한국정보전자통신기술학회논문지(jkiiect)'24-10, Vol.17 No.5
논문 24-17-05-358 http://dx.doi.org/10.17661/jkiiect.2024.17.5.358

동일 평면상에 존재하는 특징점 검출을 이용한 실시간 카메라 추적 기법

이석한*

Real-time camera tracking using co-planar feature points

Seok-Han Lee*

요 약 본 논문에서는 3차원 공간 내의 2차원 평면에 존재하는 특징점을 검출한 다음 이를 기반으로 실시간 카메라 추적을 수행하기 위한 방법을 제안한다. 제안된 방법은 두 단계를 거쳐서 수행된다. 먼저 3차원 공간 내에서 다수의 특징점을 검출한 다음 동일 평면 상에 존재하는 특징점들만을 선별한다. 그리고 특징점들이 존재하는 평면 객체와 카메라 영상 평면 사이의 사영기하(projective geometry) 관계를 추정한 다음 이를 기반으로 카메라의 외부 파라미터(extrinsic parameter)를 추정한다. 실험 과정에서 별도의 특수 조명은 이용하지 않으며 일반적인 조명이 갖추어진 실내 환경 내에서 실험을 수행하였다. 제안된 방법은 기존의 방법과는 다르게 실시간으로 평면 객체 상의 새로운 특징점을 검출하여 카메라 추적에 이용함으로써 카메라 자세정보 초기화를 위한 기준 특징점(reference feature)이 검출 불가능한 상황에서도 카메라의 추적이 가능하다는 장점을 갖고 있다. 실험을 통하여 제안된 방법은 평균 5~7픽셀 정도의 재사영 오차를 발생시킴을 확인할 수 있었다. 이는 영상의 해상도를 고려했을 때 상대적으로 작은 값으로 간주될 수 있으며, 기준 특징점이 영상 내에 존재하지 않은 환경에서도 카메라 추적을 수행할 수 있음을 알 수 있다.

Abstract This paper proposes a method for the real-time camera tracking which detects and employs feature points located on a planar object in 3D space. The proposed approach operates in two stages. First, multiple feature points are detected in the 3D space, and then only those that exist on the planar object are selected. The camera's extrinsic parameters are then estimated using the projective geometry relationship between the feature points of the plane and the camera's image plane. The experiments are conducted in a typical indoor environment with regular lighting, without any special illumination setups. In contrast to conventional approaches, the proposed method can detect new feature points on the planar object in real-time and employ them for the camera tracking. This allows for continuous tracking even when the reference features for the camera pose initialization are not available. The experimental results show an average re-projection error of about 5 to 7 pixels, which is relatively small given the image resolution, and demonstrating that camera tracking is possible even in the absence of reference features within the image.

Key Words : Real-time camera tracking, planar homography, camera parameter estimation, visual tracking, augmented reality

1. 서론

2차원 영상 카메라의 자세(pose) 정보를 실시간으로 추적하고 카메라 자세 파라미터를 정확히 추정하기 위한 카메라 추적 기법은 증강현실 (AR; augmented reality), 사물인터넷(internet of things), 자연 인터페이스(natural interface) 등의 분야를 구현하기 위한 기반 기술이다. 특히 AR 영상의 경우 사용자에게 제공되는 콘텐츠의 형태가 실사 영상에 가상의 부가정보가 정합되는 형태로 구성이 되며 이를 통하여 사용자는 실제

*Department of Information and Communication Engineering, Jeonju University
Received October 04, 2024 Revised October 12, 2024 Accepted October 14, 2024

3차원 환경과 가상 정보가 혼합된 인터페이스를 제공받을 수 있다. 한편 부가적인 장비나 센서가 없이 2차원 영상만을 이용하는 영상 기반 증강현실은 카메라의 영상으로부터 카메라 자세를 실시간으로 추정한 다음 가상의 부가적인 정보를 하나의 기준 좌표계안에 정합시키는 기술로, 증강현실 콘텐츠의 다른 요소들은 안정적인 카메라 자세 추정 결과를 토대로 구현된다. 이를 위해 AR 콘텐츠를 이용하는 사용자의 위치 및 시선 정보는 실시간으로 변화하기 때문에 사용자의 무작위적인 움직임까지 고려한 환경에서의 실시간 카메라 추적 방법까지 반드시 고려되어야 한다. 본 논문에서는 평면 객체 상의 특징점을 검출한 다음 이를 실시간 카메라 추적에 이용하는 카메[라 추적 기법을 제안한다. 제안된 방법은 2차원 카메라 영상과 3차원 공간 내에 존재하는 2차원 평면 객체 사이의 2차원 사영관계를 이용해서 카메라 자세 정보를 추정한다. 제안된 방법은 두 단계로 구성이 되어 있다. 먼저 3차원 공간 내의 특징점들을 실시간으로 검출한 다음 동일 평면 상에 존재하는 특징점들만을 선별하고 그 외의 특징점들은 제거한다. 그리고 카메라 영상과 3차원 공간 내의 평면 객체 간의 2차원 평면 사영관계를 이용하여 카메라의 자세 정보를 추정하는 과정을 거쳐서 실시간 카메라 추적을 수행한다. 또한 제안된 방법은 실시간으로 새로운 특징점들을 지속적으로 검출하고 이를 카메라 추적이 이용함으로써 카메라 자세정보 초기화에 사용되는 기준 특징점(reference feature)이 카메라 장면 내에서 사라진 경우에도 카메라 추적을 계속 수행할 수 있다는 장점이 있다. 본 논문의 2장에서는 카메라 추적 관련 연구 동향에 대해 설명하고 3장에서는 동일평면 상의 특징점 검출 및 카메라 파라미터 계산 과정에 대해 설명한다. 4장에서는 제안된 방법의 실험 및 결과 분석에 대한 내용을 제시하고 5장에서 최종 결론 및 향후 연구 계획에 대해 설명한다.

2. 관련 연구 동향

기존의 평면 객체 기반 실시간 카메라 추적에 관한 관련 연구로는 A.H. Lee 등의 연구[1]와 S.H. Lee[2] 등이 수행한 논문이 있다. A.H. Lee[1]의 연구에서는 카메라 자세의 초기화를 위해 카메라의 초기 위치가 기준 바닥 평면과 수직인 방향에 위치한다고 가정하였다. 그리고 카메라 초기 위치에서 취득된 영상으로부터 기준 특징점들을 검출한 다음 카메라 영상 내에서 기준 특징점들의 좌표 변화를 실시간으로 검출하고 3차원 공간에서의 기준 특징점들과 영상 내의 특징점 좌표를 이용하여 카메라 자세 정보를 추정하였다. 이 방법은 3차원 공간 내에서 새롭게 검출되는 특징점들이 모두 동일 평면 상에 존재한다고 가정하였으며 따라서 평면 상에 존재하지 않는 특징점이 검출되었을 경우 카메라 자세 추정에 심각한 오차가 발생한다는 문제가 있다. S.H. Lee[2] 등의 연구에서는 카메라 자세 정보의 초기화를 위해 사전에 정의된 크기를 갖는 검은색 사각형 형태의 패턴의 4개의 모서리를 이용하여 기준 특징점을 정의하였다. 그리고 카메라 자세 추정에 대한 예측값을 추정한 다음 이에 대한 측정값을 카메라 영상 내의 기준 특징점 좌표으로부터 계산한 다음 측정값으로부터 취득된 오차값을 기반으로 카메라의 자세정보의 예측 값을 수정하는 과정을 수행함으로써 카메라 자세 추정을 수행하였다. S.H. Lee[2] 등의 연구에서는 비교적 강건한 카메라 추적 결과를 제공하였으나 새로운 특징점을 추가적으로 검출하는 과정이 포함되어 있지 않으며 따라서 초기에 설정된 기준 특징점이 카메라 영상 좌표계의 밖으로 이동해서 검출이 아예 불가능한 상태인 경우 카메라 추적 또한 실패하게 되는 단점이 있다. 이와 유사한 연구로 L. Alvarez[3] 등이 수행한 연구가 있다. L. Alvarez[3] 등의 연구에서는 카메라가 3차원 공간 내의 평면 객체방향을 향하고 있다고 가정한 다음 직선, 사각형, 원 등과 같은 사전에 정의된 형태를 갖는 도형을 이용하여 카메라 영상과 평면 객체 사이의 사영관계를 계산하고 이를 기반으로 카메라 자세 정보를 추정하였다. 이러한 방법은 사전에 정의된 형태의 패턴 검출을 위해 상대적으로 복잡한 과정과 많은 계산량을 필요로 하며, 패턴 검출의 정확도에 의해 사영관계의 계산 결과 또한 크게 영향을 받게 되는 문제점이 발생할 수 있다. 그 밖에 X. Liu[4], Z. Zhang[5] 등이 수행한 평면 객체 추적에 관현 연구가 있다. 이들 연구는 기본적으로 2차원 영상 내에 존재하는 평면 객체의 강건한 추적 방법에 대한 연구이며 평면 검출 알고리즘에 대한 연구 과정 및 결과가 소개되어있으나 카메라 추적과 관련된

내용은 다루고 있지 않다. 제안된 방법은 평면 객체 상에 존재하는 새로운 특징점을 지속적으로 검출하고 이를 이용하여 카메라 행렬을 추정함으로써 카메라 자세 초기화를 위한 기준 특징점들이 영상 밖으로 완전히 이동하여 검출이 불가능한 상황에서도 새로운 특징점들을 이용하여 카메라 추적이 가능하다는 점에서 기존의 방법들과는 차별된다.

3. 제안된 방법

3.1 새로운 특징점 검출 과정

본 논문에서는 카메라 추적을 위한 새로운 특징점 검출을 위해 S.H. Lee[6] 등이 제안한 방법을 사용한다. 즉, 두 개의 평면 객체 사이의 사영변환(projective transformation) 관계를 나타내는 2차원 호모그래피(homography)를 계산한 다음 이를 이용하여 평면 객체 상에 존재하지 않는 특징점들을 제거하고 평면상의 특징점들만을 검출하여 카메라 추적에 이용한다. 그림 1은 3차원 공간에서 한 점이 평면 객체 상에 존재할 때와 그렇지 않은 경우의 사이의 평면 호모그래피 관계를 보여준다. 그림 1에서 회색 사각형은 3차원 공간 내의 평면 객체를 나타내며 나머지 두 개의 사각형은 각각 두 개의 카메라 영상에 대한 평면을 의미한다. C와 C'은 각각 두 개의 카메라 중심이다. 그림 1에서 첫 번째 카메라의 중심 C와 $\mathbf{X}_1$을 통과하는 직선은 평면 $\boldsymbol{\pi}$ 상에 존재하는 한 점 X_π에서 교차하며 두 점 X_π과 $\mathbf{X}_1$는 첫 번째 영상의 2차원 점 x로 사영된다는 것을 보인다. 한편 세 개의 점 $\mathbf{X}_1$, $\mathbf{X}_2$, 및 $\mathbf{X}_\pi$는 각각 두 번째 영상의 x_1', x_2', $\tilde{x}$로 사영되며 두 카메라 영상 사이의 평면 호모그래피를 H라고 할 때 $\tilde{x}$는 $\tilde{x}$= Hx의 관계를 이용하여 계산할 수 있다. 그림에서 보이는 바와 같이 두 번째 영상의 사영점들은 첫 번째 영상에서와는 달리 하나의 점이 아닌 세 개의 개별적인 점으로 사영됨을 알 수 있다. 두 번째 영상에서 x_1'과 $\tilde{x}$ 사이의 길이인 l_1과 x_2'와 $\tilde{x}$ 사이의 길이 l_2는 평면 객체 $\boldsymbol{\pi}$와 3차원 좌표점 $\mathbf{X}_1$, $\mathbf{X}_2$ 사이의 거리 때문에 발생되는 시차값(parallax)이며 그림에서 알 수 있듯이 평면 상에 존재하는 점들은 카메라의 위치 및 방향이 변화해도 시차값이 발생되지 않는다. 평면 객체 $\boldsymbol{\pi}$이 동차좌표계(homogeneous coordinate) 내에서 식 (1)과 같이 주어졌다고 할 때 3차원 공간의 한 점 $x_i = (\tilde{x}_i^{\mathrm{T}}, \rho)^{\mathrm{T}}$는 식 (2)의 관계를 만족한다[7, 8].

$$\boldsymbol{\pi} = (\mathbf{n}^{\mathbf{T}}, d)^{\mathbf{T}} = (\mathbf{n}^{\mathbf{T}}/d, 1)^{\mathbf{T}} = (\mathbf{v}^{\mathbf{T}}, 1)^{\mathbf{T}} \tag{1}$$

$$\begin{aligned} \boldsymbol{\pi}\cdot\mathbf{X}_i^{\mathbf{T}} &= (\mathbf{v}^{\mathbf{T}}, 1)\cdot(\tilde{\mathbf{x}}_i^{\mathbf{T}}, \rho)^{\mathbf{T}} \\ &= \mathbf{v}^{\mathbf{T}}\tilde{\mathbf{x}}_i + \rho = 0 \end{aligned} \tag{2}$$

그림 1에서 두 번째 영상의 한 점 $\tilde{x}$는 평면 $\boldsymbol{\pi}$의 $\mathbf{X}_\pi$에 의해서 생성된 점이며 따라서 평면 호모그래피 관계에 의해 카메라 자세(pose)가 변화되는 조건에서도 역호모그래피(inverse homography) $\mathbf{H}^{-1}$를 이용하여 항상 평면 객체 $\boldsymbol{\pi}$의 점 $\mathbf{X}_\pi$로 역변환 시키는 것이 가능하며, 이 경우 점 $\mathbf{X}_\pi$는 두 번째 카메라의 위치 및 방향에 관계없이 항상 고정된 위치를 갖는다. 그리고 두 점 $\mathbf{X}_1$과 $\mathbf{X}_2$에 의한 x_1', x_2'를 $\mathbf{H}^{-1}$를 이용하여 역사영하게 되는 경우 고정된 점 $\mathbf{X}_\pi$와 큰 값이 거리값인 역사영 오차(bakc-projection error)가 발생된다. 식 (2)의 ρ및 영상 내의 특징점들의 역사영 오차로부터 특정 좌표 점이 $\boldsymbol{\pi}$ 위에 존재하는지 아닌지 판별 가능하다.

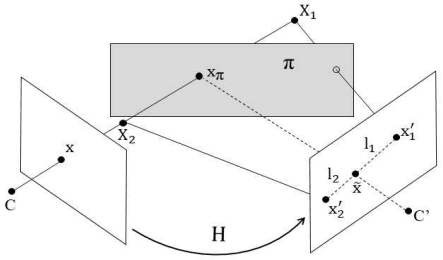


그림 1. 카메라 영상과 평면 객체 간의 평면 호모그래피
Fig. 1. Planar homography of the camera imagers and the planar object

3.2 2차원 사영관계를 이용한 카메라 자세추정

3차원 좌표계 내에 존재하는 한 평면이 기준 좌표계(world coordinate)의 *xy* 평면이라고 가정할 때 평면 위에 존재하는 하나의 점 $\mathbf{X}_n$과 영상 좌표계 상의 점 $\mathbf{x}_n$

간의 평면 호모그래피 관계는 다음과 같이 주어진다.

$$\mathbf{x}_n = \mathbf{H}\cdot\mathbf{X}_n \begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} = \mathbf{H}\cdot\begin{pmatrix} X_n \\ Y_n \\ 1 \end{pmatrix} \tag{3}$$

식 (3)에서 H는 3×3 호모그래피 사영행렬이며 평면 위의 하나의 좌표점을 2차원 영상 좌표계로 사영 시키는 사영변환이다. 식 (3)을 이용해서 다음의 2개의 선형 방정식을 얻는다.

$$x_n(X_n h_{31} + Y_n h_{31} + h_{33}) = X_n h_{11} + Y_n h_{12} + h_{13}$$
$$y_n(X_n h_{31} + Y_n h_{31} + h_{33}) = X_n h_{21} + Y_n h_{22} + h_{23}$$

위의 h_{ij}는 호모그래피 H의 i행 j열에 위치한 원소이며 $(X_n,\ Y_n)$과 $(x_n,\ y_n)$은 평면 객체와 2차원 영상 내의 n번째 좌표를 의미한다. 위의 관계식을 Ah=0의 선형 방정식 형태로 정리한 다음 최소자승추정법(least square estimation)을 이용하여 호모그래피 H를 계산한다. 여기서 h는 윗 식의 h_{ij}로 구성된 9×1 열벡터를 의미한다. 주어진 N개의 점들에 대하여 행렬 A의 크기는 $2N$×9가 되며 다음과 같은 행렬로 정리할 수 있다.

$$\mathbf{A} = \begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -X_1x_1 & -X_1y_1 & -X_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -Y_1x_1 & -Y_1y_1 & -Y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -X_2x_2 & -X_2y_2 & -X_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -Y_2x_2 & -Y_2y_2 & -Y_2 \\ & & \cdots & & & \cdots & & & \\ x_N & y_N & 1 & 0 & 0 & 0 & -X_Nx_N & -X_Ny_N & -X_N \\ 0 & 0 & 0 & x_N & y_N & 0 & -Y_Nx_N & -Y_Ny_N & -Y_N \end{pmatrix} \tag{4}$$

한편 Ah=0의 해인 h는 행렬 A의 우측 영공간(right null space)이며 행렬 A에 대한 SVD(singular value decomposition)를 통하여 얻는다. 호모그래피 행렬 H의 크기는 3×3이며 3행 3열에 위치한 원소는 스케일 요소(scale factor)로 일반적으로 h_{33}=1이 되도록 행렬 원소 전체가 정규화 된다. 따라서 H의 자유도는 8DOF(degree of freedom)이며 자유도가 2DOF인 2차원 평면 상의 점 4개 이상의 좌표값이 주어져야 H의 계산이 가능하다. [7, 8].

3.3 호모그래피 최적화

호모그래피의 정확도를 높이기 위해서 H 행렬은 별도의 최적화 과정을 거친다. 호모그래피 H는 RANSAC (RANdom SAmple Consensus)을 통하여 평가함수(cost function)인 식 (5)가 최소값을 갖도록 반복적으로 갱신된다. 식 (5)는 H를 이용해서 계산된 평면 상의 한 점 $\mathbf{X}_n$의 재사영 오차(reprojection error)를 의미한다. 3차원 좌표 $\mathbf{X}_n$와 영상 내의 한 점 사이의 호모그래피 H를 계산한 다음 H에 의해 3차원 공간으로 재사영된 평면 객체 상의 $\mathbf{X}_n$의 좌표 오차값이 최소값을 갖도록 H의 원소들이 반복적으로 수정된다[8].

$$\varepsilon(h_{ij}) = \sum_{k=1}^{N} d^2(\mathbf{x}_n,\ \mathbf{H}\cdot\mathbf{X}_n) \tag{5}$$

윗 식에서 d는 다음의 관계를 나타낸다.

$$d^2\left(\begin{pmatrix} x_i \\ y_i \\ w_i \end{pmatrix}, \begin{pmatrix} x_j \\ y_j \\ w_j \end{pmatrix}\right) = \left(\frac{x_i}{w_i} - \frac{x_j}{w_j}\right)^2 + \left(\frac{y_i}{w_i} - \frac{y_j}{w_j}\right)^2. \tag{6}$$

3.4 호모그래피를 이용한 카메라 자세 추정

3차원 공간의 한점 $\widetilde{\mathbf{M}} = (X,\ Y,\ Z)$와 영상의 한점 $\widetilde{\mathbf{m}} = (x, y)$의 동차좌표계(homogeneous coordinate) 좌표를 $\mathbf{M} = [X\ Y\ Z\ 1]^{\mathrm{T}}$, $\mathbf{m} = [x\ y\ 1]^{\mathrm{T}}$라고 할 때 이들 사이의 사영 관계는 3X4 행렬인 P에 의해 다음과 같이 계산된다.

$$\begin{aligned} \mathbf{m} &= \lambda\mathbf{P}\mathbf{M} = \lambda\mathbf{K}\left[\mathbf{R}\,|\,\mathbf{t}\right]\mathbf{M} \\ &= \lambda\mathbf{K}\left[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3\,|\,\mathbf{t}\right]\mathbf{M} \end{aligned} \tag{7}$$

식 (7)에서 λ는 사영 행렬 P의 스케일 값이며, R은 카메라의 회전 변위인 3 X 3 행렬이다. 그리고 $\mathbf{r}_i$는 행렬 R의 i번째 열(column)을 의미하며, t는 카메라의 이

동변위를 나타내는 3 X 1 이동 벡터 (translation vector)이다. 또한 $\mathbf{K}$ 행렬은 카메라의 내부 파라미터 (intrinsic parameter)가 원소인 3 X 3 카메라 교정 행렬(camera calibration matrix)로, 식 (8)과 같은 정칙 행렬 (non-singular matrix)이다.

$$\mathbf{K}=\begin{bmatrix} f_x & s & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8}$$

식 (8)에서 f_x, f_y는 영상 좌표 축 방향으로의 스케일 값이고, s는 카메라 영상의 비틀림 파라미터(skew parameter)이다. 그리고 $(x_0,\ y_0)$는 영상의 주점 (principal point)을 의미한다. 식 (8)의 카메라 행렬은 별도의 카메라 교정을 거쳐서 계산되며 본 논문에서는 Zhang의 카메라 교정 방법을 사용하여 계산하였다[9]. 식 (7)과 카메라 행렬 $\mathbf{K}$로부터 3차원 공간 내의 평면 $\boldsymbol{\pi}$와 2차원 영상 평면과 사이의 사영 관계를 구할 수 있으며, 이는 영상과 $\boldsymbol{\pi}$사이의 2차원 호모그래피로 계산된다. 평면 $\boldsymbol{\pi}$가 3차원 기준 좌표계 상의 xy평면이라고 가정하면 z 축 좌표값은 0이 되며 평면 $\boldsymbol{\pi}$ 상의 임의의 한 점의 좌표는 $\mathbf{X}_i=[X_i\ Y_i\ 0\ 1]^{\mathrm{T}}$가 된다. 평면 상의 임의의 점 $\mathbf{X}_i=[X_i\ Y_i\ 0\ 1]^{\mathrm{T}}$가 영상 좌표계 내의 한 점 $\tilde{\mathbf{x}}_i=[x_i y_i\ 1]^{\mathrm{T}}$로 사영되었다고 할 때 이들 사이의 관계는 다음과 같이 나타낼 수 있다.

$$\begin{aligned}\mathbf{x}_i&=\mathbf{K}[\mathbf{R}\,|\,\mathbf{t}]\mathbf{X}_i=\mathbf{K}[\mathbf{r}_1\quad \mathbf{r}_2\quad \mathbf{r}_3\,|\,\mathbf{t}]\mathbf{X}_i \\ &=\mathbf{K}[\mathbf{r}_1\quad \mathbf{r}_2\,|\,\mathbf{t}]\begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix}=\mathbf{K}[\mathbf{r}_1\quad \mathbf{r}_2\,|\,\mathbf{t}]\tilde{\mathbf{X}}_i\end{aligned} \tag{9}$$

한편, 3차원 공간에 존재하는 평면 객체 $\boldsymbol{\pi}$와 2차원 영상 평면 사이의 평면 호모그래피 관계에 의해서 $\boldsymbol{\pi}$ 상에 존재하는 점들 $\mathbf{X}_i$와 이들의 영상 $\mathbf{x}_i$ 사이에는 다음의 관계가 성립한다.

$$\begin{aligned}\mathbf{x}_i&=\mathbf{K}[\mathbf{r}_1\quad \mathbf{r}_2\,|\,\mathbf{t}]\tilde{\mathbf{X}}_i=[\mathbf{K}\mathbf{r}_1\quad \mathbf{K}\mathbf{r}_2\,|\,\mathbf{K}\mathbf{t}]\tilde{\mathbf{X}}_i \\ &=[\mathbf{h}_1\quad \mathbf{h}_2\quad \mathbf{h}_3]\tilde{\mathbf{X}}_i=\mathbf{H}\tilde{\mathbf{X}}_i\end{aligned} \tag{10}$$

따라서 식(10)에 의해서 평면 호모그래피 $\mathbf{H}$로부터 카메라의 회전 및 이동 변위 파라미터를 계산할 수 있다. 한편 회전 행렬의 직교 조건(orthogonal condition)에 의해서 $\mathbf{R}$행렬의 세 번째 열벡터는 $\mathbf{r}_1$과 $\mathbf{r}_2$의 외적(cross product), 즉 $\mathbf{r}_3=\mathbf{r}_1\times\mathbf{r}_2$ 이다. 위의 관계로부터 계산한 회전행렬 $\mathbf{R}$을 SVD(singular value decomposition)를 통하여 $\mathbf{R}=\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}}$로 분해한 다음 회전행렬 $\mathbf{R}=\mathbf{U}\mathbf{V}^{\mathrm{T}}$로 근사화시킨다 [10].

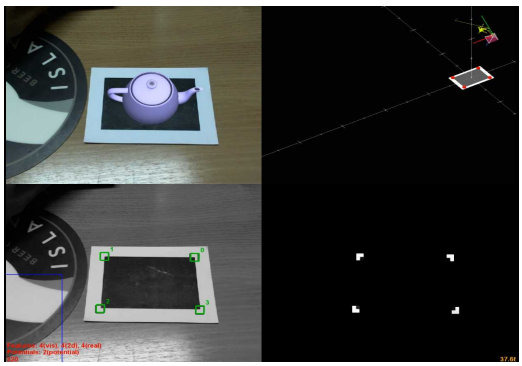


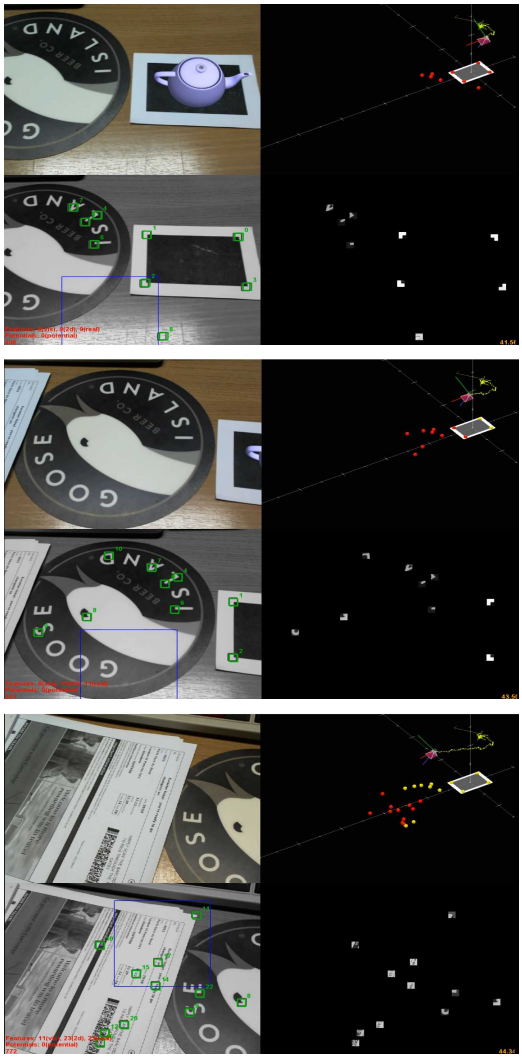
그림 2. 실시간 카메라 추적 및 새로운 특징점 검출 예제
Fig. 2. An example of the camera tracking and new feature detection

4. 실험 결과

제안된 방법에 대한 실험을 수행하기 위해 800 x 600 해상도의 영상을 1초당 30프레임으로 출력할 수 있는 USB 카메라를 이용하였다. 또한 제안된 방법을 구현하기 위해 Intel Xeon E5 CPU 및 128GB의 메모리가 내장되었으며 Windows10 운영체제가 설치된 PC를 사용하였다. 실험 과정에서 별도의 특수 조명은 이용하지 않고 일반적인 조명이 갖추어진 실내 환경 내에서 실험을 수행하였다. 그림 2에 실시간 카메라 자세 추정 결과와 특징점 검출 및 추적 결과의 예를 보인다. 카메라 자세 추정의 초기 단계에서 카메라의 위치와 방

향 등에 대한 초기화 과정을 거친다. 카메라 자세 초기화를 위해서 사전에 정의된 크기를 갖는 검은색 사각형의 모서리 4개의 좌표를 기준 특징점으로 이용하며 특징점의 좌표 추적을 위해서는 NCC(normalized cross correlation)를 이용한 템플릿 매칭을 수행한다. 한편 카메라 추적에 대한 시각적인 효과를 제공하기 위해 별도의 3D 모델과 실사 영상을 카메라 행렬을 이용하여 정합하였다. 그림 2의 좌측 상단은 카메라 행렬을 이용하여 3D 모델을 정합한 결과이며 좌측의 하단의 영상은 새로운 특징점 검출 및 특징점 추적 수행 과정을 보인다. 우측 상단은 기준 좌표계 내의 기하학적 관계를 조망한 장면(bird's eye view)을 실시간으로 보여주며, 3차원 공간에서의 카메라 움직임과 카메라 좌표, 검출된 3차원 특징점을 보이고 있다. 특히 빨간색으로 표시되는 공 모양의 객체는 현재 카메라에 검출되는 특징점들을 나타내며 노란색의 구는 현재 카메라에 검출되지 않는 특징점이다. 제안된 방법에서 새로운 특징점의 검출을 위해서 FAST 연산자를 이용한다[11, 12]. 새로운 특징점 검출 과정은 다음과 같다. 영상 내에서 정해진 크기를 갖는 임의의 영역을 선택한 다음 FAST 연산자를 이용하여 강건한 특징점 및 템플릿을 검출한다. 그리고 다음 추적 주기에서 새로 검출된 특징점의 평면 호모그래피 관계가 지속적으로 만족되면 평면에 존재하는 특징점으로 간주하고 다음 추적부터 카메라 자세 추정에 이용된다. 만약 특징점의 템플릿 매칭 결과가 불안정하거나 특징점 검출이 불가능하면 불안정한 특징점으로 간주하고 특징점을 목록에서 제거한다. 또한 특징점의 평면 호모그래피 관계가 성립하지 않는 경우에도 특징점을 목록에서 제거한다. 그림 3에 새로운 특징점 검출 및 카메라 자세 추정의 예를 보인다. 제안된 알고리즘에서는 3차원 공간 내의 기준 평면상에 존재하지 않는 특징점은 모두 제거하고 평면 위의 특징점들만을 이용하며, 따라서 그림에 보이는 특징점들은 모두 기준 평면 상에 존재한다. 카메라 추적 초기에 사용되는 기준 특징점들이 카메라 영상 밖으로 모두 벗어난 경우에도 새로운 특징점을 검출하고 카메라 자세 정보를 추정함으로써 실시간으로 카메라 추적이 수행되는 것을 알 수 있다. 그림 4와 그림 5는 각각 특징점들의 재사영 오차(reprojection error)와 역사영 오차(back-projection error)에 대한 그래프를 보여준다. 재사영 오차는 3차원 공간의 특징점을 카메라 행렬을 이용해서 영상으로 재사영시킨 점의 좌표와 영상에서 검출된 특징점 좌표 사이의 오차이며 카메라 행렬이 정확할수록 0에 가까운 값을 갖는다. 역사영 오차는 영상 내의 특징점을 평면 호모그래피를 이용하여 3차원 공간 내의 평면 위로 역사영 시킨 후 계산한 값이며, 3차원 공간 내의 특징점과 역사영 된 특징점 사이의 실제 거리 값이다. 그림 4는 추적 중인 30개의 특징점에 대한 재사영 오차를 보여주고 있으며 각각의 특징점들을 구분하기 위하여 특징점 별로 별도의 색상을 할당하였다. 재사영 오차의 경우 10픽셀 이하의 값을 갖는다는 것을 알 수 있으며 영상의 해상도를 고려할 때 상대적으로 작은 값이라고 판단할 수 있다. 하지만 카메라가 초기 위치에서 멀어질수록 오차가 증가하는 것을 확인할 수 있으며 특히 400프레임부터 오차 값이 급격히 증가하기 시작하는 것을 알 수 있다. 이는 제안된 방법이 호모그래피와 RANSAC 만을 이용하기 때문에 자세 추정 오차가 지속적으로 누적되면서 발생되는 현상이라고 판단된다. 이는 향후 별도의 최적화 과정을 추가하면 상당부분 개선될 수 있을 것으로 예상된다. 표 1은 특징점 10개의 재사영 오차 평균값이다. 평균 오차 또한 영상의 해상도 대비 상대적으로 작은 값을 갖는 것을 알 수 있다. 그림 5는 역사영 오차의 예를 나타내기 위해 5개의 특징점이 추적되는 과정에서의 역사영 오차 그래프를 보여주며 그림 4와 마찬가지로 각 특징점을 구별하기 위해 별도의 색상을 할당하였다. 그림 5의 역사영 오차에서 추적 중간에 단절된 그래프들이 확인된다. 이는 평면상에 위치하지 않은 점들의 경우 큰 값의 역사영 오차를 갖기 때문에 이러한 특징점들을 검출 과정에서 제거함으로써 발생 되는 현상이다. 그림 6은 3차원 공간 내에서의 카메라 좌표이며 카메라의 실제 3차원 좌표값으로 주어진다. 그림 3에서 600~700 프레임 사이에 급격한 흔들림 현상이 발생한 것을 확인할 수 있으며, 이는 앞서 언급한 호모그래피 오차의 누적으로 인한 것으로 판단된다. 그림 7은 동일 평면상의 특징점 검출 기법이 적용되지 않은 경우의 카메라 추적을 보인다. 4개의 기준 특징점이 영상 내에 존재하는 경우 정상적인 카메라 자세 추정이 가능하나 기준 특징점들이 영상

밖으로 벗어나면카메라 추적 및 자세 추정 또한 불가능하다.



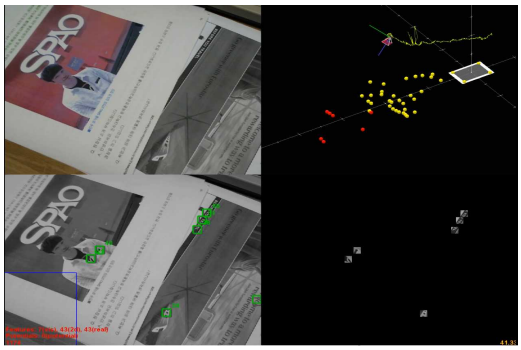


그림 3. 실시간 카메라 추적 및 새로운 특징점 검출 과정
Fig. 3. The camera tracking and new feature detection

표 1. 재사영 오차의 평균값 (단위: 픽셀)
Table 1. Averages of the reprojection errors

feat.1	feat.2	feat.3	feat.4	feat.5
1.534844	2.923790	3.864044	1.485394	1.912778
feat.6	feat.7	feat.8	feat.9	feat.10
7.335517	4.348311	4.121444	1.508944	1.683497

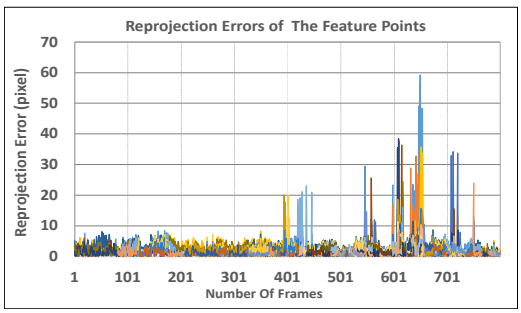


그림 4. 특징점의 재사영 오차
Fig. 4. Reprojection errors of the feature points

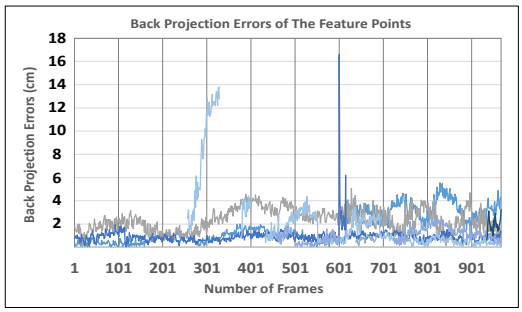


그림 5. 특징점의 역사영 오차
Fig. 5. Back-projection errors of the feature points

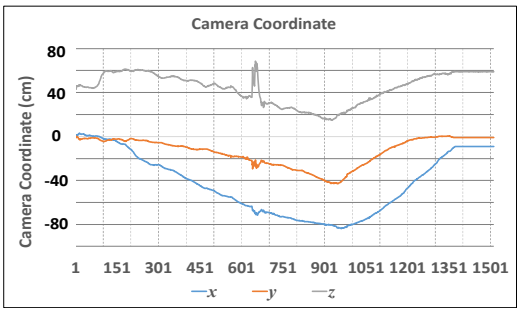


그림 6. 카메라의 좌표
Fig. 6. Camera corrdinates

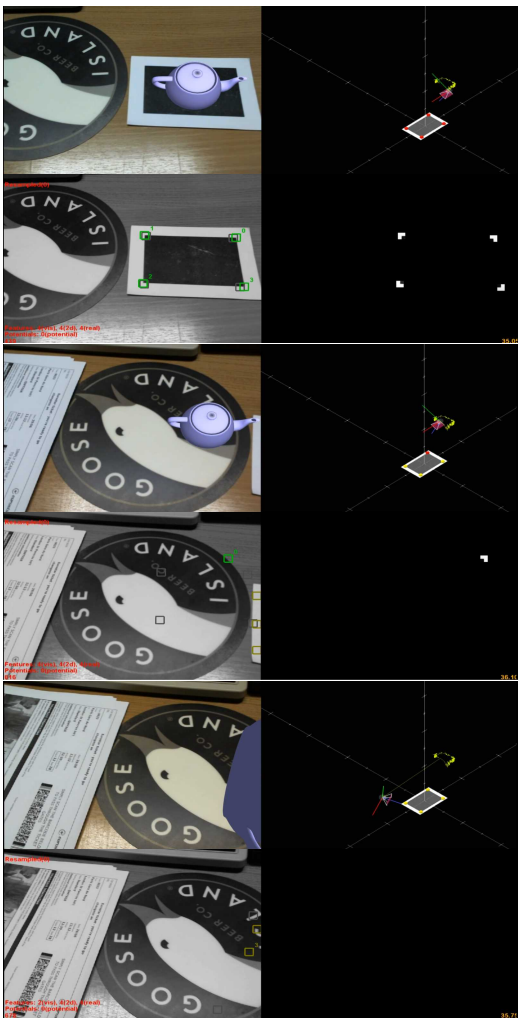


그림 7. 평면 객체 상의 새로운 특징점 검출을 이용하지 않는 기존의 방법

Fig. 7. Conventional method which does not employ the new feature detection on the planar object

5. 결론

본 논문에서는 3차원 공간 내에서 동일 평면 객체 상에 존재하는 특징점들 만을 검출한 다음 이를 이용하여 실시간 카메라 추적을 수행하기 위한 방법을 제안하였다. 제안된 방법은 기존의 방법들과는 다르게 평면 객체 상에 존재하는 새로운 특징점을 지속적으로 검출한 다음 이를 카메라 추적에 이용함으로써 카메라 자세 초기화를 위해 사용되는 기준 특징점들이 카메라 영상 밖으로 완전히 벗어난 상황에서도 카메라 추적이 가능하다는 장점이 있다. 또한 제안하는 방법은 카메라 영상 이외에 별도의 가속도센서, 자이로스코프, 기울기 센서 등과 같은 장비를 사용하지 않고 영상에서 취득되는 2차원 시각 정보만을 이용해서 카메라의 자세정보를 추정한다. 실험을 통하여 제안된 방법의 성능 및 결과를 확인하였으며, 특히 제안된 방법에서 평면 호모그래피 계산에서 발생하는 오차가 카메라 자세 추정 과정에서 지속적으로 누적되는 현상이 있음을 확인하였다. 이러한 오차는 카메라가 초기 위치에서 멀어질수록 카메라 자세 추정과 특징점 추적에 더욱 많은 영향을 미칠 수 있음 또한 확인하였으며, 좌표계의 기준점과 카메라 사이의 거리가 길어지는 경우 지속적으로 증가함 또한 확인하였다. 이는 제안된 방법이 호모그래피와 RANSAC 이외에 다른 방법을 사용하지 않았기 때문에 발생하는 현상으로 판단된다. 향후 연구에서는 제안된 방법에 추가적으로 별도의 최적화 알고리즘 및 새로운 특징점 검출 기법 등을 결합하여 더욱 강건하고 안정적인 카메라 추적 기법을 구현할 계획이다.

REFERENCES

[1] Lee, A.-H., Lee, S.-H., Lee, J.-Y., Kim, T.-E., & Choi, J.-S., "Real-time camera tracking using planar object detection," Lecture Notes in Electrical Engineering (LNEE), Vol. 120, pp. 373-378., 2011

[2] S.H. Lee, "Real-time Monocular Camera Pose Estimation using a Particle Filiter Intergrated

with UKF," Journal of Korea Institute of Information, Electronics, and Communication Technology, Vol. 16, No. 5, pp. 315-324, 2023

[3] Alvarez, L., Gomez, L., Henriquez, P., & Mazorra, L., Automatic camera pose recognition in planar view scenarios. In Lecture Notes in Computer Science (LNCS), Proceedings of the Iberoamerican Congress on Pattern Recognition, pp. 406-413, Springer, 2012

[4] Liu, X., Liu, X., Yi, Z., Zhou, X., Le, T., Zhang, L., Huang, Y., Yang, Q., & Fan, H.. PlanarTrack: A Large-scale Challenging Benchmark for Planar Object Tracking. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 20449-20458, 2023

[5] Zhang, Zhicheng, Liu, Shengzhe, & Yang, Jufeng, Multiple Planar Object Tracking. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 23460-23470, 2023

[6] S.H. Lee, "Detection of the co-planar feature points in the three dimensional space," Journal of Korea Institute of Information, Electronics, and Communication Technology, Vol. 16, No. 5, pp. 499-508, 2023

[7] R .Hartely and A. Zisseman, Multiple View Geometry, Cambridge, 2003.

[8] O. Faugeras, Three-Dimensional Computer Vision, Cambridge, 1993

[9] Z. Zhang, "A Flexible New Technique for Camera Calibration and the Application," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 11, pp. 1330-1334, Nov. 2000

[10] G. Golub and C. Loan, "Matrix Computations: 3rd ed.," Johns Hopkins Univ. Press, 1996.

[11] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," IEEE International Conference on Computer Vision, Vol. 2, pp. 1508-1511, Oct. 2005

[12] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," European Conference on Computer Vision, Vol. 1, pp.430-443, May. 2006

[13] S. Thrun, W. Burgard, D. Fox, Probabilistic Robotics, pp. 65-67, MIT Press, 2008

저자약력

이 석 한 (Seokhan Lee) **[정회원]**

- 1999년 2월 중앙대학교 전자공학과 졸업 (공학사)
- 2001년 2월 중앙대학교 첨단영상대학원 졸업 (공학석사)
- 2001년~2004년 ㈜LG전자 주임연구원
- 2009년 2월 중앙대학교 첨단영상대학원 졸업 (공학박사)
- 2013년~현재 전주대학교 정보통신공학과 교수

〈관심분야〉 실시간 카메라 추적, 증강현실, 컴퓨터 비전