

## 모션 추정과 객체 추적을 이용한 이미지 깊이 검출기법

조 범 석\* · 김 영 로\*

### *A Technique of Image Depth Detection Using Motion Estimation and Object Tracking*

Joh, Beom-Seok · Kim, Young-Ro

#### 〈Abstract〉

In this paper, we propose a new algorithm of image depth detection using motion estimation and object tracking. In industry, robots are used for automobile, conveyer system, etc. But, these have much necessary time. Thus, in this paper, we develop the efficient method of image depth detection based on motion estimation and object tracking.

Key Words : Motion Estimation, Object Extraction

## I. 서론

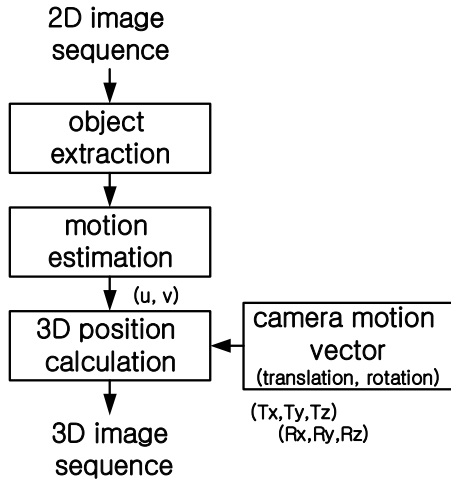
현재 발표된 이미지 깊이 검출 알고리즘은 특징점 추출과 이에 따른 이동 거리 추정을 이용하는 방법 등이 있다[1, 2, 3]. 그러나, 잡음에 따른 잘못된 특징점 추출과 이에 따르는 이동 거리 추정의 오류가 발생하는 문제점이 발생하는 단점이 있다. 또한 필터링을 이용하여 문제점을 처리할 수 있다고 하지만 고주파 성분이 많거나, 모션이 많은 이미지의 경우에는 효율이 더욱 떨어진다. 따라서 이 연구에서는 영역을 구분하여 객체를 추출하고 모션을 추정함으로써 객체의 이미지 깊이를 검출할 수 있도록 한다.

## II. 이미지 깊이 검출 시스템

이미지 깊이 검출 시스템의 블록도는 다음 그림과 같다. 먼저 이미지 센서를 통하여 2D 이미지를 획득하고 이에 따른 객체를 추출하는 객체 추적(object extraction)부, 2D 이미지들에서 모션을 추정하는 모션 추정부, 카메라 또는 로봇의 이동 또는 회전을 나타내는 카메라 모션 벡터부, 마지막으로 두 부분에서 획득한 정보를 이용하여 객체의 3D 위치를 검출하는 3D 위치 계산부가 있다.

객체 추적(object extraction)부에서 이미지 분할(segmentation)프로세스를 통하여 영역간의 경계선을 객체들 간의 경계점으로 인식하여 구분된 영역을 객체로써 추출한다. 이미지 분할은 이미지를 선택하여 비슷한 픽셀 값끼리 영역을 분할하는 방법이다.

\* 명지전문대학 컴퓨터정보과 교수



<그림 1> 2차원 정보의 3차원 정보 변환

이미지 분할 프로세스는 간략화, 경사도, 분수령의 단계를 거쳐 구분된 영역을 만들어준다[4, 5].간략화(simplification)부에는 형태학 여과기(morphological filter), 평균 여과기(average filter), 중간값 여과기(median filter), 가우시안 여과기(gaussian filter) 등을 사용하는 방법들이 있다. 형태학 여과기는 opening, closing을 함으로써 이미지를 단조롭게 하여 국부적인 최소값을 없애준다. 평균 여과기는 현재 위치값과 주변 화소 값의 평균값으로 새롭게 구할 수 있다. 따라서 국부적인 최소값들을 제거한다. 중간값 여과기는 주변 픽셀 값과 현 위치의 원래 픽셀값과의 중간값을 택한다.

경사도(gradient)부에서는 이전 단계에서 간략화된 이미지로부터 공간 경사도를 구하게 된다. 이때 비슷한 픽셀값들을 지닌 영역들간의 경계는 작은 값의 경사도를 가지고 있고 현저하게 다른 영역들간의 경계는 큰 값의 경사도를 가진다. 따라서 공간 경사도는 다음 단계인 이미지를 영역분할 하는 분수령(watershed)부의 입력 값으로 사용할 수 있다. 이 값은 각 영역 간들의 경계를 나타내며 이에 분수령 알고리즘을 이용하여 영역들의 경계를 보다 뚜렷하게 나눌 수 있다. 따라서 경사도 이미지는 원 이미지보다 분수령 알고리즘(watershed algorithm)이 효

과적으로 영역을 구분할 수 있도록 한다.

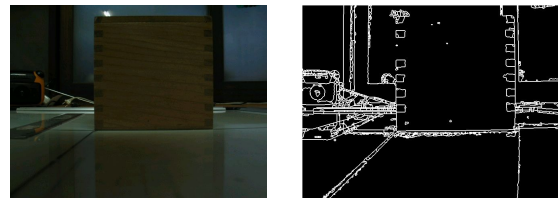
형태학 연산자를 이용하여 경사도를 구하는 방법은 다음 식과 같다.

여기서  $G, D, E$ 는 각각 경사도, 팽창(dilation), 침식(erosion)을 의미한다. 이 연산자는 경사도 값이 항상 양의 값을 지니고 있어 계산량이 적은 장점이 있다. 이밖에 경사도 연산자(operator)의 종류로서 다음과 같이 sobel, roberts, cubic 연산자가 있다.

그레이(gray) 이미지 정보만으로 영역의 경계를 구분하기 어려운 점을 극복하기 위하여 화소의 컬러값들  $Y, C_b, C_r$ 로부터 경사도  $G_Y, G_{C_b}, G_{C_r}$  또는 화소의 컬러값  $R, G, B$ 로부터 경사도  $G_R, G_G, G_B$ 를 구하고 이에 대한 각각의 가중치를 주어 이 값들에 대한 최대값을 그 화소의 경사도로 이용하기도 한다.

$$G = \max \{ W_1 G_R, W_2 G_G, W_3 G_B \}$$

실제로 이렇게 구한 값들을 이용하여 분수령 알고리즘을 적용하면 과잉 영역 분할이 발생한다. 그러므로 이러한 문제점을 해결하기 위하여 위 수식에서 구한 각 화소의 경사도를 분수령 알고리즘에 적용하기 앞서서 적절한 임계치를 주어 절삭(clipped out)한다. 즉 임계치보다 작은 값은 0으로 놓는다. 그림 2는 카메라에서 획득한 이미지와 이에 따른 영역 분할한 결과를 나타낸 이미지이다.



<그림 2> 카메라에서 획득한 이미지 및 영역 분할 이미지

모션 추정(motion estimation)부에서는 2장의 이미지를 가지고 optical flow 방법(horn-shunck, robust 수식 등)을 이용하여 모션 벡터(u,v)를 찾는다[6]-[8].

$$I_x \cdot u + I_y \cdot v + I_t = 0$$

여기서  $I_x, I_y, I_t$ 는 각각 특정 픽셀에서 수평, 수직, 시간적 기울기를 나타낸다. 따라서 다음과 같이 나타낼 수 있다.

$$I_x = \frac{1}{4} \{f(x+1, y, k) - f(x, y, k) + f(x+1, y+1, k) - f(x, y+1, k) + f(x+1, y, k+1) - f(x, y, k+1) + f(x+1, y+1, k+1) - f(x, y+1, k+1)\}$$

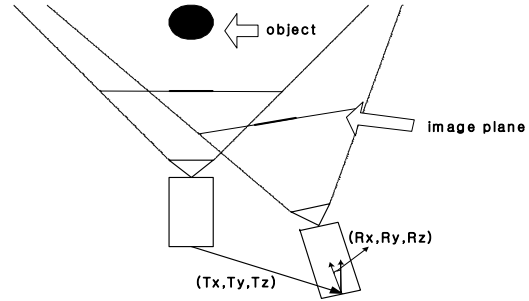
$$I_y = \frac{1}{4} \{f(x, y+1, k) - f(x, y, k) + f(x+1, y+1, k) - f(x+1, y, k) + f(x, y+1, k+1) - f(x, y, k+1) + f(x+1, y+1, k+1) - f(x+1, y, k+1)\}$$

$$I_t = \frac{1}{4} \{f(x, y, k+1) - f(x, y, k) + f(x+1, y, k+1) - f(x+1, y, k) + f(x, y+1, k+1) - f(x, y+1, k) + f(x+1, y+1, k+1) - f(x+1, y+1, k)\}$$

여기서  $f(x, y, k)$ 는 k번째 프레임에서  $(x, y)$  위치의 픽셀값을 나타낸다.

카메라 모션 벡터부에서는 그림 3과 같이 이미지 센서가 장착된 로봇의 위치 및 회전 이동 정보를 적외선 센서, 또는 실측을 통하여 획득한다. 즉 이미지를 획득할 때 로봇의 모션 벡터  $(T_x, T_y, T_z), (R_x, R_y, R_z)$ 를 구한다.

3D 위치 계산부에서는 이미지에서 획득한 화소들의 실제 3D 위치 즉  $(X, Y, Z)$ 을 계산한다. 다음 수식들과 모션 추정부에서 구한 모션 벡터  $(u, v)$ 를 이용하여  $(X, Y, Z)$ 를 구한다.



<그림 3> 로봇에 장착된 이미지센서의 위치 이동 및 회전

$$x_a = \frac{fX}{s_x Z} + c_x, y_a = \frac{fY}{s_y Z} + c_y$$

여기서  $x_a, y_a$ 는 2D 이미지에서 화소의 위치를 나타내고,  $c_x, c_y$ 는 2D 이미지의 가운데 위치를 나타낸다. 그리고  $f$ 는 focal length를,  $s$ 는 카메라 샘플링과 화소길이에 의한 스케일을 나타낸다.

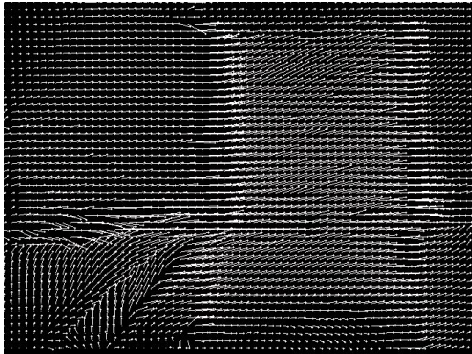
실제 3D 위치 P 즉  $(X, Y, Z)$ 에서의 모션 벡터  $u, v$ 는 다음과 같이 나타낼 수 있다.

$$u = [x \frac{T_z}{Z} - f \frac{T_x}{Z s_x}] + [xy \frac{s_y R_y}{f} - (\frac{f}{s_x} + x^2 \frac{s_x}{f}) R_y + y \frac{s_y}{s_x} R_z]$$

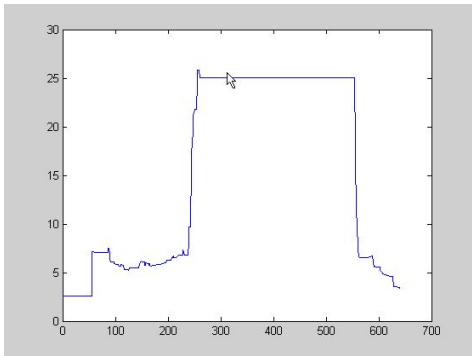
$$v = [y \frac{T_z}{Z} - f \frac{T_y}{Z s_y}] + [(\frac{f}{s_y} + y^2 \frac{s_y}{f}) R_x - xy \frac{s_x}{f} R_y - x \frac{s_x}{s_y} R_z]$$

따라서 optical flow에서 구한 모션 벡터와 카메라의 위치 이동 및 회전에 의하여 구한 모션 벡터들  $(T_x, T_y, T_z), (R_x, R_y, R_z)$ 를 이용하여 실제 3D 위치  $(X, Y, Z)$ 를 구한다. 그림 2는 카메라가 이동하면서 생기는 모션 벡터들을 나타낸다. 그림 4는 이미지의 하나의 줄에서(카메라 렌즈와 같은 높이) 구한 모션 벡터(u)를 나타낸다. 그림 6은 그림 5에서 구한 모션 벡터(u)와

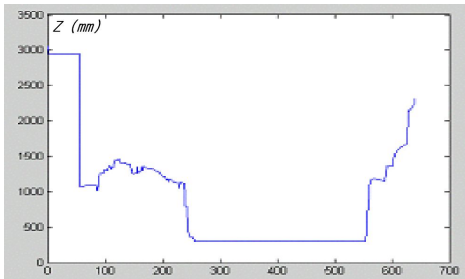
카메라의 이동, 회전, 이미지의 해상도 정보를 이용하여 구한 물체의 위치(여기서는 카메라와 떨어진 위치)  $Z$ 을 나타낸다.



<그림 4> 모션 벡터



<그림 5> 수평 모션 벡터( $u$ )



<그림 6> 물체의 위치  $Z(mm)$

### III. 결론

기존의 이미지 깊이 검출 방법은 획득한 이미지의 잡음에 따라 부정확한 위치를 검출하는 어려움이 있다. 이 논문에서는 모션을 적용할 때 제한된 객체 추적 영역을 사용하여 기존의 이미지 깊이 검출 알고리즘 보다 향상된 결과를 얻으며 실제 위치와 유사한 위치 결과를 얻을 수 있다.

### 참고문헌

- [1] I. Shimshoni, Y. Moses, and M. Lindenbaum, "Shape reconstruction of 3D bilaterally symmetric surfaces," *Int'l J. Computer Vision*, vol. 39, no. 2, Sept. 2000, pp. 97-110.
- [2] B. J. Super and A. C. Bovik, "Shape from texture using local spectral moments," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 17, no. 4, Apr. 1995, pp. 333-343.
- [3] A. Pentland, "A new sense for depth of field," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 9, no. 4, Jul. 1987, pp. 523-531.
- [4] L. Vincent and P. Soille, "Watershed in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, Jun. 1991, pp. 583-598.
- [5] R. Castagno, T. Ebrahimi, and M. Kunt, "Video segmentation based on multiple features for interactive multimedia application," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 5, Sept. 1998, pp. 562-571.
- [6] B. K. P. Horn and B. G. Shunck, "Determining optical flow," *Artif. Intell.*, vol. 17, no. 1, Aug. 1981, pp. 185-203.

- [7] M. J. Black, "The robust estimation of multiple motions: parametric and piecewise-smooth flow fields," Computer vision and image understanding, vol. 63, no. 1, Jan. 1996, pp. 75-104.
- [8] J. Y. A. Wang and E. H. Adelson, "Representing moving images with layers," IEEE Trans. On Image Processing, vol. 3, no. 5, Sept. 1994, pp. 625-638.

■ 저자소개 ■



조범석  
Joh, Beom Seok

2002년~현재 명지전문대학 컴퓨터정보과 교수  
2001년 명지대학교 대학원 컴퓨터공학과 공학박사  
1988년 명지대학교 대학원 컴퓨터공학과 공학석사  
1986년 명지대학교 전자계산학과 공학사  
관심분야 : 멀티미디어 데이터베이스  
검색기법, Ubiquitous Computing



김영로  
Kim, Young Ro

2003년~현재 명지전문대학 컴퓨터정보과 교수  
2001년~2003년 삼성전자 시스템LSI 책임연구원  
2001년 고려대학교 전자공학과 컴퓨터공학 박사  
1996년 고려대학교 전자공학과 컴퓨터공학 석사  
1993년 고려대학교 전자공학과 학사  
관심분야 : 신호 및 영상처리, 멀티미디어 통신

논문접수일 : 2008년 4월 21일, 수정일 : 2008년 6월 5일(1차)  
게재확정일 : 2008년 6월 13일