

3차원 체스보드 패턴을 이용한 효율적인 카메라 캘리브레이션 방법

고은상, 호요성
광주과학기술원 정보통신공학부
esko@gist.ac.kr; hoyo@gist.ac.kr

Efficient Camera Calibration Method Using 3D Chessboard Pattern

Eunsang Ko and Yo-Sung Ho
School of Information and Communications
Gwangju Institute of Science and Technology (GIST)

요약

카메라 캘리브레이션은 3차원 영상처리를 수행하기 전에 필수적으로 진행해야 하는 전처리 과정이다. 현재까지 카메라 캘리브레이션 방법에 대해 많이 연구되어 왔으며, 지금은 2차원 체스보드를 촬영하고 패턴의 꼭짓점 4개의 좌표를 수동으로 입력하는 카메라 캘리브레이션을 수행하는 방법이 가장 많이 사용되고 있다. 하지만 정확도를 향상을 위해, 여러 장 촬영된 모든 체스보드 패턴에 대해 꼭짓점 좌표를 입력해야 하므로 번거로운 작업이 필요하다. 본 논문에서는 3차원 체스보드 패턴을 이용하여 보다 빠르고 정확하게 카메라 캘리브레이션을 진행할 수 있는 방법을 소개한다. 동일한 환경에서 다른 카메라 캘리브레이션 방법과 재투영 오차의 평균을 비교했을 때 0.24 픽셀 감소하였고 카메라 캘리브레이션을 수행하는 시간도 크게 줄일 수 있었다.

1. 서론

사람이 보는 세상은 3차원이지만, 카메라로 영상을 촬영하면 3차원의 점들이 2차원 이미지 상에 맺히게 된다. 이때, 사용된 카메라 렌즈와 카메라 이미지 센서까지의 거리 등의 카메라 내부적인 값에 영향을 받는다. 따라서 이러한 내부적인 값과 3차원 공간에서 카메라의 위치인 외부적인 값을 알아야 촬영한 영상을 가지고 3차원 영상처리에 활용할 수 있다. 예를 들어, 가상 시점의 카메라 파라미터를 구하여 가상 시점의 영상을 보간(interpolation)하거나 다른 카메라 시점으로 카메라 파라미터를 이용해서 3차원 워핑(warping)을 수행할 수 있다. 이 때 카메라 파라미터는 영상의 결과에 큰 영향을 준다 [1].

카메라 캘리브레이션(calibration)은 카메라의 내부 파라미터(intrinsic parameter)와 외부 파라미터(extrinsic parameter)를 구하는 과정이다. 내부 파라미터는 렌즈 중심에서 이미지 센서까지의 거리 값인 초점거리(focal length)와 카메라 렌즈의 광학축이 이미지 센서와 만나는 점의 영상 좌표를 가리키는 주점(principal point), 이미지 센서의 기울기를 타나내는 비대칭계수(skew coefficient) 값으로 구성된다. 반면에 외부 파라미터는 3차원 공간에서 카메라의 회전(rotation) 행렬과 평행이동(translation) 벡터로 구성된다. 렌즈의 줌 값이나 초점거리가 달라지면 내부 파라미터 값에 영향을 받고, 카메라의 위치가 달라지면 외부 파라미터 값이 달라지므로 렌즈 값에 변화를 줄 수 있는 모든 자동 옵션을 해제하고 카메라가 움직이지 않도록 고정해야 한다 [2, 3].

한편, 최근에 휴대 가능한 카메라나 노트북을 이용하여 3차원 영상처리를 활용하는 사례가 증가하고 있다. 하지만 휴대할 수 있는 만큼 카메라 파라미터가 달라지기 더 쉽기 때문에, 부득이하게 영상 처리를 수행하기 전에 카메라 캘리브레이션을 매번 수행해야 한다. 지금까지 카메라 캘리브레이션 방법에 대해 많이 연구되어 왔고, 카메라 캘리브레이션을 수행하기 위한 프로그램은 많이 있지만, 현재는 대부분의 사람들이 Caltech에서 MATLAB으로 구현한 Zhang 방법을 사용하고 있다 [4]. 이 방법은 MATLAB을 통해 실행해야 하고, 여러 장 촬영된 모든 체스보드 패턴에 대해 4개의 꼭짓점 좌표를 입력해야 한다. 따라서 다시 점 카메라에 대해 캘리브레이션을 수행할 경우 많은 시간과 번거로운 작업이 요구된다.

본 논문에서는 3차원 체스보드 패턴을 이용하여 카메라 캘리브레이션을 수행할 수 있는 방법을 소개한다. 또한 카메라 캘리브레이션의 정확도를 측정하기 위해, 기존의 Zhang 방법과 비교하여 재투영 오차의 평균을 비교하고 분석했다.

2. 카메라 캘리브레이션 방법

2.1 특징점 검출 및 3차원 좌표와 2차원 좌표 맵핑

카메라 캘리브레이션을 수행하기 위한 첫 단계는 사람 눈에 보이는 3차원 점과 2차원 이미지 상에 맺힌 영상의 2차원 좌표를 맵핑(mapping)하는 것이다. 따라서 이미지 상의 한 점을 3차원 원점으로 설정하게 되는데 3차원 좌표를 구분하기 위해 카메라 캘리브레이

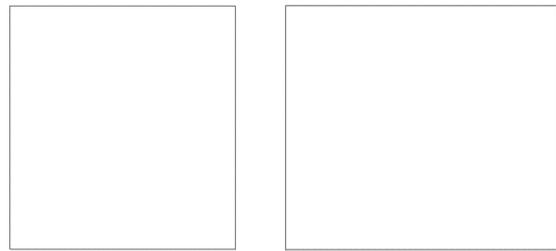
선 방법의 대부분은 그림 1(a)의 체스보드 패턴을 사용한다. 다른 패턴으로는 그림 1(b)과 그림 1(c)과 같이 교차하지 않는 사각형이나 원형 패턴, 본 논문에서 소개하는 그림 1(d)인 3차원 체스보드 패턴이 있으며, 사용하는 패턴에 따라 특징점 검출 방법과 카메라 캘리브레이션의 수행 방법도 약간 달라진다.



(그림 1) 여러 가지 캘리브레이션 패턴

그림 1(a) 그림 1(b), 그림 1(c)의 패턴을 사용하여 카메라 캘리브레이션을 수행할 때는 다중평면(multiplane) 방법을 이용한다. 다중평면 방법은 동차 선형 방정식(homogeneous linear system)을 기반으로 같은 평면에서 여러 시점 사이의 관계를 찾는 방법이다 [5]. 따라서 기존 방법에서는 최소 20장의 패턴에 대해 수행하는 것을 권장하며, 각 패턴에 대해 가장 처음에 입력한 좌표가 3차원 공간의 원점이 된다. 이어서 입력한 3개의 좌표로 큰 사각형을 만들고, 큰 사각형 안의 사각형 개수를 입력받아 모든 특징점을 찾는다. 다음으로는 사각형의 실제 크기를 가지고 그림 2(a)와 같이 특징점에 대해 3차원 좌표와 2차원 좌표의 맵핑 쌍을 구할 수 있다. 사용한 패턴은 2차원이므로 맵핑한 3차원 좌표의 Z 값은 항상 0이며 사각형 길이에 따라 3차원 좌표의 X 와 Y 값만 증가한다. 예를 들면, 그림 2의 작은 사각형 크기가 30 mm 일 때, 맨 왼쪽 위 특징점의 3차원 좌표는 (0, 150, 0)이고, 맨 오른쪽 아래는 (150, 0, 0)이다.

반면에, 그림 1(d) 패턴을 사용할 때는 직접 선형 변환(direct linear transform) 방법을 이용한다. 다중평면 방법과 다른 점은 체스보드 3개의 평면이 한 점에서 만나는 점을 3차원 공간의 원점으로 설정하며 각 특징점에 대해 3차원 좌표의 Z 값이 달라진다. 예를 들어, 맨 왼쪽 위 특징점의 3차원 좌표는 (150, 0, 150)이고, 맨 아래는 (150, 150, 0)이다.



(a) 다중평면 방법 (b) 직접 선형 변환 방법
(그림 2) 패턴에서 3차원 좌표와 2차원 좌표의 맵핑 쌍 획득

2.2 카메라 프로젝션 행렬 획득

카메라 프로젝션 행렬은 3행 4열로 이루어지며 핀홀 카메라 모델을 표현한다. 식 (1)과 같이 카메라 프로젝션 행렬에 3차원 좌표를 동차 좌표(homogeneous coordinates)로 만들어 곱하면 2차원 이미지 상에 투영된 2차원 좌표를 얻을 수 있다. 역으로 2.1 절에서 3차원 좌표와 2차원 좌표를 맵핑한 쌍을 알고 있으면 카메라 프로젝션 행렬을 획득할 수 있다. 식 (2)는 최소 6개의 3차원 좌표와 2차원 좌표 맵핑 쌍을 가지고 카메라 프로젝션 행렬을 얻기 위해 G 행렬 만드는 방법을 보여준다 [6].

$$s \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = A[R|t] \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (1)$$

$$\begin{pmatrix} X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & u_i X_i & u_i Y_i & u_i Z_i & u_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & v_i X_i & v_i Y_i & v_i Z_i & v_i \end{pmatrix} p = 0 \quad (2)$$

where $p = (p_{11} p_{12} \dots p_{34})^T$ and $0 = (00)^T$

$$\min_p \|Gp\|^2 \text{ subject to } \|p\| = 1 \quad (3)$$

식 (2)에서 i 는 3차원 좌표와 2차원 좌표 맵핑 쌍의 인덱스, X, Y, Z 는 3차원 좌표, u, v 는 각각 2차원 좌표를 의미한다. 식 (2)와 같이 G 행렬을 만든 다음에는, 식 (3)처럼 $\|p\|$ 가 1이 되도록 전제로 하여 $\|Gp\|$ 를 최소화 하면 된다. 쉽게 말해, p 는 가장 작은 특이 값에 해당하는 단위 특이벡터가 된다. 따라서 p 는 G행렬을 특이값 분해(singular value decomposition)하여 $G = UDV^T$ 에서 V 행렬의 마지막 열의 값들이다. 특이값 분해를 하면 G가 $m \times n$ 행렬일 때 $m \times m$ 직교 행렬(orthogonal matrix) U와 $m \times n$ 대각 행렬(diagonal matrix) D, $n \times n$ 직교 행렬인 V^T 를 얻게 된다. 특이값 분해로 구한 p 를 3행 4열에 맞게 배열하면 카메라 프로젝션 행렬이 된다.

2.3 카메라 내부 파라미터와 외부 파라미터 획득

이전 과정에서 구한 카메라 프로젝션 행렬을 가지고 카메라 캘리브레이션의 최종 목적인 카메라 내부 파라미터와 외부 파라미터를 획득할 수 있다. 프로젝션 행렬에서

3×3 부분 행렬은 카메라 내부 파라미터와 카메라 외부 파라미터의 회전 행렬이 곱해진 행렬과 같으므로 3×3 부분 행렬에 대해 RQ 행렬 분해(RQ decomposition)를 수행한다. RQ 행렬 분해는 QR 행렬 분해와 순서만 다르고, 직교 행렬과 상삼각 행렬(right-triangular matrix)로 분해하는 것은 같다. 따라서 R 행렬이 카메라의 내부 파라미터고, Q 행렬이 카메라 외부 파라미터의 회전 행렬이 된다. 이렇게 구한 카메라 내부 파라미터에서 초점거리와 주점을 확인하기 위해서는 R₃₃ 값을 1로 맞춰주는 추가적인 과정도 필요하다.

아직 구하지 못한 카메라 외부 파라미터의 평행이동 벡터는 카메라 프로젝션 행렬을 다시 특이값 분해하여 카메라 중심점을 구하는 과정이 필요하다. 여기서 말하는 카메라 중심점은 3차원 공간으로 표현되며 카메라 내부 파라미터의 주점과는 전혀 다른 개념으로, 식 (4)와 같이 카메라 프로젝션 행렬의 널 벡터를 구하면 된다. 카메라 프로젝션 행렬의 널 벡터는 카메라 프로젝션 행렬을 특이값 분해하여 구할 수 있다. 따라서 2.2절에서와 같이 $P=UDV^T$ 에서 V 행렬의 마지막 열의 값들이 카메라 중심점이 된다. 카메라 중심점을 구했다면 식 (5)를 이용하여 카메라 외부 파라미터의 회전 행렬에 곱하여 평행이동 벡터를 구할 수 있다.

$$P\tilde{C}=0 \text{ where } \tilde{C}=(XYZ1)^T \quad (4)$$

$$\begin{aligned} \tilde{X}_{\text{cam}} &= R(\tilde{X}-\tilde{C})=R\tilde{X}+t \\ t &=-RC \end{aligned} \quad (5)$$

2.4 카메라 파라미터와 왜곡 보정

지금까지 카메라 프로젝션 행렬을 구하고 특이값 분해하여 카메라 파라미터를 얻을 수 있었다. 하지만 카메라 파라미터의 정확도를 향상시키기 위해 기존 방법과 같이 여러 장의 패턴에 대해 카메라 캘리브레이션을 수행한다. 첫 번째 패턴으로 수행한 카메라 캘리브레이션은 이전 과정과 같이 수행하고, 다른 각도로 촬영한 두 번째 이후 패턴으로 카메라 캘리브레이션을 수행할 때는 식 (6)과 같이 이전 패턴에 대해 얻은 카메라 프로젝션 행렬을 사용하여 점차 오차를 줄여나가는 방법으로, 다른 각도로 촬영한 패턴이 많을수록 정확도를 향상시킬 수 있다.

다른 방법은 카메라 렌즈 왜곡을 고려하여 카메라 프로젝션 행렬을 보다 정확하게 획득하는 방법이다. 카메라 렌즈 왜곡은 빛을 모아주는 렌즈 특성 때문에 발생하는 현상으로 방사형(radial) 왜곡과 탈중심(decentering) 왜곡이 있다. 카메라 프로젝션 행렬을 구하는 방법 편환 카메라 모델에서 유도된 방법이므로 식 (7)과 같이 왜곡 계수 k₁과 k₂를 고려하여 2차원 이미지 좌표와 투영되는 좌표의 거리를 최소로 하여 카메라 파라미터의 정확도를 향상시킬 수 있다.

$$\min_p \sum_i \|m_i - \Phi(P, M_i)\|^2 \quad (6)$$

$$\min_{A,R,t,k_1,k_2} \sum_i \|m_i - \tilde{m}(A,R,t,k_1,k_2, M_i)\|^2 \quad (7)$$

3. 실험결과

3차원 체스보드 패턴을 이용한 카메라 캘리브레이션 정확도를 기존 방법과 비교하기 위하여 Kinect v2 카메라를 사용했다. 약 80-100 cm 거리에서 직접 패턴을 촬영하고 각 방법으로 카메라 캘리브레이션을 수행하여 재투영 오차의 평균값을 비교했다. Kinect v2는 카메라 렌즈와 그 위치가 고정되어 있어 내부 파라미터가 변경되지 않으므로 제공되는 SDK를 통해 색상 카메라와 깊이 카메라 각각의 내부 파라미터 지상검증자료(ground truth)를 획득하고 비교했다. 표 1은 각각 색상 카메라와 깊이 카메라 내부 파라미터를 보여주며, 깊이 카메라의 경우에는 왜곡 계수까지 제공된다. 패턴을 촬영할 때는 1920×1080 해상도의 색상 카메라와 512×424 해상도의 적외선(infrared) 카메라를 사용했다. 3차원 체스보드 패턴을 이용할 때는 3차원 좌표와 2차원 좌표의 맵핑 쌍을 그림 2(b)와 같이 7개를 입력했고, 기존 방법은 그림 2(a)와 같이 36개의 특징점을 사용하면서 다른 각도 촬영한 15개의 패턴을 가지고 카메라 캘리브레이션을 수행했다. 두 패턴 모두 작은 사각형의 크기는 30 mm이다.

(표 1) Kinect v2 색상 카메라와 깊이 카메라의 지상검증자료

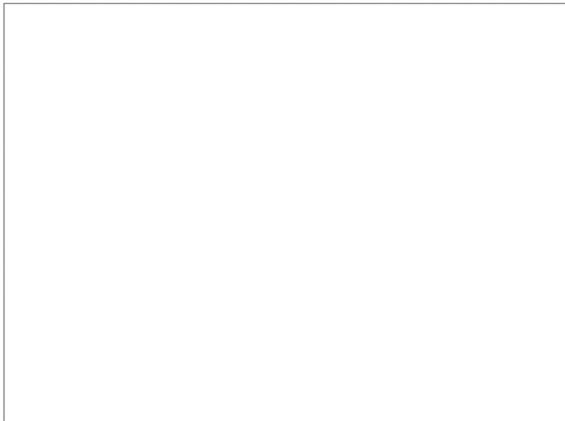
내부 파라미터	색상 카메라	깊이 카메라
<i>f_x</i>	1081.37207	364.542786
<i>f_y</i>	1081.37207	364.542786
<i>c_x</i>	959.5	261.347290
<i>c_y</i>	539.5	209.129196
<i>k₁</i>	-	0.0906964317
<i>k₂</i>	-	-0.266456813
<i>k₃</i>	-	0.0926424488

(표 2) 3차원 체스보드 패턴의 맵핑 쌍과 재투영 오차

맵핑 쌍	2차원 실제 좌표	2차원 투영 좌표	재투영 오차
1	(1223, 565)	(1223.882, 564.678)	0.881606
2	(1097, 623)	(1096.119, 622.842)	0.800810
3	(1097, 412)	(1097.178, 412.541)	0.324855
4	(1236, 378)	(1235.726, 378.031)	0.075856
5	(1435, 414)	(1435.139, 413.411)	0.365769
6	(1404, 626)	(1403.143, 626.874)	1.498492
7	(1281, 707)	(1281.852, 706.601)	0.885659

3차원 체스보드를 이용한 캘리브레이션 방법은 아직 여러 장의 패턴에 대해 카메라 캘리브레이션을 수행하는 방

법을 구현하지 못했다. 표 2는 입력한 7개의 3차원 좌표와 2차원 좌표의 맵핑 쌍과 재투영 오차 값이다. 기존 방법을 이용한 캘리브레이션 방법은 여러 장의 패턴을 사용하면 3차원 좌표와 2차원 좌표의 맵핑 쌍 개수도 다르기 때문에 그림 3의 재투영 오차 분포를 보고 카메라 파라미터의 정확도를 판단할 수 있다. 두 방법을 비교하기 위해 재투영 오차의 평균을 비교한 결과는, 15장의 패턴을 사용한 기존 방법이 3.3784, 1장의 패턴을 사용한 기존 방법이 0.9364이며, 3차원 체스보드 패턴을 사용한 방법이 0.6904로, 두 방법으로 1장의 패턴을 사용할 때의 오차는 3차원 체스보드 패턴을 사용한 방법이 0.24 픽셀 감소하였다.



(그림 3) 기존 방법을 이용한 카메라 캘리브레이션의 재투영 오차 분포

표 3은 위에서 실험한 세 가지 방법으로 획득한 Kinect v2 색상 카메라 내부 파라미터 값이다. 해당 내부 파라미터를 표 1인 지상검증자료와 비교하여 3차원 체스보드 패턴을 이용한 카메라 캘리브레이션 방법이 기존 방법보다 정확도가 얼마나 높은지, 이전에 획득한 카메라 프로젝션 행렬을 다른 각도로 촬영한 패턴에 대해 누적하여 사용하고 렌즈 왜곡 보정을 하면 정확도가 얼마나 향상되는지 비교할 수 있다. 실험 결과, 1장의 패턴을 사용한 기존 방법에서 가장 정확하게 주점을 획득했고, 초점 거리는 15장의 패턴을 사용한 기존 방법이 지상검증자료와 가장 근접한 값을 얻었다.

(표 3) 획득한 Kinect v2 색상 카메라의 내부 파라미터

내부 파라미터	기존 방법(15장)	기존 방법(1장)	3차원 체스보드
f_x	1065.10804	804.47058	919.066
f_y	1063.04329	802.95221	922.285
α	973.13102	959.5	1023.42
c_y	506.08102	539.5	558.288

표 4는 위에서 실험한 세 가지 방법으로 획득한 Kinect v2 깊이 카메라 내부 파라미터 값이다. 표 1의 지상검증자료에서 왜곡 계수도 비교할 수 있어 보다 자세하게 카

메라 캘리브레이션의 정확도를 비교할 수 있다. 실험 결과, 15장의 패턴을 사용한 기존 방법이 전체적으로 근사한 내부 파라미터 값을 획득했고, 3차원 체스보드 패턴을 사용한 방법은 1장의 패턴을 사용한 기존 방법보다 더 근사한 주점 값을 획득했다.

(표 4) 획득한 Kinect v2 깊이 카메라의 내부 파라미터

내부 파라미터	기존 방법(15장)	기존 방법(1장)	3차원 체스보드
f_x	370.04937	338.5	329.281
f_y	368.87198	339.5	322.419
α	263.41621	255.5	262.234
c_y	190.92241	211.5	209.643
k_1	0.10892	0.10913	-
k_2	-0.07498	-0.13236	-
k_3	-0.00412	0.00185	-

4. 결론

본 논문에서는 3차원 체스보드 패턴을 이용한 카메라 캘리브레이션 방법을 알아보고 기존 방법의 차이와 정확도를 비교했다. 3차원 체스보드 패턴을 이용한 카메라 캘리브레이션 방법은 3차원 좌표에서 Z 값이 달라지므로 보다 정확한 카메라 프로젝션 행렬을 구할 수 있었다. 기존 방법은 카메라 렌즈 왜곡을 보정하고, 다른 각도로 촬영한 패턴에서 카메라 프로젝션 행렬을 획득할 때 이전 정확도를 향상 시켰다. 3차원 체스보드 패턴을 이용하는 방법도 기존 방법의 카메라 파라미터와 왜곡 보정을 적용하면 더 향상된 정확도의 카메라 파라미터를 보다 쉽고 빠르게 획득할 수 있다.

감사의 글

본 연구는 미래창조과학부 ‘범부처 Giga KOREA 사업’의 일환으로 수행하였음. [GK13C0100, 기가급 대용량 양방향 실감 콘텐즈 기술 개발]

참고 문헌

- [1] 이상범, 이천, 호요성, “3차원 영상 생성을 위한 깊이 맵 추정 및 중간시점 영상합성 방법,” 한국통신학회 논문지, 제 34권, 제 10호, pp. 1070-1075, 2009.
- [2] Z. Zhang, “A Flexible New Technique for Camera Calibration,” IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 22, No. 11, pp. 1330-1334, 2000.
- [3] <http://darkpgmr.tistory.com/32>
- [4] http://www.vision.caltech.edu/bouguetj/calib_doc
- [5] http://en.wikipedia.org/wiki/Chessboard_detection
- [6] Z. Zhang, “Camera Calibration”, Chapter 2, pages 4-43, in G. Medioni and S.B. Kang, eds., Emerging Topics in Computer Vision, Prentice Hall Professional Technical Reference, 2004.