

# 움직이는 사람 추적을 통한 카메라 캘리브레이션 기법 조사

최정원, 권동우, 지영민  
한국전자기술연구원

{jungwon, dwkwon, ym.ji}@keti.re.kr

## A survey on camera calibration from tracking of moving persons

Jungwon Choi, Dongwoo Kwon and Youngmin Ji  
Korea Electronics Technology Institute (KETI)

### 요약

3 차원 공간의 점들이 카메라를 통해 2 차원 평면으로 투영될 때 영향을 주는 카메라의 내부 파라미터와 외부 파라미터를 계산하는 과정을 카메라 캘리브레이션이라고 한다. 카메라 캘리브레이션에는 다양한 기법이 있는데 본 논문에서는 보정의 기준이 되는 물체가 없는 상태에서 움직이는 사람을 추적하고 사람의 머리 및 발의 위치를 이용하여 파라미터를 추정하는 카메라 캘리브레이션 기법에 대해 살펴보고자 한다.

### I. 서론

카메라를 통해 3 차원 공간이 2 차원 평면으로 투영될 때에는 카메라의 방향과 위치, 카메라의 초점거리나 렌즈의 왜곡계수 등에 의하여 영향을 받는다. 카메라 캘리브레이션은 이러한 투영 과정에서 영향을 주는 카메라의 파라미터들을 추정하는 과정이다. 이를 통해 영상에서 발생하는 왜곡과 오차를 보정할 수 있다. 앞서 말한 카메라의 방향과 위치는 외부 파라미터에 해당하고, 카메라의 초점거리와 렌즈의 왜곡계수는 카메라의 내부 파라미터에 해당한다. 카메라 캘리브레이션은 크게 두가지 범주로 분류할 수 있는데 첫번째는 체커보드와 같은 보정의 기준이 되는 물체가 있는 상황에서 해당 물체에 대한 정보는 이용하는 방법이고, 두번째는 카메라 장면에 대한 사전 정보가 없는 카메라 셀프 캘리브레이션이다. 본 논문에서는 두번째 범주에 해당하는, 움직이는 사람을 추적하여 머리와 발의 위치를 활용하는 카메라 캘리브레이션에 대해 살펴본다.

### II. 본론

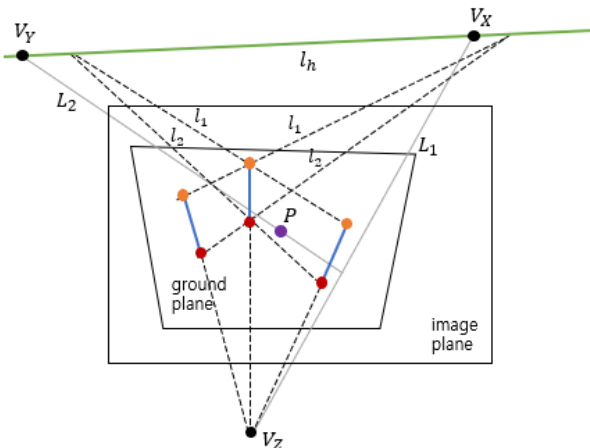
#### 2.1 소실점

3 차원 공간에서 평행한 직선들이 영상에 투영되어 원근효과에 의해 2 차원 평면에서는 한 점에서 만나는 것처럼 보일 때, 그 점을 소실점이라고 한다. 카메라 파라미터를 추정할 때 사용될 소실점들을 2 차원 이미지에서 나타낸 것이 <그림 1>이다. 파란색 선은 사람을 나타내고 머리는 주황색 점, 발은 빨간색 점으로 표시되어 있다. <그림 1>의 소실점  $V_Z(u_z, v_z)$  를 찾기 위해서 움직이는 사람은 지면과 수직이라고 가정하고, 사람의 머리와 발을 잇는 직선들의 교점이기 때문에 최소한 2 개 이상의 움직이는 사람이 특정되어야 한다. 임의의 한 사람의 머리에 해당하는 점을  $(u_1, v_1)$  이라고 하고 발에 해당하는 점을  $(u_2, v_2)$  라 하면  $(u_1, v_1)$ ,  $(u_2, v_2)$ ,  $(u_z, v_z)$  는 한 직선 위에 있기 때문에 기울기를 이용하면 다음과 같은 식이 나온다.

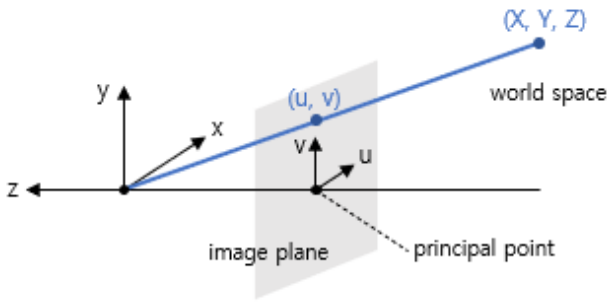
$$u_z(v_1 - v_2) + v_z(u_2 - u_1) = u_2v_1 - u_1v_2 \quad (1)$$

수식 (1)의 미지수는 2 개이기 때문에 미지수를 구하기 위한  $Ax = b$  를 구성하기 위해서는 최소한 2 개의 식이 필요하다. 총 N 개의 식이 있을 때 A는 Nx2 행렬, x는  $[u_z, v_z]^T$  이고 b는 Nx1 벡터이다.  $x = (A^T A)^{-1} A^T b$  에서 최소 제곱법을 이용하여  $\|Ax - b\|$  의 값을 최소화 하는 x를 구할 수 있다.[3] 2 개의 움직이는 물체만 이용해도 소실점  $V_Z$ 를 구할 수 있지만 더 확실한 계산을 위해서는 2 개보다 많은 물체가 필요하다.

<그림 1>의 소실선  $l_h$ 는  $l_1$ 과  $l_2$ 의 교점들로 이루어 지는데  $l_1$ 은 머리에 해당하는 점들을 이은 직선이고  $l_2$ 는 발에 해당하는 점들을 이은 직선이다. 소실선  $l_h$ 를 구하기 위해 사용되는  $l_2$ 는 2 차원 이미지의 각 열의 발의 위치 중 최상단과 최하단에 해당하는 픽셀을 지정하고 각도가 15~65 도 사이가 되도록 최상단



<그림 1> 소실점과 소실선의 기하학적 구조



<그림 2> 3 차원 공간에서 2 차원 평면으로의 투영

픽셀들과 최하단 픽셀들을 잇는 직선을 생성하여 구할 수 있다. 그리고 해당 발 위치에 대응하는 머리의 위치를 잇는 직선이  $l_1$  이 된다. 이렇게 구한  $l_1$  과  $l_2$  를 이용하여 최소 제곱 오차 선형 회귀로  $l_h$  를 추정할 수 있다.[2]

소실점  $V_x$  는 소실선  $l_h$  위의 임의의 점으로 선택한다.  $V_x$  와  $V_z$  를 지나는 직선을  $L_1$  이라 하고 이미지 평면의 주점인 P 를 지나면서  $L_1$  과 수직인 직선을  $L_2$  라 하면 주점 P 는  $V_x, V_y, V_z$  로 이루어진 삼각형의 수심이라는 성질을 이용하여  $L_2$  와  $l_h$  의 교점으로 소실점  $V_y$  를 구할 수 있다.[1]

## 2.2 카메라 캘리브레이션

소실점들을 구한 다음에는 이를 이용하여 카메라 파라미터를 추정한다. <그림 2>에서 3 차원 공간(world space)의 점  $(X, Y, Z)$  가 2 차원 평면(image plane)의 점  $(u, v)$  로 투영될 때 사용되는 투영 행렬을  $P$  라고 하면,  $[u, v, 1]^T \sim P \cdot [X, Y, Z, 1]^T$  이 성립한다. 이때 행렬  $P$  는 5 개의 카메라 내부 파라미터로 이루어진 행렬  $K$ , 외부 파라미터 3 개로 이루어진 행렬  $R$ , 또 다른 외부 파라미터 3 개로 이루어진 행렬  $t$  로 나타낼 수 있다.

$$P = K \cdot [R|t]$$

$$\text{where } K = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, t = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}, R = R_z \cdot R_x \cdot R_y,$$

$$R_z = \begin{bmatrix} \cos(\text{roll}) & -\sin(\text{roll}) & 0 \\ \sin(\text{roll}) & \cos(\text{roll}) & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\text{pitch}) & -\sin(\text{pitch}) \\ 0 & \sin(\text{pitch}) & \cos(\text{pitch}) \end{bmatrix},$$

$$R_y = \begin{bmatrix} \cos(\text{yaw}) & 0 & -\sin(\text{yaw}) \\ 0 & 1 & 0 \\ \sin(\text{yaw}) & 0 & \cos(\text{yaw}) \end{bmatrix} \quad (2)$$

카메라 파라미터를 구할 때는 수식 (2)의 관계를 이용하고 EDA(Estimation of Distribution Algorithm)를 활용하여 최적 파라미터를 찾아낸다.[1] 이때  $f_x$  와  $f_y$  는 x 축, y 축 방향으로의 초점 거리(focal length),  $(c_x, c_y)$  는 주점(principal point)의 좌표,  $s$  는 왜곡(skew),  $pitch, yaw, roll$  은 각각 x 축, y 축, z 축 방향으로의 회전,  $t_x, t_y, t_z$  는 x 축, y 축, z 축 방향으로의 이동을 의미한다. 일반적으로 직사각형의 픽셀을 가정하기 때문에  $s$  는 0 으로 설정한다. 또한 원점  $(0, 0, 0)$  이 지상면과 카메라를 통과하는 수직선과의 교차점에 위치한다고 가정할 수 있기 때문에  $t_x$  와  $t_z$  도 0 으로 설정한다. 일부 카메라 파라미터의 초기값들은 앞에서 구한 소실점들을 이용하여 계산될 수 있다. EDA 를 이용한 카메라 파라미터의 최적화 과정은 다음과 같다.

1. 카메라 캘리브레이션에 사용될 파라미터들의 초기 범위를 지정하고 이 범위를 기반으로 초기 모집단 생성
2. 초기 모집단을 기반으로 probabilistic model 을 생성하고 이를 기반으로 후보 파라미터 조합을 샘플링
3. 각 후보 파라미터 조합을 사용하여 보정된 이미지를 생성하고 실제 이미지와 비교하여 reprojection error 를 계산
4. 평가된 후보 파라미터 조합을 기반으로 probabilistic model 업데이트
5. 업데이트된 probabilistic model 을 기반으로 새로운 후보 파라미터 조합 샘플링
6. 3~5 의 과정을 종료 조건이 충족될 때까지 반복하여 최적 파라미터 조합 추정

이러한 카메라 캘리브레이션 알고리즘은 입력 데이터로 2 개의 소실점  $V_x, V_y$  와 카메라 높이의 대략적인 범위만 필요로 한다.

## III. 결론

본 논문에서는 컴퓨터 비전 분야에서 꼭 필요한 과정인 카메라 캘리브레이션 중에서 움직이는 사람을 추적하여 머리와 발의 위치를 활용하는 기법에 대하여 조사하였다. 해당 기법은 보정의 기준이 되는 물체를 필요로 하지 않고 2 개의 소실점과 카메라 높이의 대략적인 범위만 필요로 하기 때문에 다른 많은 응용 프로그램으로 확장할 수 있을 것으로 기대된다.

## ACKNOWLEDGMENT

본 연구는 산업통상자원부(MOTIE)와

한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구

과제입니다. (No. 20202000000010)

## 참 고 문 헌

- [1] Z Tang, YS Lin, KH Lee, JN Hwang, JH Chuang and Z Fang. "Camera Self-Calibration from Tracking from Moving Persons," 2016 23<sup>rd</sup> International Conference on Pattern Recognition, 2016.
- [2] W Kusakunniran, H Li and J Zhang. "A Direct Method to Self-Calibrate a Surveillance Camera by Observing a Walking Pedestrian," 2009 Digital Image Computing: Techniques and Applications, 2009.
- [3] M Hodlmoser, B Micusik and M Kampel. "Camera Auto-Calibration Using Pedestrians and Zebra-Crossings," 2011 IEEE International Conference on Computer Vision Workshops, 2011.