

스테레오 영상 매칭을 활용한 기초행렬 추정

최현지, 이덕우*
 계명대학교 컴퓨터공학부
 tyg06229@naver.com, dwoolee@kmu.ac.kr

Estimation of Fundamental Matrix Using Stereo Image Matching

Hyunji Choi Deokwoo Lee &
 Department of Computer Engineering, Keimyung University

요약

본 논문에서는 스테레오 영상으로부터 얻은 특징점들을 활용하여 기초행렬(Fundamental matrix)을 추정하는 실험을 한다. 획득한 영상들은 보정이 되어 있으며, 특징점 추출 후 매칭은 RANSAC 등의 기존 알고리즘을 활용한다. 기초 행렬을 얻기 위해 스테레오 영상으로부터 정의되는 에피폴라 점, 에피폴라 선, 에피폴라 평면을 정의하고, 이들로부터 얻을 수 있는 기하학적 관계식을 활용하여 기초행렬을 수학적으로 추정해 보고, 실험으로 수학적 이론을 검증해 본다.

1. 서론

컴퓨터 비전(computer vision) 분야가 발달함에 따라 2D 이미지에서 3D 정보를 찾기 위한 연구는 지속해서 진행되어왔다 [1,2,3,4]. 그로 인해 생긴 시각센서 중 하나 스테레오 카메라이다. 스테레오 비전은 두 대 이상의 카메라를 사용하여 이미지의 시차를 이용하여 이미지의 깊이 정보를 획득하는 기법이다. 이것은 간단한 하드웨어 장치를 바탕으로 3D 정보를 추출할 수 있어 많은 분야에서 활용되는 방법이다.

2. 용어 정리

두 개의 카메라가 존재한다는 가정하에 두 카메라를 O 카메라, O' 카메라로 정의한다. 그리고 카메라 O와 O'로 보여지는 이미지는 각각 A 이미지와 B 이미지로 정의한다. 그리고 O 카메라와 O' 카메라는 3D 좌표계의 한 점 X를 각각 투영한다고 할 때, A 이미지에 투영된 X 점은 x로 표현되고, B 이미지에 투영된 X는 x'로 표현될 수 있다.

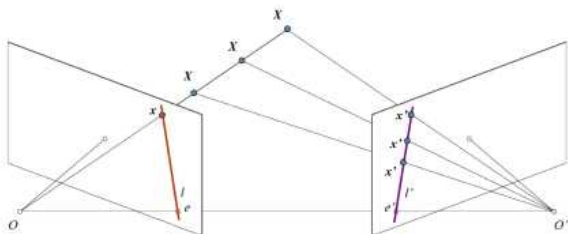


그림 1 출처: openCV 홈페이지(docs.opencv.org)

그림 1에서 확인할 수 있듯이 3D 좌표계의 한 점은 두 이미지에 각각 투영된다. 이 현상이 두 2D 이미지에서 깊이 정보를 획득하는 중요한 열쇠인데, 만약 하나의 카메라만 존재한다고 가정하면, 위의 x 픽셀은 3D 좌표계에서 깊이를 특정 지을 수 없게 된다.

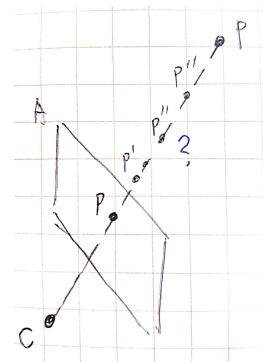


그림 2 한 대의 카메라

그 이유는 그림 2와 같이 이미지 A(그림 1에서는 O)에 투영되는 p가 3D 좌표계에서 p'인지 p''인지 특징지을 수 없기 때문이다. 이러한 이유로 스테레오 비전에서는 두 대의 카메라가 중요한 역할을 하게 된다.

이 스테레오 카메라가 만드는 에피폴라 기하(Epipolar Geometry)에서는 여러 용어가 사용되는데, 대표적인 용어가 에피폴(Epipole)이다. 에피폴은 두 카메라를 선으로 연결했을 때 이미지 면(image plane)에 생기는 교점을 뜻한다. 이때 이 점은 영상 안에 그려질 수 있고 또는 영상밖에 그려질 수 있다. 이때 두 카메라 사이에 연결된 선을 기준선

(baseline)이라고 하며, 기준선(baseline)을 포함하는 평면을 에피폴라 평면(Epipolar plane)이라고 한다. 각각의 카메라에 대한 이미지 면과 에피폴라 평면의 교차를 에피폴라 선(Epipolar line)이라고 한다. 이 에피폴라 선은 각각의 이미지 평면에 존재한다.

그림 1을 살펴보면 3D 좌표계의 X는 A 이미지와 B 이미지의 각각의 에피폴라 선 위에 투영된다. 그러므로 A 이미지의 x를 B 이미지에서 찾기 위해서는 B의 에피폴라 선을 탐색하면 된다. 이를 에피폴라 제약(Epipolar constraint)이라고 한다.

3. 기초행렬

그림 1의 x와 x'는 서로 연관될 수 있는데, x와 x'를 연관 지어 줄 수 있는 행렬이 기초 행렬(Fundamental matrix)이다. 우선 기초 행렬(Fundamental matrix)에 대해 알아보기 전 기본 행렬(Essential Matrix)에 대한 두 점의 관계를 설명한다.

$$p'^T E p = 0 \quad \text{수식 1}$$

$$[u' \ v' \ 1] E \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = 0 \quad \text{수식 2}$$

위의 수식이 관계가 만족하는 E 행렬이 항상 존재한다는 가정하에 E 행렬이 기본 행렬이다. 수식 1의 p'와 p는 3D 좌표계의 한 점이 정규화된 이미지에 투영된 각각의 점이다. 임의의 두 이미지 좌표 사이의 관계는 회전, 평행 이동으로 관계지을 수 있기에 기본 행렬은 다음과 같이 표현될 수 있다. 수식 3의 표현은 R로 회전시킨 다음 t와 외적의 과정을 나타내는 행렬이다.

$$E = [t]_x R \quad \text{수식 3}$$

기본 행렬은 정규화된 이미지 평면에서의 매칭 쌍 사이의 기하학적 관계를 나타내는 행렬이다. 즉, 카메라 파라미터에 대한 정보는 제외한 행렬인 것이다. 여기에 카메라 파라미터 정보까지 포함한 두 이미지의 실제 픽셀(pixel) 좌표 사이의 기하학적 관계를 나타낸 행렬이 기초 행렬(Fundamental matrix)이다.

$$p_{img}'^T F p_{img} = 0 \quad \text{수식 4}$$

$$[x' \ y' \ 1] F \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0 \quad \text{수식 5}$$

이미지 A에 대한 카메라 내부 파라미터 행렬이 K이고, 이미지 B에 대한 카메라 내부 파라미터 행렬 K'라고 할 때 이미지 A와 이미지 B 사이의 기초 행렬은 다음과 같다.

$$E = K'^T F K \quad \text{수식 6}$$

또한, 이미지 픽셀 좌표 p_{img} 와 정규 좌표계 좌표 p 사이의 카메라 파라미터 행렬 K에 대하여 다음의 수식이 성립한다.

$$p_{img} = K p \quad \text{수식 8}$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad \text{수식 9}$$

수식 6과 수식 8을 이용하면 수식 4를 증명해낼 수 있다.

$$\begin{aligned} p'^T E p &= 0 \\ (K'^{-1} p_{img}')^T E (K^{-1} p_{img}) &= 0 \\ p_{img}'^T (K'^T)^{-1} E (K^{-1}) p_{img} &= 0 \\ p_{img}'^T F p_{img} &= 0 \end{aligned}$$

이렇게 증명한 수식 4를 풀어서 다시 계산하면 AF=0으로 표현할 수 있다. 이때의 A와 F 행렬은 다음과 같이 표현된다.

$$\begin{bmatrix} x_m x'_m & x_m y'_m & \dots & x'_m y'_m & m & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ \dots \\ f_8 \\ f_9 \end{bmatrix} = 0 \quad \text{수식 14}$$

수식 14로 알 수 있듯이 기초행렬은 8개의 매칭 점으로 유추할 수 있다.

4. 실험 방법

위 실험은 Visual Studio 2019를 사용하여 C++ 언어로 openCV를 사용하여 진행하였다. 실험에서 사용할 이미지는 <https://vision.middlebury.edu/>에서 제공해주는 스테레오 데이터셋을 사용한다. 두 이미지의 매칭점을 찾기 위해 ORB를 사용하여 특징점을 검출하고 기술자를 계산한 다음 기술자를 바탕으로 서로 매칭한 다음 두 이미지 간의 거리(distance)가 가장 짧은 순으로 50개의 특징점을 뽑아 기초 행렬을 계산하는 입력 값으로 사용한다. 또한, openCV의 findFundamentalMat() 함수를 이용하여 8-point 알고리즘으로 기초 행렬을 계산한다.

4. 결과

기초 행렬을 구한 결과는 다음과 같다.

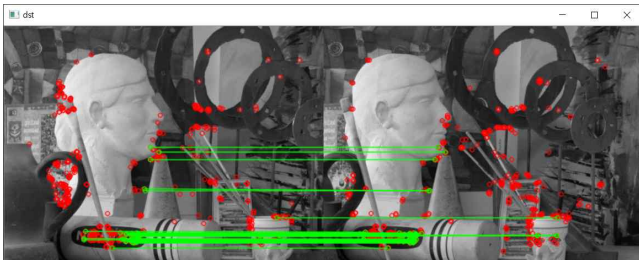


그림 3 첫 번째 스테레오 이미지 매칭 결과

```
C:\WINDOWS\system32\cmd.exe
fundamental
[9.386209201459813e-07, -9.421229954422635e-05,
0.02528589217933469;
8.570608410526948e-05, 6.297331052049493e-06,
0.05665912147936494;
-0.02263578266040026, -0.06268496659431902, 1]
```

그림 4 첫 번째 스테레오 이미지 기초행렬 추정

$$F = \begin{bmatrix} 0.00 & 0.00 & 0.03 \\ 0.00 & 0.00 & 0.00 \\ -0.02 & -0.06 & 1.00 \end{bmatrix} \quad \text{수식 15}$$



그림 5 두 번째 스테레오 이미지 매칭 결과

```
C:\WINDOWS\system32\cmd.exe
fundamental
[-3.122131601157945e-07, 0.0001705162378379057,
0.001996818134408293;
-0.00016341521242849, 2.071965296850581e-05,
-2.293256566347483;
-0.00605618769157721, 2.299065667090443, 0.9999999999999999]
```

그림 6 두 번째 스테레오 이미지 기초 행렬 추정

$$F = \begin{bmatrix} 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & -2.29 \\ -0.01 & 2.30 & 1.00 \end{bmatrix} \quad \text{수식 16}$$



그림 7 세 번째 스테레오 이미지 매칭 결과

```
C:\WINDOWS\system32\cmd.exe
fundamental
[4.784959024148408e-08, -0.0002262308130651086,
0.03090244055915631;
0.0002304821556423547, -2.221800430627928e-05,
0.2869769505224241;
-0.03143766564390958, -0.2914142119102205, 1]
```

그림 8 세 번째 스테레오 이미지 기초행렬 추정

$$F = \begin{bmatrix} 0.00 & 0.00 & 0.03 \\ 0.00 & 0.00 & 0.29 \\ -0.03 & -0.29 & 1.00 \end{bmatrix} \quad \text{수식 17}$$

대체로 우수한 성능으로 기초행렬이 추정된 것을 볼 수 있으나 가끔 매칭이 잘못된 값을 선택하여 기초행렬을 추정할 경우 기초행렬의 정상 범위에서 벗어나는 행렬이 생성된다.

위 실험을 바탕으로 두 이미지를 이용하여 기초행렬을 추정하는 실험을 진행하였다. 위 실험을 통해 기초행렬을 추정하면 이를 이용하여 기초행렬을 이용한 2D 이미지에서 깊이 정보를 얻어 낼 수 있을 것이다.

Acknowledgement

본 논문은 교육부와 한국교육재단의 계명대학교 대학혁신 지원사업비를 지원받아 수행된 연구의 결과입니다.

참고문헌

[1] R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, 2nd Ed., Cambridge University Press, Mar. 2004.

[2] 김영성, 임종우, 평면 위 직선과 호모그래피를 이용한 스테레오 카메라 교정. 한국정보과학회 2015년 동계학술발표논문집, 2015년 12월, 1143-1145.

[3] 윤용인, 옥형수, 최종수, 오정수, 평면 호모그래피를 이용한 3차원 재구성. 한국통신학회논문지, 제 31권 4C호, 2006, 381-390.

[4] J. Kaminski and A. Sashua, On Calibration and Reconstruction from Planar Curves, European Conference on Computer Vision (ECCV2000), 2003, pp. 678-694.