논문 2010-47CI-4-4

사용자의 얼굴과 카메라 영상 간의 호모그래피를 이용한 실시간 얼굴 움직임 추정

(Online Face Pose Estimation based on A Planar Homography Between A User's Face and Its Image)

구 떠올라*, 이 석 한**, 두 경 수**, 최 종 수***

(Deo-olla Koo, Seok-Han Lee, Kyung-Soo Doo, and Jong-Soo Choi)

요 약

본 논문에서는 단일 카메라를 이용하여 얼굴의 움직임 정보를 추정하고 3차원 모델을 합성하기 위한 기법을 제안한다. 먼저 단일 카메라 입력 영상에서 사용자의 얼굴 영역 특징 점 취득을 위한 4개의 하부 이미지를 획득한다. 획득된 4개의 하부 이미지를 템플릿으로 사용하여 사용자 얼굴 영역의 정보를 추출하며, 이들 4개의 특징 점을 사용하여 사용자 얼굴과 카메라 영상 평면 사이의 사영 관계를 계산한다. 취득된 카메라 행렬로부터 얼굴의 움직임 정보인 이동과 회전 성분을 추정할 수 있으며, 이를 기반으로 3차원 모델의 자세 정보를 설정한 다음 이를 사용자 얼굴에 가상의 객체를 합성하기 위한 정보로 이용한다. 다양한 실험을 통하여 사용자 얼굴의 움직임에 대한 정보 추출의 정확도를 검증하였다.

Abstract

In this paper, we propose a simple and efficient algorithm for head pose estimation using a single camera. First, four subimages are obtained from the camera image for face feature extraction. These subimages are used as feature templates. The templates are then tracked by Kalman filtering, and camera projective matrix is computed by the projective mapping between the templates and their coordinate in the 3D coordinate system. And the user's face pose is estimated from the projective mapping between the user's face and image plane. The accuracy and the robustness of our technique is verified on the experimental results of several real video sequences.

Keywords: Head pose estimation, face tracking, augmented reality

I. 서 론

최근 지속적인 발전을 하고 있는 컴퓨터의 다양한 하드웨어들과 이로 인한 다양한 주변기기들의 발전이 급속도로 진행되고 있다. 이러한 주변기기의 발전으로 인하여 다양한 입출력 장치가 개발되었고, 입력 장치중 하나인 카메라의 성능 또한 비약적인 발전을 거듭하고 있다. 카메라의 발전과 보편화에 따라 카메라를 이용한 다양한 연구와 응용 시스템들이 개발되고 있으며, 특히 카메라를 이용한 많은 연구 중 사용자의 얼굴을 추적하여 이용하고자 하는 연구가 최근 활발히 진행되고 있다[1~5]. 이들 중 Qiang[1]는 적외선 카메라를 이용하여 얼굴의 움직임을 추정해 냈는데, 먼저 적외선 카메라를 이용하여 사람의 두 눈에 대해 찾아낸 후 타원 정합을 하고 얼굴의 움직임에 대해 추정을 하는 방법이었다. 하지만 적외선 카메라를 사용한 얼굴의 움직임 추정은

* 학생회원, ** 정회원, *** 평생회원, 중앙대학교 첨단영상대학원

(Dept. of Advanced Imageing Science, Multimedia, and Film, Chung-Ang University)

※ 본 논문은 2단계 BK21사업, 서울시 산학연협력사업 기술이전지원사업(TR080601), 서울시 산학클러스터(SFCC) 사업(10570)의 지원 결과임.

접수일자: 2010년6월1일, 수정완료일: 2010년7월7일

적외선 카메라를 이용할 수 있는 환경적인 제약이 있으며, 일반적인 환경에서의 응용이 불가 하다는 단점이 있다. 한편, 3D 모델기반의 얼굴의 움직임 추정에 대한 연구 또한 활발히 진행 중이다. Davis는 3차원 모델을 생성한 다음 이를 이용하여 사용자의 얼굴의 움직임 추정에 이용하였다[2]. 모델기반 얼굴 추적기법은 사용자의 얼굴과 같은 3D모델을 만들어서 얼굴의 움직임 방향에 맞는 모델과의 정합을 통해 얼굴의 움직임 정도를 추정해내는 방법이다. 하지만 모델기반 얼굴 움직임 추정은 기존에 생성된 모델이외의 새로운 얼굴이 입력될 때는 새로운 얼굴에 대한 모델을 생성해야 하는 단점이 있다. 얼굴의 움직임을 추정하기위해 두 대의 카메라를 사용하는 연구 또한 진행 중이다[3, 10]. 하지만 두 대의 카메라를 이용한 방법은 두 대 카메라를 교정(calibration) 해야 하는 등의 전 처리 과정이 필요하며, 두 대 카메라 사이의 기하학적 조건에 대한 환경적 제약성이 따른다는 어려움이 존재할 수 있다.

본 논문에서는 단일 카메라로부터 얼굴영상을 입력받아 사용자 얼굴이 평면이라는 가정 하에 얼굴의 움직임을 추정하는 방법에 대해 제안한다. 제안된 방법은 카메라를 통해 입력된 얼굴영상이 2차원 좌표로 표현이 되고, 얼굴 내 4개 특징 점을 취득해 만든 평면 사영변환에 따라 얼굴의 움직임을 추정한다. 단일 카메라를 이용한 얼굴 움직임 추적을 수행하기 위해 먼저 카메라 입력 얼굴영상에서 4개의 하부 이미지를 획득하고, 획득된 하부 이미지를 템플릿으로 사용한 템플릿 정합(Template Matching)을 통해 얼굴의 특징 점을 취득한다. 취득된 특징 점에 칼만 필터(Kalman Filter)를 적용하여 특징 점을 추정하고, 추적되는 특징 점들을 이용하여 평면 호모그래피(Planar Homography)를 획득한다. 획득된 평면 호모그래피와 카메라 내부 파라미터(Intrinsic Parameter)를 이용하여 카메라의 회전 및 이동에 대한 정보를 갖는 카메라 사영행렬(Projective Matrix)을 취득한다. 이렇게 취득된 카메라 사영행렬은 카메라가 고정되어 있는 상태에서 얼굴의 움직임이나 얼굴이 고정된 상태에서 카메라의 움직임이나 두 움직임이 같다는 가정 하에 얼굴의 움직임 추정에 사용된다. 이후 취득된 사영행렬을 기반으로 얼굴의 회전정도와 이동정도를 추정하고, 얼굴의 움직임에 따른 가상의 3차원 물체를 얼굴에 합성해 준다. 제안된 방법은 훈련 데이터 혹은 3D 모델 없이 카메라 입력 영상으로부터의 초기화를 통한 기준 정보의 획득만으로 얼굴의 움직임을 추정할 수 있으며, 또한 3D 모델 데이터를 사용하지 않기 때문에 수행 속도 및 실제 구현 과정의 복잡도 측면에서 유리하다는 장점이 있다. 본 논문의 구성은 Ⅱ장에서는 먼저 시스템의 구성도와 얼굴의 움직임을 추정하기 위한 얼굴의 하부 이미지 획득에 대해 언급한다. 또한 획득된 하부 이미지를 이용한 특징 점 획득과 칼만 필터를 이용한 특징 점 추적에 대해 언급한다. 그리고 평면 호모그래피를 이용한 카메라 사영행렬의 획득에 대해 언급한다. Ⅲ장에서는 본 논문의 실험 결과를 기술하며, Ⅳ에서 논문의 결론을 맺는다.

Ⅱ. 본 론

1. 얼굴 하부 이미지 획득

단일 카메라 입력 영상에서 얼굴의 움직임을 추정하기 위한 시스템의 구성도는 그림 1과 같다. 카메라 입력 얼굴영상에서 얼굴의 움직임을 추정하기 위해서는 먼저 얼굴의 하부 이미지를 획득해야 한다. 얼굴의 하부 이미지는 양쪽 눈의 각 끝 부분과 입술의 양쪽 끝의 총 4개의 이미지를 획득한다. 특징은 Harris 검출자[13~14]를 이용하여 추출되었다. 먼저 얼굴의 하부 이미지를 획득하기 위해 카메라 입력 얼굴영상에서 하부 이미지를 획득하기 위한 범위에 사각형의 틀을 그려준다. 그 후 그려진 영역의 이미지들을 따로 저장한다. 이때 획득되는 하부 이미지들은 컬러영상이 아닌 그레이 영상으로 변환시킨다. 획득된 4개의 하부 이미지는 추후 템플릿 정합이 수행되는 템플릿영상으로 이용된다[19~20].

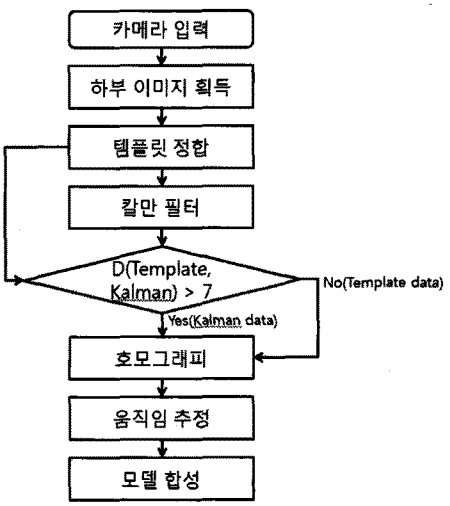
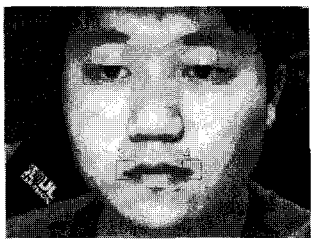


그림 1. 시스템 구성도
Fig. 1. Flowchart of the system.

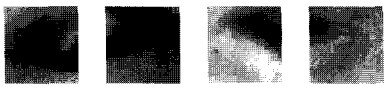
그림 2는 하부이미지 획득을 위한 사각 틀 영상과 획득된 4개의 하부 이미지이다.

2. 얼굴 특징 점 획득 및 추적

얼굴의 특징 점을 찾기 위해서는 그림 2의 b와 같은 4개의 하부 이미지를 템플릿으로 사용하는 템플릿 정합을 수행한다. 또한 찾아진 특징 점들의 추적을 위해 칼만 필터를 사용한다.



(a) 얼굴 영역에서의 템플릿 영상 취득



(b) 4개의 템플릿 영상

그림 2. 사용자 얼굴에 대한 템플릿 영상 추출
Fig. 2. Template images of the feature point.

가. 템플릿 정합

얼굴의 특징 점들을 찾기 위한 템플릿 정합은 수식 (1)을 이용하여 수행된다.

$$R(x,y)=\frac{\sum_{x'=0}^{w-1}\sum_{y'=0}^{h-1}[T(x',y')I(x+x',y+y')]}{\sqrt{\sum_{x'=0}^{w-1}\sum_{y'=0}^{h-1}T(x',y')^2\sum_{x'=0}^{w-1}\sum_{y'=0}^{h-1}I(x+x',y+y')^2}}. \tag{1}$$

수식 (1)은 NCC(Normalized Cross Correlation)방법으로 정규화 된 상관계수를 구하고, 상관계수의 값이 큰 영역에 대해 템플릿과 유사영역으로 판단한다. 수식 (1)에서 I는 카메라로부터 취득되는 영상을 나타내며, T는 미리 획득된 하부 이미지로서 템플릿 영상을 나타낸다. 또한 h와 w는 템플릿 영상의 높이와 폭을 나타낸다. 수식 (1)을 통해 템플릿 정합을 할 때 4개의 하부이미지에 대해 4번의 템플릿 정합을 수행한다. 따라서 전체 영상에서 4번의 템플릿 정합을 수행하게 되면 수행 속도가 감소할 수 있다. 이런 문제를 해결하기 위해 본 논문에서는 각각의 하부이미지들의 인접 부위에서 템플릿 정합을 수행하도록 템플릿 정합 후보 영역을 설정했다. 템플릿으로 사용되는 하부 이미지의 크기가 24X24크기이므로, 템플릿 정합의 후보 영역은 이보다 큰 36X36의 크기로 설정했다. 설정된 네 영역에 대해 각각의 템플릿 정합을 수행한 영상을 그림 3에 보인다.

(a) 얼굴의 우측면 (b) 얼굴의 좌측면

그림 3. 템플릿 정합 수행 결과
Fig. 3. Results of the template matching.

나. 칼만 필터

템플릿 정합을 통해 획득된 특징 점들에 대한 추적을 위해 칼만 필터를 수행한다. 칼만 필터는 이산 신호 처리에서 이전 상태를 기반으로 이후 상태를 예측하기 위한 것이다[12]. 수식 (2)는 칼만 필터의 시스템 방정시과 관측 방정식을 나타낸다.

$$\begin{aligned} X_{k+1} &= Ax_k + w_k, \quad p(w) = N(0,Q) \\ Z_{k+1} &= Hx_k + v_k, \quad p(v) = N(0,R) \end{aligned} \tag{2}$$

수식 (2)에서 x와 z는 각각 해당 시간에 알고자 하는 변수들의 상태 벡터와 관측 벡터를 나타낸다. W와 v는 평균값이 0인 시스템의 오차와 관측 오차를 나타낸다. 또한 분산은 Q와 R인 정규 확률 분포에 따라 결정된다. 수식 (3)은 칼만 필터를 적용하기 위해 사용되는 천이 행렬과 관측 행렬을 나타낸다.

$$A=\begin{bmatrix}1&0&1&0\\0&1&0&1\\0&0&1&0\\0&0&0&1\end{bmatrix},\ H=\begin{bmatrix}1&0&0&0\\0&1&0&0\end{bmatrix} \tag{3}$$

수식 (3)의 A는 이전 상태와 현재 상태를 정의하는 천이 행렬이고, H는 이전 상태와 현재 예측과의 관계를 정의해 주는 관측행렬이다. 칼만 필터의 측정과 예측에 관한 흐름을 그림 4에 보인다. 제안된 방법에서는 템플릿 정합을 통해 획득되는 얼굴의 특징점들과 칼만 필터를 통해 추적되는 특징 점들의 거리를 오차값으로 간주하며, 이들 사이의 거리가 일정 범위를 벗어나면 칼만 필터의 예측값을 이용하였으며, 일정 범위 내이면 영상

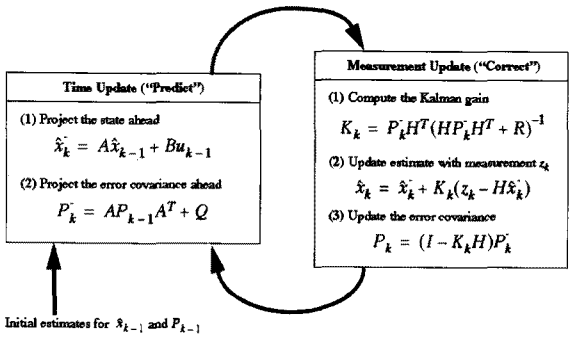


그림 4. 칼만 필터 수행 흐름도
Fig. 4. Kalman filter Framework.

으로부터의 측정값을 이용한다.

다. 평면 호모그래피를 이용한 사영 행렬 획득

얼굴의 특징 점 획득 및 추적 후 획득된 4개의 특징 점들을 이용하여 평면 호모그래피(Planar Homography)를 구한다. 즉, 평면 호모그래피를 이용하여 영상과 사용자 얼굴 영역 사이의 사영관계를 추정하며, 이에 대한 관계를 아래의 식 (4)에 보인다.

$$\mathbf{x}' = \mathbf{H}\mathbf{x}, \mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}. \quad (4)$$

수식 (4)에서 x′은 이동된 점을 나타내고, X는 이동되기 전의 점을 나타내며 H는 호모그래피 3X3행렬을 의미한다. 사람의 얼굴은 3차원 입체의 형상을 하고 있지만 카메라를 통해 입력받은 영상은 x, y좌표로 표현되는 2차원 영상이다. 또한 3개의 점 이상이면 하나의 평면을 만들 수 있으므로 얼굴에서 취득된 4개의 특징 점들을 통해 하나의 평면이 생성된다. 이들 평면 간의 이동관계를 평면 호모그래피를 이용하여 추정한다. 그림 5는 얼굴에서 취득된 4개 특징 점들을 통해 생성된 평면의 호모그래피 관계를 나타낸 것이다.

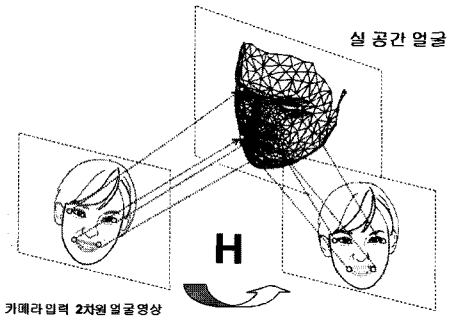


그림 5. 사용자 얼굴과 영상 사이의 호모그래피 관계
Fig. 5. 2D projective mapping between image plane and geometric cues of user's face.

호모그래피를 통해 카메라의 사영 행렬을 획득하기 위해서는 카메라의 내부 파라미터(Intrinsic parameter) 이용한다. 카메라의 내부 파라미터는 3 X 3 상삼각 행렬이며 K로 나타낸다. 수식 (5)는 카메라 내부 파라미터를 나타낸다. 수식 (5)의 카메라 내부 파라미터는 별도의 카메라 교정 과정을 통해서 320 X 240 영상으로부터 계산된다.

$$\mathbf{K} = \begin{bmatrix} a_x & s & x_0 \\ & a_y & y_0 \\ & & 1 \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} 410.0 & 0.0 & 320.0 \\ 0.0 & 420.93 & 240.0 \\ 0.0 & 0.0 & 1.0 \end{bmatrix} \quad (5)$$

특징점 좌표 및 H와 교정행렬 K를 이용하여 다음의 (6)과 같이 카메라 이동 및 회전 정보를 취득한다.[7~9]

$$\begin{aligned} &x' = \mathbf{H}\mathbf{x}, \\ &\mathbf{H} = \mathbf{K}[\mathbf{r}_1\, \mathbf{r}_2\, \mathbf{t}], \\ &\mathbf{H}' = [\mathbf{r}_1\, \mathbf{r}_2\, \mathbf{t}] = \mathbf{K}^{-1}\mathbf{H}, \\ &r_3 = r_1 \times r_2, \\ &\mathbf{P} = \mathbf{K}[\mathbf{R}|\mathbf{t}] = \mathbf{K}[\mathbf{r}_1\, \mathbf{r}_2\, \mathbf{r}_3\, \mathbf{t}]. \end{aligned} \quad (6)$$

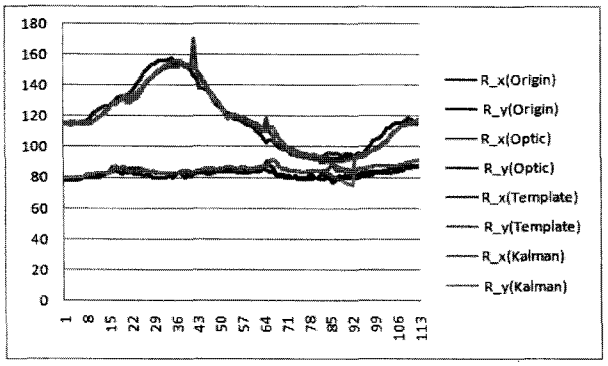
수식 (6)에서 r_3은 r_1와 r_2의 외적이다. P는 3 X 4카메라 사영 행렬을 나타낸다. 획득된 카메라 사영 행렬은 카메라가 고정된 상태에서 얼굴이 움직인다고 한다면 얼굴의 이동과 회전에 대한 상대적인 변위 정보를 의미한다. 계산된 카메라 행렬을 이용하여 얼굴 움직임 정보를 실시간으로 추적할 수 있으며 이를 실시간 인터랙션, 증강현실 구현 등에 직접 이용 가능하다[6, 16].

III. 실 험

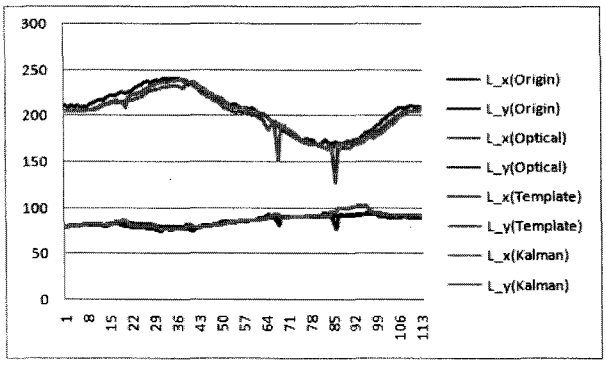
본 논문에서는 단일 카메라를 통해 입력받은 얼굴 영상에서 얼굴의 움직임을 추정하는 방법에 대해 언급하였다. 실험은 카메라는 8mm렌즈 1394카메라를 사용하였다. 먼저 카메라 입력 얼굴 영상에서 얼굴의 특징 점 취득 및 추적을 하고, 이후 평면 호모그래피를 획득하고 카메라 내부 파라미터를 이용하여 사영 행렬을 획득한다. 획득된 사영 행렬을 통해 얼굴의 움직임 및 회전에 대한 정도를 추정해 낸다.

1. 얼굴 특징 점 추적

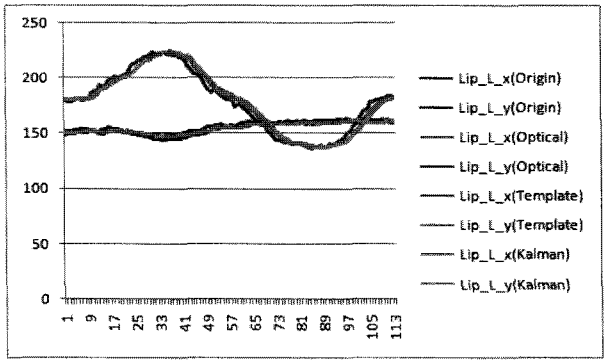
얼굴 영상에서 하부 이미지를 획득하고 이것을 템플릿으로 사용하여 템플릿 정합을 수행해 얻어진 특징 점



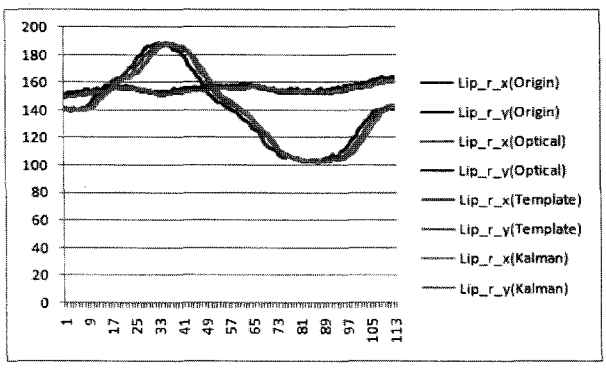
(a) 우측 눈 부분 추적



(b) 좌측 눈 부분 추적



(c) 입술의 우측 특징점 추적



(d) 입술의 좌측 특징점 추적

그림 6. 4개의 특징 점 추적
Fig. 6. Tracking results for the user face.

들에 대해 칼만 필터를 이용하여 추적을 하였다. 이때 얼굴 특징 점들에 대한 추적이 잘 이루어지는지 비교를 위해 광학적 흐름(Optical Flow)방법을 이용하였다. 획득 된 특징 점들에 대해 칼만 필터를 적용할 때 칼만 필터의 예측 값과 측정 값을 일정 조건에 따라 바꿔서 주어줬다. 이때 조건이라 하면 칼만 필터를 통해 나온 예측 값과 측정 값 사이의 길이를 비교하여 길이가 7이상이면 칼만 필터의 예측 값을 7미만이면 칼만 필터의 측정 값을 이용하였다. 그림 6은 얼굴의 4개 특징 점에 대한 추적의 결과를 그래프로 보여준다. 그림 6의 (a)와 (b)는 눈의 양끝에 대한 두 점의 추적 결과를 나타낸 것이다. (a)는 오른쪽 바깥쪽 점이고, (b)는 왼쪽 바깥쪽 점이다. 두 그래프를 보면 중간 중간 그래프의 선이 튀는 것을 볼 수 있다. 이는 실험 중 불가피한 눈의 깜빡임으로 인해 광학적 흐름의 추적이 순간순간 튀는 것이다. 하지만 템플릿 정합과 칼만 필터를 적용한 경우 그래프의 튀는 현상이 적은 것을 볼 수 있다. 그림 6의 (c)와 (d)는 입술의 양쪽 끝 점들에 대한 그래프를 나타낸 것이다. 이때는 모든 그래프들에서 급격한 변화가 없는 것을 볼 수 있다. 또한 그림 6에서 템플릿 정합의 그래프와 칼만 필터의 그래프가 겹쳐져 보이는 경우가 있는데 이는 칼만 필터의 예측 값과 측정값의 차이가 7미만이므로 측정값을 적용했기 때문이다.

2. 얼굴 움직임 추정

얼굴 움직임 추정을 위해 획득되는 4개의 특징점들을 이용하여 평면 호모그래피를 취득해 낸다. 이후 취득된 평면 호모그래피와 미리 취득한 카메라 내부 파라미터를 이용하여 카메라 사영 행렬을 획득한다. 얼굴이 평면이라는 가정 하에서 획득된 카메라 사영행렬은 카메라는 고정되어 있고 얼굴이 움직인다고 한다면 사영행렬의 값이 곧 얼굴이 움직인 이동과 회전의 값이 된다. 따라서 획득된 카메라 사영행렬의 값을 이용하여 얼굴의 움직임을 추정한다. 얼굴 움직임 추정의 정확도

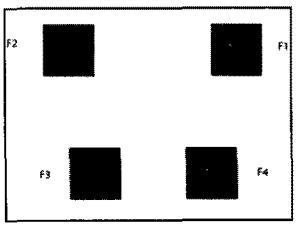
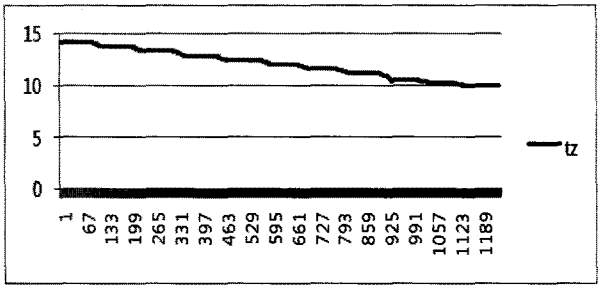
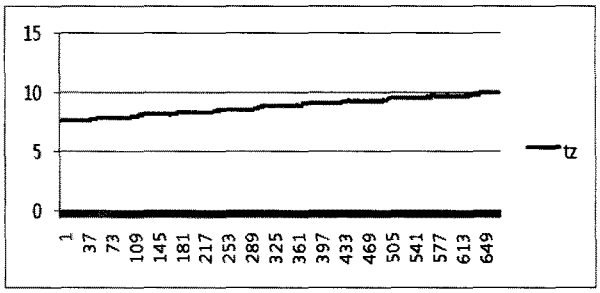


그림 7. 정확도 판별을 위해서 사용된 테스트 패널
Fig. 7. The test panel for accuracy evaluation.



(a) 테스트 패널 이동 (전진 방향)

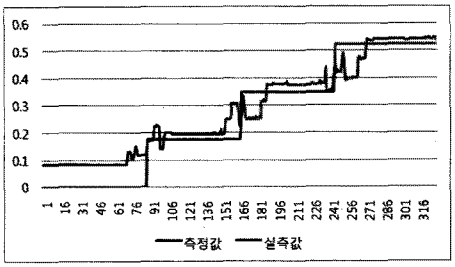


(b) 테스트 패널 이동 (후진 방향)

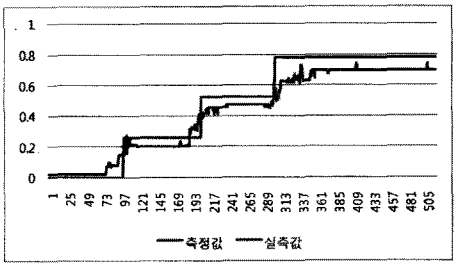
그림 8. 실험 판 전/후진 시 tz의 변화량
Fig. 8. Position estimation of the test panel.

를 실험하기 위해 본 논문에서는 카메라 교정에 많이 이용되는 격자무늬 테스트 패널을 사용하였다.

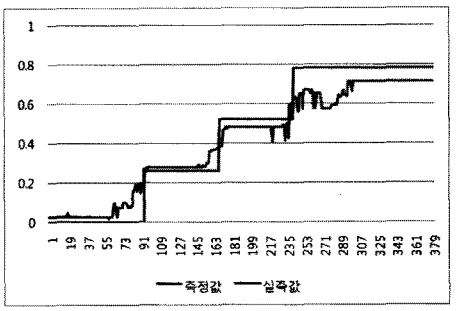
그림 7의 실험 판을 이용한 움직임 추정 실험은 카메라 전, 후로 각각 10cm을 1cm 간격으로 움직이면서 사영행렬 P에서 tz의 변화를 본다. 이때 tz의 값은 P에서 카메라 중심의 z값을 나타낸다. 또한 3차원 공간의 x, y, z축의 각 축들을 기준으로 다음과 같이 회전시키면서 움직임의 변화량과 정확도에 대해 살펴본다. 먼저 x축 기준으로 상위 방향 30도를 10도씩 회전했을 때, y축을 기준으로 15도씩 45도 좌, 우 회전했을 때, 그리고 z축 기준으로 10도씩 30도 좌우 회전했을 때의 변화량과 정확도를 살펴본다. 그림 8은 실험 판을 앞, 뒤로 움직였을 때 z값의 변화량에 대해 나타낸 것이다. 그림 8을 보면 그래프가 계단 형식을 띄는 것을 볼 수 있다. 이것은 10cm의 간격을 이동하는 동안 1cm씩 이동하고 3-5초 정도 멈췄다가 다시 이동하는 방식으로 실험을 했기 때문이다. 또한 실험 판을 앞으로 움직이면 tz의 값이 줄어들고 반대로 뒤로 움직이면 tz의 값이 커지는 것을 볼 수 있는데 이는 프로그램상의 좌표계가 오른손 좌표계이기 때문이다. 실험 판을 이용한 회전에 대한 실험을 그림 9에 보이며, 각 그래프의 y축에 대한 변위각을 나타낸다. 그림 9의 그래프도 그림 8의 그래프와 같이 계단 형태을 띄고 있는데 이 또한 일정 간격 마다



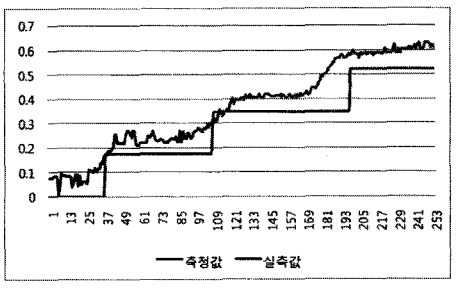
(a) x축 기준 상위 방향 회전



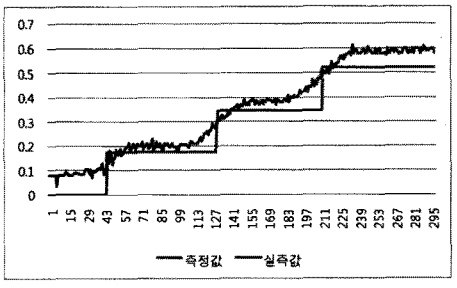
(b) y축 기준 우측 방향 회전



(c) y축 기준 좌측 방향 회전



(d) z축 기준 우측 방향 회전

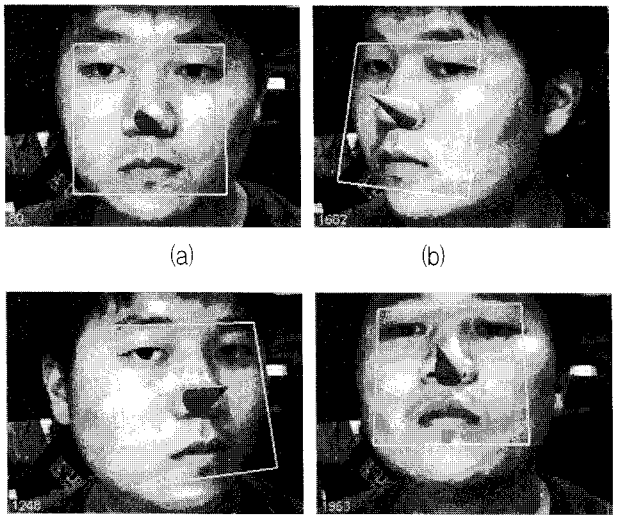


(e) z축 기준 좌측 방향 회전

그림 9. 회전에 따른 변화량
Fig. 9. Estimation of the rotational motion of the test panel.

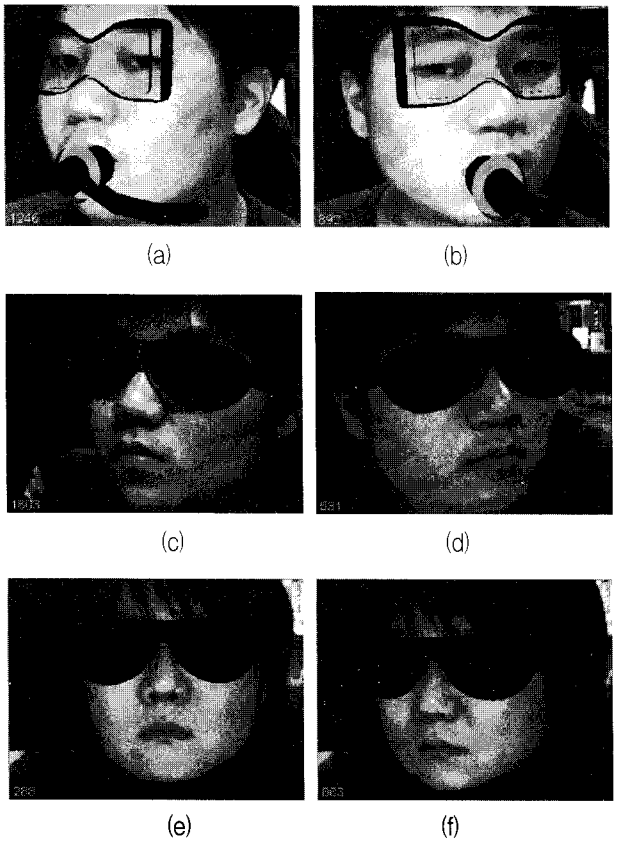
잠시 멈췄다가 재측정 했기 때문이다. 그림 9에서 측정값들의 미세한 잡음이 보이는데 이는 실험 중 미세한 떨림, 조명 변화로 인한 특징점 검출 오차이다.

또한 카메라의 화각과 방사왜곡의 영향이라고 고려



(a) (b) (c) (d)

그림 10. 얼굴 움직임에 따른 가상 물체 합성 영상(1)
Fig. 10. Augmentation of virtual object by the proposed system.



(a) (b) (c) (d) (e) (f)

그림 11. 얼굴 움직임에 따른 3차원 모델 합성 영상(2)
Fig. 11. Application examples for the proposed system.

된다. 그림 9의 그래프들을 보면 측정값들의 불규칙적인 떨림이 보이는데 이는 실험 중 미세한 떨림과 조명의 변화로 인한 특징 점들의 떨림에 따른 오차이다. 또한 카메라의 화각과 방사왜곡의 영향이라고 고려된다. 그림 9의 a는 움직임이 없는 경우 초기 오차가 5도 정도지만 10, 20, 30도씩의 움직임에 따른 평균 오차는 1.3도 정도였다. 그림 9의 b, c의 경우 15, 30, 45도 회전 시 우측 및 좌측 회전에 따른 각각의 오차가 15도에선 약 2도, 3도의 차이가 났다. 그리고 30도 회전은 각각 3도, 2도의 차이가 났으며, 45도 회전 오차가 커져서 양방향의 오차는 각각 6도 정도였다. 마지막으로 그림 9의 d와 e의 경우인 z축을 기준으로 우, 좌 방향으로 10도씩 30도 회전시켰을 때 10도에선 각각 약 1도, 3도의 오차가 있었다. 또한 20도 회전 시 약 3도, 6도의 오차가 있었으며 30도 회전 시 약 4도, 5도의 오차가 발생했다. 이런 오차의 발생은 실험 중 생기는 미세한 떨림과 조명의 변화 때문이라고 생각된다. 하지만 실측값과 측정값에 대한 모든 각도에서의 평균 오차는 약 4도 정도였다. 이와 같이 실험을 통해 얼굴 움직임에 따른 카메라 사영행렬의 각 인자들에 대한 변화를 검증해 보았다. 카메라 사영 행렬을 통한 얼굴의 움직임에 대해 위의 실험과 같이 검증을 해 보았다. 제안된 시스템의 증강현실 응용 예를 그림 10, 그림 11에 보인다. 그림 10은 얼굴에 사각형과 원뿔 객체를 합성하여 카메라 행렬을 통한 얼굴 움직임 추정이 정확히 되었는지에 대한 시각적인 확인을 위한 영상을 보인다. 그림 11은 얼굴의 움직임에 따라 취득된 카메라 사영 행렬을 기준으로 얼굴에 3차원 모델을 합성하는 예를 보여준다.

Ⅳ. 결 론

본 논문에서는 카메라 입력 얼굴 영상에서 얼굴의 움직임 및 모델 자세 추정에 대해 논하였다. 먼저 카메라 입력 얼굴 영상에서 얼굴의 4개 하부 이미지를 획득 하고, 획득된 하부 이미지들을 템플릿으로 사용해 템플릿 정합을 수행하였다. 템플릿 정합을 통해 획득된 4개의 특징 점에 대해 칼만 필터를 사용해 추적을 하였고, 추적된 4개 특징 점을 이용하여 호모그래피를 획득했다. 획득된 호모그래피를 사용하여 얼굴의 움직임 추정을 위한 카메라 사영 행렬을 취득 했으며, 카메라 사영 행렬을 통해 얼굴의 움직임에 따른 모델 자세를 추정 했고, 추정된 3차원 모델을 얼굴에 합성을 수행했다.

제안된 방법은 특정한 훈련 데이터 혹은 모델 데이터 없이 카메라 입력 영상에서 초기화를 통한 4개의 하부 이미지 획득만으로 얼굴의 움직임을 추정할 수 있으며, 또한 3D 모델 데이터를 사용하지 않기 때문에 처리과정이 간단하다는 장점이 있다. 실험을 통해 단일 카메라로부터 얼굴의 움직임을 추정해 내고, 3차원 모델의 합성이 가능하다는 것을 확인했다. 향후 연구에서는 카메라의 왜곡 보정과 조명의 영향에 따른 초기 취득 특징 점들의 떨림에 대한 수정을 통해 더욱 강건한 움직임 추정을 하는 연구를 수행할 계획이다. 제안된 방법에서는 연산량 및 수행 시간을 고려하여 사용자 얼굴에 대한 4개의 템플릿 영상만을 추출하여 이용하였으나 더 많은 수의 특징점 및 템플릿을 이용하였을 경우에 대한 보다 객관적이고 상세한 분석이 필요하다고 생각된다. 또한 다양한 사용자의 얼굴 변화 및 보다 열악한 환경 하에서의 강건한 추적 성능 구현 등에 대한 연구 또한 진행될 예정이다.

참 고 문 헌

[1] Q. Ji, R. Hu, "3D Face pose estimation and tracking from a monocular camera," *Image and Vision Computing*, Vol. 20, Issue 7, pp. 499-511, May 2002.

[2] T. Horprasert, Y. Yacoob and L. S. Davis, "Computing 3-D Head Orientation from a Monocular Image Sequence," In Proc. IEEE International Conference on Face and Gesture Recognition, Killington, pp. 242-247, VT, 1996.

[3] R. Yang and Z. Zhang, "Model-based head pose tracking with stereovision," In Proc. The 5th International Conference Automatic Face and Gesture Recognition, pp. 255-260, Washington, D.C., 2002.

[4] 정민교, 박지숙, 엄성용, 조현희, "간단한 얼굴 방향성 검출," 멀티미디어학회 논문지 Vol. 9 No. 2 pp. 234-243 2006.

[5] P. Marco "Vision-based user interfaces: methods applications," *International Journal of Human-Computer Studies*, Vol. 57, pp. 27-73, 2002.

[6] R. Azuma, "A Survey of Augmented Reality," *Teleoperators and Virtual Environments*, Vol. 6, No. 4, pp. 355-385, Aug. 1997.

[7] H. M. Park, S. H. Lee, J. S. Choi, "Wearable Augmented reality System using Gaze Interaction," The 7th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR2008), pp. 175-176, Cambridge, UK, Sep. 2008.

[8] 이석한, "증강현실 구현을 위한 실시간 카메라 추적 및 3차원 기하학 정보 추정에 관한 연구," 중앙대학교 박사학위 논문, 2008.

[9] R. Hartely, and A. Zisserman, "Multiple View Geometry in Computer Vision," Cambridge University Press, Cambridge, 2003.

[10] Y. Matsumoto and A. Zelinsky, "An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement," Proc. IEEE Fourth Int. Conf. Face and Gesture Recognition, pp. 499-505, 2000.

[11] R. S. Feris, J. Gemmell, K. Toyama, and V. Kruger, "Facial Feature Detection Using A Hierarchical Wavelet Face Database," MSR-TR-2002-05, University of Maryland, 2002.

[12] G. Welch and G. Bishop, "An introduction to the Kalman filter," Dept. Comp. Sci., Univ. North Carolina, Chapel Hill, TR95-041.

[13] C. Harris and M. J. Stephens, "A combined corner and edge detector," In Alvey Vision Conference, pp. 147-152, 1988.

[14] M. Nixon and A. Aguado, "Feature Extraction and Image Processing Second edition," Academic Press is an imprint of Elsevier, pp. 186-196, 2008.

[15] R. Poppe, "Vision-based human motion analysis: An overview," Computer Vision and Image Understanding, Vol. 108, No. 1-2,pp. 4-18, 2007.

[16] P. Marco "Vision-based user interfaces: methods applications," *International Journal of Human-Computer studies*, Vol. 57, pp. 27-73, 2002.

저 자 소 개

구떠올라(학생회원)
2008년 남서울대학교 멀티미디어 학과 졸업(공학사)
2010년 중앙대학교 첨단영상대학원 졸업(공학석사)
<주관심분야: 증강현실, HCI, 3D 컴퓨터비전>

이 석 한(정회원)
1999년 중앙대학교 공과대학 전자공학과 졸업(공학사)
2001년 중앙대학교 첨단영상대학원 졸업(공학석사)
2001년~2004년 LG전자 주임연구원
2009년 중앙대학교 첨단영상대학원 졸업 (공학박사)
2009년~2010년 중앙대학교 첨단영상대학원 CT연구소 전임연구원 (post Doc.)
2010년~현재 중앙대학교 첨단영상대학원 연구전담교수
<주관심분야: 실시간카메라추적, 증강현실, 3D 컴퓨터비전>

두 경 수(정회원)
대한전자공학회 논문지
제47권 SP편 제3호 참조

최 종 수(평생회원)-교신저자
대한전자공학회 논문지
제46권 SP편 제2호 참조