

두개의 영상으로 부터 Epipolar Line과 Matching을
이용한 3차원 물체의 위치 인식

조석제 박길흠 이광호 김영모 하영호
경북대학교 공과대학 전자공학과

3-D Recognition of Position using Epipolar Line and Matching
from Stereo Image

Seok-Je Cho, Kil-Houm Park, Kwang-Ho Lee, Young-Mo Kim and Yeong-Ho Ha
Dept. of Electronics, Kyungpook National Univ.

ABSTRACT

Extraction of depth information from stereo image uses the matching process between them and this requires a lot of computational time. In this paper, a matching using the feature points on the epipolar line is presented to save the computations. Feature points are obtained in both image and correlated each other. With the coordinates of the matched feature points and camera geometry, the position and depth informations are identified.

I. 서론

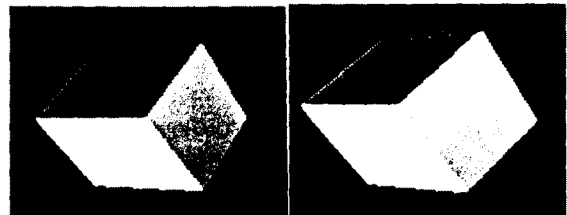
거리 정보 추출은 물체 인식에서 중요한 문제이며 2차원 영상 즉, 한장의 사진으로도 물체의 거리 정보를 얻을 수 있다. 이는 폐색(Occlusion), 그림자(Shadows), 조직 경사도(Texture gradient), 물체의 크기 등과 같은 많은 단서들(Cues)을 이용하기 때문이라 할 수 있다. 그러나 카메라와 같은 기계적 장치에 의해서 3차원 물체를 해석하려면 Stereo 방법이 유용하다.

Stereo 방법에서 가장 중요한것은 좌우 영상에서 대응점을 찾는 것이다. 두 영상에서 특징을 대응시키는 방법으로는 특징점을 추출하여 대응시키는 방법과 특징점을 규정하지 않고 작은 영역을 Cross correlation 방법이나 Intensity 차와 같은 척도(Measure)에 의해 대응하는 방법⁽¹⁾⁽²⁾이 있다.

특징점 Matching은 일반적으로 특징점의 수가 적은 경우에는 효율적이며 사용되는 특징점은 작은 시각 변화에서 특성이 현저히 변화하지 않아야

하며 특징점이 많은 경우에는 부적절하다. 가령 선(Line segment)을 특징점으로 한다면 특징점이 회전(Rotation)하면 특성이 현저히 변화하나 병진이동(Translation)은 그렇지 않다. 또한 Corner는 회전과 병진이동 모두에도 적합하지 않다.

정확하고 신속하게 두 영상에서 대응점을 찾는 것은 Stereo 영상 처리에서 매우 중요한 문제이다. 본 연구에서는 합성 육면체 영상(그림1)을 사용해서 좌우 두영상 모두에서 윤곽선을 검출하여 꼭지점을 특징점으로 하고, Matching시 탐색 범위를 더 줄이기 위해 Epipolar line을 구하였다. 그리고 좌영상의 한 특징점과 우영상에서 이에 대응하는 Epipolar line상의 특징점들에서만 Cross Correlation방법에 의해 그 대응 특징점을 Matching하였고, 또한 좌우 영상의 대응점 좌표와 각 카메라의 위치와 회전 각도로 부터 특징점의 World 좌표계상의 위치를 구하였다.



(a) (b)

그림1. 육면체 좌우영상

a) 좌 영상

b) 우 영상

II. 대응점 결정 및 위치 인식

1. 전처리 과정

육면체의 꼭지점을 특징점으로 추출하기 위해 Laplacian operator로 육면체의 윤곽선을 검출하고 (그림 2-a) 이를 세선화하여 (그림 2-b) Chain code를 구성하였다. Chain code로 부터 특징점을 찾는 방법은 Johnston과 Rosenfeld가 제안한 방법을 이용하였다. 즉 이 방법은 각점 $\{(X_i, Y_i)\}_i^m$ 들이 $(X_1, Y_1) = (X_n, Y_n)$ 인 폐곡선을 따라 높은 곡률을 가지는 부분을 추출하는 방법이다.

$$\vec{A}_{ik} = (X_i - X_{i+k}, Y_i - Y_{i+k}) \quad (1)$$

$$\vec{B}_{ik} = (X_i - X_{i-k}, Y_i - Y_{i-k}) \quad (2)$$

$$C_{ik} = (\vec{A}_{ik} \cdot \vec{B}_{ik}) / \|\vec{A}_{ik}\| \|\vec{B}_{ik}\| \quad (3)$$

여기서 C_{ik} 는 벡터 \vec{A}_{ik} 와 \vec{B}_{ik} 가 이루는 각의 Cosine값이다. C_{ik} 값이 부분 최대치가 되는 곳이 높은 곡률을 가지는 꼭지점 이다. 본 논문에서는 이 꼭지점을 특징점으로 한다.

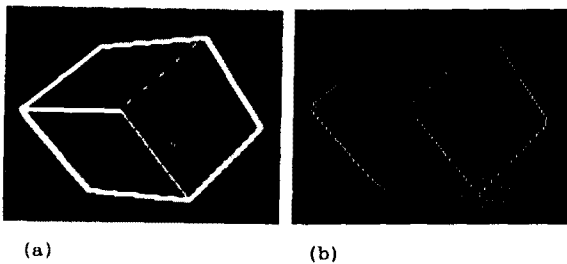


그림 2. 전처리 영상

- a) 라플라시안 적용후 영상
- b) 세선화된 영상

2. 대응점 결정

두 카메라가 서로 떨어져 있을때 그림 3에서와 같이 두 카메라의 중심점 C_L, C_R 과 물체점 O 가 만나는 평면을 Epipolar Plane이라 하며 이 Epipolar Plane과 두 영상이 교차 되는 곳에서의 선을 Epipolar Line이라 한다. 좌 영상의 한 특징점에 대응하는 우 영상의 점은 반드시 Epipolar Line상에 존재한다. 그래서 두 영상에서의 2차원 Matching 문제가 Epipolar Line위의 1차원 Matching 문제로 되어 탐색 범위가 좁게된다.

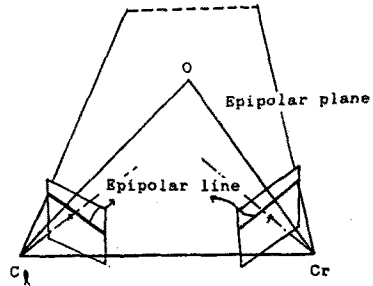


그림 3. Epipolar plane 과 Epipolar line

우 영상에 있는 Epipolar Line을 구하기 위해 좌 영상의 각 특징점과 카메라의 위치를 오른쪽 카메라 좌표계로 변환한다. 변환 과정은 식 (4), (5)와 같으며, 왼쪽 카메라 좌표계의 특징점과 원점(벤즈의 중심)을 먼저 World좌표계로 변환한후 이를 오른쪽 카메라 좌표계로 변환한다.

$$R_o \cdot V_l + T_l = V_o \quad (4)$$

$$R_r [V_o - Tr] = V_r_l \quad (5)$$

여기서 R_o 와 T_l 은 각각 왼쪽 카메라 좌표계에서 World 좌표계로의 회전과 병진 이동을 나타내며, R_r 과 Tr 은 World 좌표계에서 오른쪽 카메라 좌표계로의 회전과 병진 이동을 나타낸다. 또한 V_l 은 왼쪽 카메라 좌표계를, V_o 는 World 좌표계에서 좌영상의 특징점을 V_r_l 은 오른쪽 카메라 좌표계에서 좌영상의 특징점을 나타낸다.

오른쪽 카메라 좌표계에서 왼쪽 카메라의 원점과 특징점 그리고 오른쪽 카메라의 원점을 연결하는 평면 방정식 (6)을 구한다.

$$Z_r = -pX_r - qY_r - c \quad (6)$$

Z_r 축이 광축 이므로 영상은 $Z_r = dz$ 에 위치해 있다고 할 수 있다. 여기서 dz 는 초점 거리(f)와 Scale factor(S)의 곱이다. 그러므로 Epipolar Line은 식 (7)이 된다.

$$pX_r + qY_r + c + dz = 0 \quad (7)$$

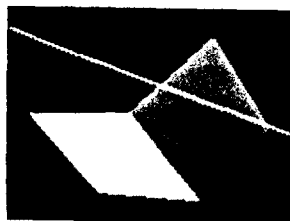


그림 4. Epipolar line

Epipolar Line과 만나는 우영상 상의 특징점이 곧 대응점이다. 그러나 Epipolar Line상에 두개 이상의 특징점이 있을경우에는 그 특징점 각각에서만 Template Matching을 수행해서 대응점을 찾는다. Template Matching은 좌 영상에서 Template를 취해서 우 영상의 Epipolar Line과 만나는 특징점과 Correlation을 해서 최대가 되는 특징점을 대응점으로 한다.

3. 심도 및 위치 정보 추출

특징점들의 심도와 위치 정보를 추출하기 위해 우선 World 좌표계와 카메라 좌표계의 관계를 알아야 한다. World 좌표계의 좌표점들을 카메라 좌표계로 변환하려면 원점의 위치 변환 (r1,r2,r3)와 Z축과 Y축의 회전 즉 Pan과 Tilt (그림5)에 의해 식 (8)~(10)과 같은 변환을 한다. 여기서 광축은 카메라 좌표계의 Z축이다.

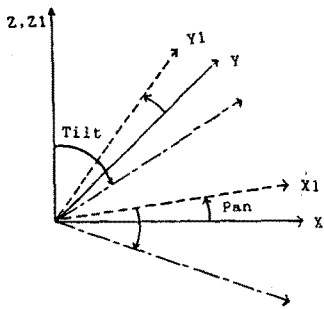


그림5. Pan과 Tilt에 의한 좌표축의 변환

$$T = (r1 \ r2 \ r3)^T \quad (8)$$

여기서 T는 (r1,r2,r3)에 의한 각 축의 병진 이동 Matrix이다.

$$R = \begin{pmatrix} \cos\phi\cos\theta & \cos\phi\sin\theta & -\sin\phi \\ -\sin\theta & \cos\theta & 0 \\ \sin\phi\cos\theta & \sin\phi\sin\theta & \cos\phi \end{pmatrix} \quad (9)$$

여기서 R은 Pan angle θ , Tilt angle ϕ 에 의한 회전 Matrix이며 $R^T R = I$ 이다.

병진 이동(8)과 회전(9)에 의해 카메라 좌표계로의 변환은 식(10)이 된다.

$$Vc = R (Vo - T) \quad (10)$$

여기서 Vc는 카메라 좌표계 $(Xc, Yc, Zc)^T$ 이며 Vo는 World 좌표계 $(Xo, Yo, Zo)^T$ 이다.

카메라와 영상을 모델링 하면 그림6과 같이 카메라 중심 C와 초점 거리 f, 영상면 I로 구성된다.

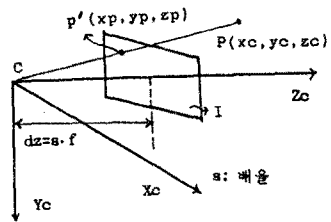


그림6. 영상면 상에 점 P의 투영

(xc, yc, zc) 에 위치한 점 P는 영상면에서 (xp, yp, zp) 로 주어지며 그 관계는 식(11)~(13)과 같다.

$$Xp = dz \cdot Xc / Zc \quad (11)$$

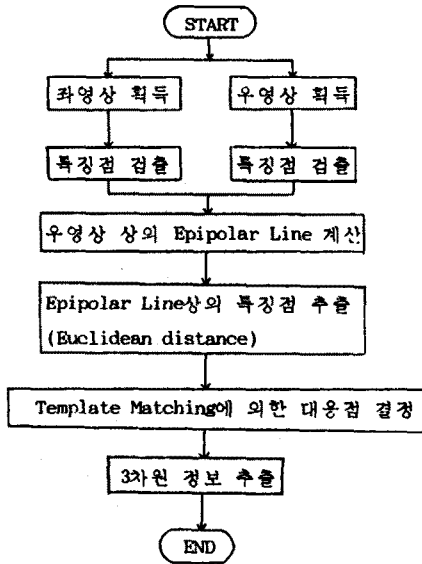
$$Yp = dz \cdot Yc / Zc \quad (12)$$

$$Zp = dz \quad (13)$$

좌 영상의 특징점과 이에 대응하는 우 영상의 특징점의 좌표를 알고, 두영상의 병진 이동 정도와 회전 각도와 초점거리를 알면 식 (10)~(13)의 관계에 의해 World 좌표계에서 특징점의 좌표를 구할수 있다.

III. 실험 및 고찰

본실험에서는 서로 다른 위치에서 얻은 두개의 육면체 합성 영상을 사용하여 Simulation하였다. 실험 순서도는 다음과 같다.



World 좌표계에서 특징점의 위치와 본 Simulation에서 구한 결과는 표1과 같다.

표1. 특징점 위치의 Simulation 결과 및 오차

특징점의 위치	Simulation 결과	오차
-250, -250, 250	-250, -250, 248	0, 0, 2
-250, 250, -250	-255, 248, -257	5, 2, 7
-250, 250, 250	-242, 257, 255	8, 7, 5
250, -250, 250	260, -238, 258	10, 12, 8
250, -250, -250	252, -249, -251	2, 1, 1
250, 250, 250	244, 244, 242	6, 6, 8
250, 250, -250	255, 255, -248	5, 5, 2

본 Simulation에서 사용된 두 카메라의 Geometry는
 왼쪽 카메라 ; $T1 = (2000, 2000, 2000)$
 $(\theta_1, \phi_1) = (45^\circ, -125.3^\circ)$
 오른쪽 카메라 ; $T2 = (1500, 1800, 2300)$
 $(\theta_2, \phi_2) = (50.2^\circ, -134.5^\circ)$
 로 하였다.

위 결과에서와 같이 실제 위치와 Simulation 결과와의 오차가 약 5% 이내였다.

IV. 결론

본 논문에서는 3차원 물체 인식에서 중요한 요소인 물체의 거리와 위치 정보를 인식하고자 하며 거리와 위치 정보를 구하기 위해서 먼저 두 영상에서 경계선을 추출 세선화하여 Chain code를 이용해서 특징점을 각각 구하였다. 좌우 영상의 대응특징점은 Epipolar Line상의 특징점을 대응점으로 하는 방법을 사용 하였으며 이는 Correlation에 의해 Matching함으로써 보다 신속하고 정확하게 Matching 할수 있다. Matching한 두 영상좌표와 카메라의 Geometry로부터 특징점들의 World좌표를 구해낼수 있었다.

참고 문헌

- Gerard Medioni and Ramakant Nevatia, "segment-Based Stereo Matching," CVGIP, vol 31, pp2-18, 1982.
- Ramakant Nevatia, Machine Perception, Prentice-Hall, 1982.
- Azriel Rosenfeld and Avinash C. Kak Digital Picture Processing, Academic Press 1982
- Berthold Klaus Paul Horn, Robot Vision, MIT Press, 1986
- Azriel Rosenfeld and Emily Johnston, "Angle Detection on Digital Curves," IEEE Trans. Comput. vol c-22, pp 875-878, Sep., 1973.
- R.O. Duda and P.E. Hart, Pattern Classification and Scene Analysis, Wiley-interscience, 1973.