌표계와 얼굴 중심 좌표계로 나누어 나타냈다. 이때 얼굴의 이동은 카메라 좌표계를 기준으로 하여 얼굴 좌표계 원점의 이동으로, 얼굴의 회전은 얼굴 좌표계에서 얼굴 축을 기준으로 한 얼굴 특징점의 회전에 정의한다. p_i 를 얼굴 좌표계에서 파악된 특징점의 위치 벡터(x_i, y_i, z_i , 왼쪽 눈: $i=1$, 오른쪽 눈: $i=2$, 왼쪽 콧구멍: $i=3$, 오른쪽 콧구멍: $i=4$, 왼쪽 입끝점: $i=5$, 오른쪽 입끝점: $i=6$)라 하고 s 를 카메라 좌표계에 대한 얼굴 좌표계 원점의 위치 벡터라 하면 다음과 같은 관계식을 얻을 수 있다.

$$P_i = s + R p_i \quad (3)$$

단, $i=1, 2, \dots, 6$ 는 각각 왼쪽 눈, 오른쪽 눈, 왼쪽 콧구멍, 오른쪽 콧구멍, 왼쪽 입끝점, 오른쪽 입 끝점

P_i 는 카메라 좌표계에서 관측된 얼굴 특징점의 좌표이고, R 은 3×3 회전 행렬이다.

3.3 이동 움직임 모델(model of translational motion)

이때 식 (3)에서 얼굴의 이동(s)은 짧게 정한 시간 간격(Δt)동안 등가속도 운동으로 가정할 수 있으므로 다음 식 (4)와 같이 등가속도 운동 모델로 나타낸다.

$$s = s_0 \cdot \Delta t + \frac{1}{2} \cdot a \cdot \Delta t^2 \quad (4)$$

s_0 는 카메라 좌표계에서 초기 얼굴 좌표계 원점의 좌표 벡터이고, a 는 원점의 이동에 대한 가속도 벡터이다.

3.4 회전 움직임 모델(model of rotational motion)

얼굴의 회전은 3차원 상에서 얼굴 좌표계를 기준으로 하

여 얼굴 특징점의 회전으로 정의하였다. 이 논문에서는 오일러 각도(euler angle)를 이용하여 시간에 따른 얼굴의 회전을 나타내었다. 식 (4)와 마찬가지로 얼굴의 회전 역시 짧은 단위 시간동안의 등가속도 운동으로 가정하여 회전을 모델링하였다.

4. 확장 칼만 필터

본 논문에서는 앞에서 기술한 카메라 투영 모델과 얼굴의 3차원 움직임 모델을 바탕으로 반복적 확장 칼만 필터를 이용하여 얼굴의 3차원 움직임량을 추정하였다. 반복적 확장 칼만 필터의 출발점이 되는 확장 칼만 필터를 먼저 살펴보기로 한다.

확장 칼만 필터에 의한 삼차원 움직임량 추정방법은 예측(prediction), 측정(measurement) 및 보정(update)의 3단계를 반복하면서 얼굴의 삼차원 움직임량을 추정한다[11-13]. 즉, 이차원 카메라 영상에서 추출된 얼굴 특징점의 위치와 등가속도 운동 모델에 기반하여 얼굴의 삼차원 이동(translation) 및 회전(rotation)량을 구하는 것이다. 확장 칼만 필터에 의해 삼차원 움직임량을 추정하기 위하여 본 논문에서는 식 (5)와 같이 18×1 크기의 상태 벡터(state vector)를 정의하였다.

$$\mathbf{x}(t) = (\mathbf{p}(t), \mathbf{p}(t), \mathbf{v}(t), \mathbf{w}(t), \mathbf{a}(t), \mathbf{a}(t))^T \quad (5)$$

여기서 $\mathbf{p}(t)$ 는 카메라 좌표계에 대한 얼굴 좌표계 원점의 X, Y, Z축 방향의 삼차원 이동량(translation)을 나타내는 3×1 벡터이고, $\mathbf{p}(t)$ 는 얼굴 좌표계내에서 얼굴 특징점의 X, Y, Z축 중심의 삼차원 회전각도(rotation angle)를 나타내는 3×1 벡터이다. 또한, $\mathbf{v}(t), \mathbf{w}(t)$ 는 각각 삼차원 이동 및 회전 속도를 나타내는 3×1 벡터들이고, $\mathbf{a}(t), \mathbf{a}(t)$ 는 각각 삼차원 이동 및 회전 가속도를 나타내는 3×1 벡터들이다. 이러한 상태 벡터는 3.3 및 3.4절에서 전술한 바와 같이 다음과 같은 등가속도 운동 모델에 기반한 상태 천이 행렬(state transition matrix), $\Phi(\Delta t)$ 에 의해 식 (6)과 같이 다음 단계의 삼차원 움직임량으로 예측된다.

$$\hat{\mathbf{x}}(t+1)^- = \Phi(\Delta t) \hat{\mathbf{x}}(t) + \mathbf{u}(t) \quad (6)$$

또한, 3차원 상에서의 얼굴 특징점의 좌표를 2차원 이미지로 투영시키는 과정에는 아파인 변환(affine transformation)과 원근 투영 변환(perspective projection transformation)이 함께 포함되어 있으므로, 선형적인 행렬 형태로 나타내지 못하고 식(8)과 같이 비 선형적인 함수($h(\cdot)$)의 형태로 나타낸다. 여기서 $\mathbf{u}(t)$ 는 제로 평균(zero mean)과 분산값(U)을 가지는 가우시안 잡음(gaussian noise)이다. 이러한 과정에서 측정 잡음이 섞이게 되는데 본 논문에서는

잡음 변수를 독립적 평균이 0인 가우스 분포로 가정하였다.

$$\mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{v}_k \quad (7)$$

식 (6)과 (7)로부터, 예측된 삼차원 움직임량 ($\hat{\mathbf{x}}(t+1)^-$)은 식 (8)에 의해 보정되어, 정확한 삼차원 움직임량을 구하게 된다.

$$\begin{aligned} \hat{\mathbf{x}}(t+1) &= \hat{\mathbf{x}}(t+1)^- \\ &+ K(t+1) (\mathbf{y}(t+1) - h(\hat{\mathbf{x}}(t+1)^-)) \end{aligned} \quad (8)$$

여기서 $K(t+1)$ 는 칼만 게인(Kalman gain)이며, 이때 칼만 게인은 다음 식 (9), (10), (11)과 같이 추정오차 공분산 행렬로 나타낼 수 있다.

$$K(t+1) = \mathbf{P}_k^+ H_k^T (H_k \mathbf{P}_k^+ H_k^T + \mathbf{R}_k)^{-1} \quad (9)$$

$$\mathbf{P}_k = (\mathbf{I} - K(t+1) H_k) \mathbf{P}_k^+ \quad (10)$$

$$\mathbf{P}_{k+1} = \Phi(\Delta t) \mathbf{P}_k \Phi(\Delta t)^T + \mathbf{Q}_k \quad (11)$$

$\mathbf{P}_k$ 는 추정오차 공분산 행렬(estimation error covariance matrix)이고, $\mathbf{R}_k, \mathbf{Q}_k$ 는 각각 측정(measurement) 잡음과 과정(process) 잡음이다. H_k 는 $h(\cdot)$ 함수를 상태 벡터에 관하여 편미분한 자코비안 행렬이다[7-9, 15].

$$H_k = \left[\frac{\partial h}{\partial \mathbf{x}} \right]_{\mathbf{x}=\hat{\mathbf{x}}(t)} \quad (12)$$

5. 반복적 확장 칼만 필터

이러한 확장 칼만 필터는 정확한 초기값을 사용하지 못하는 경우 필터가 수렴하지 못하고 발산하는 경우가 종종 발생한다. 동시에 식 (6)에서처럼 구간 등가속도 모델에 의해 예측하므로 얼굴의 급격한 방향 변환이 생기는 경우 추적하던 얼굴을 놓쳐버리는 경우가 종종 발생한다. 이러한 문제를 해결하기 위하여 본 논문에서는 반복적인 확장 칼만 필터 방법을 사용하였다.

반복적 확장 칼만 필터(iterated extended Kalman filter)는 이러한 확장 칼만 필터가 가지고 있는 문제점을 보완하기 위해 (10)의 식에 의해 구해진 추정 오차 공분산 행렬 값이 정해진 임계치보다 커지는 경우, 이전 값에서 현재 값을 추정하는 과정에서 반복적으로 칼만 게인을 바꾸어 가며 가장 알맞은 값을 추정하는 알고리즘이다[19, 20]. 반복적 확장 칼만 필터에서는 확장 칼만 필터 식(8)을 식(13)과 같이 변형하여 현재 값을 추정한다.

$$\begin{aligned} \hat{\mathbf{x}}(t+1) &= \hat{\mathbf{x}}(t+1)^- + K(t+1) (\mathbf{y}(t+1) - h(\hat{\mathbf{x}}(t+1)^-)) \\ &- H(\hat{\mathbf{x}}_k^+ - \hat{\mathbf{x}}(t+1)^-) \end{aligned} \quad (13)$$

이때 $\hat{x}_k^-$ 은 초기에 $\hat{x}(0)^-$ 으로 정하여 t 는 0부터 시작하여 t 의 값이 하나씩 증가하면서 $\hat{x}(t)^-$ 와 $\hat{x}(t+1)^-$ 의 차가 정한 값보다 작을 때까지 반복하여 칼만계인을 바꾸어 가며 알맞은 값을 찾는다. 여기서 얻은 $\hat{x}(t+1)^-$ 가 이전 값에서 예측한 현재 값($\hat{x}_k^-$)이다. 이러한 과정은 확장 칼만 필터의 루프에서 칼만계인을 이용하여 이전 값과 측정값을 이용하여 현재값을 추정할 때 포함되어 보다 정확한 현재 값을 추정한다. 즉, 식 (6), (7), (8)에 의해 계산된 칼만 계인을 그대로 이용하여 얼굴의 3차원 움직임량을 추정하는 것이 아니라, 식 (5)의 회전 및 이동 속도, 그리고 회전 및 이동 각속도를 변형하여 추정 오차 공분산 행렬 값이 정해진 임계치 이내의 값이 되도록 만듦으로써 이때 계산된 칼만 계인으로부터 얼굴의 움직임량을 추정한다.

6. 실험 및 결과

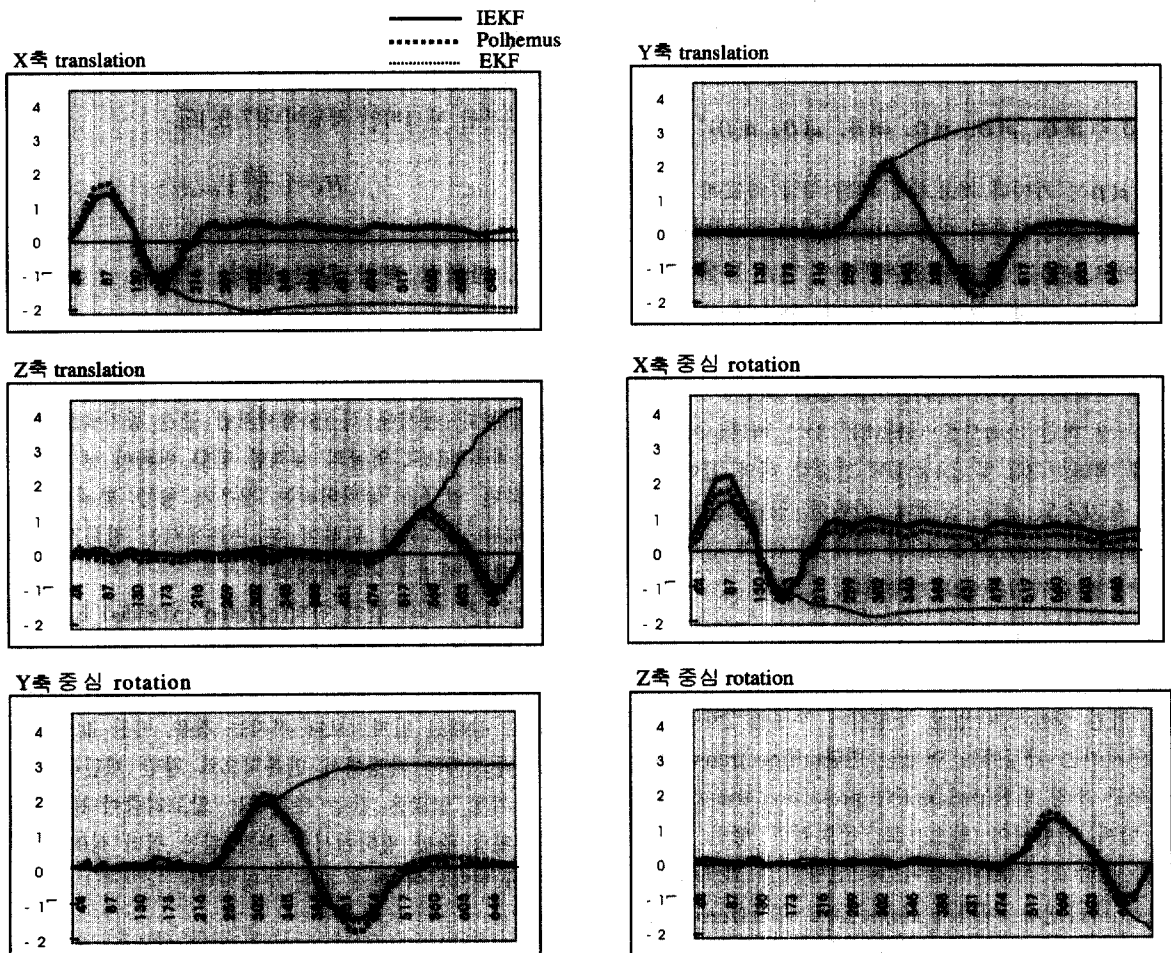
실험환경으로는 Pentium-Pro 200MHz를 사용하여 모니터 위에 설치된 CCD카메라(CCN-241IA)와 비디오 capture

board (DT3152)를 통해 사용자의 얼굴 움직임을 입력하고, 이때 입력된 영상으로부터 얼굴의 3차원 움직임량을 추정하였다. 본 논문에서는 IEKF에 의한 움직임량(X, Y, Z translation 량, X, Y, Z rotation 량) 추정 성능을 기존의 EKF뿐 아니라, 실제 가상 현실 시스템에서 많이 사용되는 위치 추적 장비인 Polhemus Fastrak과 비교 실험하였다. Polhemus Fastrak는 송신단과 수신단 사이의 전자기장 결합 방식에 의해 구동되며, 최대 60Hz의 sampling rate으로 3차원 움직임량을 나타내는 6가지 자유도 변수(X, Y, Z translation량, 3 회전각(yaw, pitch, roll))를 나타낸다. 다음(그림 5)는 실험에 사용한 데이터의 예이다.



(그림 5) 얼굴의 3차원 움직임량 추정을 위한 실험 데이터의 예

(그림 6)은 IEKF에 의한 방법과 EKF, 그리고 Polhemus



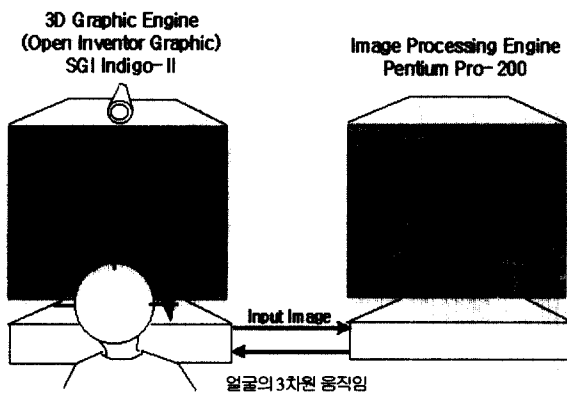
(그림 6) IEKF의 추정 정확도

Fastrak과의 성능 비교 결과이다.

(그림 6)의 translation 그래프에서의 세로축은 cm 단위이고 rotation 그래프에서의 세로축은 도(degree) 단위이다. 그래프에서 가로축은 영상 frame을 나타낸다. 실험 결과, IEKF의 경우 Polhemus Fastrak에 의해 파악된 얼굴의 실제 움직임량과 거의 유사하게 3-D 움직임량을 추정함을 알 수 있다(평균 2.25도의 회전량 추정 에러와 2.58cm의 이동량 추정 에러). 특히 (그림 6)에서 알 수 있듯이 EKF의 경우는 얼굴의 급격한 움직임 방향 변화가 생기게 되면 추적하지 못하고 발산해 버리는 결과를 나타냈지만, IEKF는 제대로 추적하는 결과를 나타냈다.

7. 얼굴의 3차원 움직임 추정을 이용한 3D view의 조정

본 논문에서는 이러한 얼굴의 3차원 움직임 추정을 이용하여 desktop VR환경에서 3차원 그래픽 view를 조정하는 시스템을 개발하였다. (그림 7)은 실제 시스템의 구성 환경이다. 3D 화면의 관심이 있는 부분을 보기 위해 얼굴을 움직이는 경우, 얼굴의 3차원 움직임량을 계속 추적하여 추출된 움직임량을 바탕으로 3D view를 조정한다. 실험 결과 본 연구의 알고리즘을 사용하여 3D view를 조정하였을 경우, 기존의 mouse 장비 등을 사용하였을 경우보다 현실감 나는 가상 현실 환경을 제공할 수 있었다. 연속적으로 입력되는 영상은 영상 처리부(Win95환경의 Pentium Pro 200MHz)에서 처리되며, 계산된 얼굴의 3차원 움직임 결과는 Unix/Window Socket을 이용하여 SGI Workstation으로 전송되어 3차원 view를 조정한다.

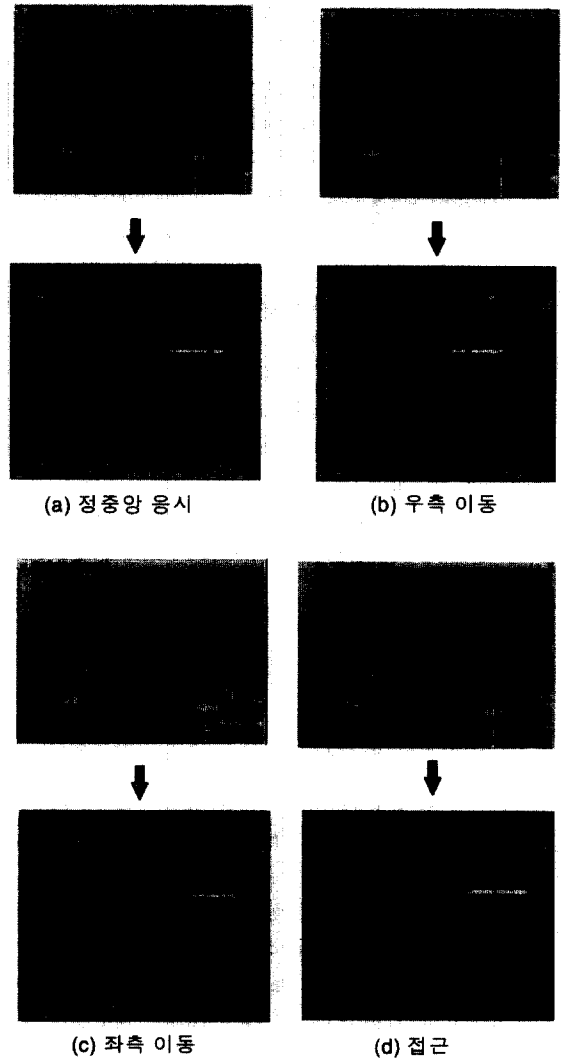


(그림 7) Desktop 가상 현실 환경에서의 3차원 화면 조정

(그림 8)은 얼굴의 움직임에 따른 3D view의 이동 결과를 나타낸 것이다.

8. 결 론

얼굴의 3차원 움직임량을 정확하게 추적하는 연구들은



(그림 8) 얼굴의 움직임에 의한 3차원 view 조정

컴퓨터 비전과 휴먼 컴퓨터 인터페이스 등에서 많은 응용 범위를 가지고 최근 활발히 연구가 진행되고 있다. 기존에 사용된 얼굴의 3차원 움직임 추정 알고리즘으로는 신경망(Neural Network), 광류 분석법(optical flow methods) 및 확장 칼만 필터(Extended Kalman Filter)등이 있는데, 이중 신경망과 광류분석법은 처리시간이나 정확도 면에서 많이 떨어지는 단점 때문에 확장 칼만 필터를 많이 이용하고 있다. 그러나 확장 칼만 필터 방법 역시 필터의 초기값에 따라 얼굴 움직임을 제대로 추정하지 못하고 발산하는 경우가 종종 발생하며, 얼굴의 회전 방향 변화시 이에 대처하지 못하는 경우도 종종 발생한다. 본 논문에서는 이러한 문제점을 해결하기 위하여 확장 칼만 필터의 변형 형태인 반복적 확장 칼만 필터를 이용하여 얼굴의 3차원 움직임량을 정확하게 추정하였다. 향후 연구에서 얼굴 특징점의 추출 정확도를 높인다면 보다 정확하게 얼굴의 3차원 움직임량을 추정할 수 있을 것으로 예상된다.

참 고 문 헌

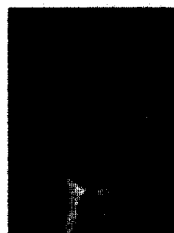
- [1] Jaihie Kim, Kang Ryoung Park, Jeoung Jun Lee, S.R.leClair, "Intelligent Process Control via Gaze Detection Technology," Engineering Applicatins of Artificial Intelligence Vol.13, No.5 August 2000, pp.577-587.
- [2] Ming XU, Takao AKATSUKA, "Detecting Head Pose from Stereo Image Sequence for Active Face Recognition," *Proc. of International Conf. Automatic Face and Gesture Recognition*, pp.82-87, 1998.
- [3] T. Horprasert, Y. Yacoob, and L.S. Davis., "Computing 3-D Head Orientation from a Monocular Image Sequence," *Proc. of International Conf. Automatic Face and Gesture Recognition*, pp.242-247. 1996.
- [4] T. Fukuhara, T. Murakami, "3D-motion estimation of human head for model-based image coding," *IEE Proc.*, Vol.140, No.1, pp.26-35, 1993.
- [5] Andrew Gee and Roberto Cipolla, "Fast visual tracking by temporal consensus," *Image and Vision Computing*, Vol.14, pp.105-114, 1996.
- [6] Ted J. Broida and Rama Chellappa, "Estimation of object motion parameters from noisy image," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(1) : 90-99, Jan 1986.
- [7] Ted J. Broida and Rama Chellappa, "Estimating the kinematics and structures of a rigid object from a sequence of monocular images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(1) : 90-99, Jan 1986.
- [8] Robert Grover Brown and Patric Y. C. Hwang, *Introduction to Random Signal Analysis and Kalman Filtering*, John Wiley & Sons, New York, 1997.
- [9] Simon Haykin, *Adaptive filter theory*, Prentice hall international editions 1996.
- [10] Mohinder S. Grewal Angus P. Andrews, *Kalman Filtering Theory and practice*, Prentice Hall Press 1993.
- [11] A. Azarbayeyani. T. Starner. B. Horowitz. A. Pentland, "Visually Controlled Graphics," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.15, No.6, pp.602-605, June 1993.
- [12] Ali Azarbayejani and Alex P. Pentland, "Recursive Estimation of Motion, Structure, and Focal Length," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.17, No.6, June 1995.
- [13] A. Azarbayejani, B. Horowitz, and Alex P. Pentland, "Recursive Estimation of Structure and Motion using Relative Orientation Constraints," 1993 IEEE Conf. on CVPR IEEE Computer Society Press, pp.294-299, June 1993.
- [14] Ted J. Broida and Rama Chellappa, "Recursive estimation of 3D motion from a monocular image sequence," *IEEE Trans. Aerospace and Electronics Sytems*, Vol.26, No.4, pp.639-656, July 1990.
- [15] Maybeck, P. S., "Stochastic Models, Estimation, and Control, Vol.2, New York : Academic Press, 1922.
- [16] Gem-sun J. Young and Rama Chellappa, "3-D Motion Estimation Using a Sequence of Noisy Stereo Images : Models, Estimation, and Uniqueness Results.s
- [17] [http : //www.polhemus.com](http://www.polhemus.com)
- [18] 박강령, "얼굴의 이차원 및 삼차원 움직임을 이용한 시선 위치 추적에 관한 연구", 연세대학교 대학원 전자공학과 박사학위논문, 2000년 2월.
- [19] Maybeck, P. S., "Stochastic Models, Estimation, and Control, Vol.2, New York : Academic Press, 1982
- [20] Yaakov Bar-Shalom and Xiao-Rong Li, *Estimation and Tracking Principles, Techniques, and Software*, Artech House Inc., 1993.



박 강 령

e-mail : parkgr@lgcit.com

1994년 연세대학교 전자공학과(공학사)
1996년 연세대학교 전자공학과(공학석사)
2000년 연세대학교 전자공학과(공학박사)
2000년~현재 LG 전자기술원 선임연구원
관심분야 : Biometric영상 처리, 컴퓨터 vision, 컴퓨터 그래픽스



김 재 희

e-mail : jhkim@bubble.yonsei.ac.kr

1979년 연세대학교 전자공학과 졸업
1982년~1984년 Case Western Reserve Univ.
전기공학과 공학 석사, 공학 박사
1984년~현재 연세대학교 기계전자공학부 교수

관심분야 : Biometric 영상 처리, 컴퓨터 vision, 패턴인식