

반복적 호모그래피 개선에 의한 강건한 평면 추적

*김가람 *박정식 **박한훈 *박종일¹⁾

*한양대학교, **부경대학교

grkim@mr.hanyang.ac.kr, nangsik@mr.hanyang.ac.kr, hanhoon_park@pknu.ac.kr,

*jipark@hanyang.ac.kr

Robust Planar Tracking Based on Iterative Homography Refinement

*Kim, Karam *Park, Jungsik **Park, Hanhoon *Park, Jong-II*

*Hanyang University, **Pukyong National University

요 약

평면 추적(planar tracking) 기반의 카메라 추적에 있어, 특징 검출자의 반복성과 특징 기술자(descriptor)의 정합 성능에 따라서 떨림 현상(jitter)이 발생한다. 특히, 모바일 환경에서와 같은 연산력이 부족한 환경에서 고속화를 위해 특징 검출 및 기술 알고리즘을 간략화 시킬 경우, 이러한 떨림 현상은 심각한 문제가 된다. 본 논문에서는 이러한 문제를 해결하기 위해 입력 영상을 워핑(warping)하여 특징 점을 재검출한 후 카메라 영상과 참조 영상(reference image) 사이의 호모그래피를 보완하는 방법을 제안한다. 실험을 통해 제안된 방법이 특징 검출 및 기술 알고리즘의 성능을 보완하여 떨림 현상을 크게(70% 이상) 감소시킴을 확인하였다.

1. 서론

미리 알고 있는 2 차원 평면에 대한 참조 영상을 가지고 있을 때, 특징 매칭(feature matching)을 통해 카메라 영상에서 평면을 추적함으로써 카메라의 3 차원 움직임을 추적할 수 있고(이하, 평면 추적 기반 카메라 추적이라고 함), 임의의 가상 객체를 카메라 영상에 증강(augmenting)할 수 있다. 이때, 안정적으로 객체를 띄우기 위해서는 2 가지의 조건이 필요하다. 그 중 한가지 조건은 고속으로 평면으로 추적할 수 있어야 하고, 다른 한 가지 조건은 영상에서 찾아진 평면 상의 특징이 현재 카메라 영상에 찾아진 평면 상의 특징과 정확하고 강건(robust)하게 매칭되어야 한다는 것이다. 그러나, 이 두 조건은 서로 상충되므로, 모두 만족시키는 것은 쉽지 않다.

영상에서 특징을 검출하는 방법은 다양하게 소개되어 왔다. 우선, 코너(corner)를 검출하는 방법으로 Harris 방법[1], Shi-Tomasi 방법[2], FAST[3] 등이 있다. 이 중에 FAST는 매우 빠르고 간단한 알고리즘을 가진다는 장점 때문에 가장 널리 사용되고 있다. 다만, 검출된 코너의 반복성(repeatability)이 떨어진다는 단점이 있다. 다음으로, 영역(region)이나 블롭(blob)을 검출하는 대표적 방법으로 SIFT[4]가 있다. SIFT는 가장 뛰어난 강건성(robustness)을 보장하기 때문에 널리 사용되어 왔지만, 연산의 복잡성 때문에

실시간 연산을 필요로 하는 응용 분야에서는 사용할 수 없다. 이러한 SIFT의 단점을 보완하기 위해서 SURF[5]는

LoG(Laplace of Gaussian) 필터를 DoG(Difference of Gaussian) 필터로 간소화하고 이를 다시 박스(box) 필터로 간소화함으로써, 적분 영상을 이용하여 고속으로 특징 영역을 검출, 기술할 수 있는 있다. 그러나, 영상의 해상도가 커지거나 모바일 환경과 같이 연산력이 부족한 환경에서는 여전히 실시간 연산을 수행할 수 없다. 이에 최근에는 보다 연산 속도를 향상시킨 BRIEF[6], ORB[7] 같은 방법들이 나오고 있지만 강건성 면에서는 많은 보완이 필요하다. 예를 들어서 BRIEF의 경우에는 회전에 영향을 받는다는 단점이 있다.

결국, 지금까지 제안된 여러 방법을 평면 추적 기반 카메라 추적에 사용할 경우, 특징점 및 영역 매칭에서 잘못되는 경우가 많이 발생하고, 정지한 카메라에 대해서도 참조 영상과 카메라 영상 사이의 호모그래피가 변하면서 떨림 현상이 발생한다. 즉, 카메라 영상을 참조 영상으로 역투영(back-projection)시켰을 때, 역투영된 영상은 참조 영상과 조금씩 차이를 보이며, 이러한 차이는 시간적으로 계속 변한다. 본 논문에서는 이러한 떨림 현상 문제를 보완하기 위해 역투영된 영상과 참조 영상 사이의 미소(infinitesimal) 호모그래피를 구해서 원래의 카메라 영상과 참조 영상 사이의 호모그래피를 갱신해줌으로써, 정확한 카메라 추적이 가능하도록 유도하고 떨림 현상을 해소하는 방법을 제안한다.

1) 교신저자

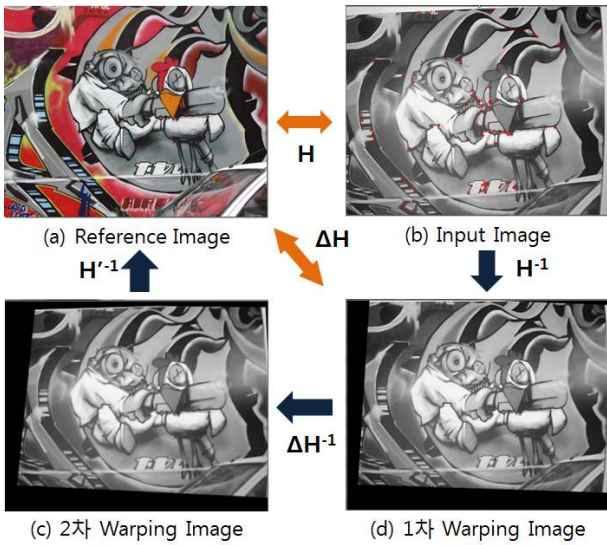


그림 1. 카메라 입력 영상과 참조 영상 사이의 호모그래피 계산. 카메라 입력 영상이 들어오면, 참조 영상과의 매칭을 통해서 초기 호모그래피(H)를 구하고, 이 H를 통해 1차 워핑된 영상과 참조 영상 사이의 호모그래피(ΔH)를 다시 계산해서 전체 호모그래피를 개선한다. 이러한 과정은 ΔH가 단위 행렬에 가까워질 때까지 반복된다

2. 반복 워핑 기반 호모그래피 개선에 의한 카메라 포즈 계산

카메라 입력 영상이 들어오면, 참조 영상과의 정합을 통해서 호모그래피(H)를 구한다. 이상적인 상황이라고 가정한다면 카메라 영상에 H⁻¹를 곱하면 참조 영상과 일치해야 한다. 하지만 블러나 노이즈 등의 이유로 워핑된 영상과 참조 영상은 정확히 정합되지 않는다. 이는 증강 현실에서, 최종 결과인 카메라 움직임에 영향을 주어 증강된 가상 객체의 떨림 현상으로 나타나는데, 이를 보완하기 위해서 워핑된 영상과 참조 영상 사이의 호모그래피(ΔH)를 구하고, 앞서 구한 H와 곱해서 다음과 같이 호모그래피를 개선한다(그림 1 참조).

$$H' = \Delta H \cdot H. \tag{1}$$

이러한 과정은 ΔH가 단위 행렬에 가까워질 때까지 반복된다. 식 (1)을 통해 구해진 호모그래피로부터 카메라의 3차원 움직임은 다음과 같이 구할 수 있다. H를 카메라의 내부 행렬과 포즈 행렬로 나타내면 다음과 같다.

$$H' \propto A [R_1 \ R_2 \ T] \tag{2}$$

여기서, R₁, R₂와 T는 각각 회전 행렬의 열 벡터와 이동 벡터를 나타낸다. 또한 A는 카메라 내부 행렬로 미리 카메라 보정(calibration)을 통해서 구할 수 있다. 식

(2)에서 카메라 포즈 행렬은 G = A⁻¹H로 유도할 수 있다.[8] 이 때, G는 크기(scale)에 독립적이지 않기 때문에 R₁과 R₂는 완벽히 직교정규(orthonormal)하지 않다. 따라서, 정규화(normalization)를 수행 하면,

$$R_k = G_k / l, \quad (k=1,2), \quad T = G_3 / l, \tag{3}$$

where $l = \sqrt{\|G_1\| \cdot \|G_2\|}$,

이고,

$$c = R_1 + R_2, \quad p = R_1 \times R_2, \quad d = c \times p, \tag{4}$$

이라고 하면,

$$R_3 = R_1' \times R_2', \tag{5}$$

where

$$R_1' = \frac{1}{\sqrt{2}} \left(\frac{c}{\|c\|} + \frac{d}{\|d\|} \right), \quad R_2' = \frac{1}{\sqrt{2}} \left(\frac{c}{\|c\|} - \frac{d}{\|d\|} \right),$$

이 된다. 그러므로, A를 안다면 카메라 포즈 행렬을 구할 수 있다.

3. 실험 결과

실험을 위해 그림 1에서 보이는 평면에 대해 웹 카메라(Logitech HD 프로 웹캠 C910, 320 x 240)를 사용하여 영상을 획득한 후, SURF를 이용하여 특징 검출, 추적을 수행하고, 2장에서 설명한 방법을 이용하여 카메라의 포즈를 추정하였다.

추정된 카메라 포즈에서의 떨림 정도를 정확하게 측정하기 위해서 카메라를 한 곳에 고정해 놓고 촬영한 영상을 대상으로 카메라 포즈를 추정한 후, 매 프레임 카메라 포즈가 얼마나 변하는지 측정하였다. 카메라를 정지시켜 놓고 촬영했기 때문에 이상적인 경우에는 추정된 카메라 포즈 역시 일정한 값으로 나와야 한다. 하지만 그림 2에서 보는 것처럼, 코너 검출과 기술에서의 에러로 인해서 카메라 포즈가 잘못 추정되어 매 프레임 변하게 되어 떨림 현상이 생긴다. 코너 검출과 기술에서의 에러가 클수록 카메라 포즈에서의 떨림이 심해진다. 그림 2에서 붉은 선은 ΔH를 구해서 개선된 호모그래피(식 (1)에서 H')를 이용하여 추정된 카메라 포즈에 대한 그래프이고, 푸른 선은 개선되지 않은 호모그래피(식 (1)에서 H)를 이용하여 추정된 카메라포즈에 대한 그래프이다. 그림 2의 결과를 통해 개선된 호모그래피를 이용했을 때 떨림 현상이 더 작다는 것을 알 수 있다. 이를 정량적으로 분석하기 위해 표 1에서 보는 것처럼, 카메라 포즈의 각 파라미터의 시간에 대한 변화량을 계산하였다. 표 2에서 제안된 방법에 의한 떨림 현상의 감소율을 보여준다. 제안된 방법을 통해 떨림 현상이 72.53%정도 감소함을 알 수 있다.

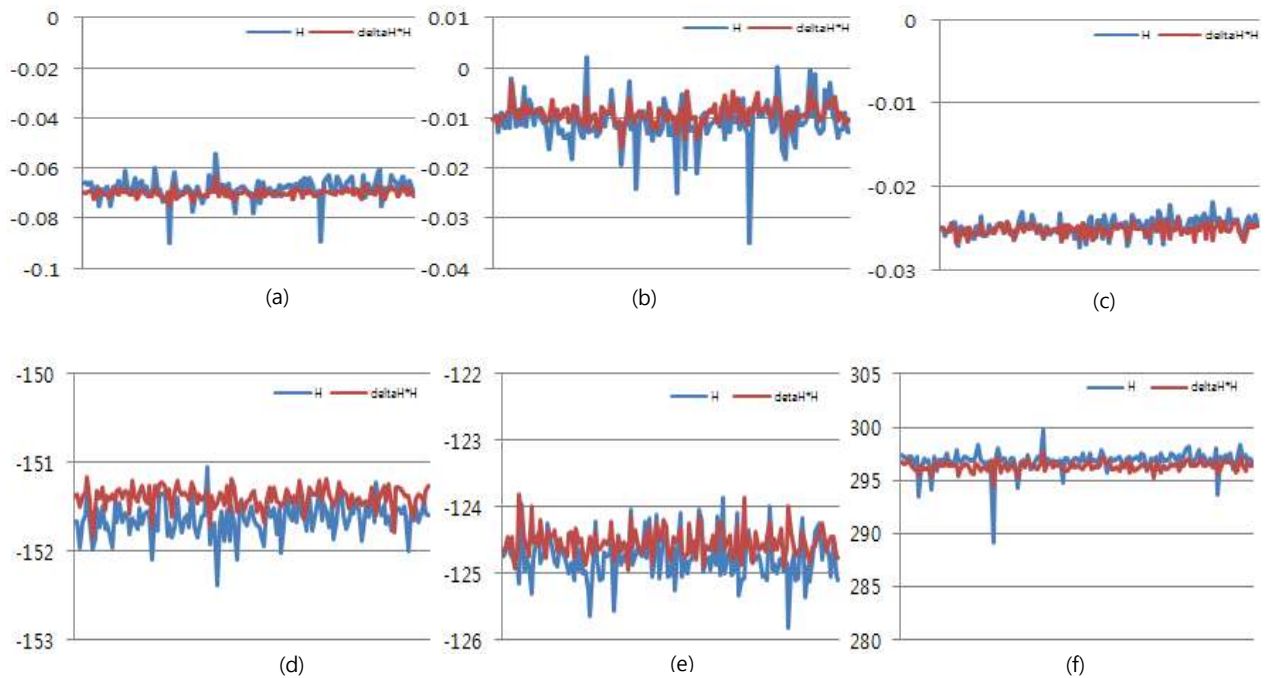


그림 2. 추정된 카메라 포즈 결과. (a)–(c): 카메라의 x, y, z 축에 대한 회전 파라미터, (d)–(f): 카메라의 x, y, z 축에 대한 이동 파라미터. 제안된 방법과 기존 방법을 사용해서 정지된 카메라에서 평면을 추적할 때 카메라 포즈 추정의 성능을 평가한 것이다. 모든 그래프에서 제안된 방법이 기본 방법에 비해 평균적으로 변화율이 감소했을 뿐만 아니라, 순간적으로 크게 요동하는 경우도 매우 감소한 것을 알 수 있다.

표 1. 추정된 카메라 포즈 파라미터의 단위 시간에 대한 변화량

	x 이동	y 이동	z 이동	x 회전	y 회전	z 회전
기존방법 (식 (1)을 수행하지 않음)	0.03470	0.11129	0.965880	1.8E-05	1.9E-05	1.1E-06
제안한 방법 (식 (1)을 한 번 수행)	0.01355	0.04158	0.171945	2.3E-06	3.9E-06	4.1E-07
제안한 방법 (식 (1)을 두 번 수행)	0.00917	0.03275	0.10388	1.02E-06	2.54E-06	3.31E-07

표 2. 식(1)을 수행 후에 분산이 줄어드는 정도(%)

	x 이동	y 이동	z 이동	x 회전	y 회전	z 회전
식 (1)을 한 번 수행	60.95%	62.63%	82.20%	87.22%	79.47%	62.73%
식 (1)을 두 번 수행	32.71%	21.25%	39.58%	55.86%	34.77%	19.17%

제안한 방법은 3 장에서 설명한 것처럼 호모그래피를 반복해서 개선하여 카메라 포즈 에러 및 떨림 현상을 점점 줄여주는 방법이다. 그래서 여러 번을 시행 할수록 H 가 점점 더 정확해 진다. 실제, 한번 더 호모그래피를 개선해주면 표 1,

2 에서 보는 것처럼 떨림 현상은 보다 감소한다는 것을 알 수 있다. 그러나 호모그래피를 개선해주는 횟수가 증가할수록 떨림 현상의 감소율은 크게 떨어지는 반면 계산 시간은 계속 증가하기 때문에 (표 3 참조) 1~2 번을 개선하는 것이 가장 효과적이다. 사실, 본 논문에서는 특징 검출 및 추적을 위해 SURF 를 사용했기 때문에, 호모그래피를 개선하는 횟수를 증가시킬수록 계산 속도가 크게 떨어지는 문제가 있었다. 그러나, 이러한 문제는 고속 특징 검출 및 기술자를 사용함으로써 완화될 것으로 판단되며, 본 논문의 목적은 호모그래피를 개선해 줌으로써 떨림 현상이 줄어드는지를 확인하는 데 있었기 때문에, 계산 속도에 대해서는 크게 고려하지 않았다.

표 3. 카메라 추적에 걸리는 시간

	계산 시간(sec/frame)
식 (1)을 수행하지 않음	0.1173
식 (1)을 한 번 수행	0.2056
식 (1)을 두 번 수행	0.3003

4. 결론

본 논문에서는 평면 추적 기반 카메라 추적에서 떨림 현상을 줄이기 위해 참조 영상과 카메라 영상 사이의 호모그래피를 반복적으로 개선해 주는 방법을 제안하였다. 제안된 방법을 이용함으로써, 특징 검출 및 기술 단계에서 다소의 에러가 발생하더라도 강건하게 카메라 포즈를 추정하는 것이 가능하였으며, 결과적으로 떨림 현상을 크게 줄일 수 있었다. 향후, 다양한 검출자 및 기술자에 대해 제안된 방법의

성능을 검증, 비교할 예정이다. 또한 떨림 현상에 최적인 검출자 및 기술자를 개발하고자 한다.

참고 문헌

- [1] C. Harris, and M. Stephens, "A combined corner and edge detector," In Proc. of the Alvey Vision Conference, pp.147-151, 1988.
- [2] J. Shi, and C. Tomashi, "Good Features to Track," In Proc. of IEEE CVPR, pp.593-600, 1994.
- [3] E. Rosten, and T. Drummond, "Fusing points and lines for high performance tracking," In Proc. of ICCV, pp.1508-1515, 2005.
- [4] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol.60, pp.91-110, 2004.
- [5] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded-up robust features," In Proc. of ECCV, 2006.
- [6] M. Calonder, V. Lepetit, C. Strecha, and P. Fua. "Brief: Binary robust independent elementary features," In Proc. of ECCV, 2010.
- [7] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: an efficient alternative to SIFT or SURF," In Proc. of ICCV, 2011.