

카메라 회전을 이용한 이미지 센터 추정 방법

유원필, 정연구, 배영래
한국전자통신연구원,
컴퓨터.소프트웨어연구소, 휴먼정보처리부
e-mail : ywp@etri.re.kr, ykchung@etri.re.kr

Image Center Estimation Method Using Camera Rotation

Wonpil Yu, Yunkoo Chung, Youngrae Bae
Human Information Processing Dept.,
Computer & Software Research Laboratory, ETRI

요 약

본 논문에서는 카메라 파라메타의 하나로서 이미지 센터를 구하는 방법에 대해 소개한다. 3 차원 게임, 그래픽스, 의료, 측정, 형상 복원 등의 응용에 거리 정보를 포함하는 3 차원 정보는 필수적이며 이러한 정보를 추출하기 위해서는 카메라 캘리브레이션을 반드시 거쳐야 한다. 이미지 센터는 대부분의 경우에 카메라 파라메타를 구하기 위한 모델 추정 방법에서 하나의 값으로 취급되는 경우가 많아 이미지 센터 외의 카메라 파라메타의 정확도에 영향을 받을 수 있다. 본 연구에서는 카메라 회전을 이용하여 이미지 센터만을 독립적으로 구하는 방법을 제안하며 그 편이성과 반복성, 정확도를 카메라 회전 모델링, 시뮬레이션과 실험을 통해 보인다.

1. 서 론

최근의 영상 처리 및 컴퓨터 비전 응용 시장에 있어서 3 차원 정보를 이용한 모델링, 설계, 해석에 대한 필요성이 크게 증가하고 있다. 가상 현실, 게임, 의료 영상, 원격 측정, 3 차원 디스플레이, 문화, 예술 등의 산업 영역에서 표현 대상에 대한 3 차원 정보의 획득은 필수적인 사항이 되어가고 있다. 이 경우에 관심 대상체의 3 차원 정보를 구하기 위해서는 영상 입력에 사용되는 카메라의 파라메타를 먼저 얻어야만 한다.

전통적인 의미에서의 카메라 캘리브레이션은 기하학적인 정보가 알려져 있는 대상체의 3 차원 정보와 동일 물체의 2 차원 이미지 상에서의 좌표값 간의 관계로부터 카메라의 외부 변수(extrinsic parameters)와 내부 변수(intrinsic parameters)를 계산하는 과정을 일컫는다. 현재에도 많은 산업 현장에서 이러한 방식으로 캘리브레이션 작업을

수행하고 있으며, 이 경우에 캘리브레이션 대상 물체에 대한 정확한 3 차원 정보를 측정하기 위해 세심한 주의와 특히 캘리브레이션 대상 물체에 대한 정교성이 요구되는 등 작업 자체의 반복성 측면에서 적용에 불편함이 많은 것이 사실이다. 또한, 줌(zooming)이나 카메라 모션, 자동 포커싱 등 카메라의 능동적인 움직임에 의한 3 차원 정보 추출이나 로봇 네비게이션, 3 차원 형상 복원 등 자동화가 필요한 부분에 있어서 각 파라메타의 변화시 마다 캘리브레이션을 다시 수행하는 것은 매우 곤란한 일이다.

최근의 소위 “셀프-캘리브레이션”으로 불리는 방식은 이러한 수작업의 불편함을 없애기 위해 카메라 영상만을 이용하여 이미지상의 특징점을 추출하고 카메라 영상을 획득한 관점(viewpoint)간의 관계를 자동으로 추출하는 방법을 다룬다[3,4]. 그 중에서 특히 주목을 끄는 것은 카메라의 움직임을 이용한 것으로서 현재 팬/틸트(pan/tilt) 회전은 손쉽게 제어할 수 있으므로 서비스, 오락, 가정용 로봇

등에서 로봇 시각 센서의 기구적인 형태로 많이 채용되고 있다.

본 논문에서는 카메라의 팬/틸트 회전만을 이용하여 획득한 영상으로부터 카메라 파라미터를 구하는 방식을 취하며 [5]에서도 장점이 언급되었듯이 카메라의 개별적인 파라미터에 대한 독립적인 추정 방법을 구현하기 위한 시도로 이미지 센터만을 구하는 방법을 제안한다. 본 방식은 이미지 센터 이외의 다른 카메라 파라미터에 대한 정보가 필요없으며 캘리브레이션 대상 물체에 대한 제한이 없이 2 차원 포인트(point) 특징점만을 이용하는 장점을 가진다. 제한하는 방법은 모터에 의해 구동되는 팬/틸트가 가능한 김발(gimbal)에 놓여진 카메라 구조에서도 적용할 수 있지만 일반 카메라 삼각대만을 이용하더라도 적용 가능하다.

본 논문의 구성은 다음과 같다. 2 절에서 제안된 이미지 센터 추정 방법에 적용되는 카메라 이미징(imaging) 모델에 대해 소개하고 3 절에서 이미지 센터 추정 방법에 대한 소개와 제안한 방법의 성능 분석을 위한 시뮬레이션 결과에 대해 설명한다. 4 절에서 캘리브레이션 방법에 적용된 영상 처리 기법과 실험 결과에 대해 소개하며 5 절에서 본 논문에서 제안한 이미지 센터 추정 방법의 장점과 제한점에 대해 설명한다.

2. 카메라 이미징(imaging) 모델

본 절에서 소개하는 카메라 회전 모델은 제안하는 이미지 센터 추정 방법의 특성을 살펴보기 위한 시뮬레이션을 위해 도입되었으며 아울러 초점거리, scale factor 등 나머지 카메라 파라미터를 계산하는 수식을 유도하는 경우에도 활용할 수 있다. 카메라 모델은 그림 1 에 나타낸 것과 같이 회전축으로부터 거리 d 만큼 떨어진 곳(카메라 렌즈의 중심)에 카메라 좌표계가 설정되어 있고 θ_p 각도 회전하는 경우에 대해 설명한다. 이 때 회전축은 x_g 가 되며 틸트 회전인 경우에 y_g 축을 중심으로 회전하는 것을 제외하면 카메라 회전 모델링을 위한 수식 유도 과정은 두 경우에 동일하다. 카메라의 초기 위치 $\{x_{c1}, y_{c1}, z_{c1}\}$ 에서 θ_p 각도를 회전한 후의 $\{x_{c2}, y_{c2}, z_{c2}\}$ 좌표계간의 관계는 식(1)로 표시된 변이를 가지게 된다.

$$T = [\Delta T_x, \Delta T_y, \Delta T_z]^T = [0, d \sin \theta_p, d(1 - \cos \theta_p)]^T \quad (1)$$

$d = 0$ 이거나 카메라가 소실점(vanishing point)을 트래킹하는 경우에는 두 좌표계간의 관계는 단순한 회전으로 모델링할 수 있으며 이 때 3 차원 공간상의 포인트 특징점은 2 차원 이미지 평면상에서 conic section 의 일종인 hyperbola 궤적을 생성하게 된다[3]. 이 경우에 2-D 이미지 평면을 가우시안 구면위로 투영하면 hyperbola 궤적은 원을 그리게 되어 카메라의

자세---축, 회전각---를 손쉽게 구할 수 있다. 소실점이 아닌 경우 $d = 0$ 을 만족하도록 카메라를 회전하는 것은 특별한 구동 장치를 이용하지 않는 경우 일반적인 경우에 구현하기 매우 어렵다[6].

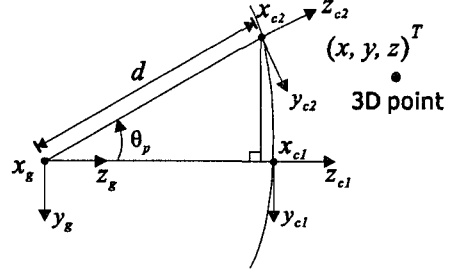


그림 1. Pan 모션시의 카메라 회전 모델링

3 차원 공간상에 위치한 포인트 특징점 θ_p 각도만큼 x_g 축을 중심으로 회전하는 경우 두 카메라 좌표계간의 관계는 다음과 같다.

$$X_{c2} = R_p(X_{c1} + T) \quad (2)$$

여기서 R_p 는 회전 매트릭스, X_{c1}, X_{c2} 는 각각 좌표계 $\{x_{c1}, y_{c1}, z_{c1}\}, \{x_{c2}, y_{c2}, z_{c2}\}$ 에서 표현된 3 차원 포인트 특징점의 좌표값이다.

식(2)에 카메라 projection 매트릭스를 곱한 후 전개하면 다음과 같다.

$$\begin{aligned} u_2 &= f_x \frac{x_c}{-(y_c + \Delta T_y) \sin \theta_p + (z_c + \Delta T_z) \cos \theta_p} + u_0, \\ v_2 &= f_y \frac{(y_c + \Delta T_y) \cos \theta_p + (z_c + \Delta T_z) \sin \theta_p}{-(y_c + \Delta T_y) \sin \theta_p + (z_c + \Delta T_z) \cos \theta_p} + v_0. \end{aligned} \quad (3)$$

여기서 (u_2, v_2) 는 $\{x_{c2}, y_{c2}, z_{c2}\}$ 위치에서 이미지 상에서의 포인트 특징점이 맺힌 좌표를 나타낸다. Tilt 모션인 경우에도 pan 모션에서와 같은 방식으로 식(4)와 같이 유도할 수 있다.

$$\begin{aligned} u_2 &= f_x \frac{(x_c + \Delta T_x) \cos \theta_t - (z_c + \Delta T_z) \sin \theta_t}{-(x_c + \Delta T_x) \sin \theta_t + (z_c + \Delta T_z) \cos \theta_t} + u_0, \\ v_2 &= f_y \frac{y_c}{-(x_c + \Delta T_x) \sin \theta_t + (z_c + \Delta T_z) \cos \theta_t} + v_0. \end{aligned} \quad (4)$$

위의 두 식에서 (u_0, v_0) 은 이미지 센터를 나타내는 양이다. 카메라 내부 변수에 해당하는 것은 (u_0, v_0) , (f_x, f_y) 로서 본 방법에서는 카메라 픽셀의 skew 파라미터는 1 로 가정하고 다루지 않도록 한다.

3. 이미지 센터 추정 방법 및 성능 분석

2 절에서 소개하였듯이 $d = 0$ 인 경우(카메라가 소실점을 트래킹하고 있는 경우)를 제외하면 2-D 이미지 상에서의 포인트 특징점의 궤적은 hyperbola 로 표현할 수 없으며 가우시안 구면 위로 투영하는 경우에도 원을 형성하지 않게 되어 카메라 자세를

해석하기가 복잡해진다. Daniilidis[5]는 d 가 0이 되는 상황을 실현하기 위하여 소실점을 이용하였지만 실제로 소실점을 안정적으로 구하는 것은 카메라의 위치와 3-D 상에서의 특징점간의 공간상의 관계가 매우 제약을 받게 되어 그렇게 효과적인 방법은 아니다. 특히 소실점의 경우에는 에지 성분을 먼저 추출해야 하며 이미지상에서 평행한 선분의 쌍을 찾아야 하는 등 전처리 과정이 복잡해 질 뿐만 아니라 에지 성분의 기술기에 대한 약간의 오차에 의해서도 결과치는 매우 민감하여 통계적인 데이터 처리 방법을 이용하여야 하는 경우가 많다.

제안된 방법의 아이디어는 실제로 d 가 0이 아니더라도 카메라의 회전에 의해 발생하는 궤적의 최저점은 소실점인 경우와 마찬가지로 이미지 평면상의 동일한 좌표 위치에서 발생한다는 사실에서 비롯되었다. 따라서 카메라의 회전에 의해 형성된 2-D 이미지상에서의 포인트 특징점의 궤적에 대한 함수 형태를 알고 있는 경우에는 최저점의 위치를 계산함으로써 이미지 센터를 손쉽게 구할 수 있는 것이다.

그림 1에 나타난 것과 같이 카메라가 pan 모션을 취함에 따라 hyperbola 궤적을 그리게 되고 이것은 d 가 0인 경우에 다음과 같은 형태를 취하게 된다[3].

$$-\frac{(v_p - v_0)^2}{f_y^2} + \frac{(u_p - u_0)^2}{f_x^2} \frac{1-g^2}{g^2} = 1. \quad (5)$$

이 때 $f_y, f_x, u_p, u_0, v_p, v_0, g$ 는 각각 v 방향, u 방향으로 환산된 초점거리, u 좌표값, u 방향으로의 이미지 센터 오프셋, v 좌표값, v 방향으로의 이미지 센터 오프셋, 3-D 포인트 특징점이 카메라 회전축과 이루는 방향 코사인값에 해당한다. 이 때, 이미지 평면의 u 축, v 축은 각각 카메라 좌표계에서 x_{ci}, y_{ci} 와 평행하다($i=1, \dots, n$). 위 식은 d 가 0인 경우에 유도된 결과식으로 이미지 센터를 구하기 위한 목적으로 식(6)과 같은 함수 형태로 변형이 가능하다.

$$u_p = a\sqrt{(v_p - v_0)^2 + b} + u_0. \quad (6)$$

2-D 이미지 좌표값으로부터 식(6)에 의해 구해지는 함수의 최저값이 발생하는 위치로부터 v_0 를 구할 수 있다. 마찬가지로 틸트 회전에 의해서 식(7)의 형태를 설정한 후 최저값 발생 위치로부터 u_0 를 얻을 수 있다.

$$v_p = a'\sqrt{(u_p - u_0)^2 + b'} + v_0. \quad (7)$$

식 (6),(7)에서 a, b, a', b' 는 식(5)와 틸트 회전시의 hyperbola 표현식을 변형하는 과정에서 f_x, f_y, g 항으로 구성되는 상수값에 해당한다.

실제 실험에서는 픽셀 잡음에 의해 정확한 좌표값이 얻어지지 않으므로 Levenberg-Marquardt 최적화 기법을 적용하여 2-D 이미지 궤적에 대한 curve fitting 과정을 거쳐 이미지 센터값을 구하였다.

2 절에서 소개한 카메라 회전 모델을 적용하고 실제 실험에 사용된 CCD 카메라의 데이터 쉬트, 렌즈 규격 등을 이용하여 시뮬레이션을 하여 제안된 방법의 성능에 대하여 살펴보기로 한다. 시뮬레이션에 사용된 파라메타는 다음과 같다.

- 카메라 회전각: $-20^\circ \sim +20^\circ, 2^\circ$ 간격
- CCD: Pulnix TMC-7DSP
- Scale factor:

$$k_v = \left(\frac{768}{640} \times 8.4 \times 10^{-6} \right)^{-1}, \quad k_u = (9.8 \times 10^{-6})^{-1},$$

여기서 k_v, k_u 는 각각 v, u 축 방향으로의 scale factor에 해당한다.

- 이미지 크기: 640×480 (Matrox Meteor-II)
- 회전축으로부터 렌즈 중심까지의 거리: 60 mm(직접 측정 결과)

시뮬레이션으로 확인한 것은 카메라 회전각이 작은 경우 픽셀 잡음 수준이 증가할수록 이미지 센터 추정치의 분산치가 증가하는 일반적인 경향과 아울러 0.5 픽셀 잡음 수준을 넘어가는 경우에 성능이 저하됨을 확인하였다. 이것은 팬(틸트) 회전각이 작은 경우에 $u(v)$ 방향으로의 값의 변화가 매우 작은 반면 픽셀 잡음은 상대적으로 매우 큰 영향을 미치기 때문이다. 그림 2는 위의 카메라 파라메타를 이용하여 시뮬레이션을 한 결과이다. 이 경우 픽셀 잡음에 의해 궤적의 저점 부근에서 왜곡이 발생하더라도 카메라의 LOS(Line Of Sight)와 3-D 포인트간의 각도 차이가 증가할수록 $u(v)$ 값의 변화가 두드러져 픽셀 잡음에 의한 효과는 약해짐을 알 수 있다. 3-D 포인트는 (0.3, 0, 1.3)m에 위치하고 있다고 가정하였고 이것은 4 절에서 소개할 실험 세팅값을 고려한 것이다.

4. 영상 처리 기법 및 실험 결과

제안된 이미지 센터 추정 기법을 실제로 적용하기 위해서는 sub-pixel 수준의 정확도를 가지면서 포인트 특징점을 추출할 수 있어야 한다. 이러한 필요성은 팬(틸트)회전의 경우 $u(v)$ 방향으로의 좌표값의 변화는 본 실험의 경우 35도 정도의 회전에 대해 10 픽셀 이내이기 때문이다. 제안된 방법에서는 cubic convolution 을 이용한 이미지 resampling 방법을 이용하여 sub-pixel 수준의 포인트 특징점 추출을 하였다[1]. 한편, 영상 획득시의 픽셀 잡음을 줄이기 위하여 이미지 중첩(image averaging)을 이용하여 잡음 제거 효과를 얻을 수 있었다.

식(5)의 파라메타를 계산하기 위해서는 최소한 5 개의 좌표값이 필요하다. 이것은 일반적인 conic section 을 구하는 경우로서 이미지 센터를 구하는 경우에도 그대로 적용 가능하다. 제안된 방법의 경우 식(6,7)에 나타난 것처럼 4 개의 포인트만으로도 충분하게 된다. 하지만 픽셀 잡음, 특징점 추출시의 에러를 고려한다면 이것보다는 더 많은 수의

좌표값을 이용해야 할 것이다.

실험에서는 일반 카메라 삼각대를 이용하여 팬 회전인 경우에 v_0 값을 구하였다. 팬 회전의 경우에 v_0 값은 78.4094 Pixel로 구해졌다. 그림 2에 팬 회전시의 카메라 영상을 나타내었다. 그림 3은 그림 2의 카메라 영상으로부터 구해진 3-D 포인트의 영상 궤적을 나타낸 것이다. **표시로 조밀하게 나타낸 것은 구해진 실험 데이터로부터 Levenberg-Marquardt 최적화 기법을 식 (6)에 적용한 결과를 증첩시켜 나타낸 것이다.

실험 결과로부터 시뮬레이션에서 예상된 3-D 포인트 특징점이 형성하는 2-D 영상 궤적을 확인하였다. 틸트 회전의 경우에는 카메라 삼각대의 구조적인 문제로 회전축과 카메라의 중심이 어긋나 있기 때문에 소실점을 이용하지 않고서는 일반 포인트 특징점을 이용할 수가 없어 실험 데이터를 구하지 못하였다. 향후 틸트 회전을 위한 기구적인 부분을 마련하여 검증을 하여야 할 것이지만 팬 회전의 경우와 마찬가지로 시뮬레이션의 예상 결과를 나타낼 것으로 여겨진다.

5. 결론

본 논문에서 소개한 이미지 센터 추정 방법은 특정 캘리브레이션 대상체를 필요로 하지 않으며 2-D 이미지 좌표값만을 이용한다는 점에서 실제 현장에서 손쉽게 적용이 가능하다. 시뮬레이션으로 제안된 방법의 반복 성능 및 잡음 대처 효과에 대해 살펴보고 실험을 통해 카메라 모델링 및 방법의 적합성을 보였다. 향후 전체 카메라 파라미터에 대한 캘리브레이션 알고리즘을 완성하여 3 차원 측정시의 제안된 이미지 센터 추정 방법의 성능을 보일 것이다. 특히, 본 방법은 로봇이나 산업 현장에서의 자동화에 적용하기에 좋은 형태이고 향후 대상 특징점(interest point)추출 과정의 자동화 및 영상간의 특징점 매칭 과정의 자동화 알고리즘을 개발하는 작업이 이루어져야 할 것이다.

참고 문헌

- [1] R. G. Keys, "Cubic Convolution Interpolation for Digital Image Processing," *IEEE Trans. ASSP*, vol. 29, no. 6, 1981.
- [2] Anup Basu, "Active Calibration of Cameras: Theory and Implementation," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 15, no. 2, pp. 256—265, 1995.
- [3] K. Daniilidis and J. Ernst, "Active Intrinsic Calibration Using Vanishing Points," *IEEE Int. Conf. On Computer Vision and Pattern Recognition*, pp. 708—713, 1996.
- [4] R. K. Lenz and R. Y. Tsai, "Techniques for calibration of the scale factor and image center for high accuracy 3-D machine vision metrology," *IEEE. Trans. Pattern Analysis and Machine Intelligence*, vol. 10, no. 5, pp. 713—720, 1988.
- [5] R. G. Willson and S. A. Shafer, "What is the Center of the Image?," *J. Opt. Soc. Am. A*, vol. 11, no. 11, pp. 2946—2955, 1994.
- [6] M. Aggarwal and N. Ahuja, "Camera Center Estimation," *IEEE Conf. on Pattern Recognition*, pp. 876—880, 2000.

[7] G. P. Stein, "Accurate Internal Camera Calibration using Rotation, with Analysis of Sources of Error," *IEEE Conf. on Computer Vision*, pp. 230—236, 1995.

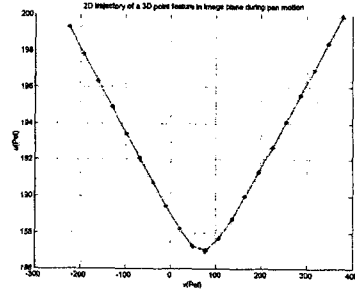


그림 2. 시뮬레이션 파라미터(3 점 참고)를 이용한 포인트 특징점의 궤적(픽셀 잡음 0.5)

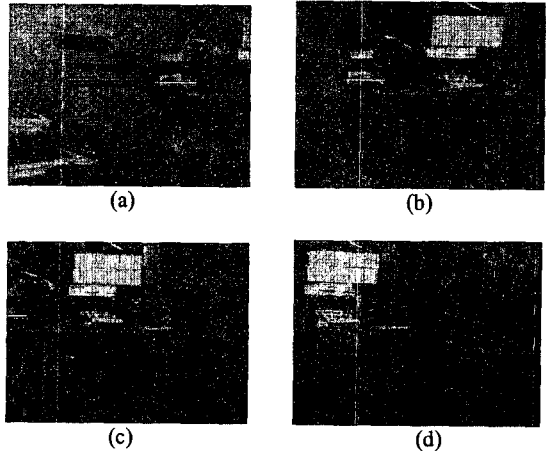


그림 3. v_0 를 구하기 위한 카메라 팬 회전시의 입력 영상. (a)-(b)-(c)-(d)의 순서이며 전체 입력 영상에서 대표적인 네 가지를 나타내었다. 카메라가 추적하고 있는 것은 흰색 박스의 상단 모서리이다.

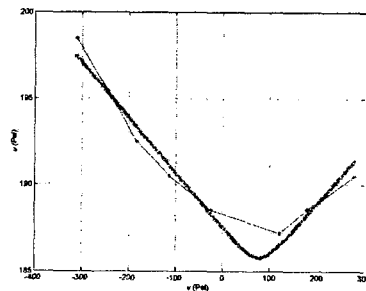


그림 4. 그림 2의 입력 영상으로부터 구해진 팬 회전시의 2-D 궤적. 조밀한 **로 표시된 것이 실험치를 바탕으로 curve fitting을 한 결과이다.