

보행자 기반 카메라 캘리브레이션 알고리즘 분석

Analysis of Camera Calibration Based on Pedestrian

서 찬 호¹ · 최 성 록^{1†}

Chanho Seo¹, Sunglok Choi^{1†}

Abstract: This study evaluates a pedestrian-based camera calibration technique that forgoes traditional vanishing points, using line segments from pedestrian paths instead. We conducted synthetic experiments adding Gaussian noise to assess the robustness of focal length estimation. Results indicate a threshold of noise tolerance, beyond which the error significantly increases, pinpointing the limitations of current methods. The paper suggests future research could improve robustness by incorporating RANSAC (Random Sampling Analysis) into the calibration process.

Keywords: Camera Calibration, Autocalibration

1. 서 론

카메라 캘리브레이션은 많은 3차원 컴퓨터비전 알고리즘에 기본적으로 필요한 과정이다. 3차원 세계(world coordinate) 와 2차원 화면(image coordinate) 간 투영 관계를 표현하기 위해 정확한 카메라 내외부 파라미터의 획득, 즉 카메라 캘리브레이션이 필요하다.

Zhnag's method [1] 가 카메라 캘리브레이션을 수행하는 가장 널리 쓰이는 대표적인 기술이다.

그러나 다양한 카메라 설치 환경에서 다양한 요구 사항 및 제반 사항이 수반되는데, 모든 경우에서 일반적인 성능을 기대하기엔 체커보드 기반 카메라 캘리브레이션은 한계를 가지고 있다. 예를 들어 CCTV 카메라가 이미 설치되어 있는 상황에서 카메라 캘리브레이션을 진행할 경우, 먼 거리의 지면에서 체커보드를 촬영하기 위해서는 매우 큰 체커보드가 필요하다. 또 체커보드 기반의 캘리브레이션은 체커보드의 코너점들을 다양한 각도와 위치에서 인식하는 것이 중요한데, 감시 카메라의 고정된 위치와 각도 때문에 이러한 다양한 포즈를 확보하는 것은 까다롭다,

보행자를 활용한 카메라 캘리브레이션 연구가 지속적으로 진행되어 왔는데, 그 이유 중 하나는 감시 카메라에서 보행자가 쉽게 관찰되는 주요 대상이기 때문이

다. 보행자를 대상으로 한 연구가 많이 이루어진 배경에는 보행자의 움직임을 통해 소실점을 획득할 수 있다는 점이 크다. 보행자의 두 임의의 시간 사이에서 움직임을 분석해 소실점을 얻을 수 있다.

반면 소실점을 사용하지 않는 방법론 또한 별도의 연구 분기로 존재하고 있다. 소실점에 의존하지 않는 다양한 방법론들은 특정 환경과 조건에서 더욱 효율적이거나 정확할 수 있기 때문이다. Nagano[2]는 전통적인 소실점 기반 방법론 대신, 보행자로부터 추출된 동일한 길이의 선분 투영식을 활용해 카메라 파라미터를 도출하는 방식을 제안하였다.

본 논문은 해당 연구에 대한 고찰 및 분석을 통해 소실점을 사용하지 않는 카메라 캘리브레이션 방법론의 한계 및 특징을 도출하고자 한다. 더 나아가서 결론에는 해결 방법에 대한 후속 연구에 대해서 고민해보고자 한다.

2. 보행자 기반 카메라 캘리브레이션

[Fig. 1] 은 카메라 캘리브레이션이 상정하는 두 좌표계 (월드 좌표계, 카메라 좌표계) 와 선분 데이터의 기하학적 관계를 표현한다. K 가 카메라 내부 파라미터 행렬, R 이 회전행렬, t 가 병진벡터를 의미할 때, 하나의 선분에서 각각 바닥점(a) 과 머릿점(b)에 대한 두 개의 투영식을 얻을 수 있다.

$$\lambda_i a_i = K(Ra_i + t)$$

$$\mu_i b_i = K(Rb_i + t)$$

※ This work was supported by an Electronics and Telecommunications Research Institute (ETRI) grant funded by the Korean government. [22ZR1210, Development of Technology to reproduce and analyze the real-world cities into digitalized intelligent urban spaces].

1. Computer Science and Engineering Department, Seoul National University of Science and Technology (SEOULTECH), Seoul, Korea

† Corresponding author (sunglok@seoultech.ac.kr)

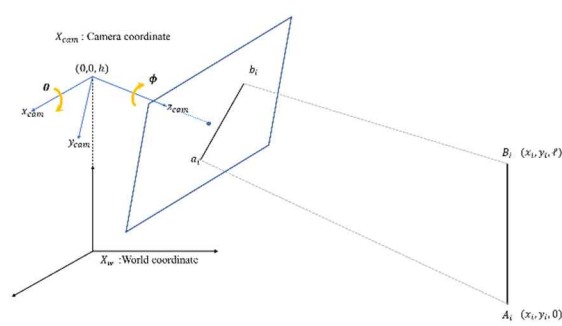
$$K = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$R = \begin{pmatrix} \cos \phi & -\cos \theta \sin \phi & \sin \theta \sin \phi \\ \sin \phi & \cos \theta \cos \phi & -\sin \theta \cos \phi \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

$$t = -R \begin{pmatrix} 0 \\ 0 \\ h \end{pmatrix}$$

이때 $c_i = \mu_i b_i - \lambda_i a_i$, $d_i = \lambda_i a_i - \lambda_1 a_1$ 로 두었을 때, 투영식에 대해서 다음과 같이 식이 전개되어 초점에 대한 closed-form을 구할 수 있다.

$$c^T K^{-T} K^{-1} d = 0$$

$$f = \sqrt{\frac{\sum (c_{i,1} d_{i,1} + c_{i,2} d_{i,2})(c_{i,3} d_{i,3})}{\sum (c_{i,3} d_{i,3})^2}}$$


[Fig. 1] Geometrical overview of camera calibration

3. 실험 결과

본 실험은 앞서 설명한 개념을 구현하여 노이즈에 대한 synthetic 성능 실험을 진행했다. 관찰한 선분 데이터에 가우시안 노이즈의 표준 편차 0, 0.1, 0.2, 0.5, 1, 2, 5, 8, 10 [pixel]을 주어 선분 데이터 개수와 성능의 상관 관계를 파악해보고자 했다.

알고리즘 및 실험 환경은 Python 3.8 환경에서 구현하였고 실험 설정 환경은 다음과 같이 정리했다.

[Table 1] Configuration of synthetic experiment

Configurations	Values
f (focal length)	1000
θ (orientation #1)	110 [degree]
ϕ (orientation #2)	0 [degree]
h (height)	3 [m]
FoV	69 [degree]
Resolution	1920 x 1080

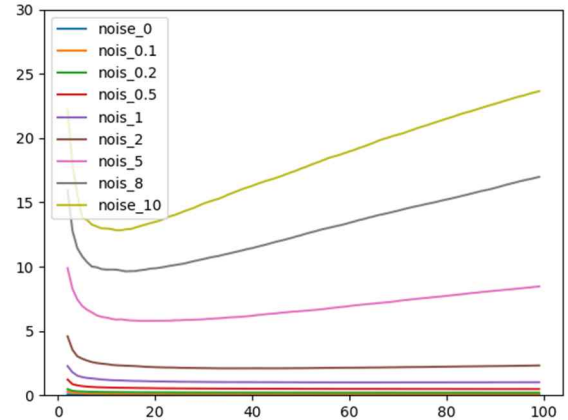
파악하고자 하는 성능은 추정된 focal length 와 참값 간의 상대 에러 값으로 설정했다. [Fig. 2] 는 해당 실험

의 결과 그래프이다. x 축은 선분 데이터 개수, y축은 상대 에러 값을 의미한다.

노이즈가 커짐에 따라 전반적으로 에러 값이 늘어나는걸 볼 수 있다. 이때 특징적인 것은, 픽셀 노이즈 표준 편차 값이 5로 설정했을 때 특정 선분 개수부터 에러 값이 반등한다는 걸 알 수 있다. 이때 에러 값이 반등하는 값은 각각 다음과 같다.

4. 결론

본 논문에서는 기존 보행자 기반 카메라 캘리브레이션 기법을 구현 및 분석하였다. 본 논문에서 진행한 실험을 통해 HD 카메라 환경에서 노이즈 레벨이 5 이상 올라 갔을 때 더욱 많은 보행자 데이터를 주어도 focal length 추정 에러가 반등한다는 것을 알 수 있었다. 이는 기존 베이스라인 코드가 특정 노이즈와 선분 개수에 대해서 강건성 측면 한계를 가지고 있다고 판단된다. 추후 RANSAC(RANdom Sampling Analysis)를 적용하여 이러한 문제점을 해결할 수 있는 연구를 진행하고자 한다.



[Fig. 2] Correlation between the number of line segments and the relative error in focal length estimation across different Gaussian noise levels, highlighting a marked error increase beyond a 5-pixel noise standard deviation.

References

- [1] Z. Zhang, "A flexible new technique for camera calibration", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no.11, 2000
- [2] Gaku Nakano, "Camera Calibration using Parallel Line Segments", International Conference on Pattern Recognition (ICPR), 2021