

스테레오 기법의 형태정보를 이용한 3차원 물체 인식

정회원 박성만*, 객원식**, 이대영*

The 3-D Object Recognition Using the Shape from Stereo Algorithm

Park, Sung-Man*, Kwak, Yoon-Sik**, Lee, Dae-Young* *Regular Members*

요약

본 논문은 3차원 물체 인식을 위한 깊이정보의 추출을 위해 스테레오 기법의 적용과 기존의 방법론에서 정합횟수에 대한 문제점을 보완하기 위한 물체의 진행 방향 벡터를 이용하는 방법론을 제안한다. 즉, 물체의 진행 방향을 이용하여 진행방향 벡터를 구하고 이를 축으로 회전 이동된 물체를 정합 시킨다. 그리고 Hough변환을 이용하여 얻은 회전 이동에 대응하는 2차원 합성영상을 참조영상으로 설정하고 입력영상과 정합 과정을 수행하였다.

ABSTRACT

In this paper, we presented the stereo algorithm for 3-D object recognition. In order to solve the problem for matching time in existed methods, we proposed the method which used the moving direction vector. On the other hand, after we extracted the moving vectors by moving direction of objects, rotated object was matched on axis of it. Using the Hough transform, we obtained the 2-D synthesesed image as reference images corresponding to the rate of moving, and then compared with the unknown input images.

I. 서론

3차원 Hough 변환을 이용한 3차원 물체인식에서 물체의 깊이 정보를 복원하기 위하여 레인지 파인더를 사용한다. 그러나 레인지 파인더를 사용하는 방법은 먼 거리에 놓인 물체와 빠르게 움직이는 물체에 대해 적용하는 것이 곤란하다는 단점을 가지고 있다. 이에 따라 Hough변환에서 물체의 깊이 정보를 얻기 위한 방법으로 인간의 시각 시스템과 유사한 스테레오 기법(shape from stereo)을 사용할 경우 이와 같은 단점을 상당량 극복할 수 있다. 또한 3차원 Hough변환에서 필요한 물체의 거리 정보 복원 문제가 해결되었다 할 지라도 물체 정합과정에서 발생하는 많은 회전 이동에 대한 문제점이 발생된다.^{[1]-[3][8][9]} 이와 같은 정합횟수에 대한 문제점을 보완하기 위해서 물체의 진행 방향 벡터를 이용

하는 방법론을 제안한다. 즉, 물체의 진행 방향 벡터를 구하고 이를 축으로 회전 이동된 물체를 정합 시킨다. 이때 Hough변환에서 얻은 회전 이동에 대응하는 2차원 합성영상을 참조영상으로 설정하고 입력영상과 정합 과정을 수행하였다.

II. 본론

스테레오 기법을 이용한 3차원 물체인식 시스템은 3부분으로 구성된다. 첫째는 스테레오 기법에 의한 물체의 거리 정보를 복원하는 부분이고 둘째는 3차원 Hough변환에 의한 영상 분류 및 정합 부분이며 셋째는 8진 트리 구조를 이용한 2차원 투영영상의 모델링과 정합 부분이다.

1. 스테레오 기법에 의한 물체의 거리 정보 복원
스테레오 기법을 사용하기 위해서는 입력영상은

* 경희대학교 전자공학과
논문번호 : 99131-0413, 접수일자 : 1999년 4월 13일

** 충주대학교 컴퓨터공학과(yskwak@gukwon.chungju.ac.kr)

다음과 같은 조건을 만족해야만 한다. 첫째, 두 카메라의 초점사이는 적절한 거리 차이가 있어야 한다. 둘째, 영상은 구별되는 특징들이 나타날 수 있도록 충분히 무늬화(textured)되어야 하며 균일한 명암 값을 가지면 곤란하다. 셋째, 거리 정보는 두 초점에서 동시에 보이는 영상점에서만 계산되어야 한다.

스테레오 기법은 일정한 거리를 두고 위치한 두 카메라에서 얻은 영상에서 동일 위치의 한 점에 대해 시각 차이(disparity)가 나타나게 되고 두 영상의 시각 차이는 물체의 거리 정보에 반비례한다. 이러한 과정을 그림1과 같은 기하학적 구조를 이용하여 설명한다.

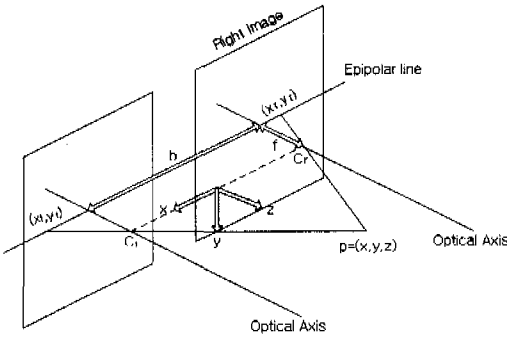


그림 1. 스테레오 영상의 기하학적 구조

그림1에서 (X_r, Y_r) 과 (X_l, Y_l) 은 오른쪽 그리고 왼쪽 영상좌표, f 는 초점으로부터 영상 평면까지의 거리이고, b 는 2대의 카메라 사이의 기본선(Baseline)이다. 위 그림으로부터 식(1)-(4)을 얻을 수 있다.

$$\frac{x_l}{f} = \frac{(x + b/2)}{z} \tag{1}$$

$$\frac{x_r}{f} = \frac{(x - b/2)}{z} \tag{2}$$

$$z(x_l - x_r) = fb \tag{3}$$

$$z = \frac{bf}{(x_l - x_r)} \tag{4}$$

식(3)에서 $x_l - x_r$ 은 두 카메라에서 검출된 시각 차이 값으로 물체의 거리 정보 z 에 반비례함을 알 수 있다. 따라서 이 식을 이용해 스테레오 영상의 거리 정보(depth map)를 복원 할 수 있다. 이를 위하여 대부분의 스테레오 기법은 다음의 단계를 포함한

다.^{[4][5]}

- 1) 왼쪽과 오른쪽 영상에서 특징점을 찾는다.
- 2) 두 영상의 특징 점간의 대응관계를 찾는다.
- 3) 대응되는 특징점으로부터 시각차이를 측정한다.
- 4) 변위에서 깊이 지도(depth map)를 얻어 3차원 표면을 재구성한다.

스테레오 영상에 의해서 물체의 거리 정보가 복원되면 시간차를 갖는 두 프레임간의 영상치를 이용해서 물체의 속도 및 진행 방향벡터를 결정할 수 있으며 이 벡터를 물체의 Hough변환시 시점벡터로 설정해 이 시점벡터를 회전축으로 한 회전과의 정합을 계산하여 계산량을 대폭 줄일 수 있다. 예를 들어 Hough변환 공간상의 시점벡터에 대한 샘플링을 약 6° 로 가정한 경우 Hough변환 시 단 위구 상에 존재하는 회전축의 수는 1102개, 이에 대해 60개의 회전이 가능하므로 비교 대상은 약 66000개가 된다. 반면에 회전축이 정해지면 비교 대상을 회전축에 대한 회전만을 고려하면 되기 때문에 비교 대상을 60개로 줄일 수 있다. 이를 그림 2에 나타내었다.

진행 방향 벡터는 식(5)-(7)에 의해서 구해진다.

$$\Delta x = \frac{z_{av}(t+1) \times (x_R(t+1) + x_L(t+1)) - z_{av}(t) \times (x_R(t) + x_L(t))}{2f} \tag{5}$$

$$\Delta y = \frac{z_{av}(t+1) \times (y_R(t+1) + y_L(t+1)) - z_{av}(t) \times (y_R(t) + y_L(t))}{2f} \tag{6}$$

$$\Delta z = z(t+1) - z(t) \tag{7}$$

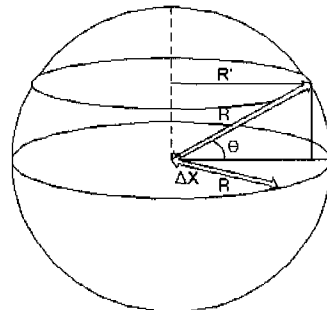


그림 2. Hough변환 공간의 단위 벡터 샘플링

- 등간격 $\Delta x = \frac{2\pi R}{60}$ 개
- 한 라인당 샘플링 수

$$n = \frac{2\pi R'}{4x} = \frac{2\pi R'}{2\pi R} = \frac{R'}{R} \cdot 60 = \frac{R \times \cos\theta}{R} \times 60 = 60 \cos\theta$$

● 전체 샘플링 수 = $2 \times \sum_{\theta=0}^{90} 60 \cos\theta - 60$

2. 3차원 Hough변환에 의한 영상 분류 및 정합

3차원 Hough변환은 2차원 Hough변환을 3차원 공간상으로 확장한 개념으로 한 벡터를 3차원 X-Y-Z공간상의 크기와 방향으로 변환시켜 다룬다. 3차원 Hough변환에서는 공간상의 방향을 표현하기 위해 단위 구면에 균일하게 분포하는 단위 벡터들을 설정하고 이를 Un이라 하였을 때 주어진 데이터 벡터 X와 단위 벡터 사이의 내적으로 3차원 공간상의 한 벡터의 크기를 표현한다. 이를 나타낸 것이 그림 3이다. 여기서 x,y,z는 3차원 좌표공간, Un은 단위벡터, P는 내적이다.

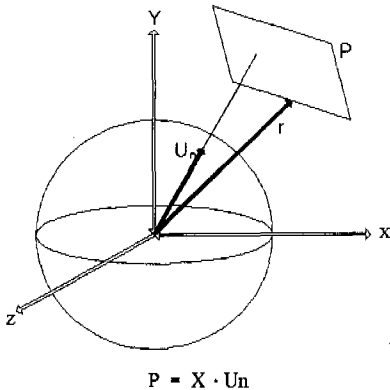


그림 3. 3차원 Hough변환 공간

모든 단위 벡터 Un과 모든 데이터 벡터 X에 대해 P값을 계산하여 2차원 공간인 H(n,p)에 누적시키면 3차원 Hough변환 공간 H(n,p)가 생성된다. 이러한 과정으로 생성된 H(n,p)는 단위 벡터 Un의 방향으로 P만큼 떨어진 곳에 위치하는 평면의 크기를 나타내게 되는데 그 이유는 한 평면의 법선 벡터에 해당되는 단위 벡터 n에 대해서는 모든 P값이 같은 값을 가지게 되기 때문이다. 따라서 3차원 Hough변환 공간에서 주어진 단위벡터 방향으로 면이 존재하게 된다면 큰 극대치를 나타낼 것이고 반대의 경우에는 0에 가까운 값을 가질 것이다. 이를 통해 쉽게 물체의 면을 검출할 수 있다. 또한 H(n,p)값이 면의 면적을 나타내고 있을 뿐만 아니라 극대치가 나타나는 P의 위치 값을 통해 면의 위치에 관한 정보도 알 수 있다. 이 정보를 통해서 물체의 평행 이동량에 관한 정보도 구할 수 있다.^{[6][9]}

3. 8진 트리 구조를 이용한 2차원 투영 영상의 모델링

8진 트리로 물체를 모델링하는 방법은 소수의 2차원 영상으로부터 3차원 물체를 묘사하는 방법으로 2차원 물체 표현 기법중 하나인 4진 트리구조 방법을 3차원 물체로 확장한 표현 기법이다.

이 방법의 장점은 임의의 직교하는 방향에서 본 3개의 3차원 물체에대한 2차원 투영영상의 4진 트리 구조로부터 8진 트리 구조를 만든 후 영상을 합성하여 임의의 방향에서 바라 본 3차원 영상을 추론할 수 있다는데 있다.^{[7][8][10]}

이 구조는 2차원 2진 영상의 4진 트리 구조방법을 3차원 물체로 확장한 표현 방법으로 물체를 포함하는 공간상의 큰 구획인 voxel을 같은 크기의 8개의 부분-voxel 혹은 8분면체로 세분하여 8진 트리 구조(octree)를 만든다.(그림 4)

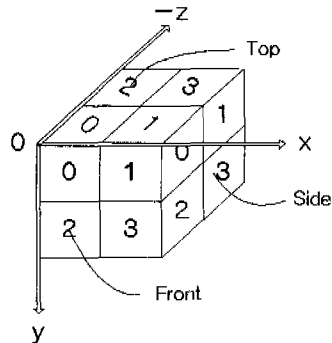


그림 4. 8진 트리 구조의 블록 분할 구조

이들 각 voxel은 동질색소(white, black)를 갖거나 혼합색소(gray)를 가지게 되는데 혼합색소의 8분면체는 부분 8면체로 더 나누어지며 이런 과정은 동질 색소의 블록을 얻을 때까지 반복되며 8개의 자노드를 갖는 트리로 표현되기 때문에 8진 트리구조로 나타난다.

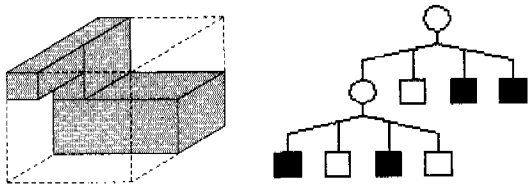
실제로 물체에 대한 3개의 직교방향 영상들(top view, front view, side view)로부터 8진트리를 생성하기 위해서는 입체교집합(volume intersection algorithm) 연산기법을 사용한다.

3개의 직교하는 4진 트리구조의 각 노드를 공간상 좌표로 교차시켜 8진트리 구조의 노드를 결정하는데 3개의 4진 트리구조의 노드가 모두 Black인 경우만이 8진 트리구조에서 black노드가 되고 4진 트리 구조중 1개 이상의 노드가 white라면 이 노드

는 white가 된다.

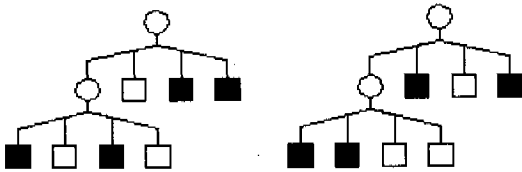
그 외의 경우는 gray노드로 voxel의 분할 과정을 반복적으로 진행해 나가며 8진 트리 구조의 노드를 결정해 간다.

8진 트리 구조의 각 노드에 해당하는 voxel을 공간상의 좌표로 바꾸어 물체를 생성하고 이 voxel의 6개의 표면중 물체의 표면에 속하는 면을 검출하여 물체 표면의 법선벡터와 입사광선이 이루는 입사각에 따라 물체의 의사 그레이 레벨값을 발생시켜 임의의 시각방향에서 본 3차원 물체를 합성해 낸다. (그림 5)



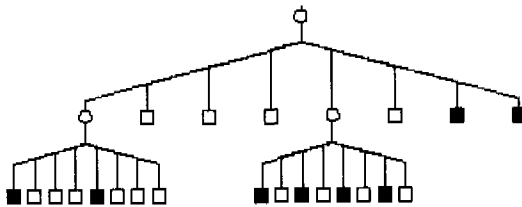
(a) 3D 물체

(b) 정면도

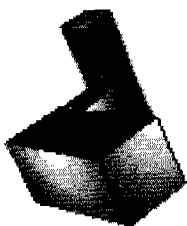


(c) 평면도

(d) 측면도



(e) 4진 트리 구조로부터 생성된 8진 트리 구조



(f) 8진 트리로 합성된 영상

그림 5. 합성 영상

III. 모의 실험 및 결과

본 논문에서는 Hough변환 과정에서 필요한 스테레오 영상에 의한 깊이 정보 복원 과정과 정합대상을 줄이기 위한 회전축 설정에 관한 부분에 대해서 모의 실험하였다. 모의 실험에 사용된 영상은 3차원 물체의 모델링에 의해 합성된 영상을 사용하였고(그림 6과 그림 7) 실험에 사용된 영상의 합성방법은 앞에서 소개했던 8진 트리 구조의 자료구조를 이용하여 합성하였다.



(a) 평면도



(b) 측면도



(c) 정면도



(d) (158,191,-203)



(e) (227, -119, -214)



(f) (212, 62, 273)

그림 6. 실험 영상(I)



(a) 평면도



(b) 측면도



(c) 정면도



(d) (158,191,-203)



(e) (227, -119, -214) (f) (212, 62, 273)

그림 7. 실험 영상 (II)

이렇게 합성된 영상에서 스테레오 기법을 이용해 물체의 깊이 정보(그림8)를 복원한 후 식(4)을 이용해 물체의 진행방향 벡터와 물체의 속도를 구하였다. 이 벡터는 항공기와 같은 경우 물체의 진행방향이 물체의 회전축이 될 수 있으므로 이 벡터를 물체의 회전축으로 삼고 이 축에 대해 회전방향을 결정해 물체의 정합에 이용할 수 있었다. 주어진 진행방향 벡터를 A, 계산된 진행방향 벡터를 B라 했을 때, 각도 오차는

$$\Delta\theta = \frac{\cos^{-1}(A \cdot B)}{|A||B|}$$

로 계산된다. 이를 나타낸 것이 표1이다.

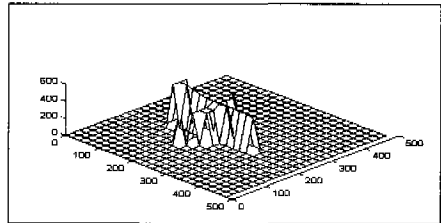
표 1. 각 항공기에 대한 3개의 시각방향에서의 결과

모델	진행 방향	계산된 진행방향	각도 오차
실험 영상 (I)	158, 191,-203	176, 197,-251	4.532
	227,-119,-214	254,-134,-256	1.856
	212, 62, 273	207, 71, 350	7.124
실험 영상 (II)	158, 191,-203	174, 205,-223	0.714
	227,-119,-214	263,-123,-285	4.993
	212, 62, 273	223, 67, 301	1.277

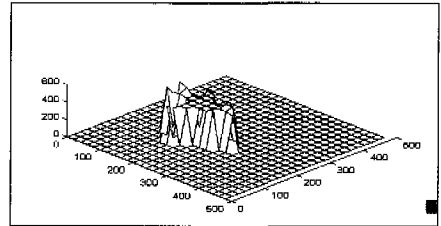
모의 실험의 오차계산 및 타당성의 검증은 위해 난수 발생기에 의한 X,Y,Z축의 난수를 발생시켜 20개의 임의의 공간상에 이동 물체를 위치시키고 표준 시각방향에서 바라보았을 때의 투영 영상을 얻어내어 위 알고리즘을 적용시켜 모의 실험한 결과 실험영상 I의 경우는 평균각도 오차, 최대각도 오차 그리고 최소각도 오차가 각각 3.606, 11.694, 0.931로 나타났다. 또한 실험 영상 II의 경우는 3.435, 7.831, 0.714로 나타났다. 여기서 알 수 있듯이 평균 오차가 샘플링 오차보다 작으므로 대체로 Hough변환을 이용한 정합 시 비교대상을 회전축을 기준으로 6도 내에 있는 5개의 회전축에 관한 회전

에 대해서만 정합을 시도해도 될 것이다.

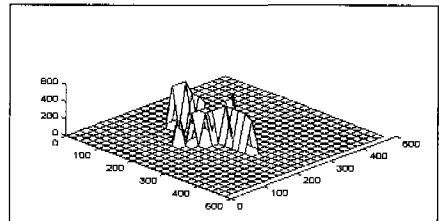
또한 오차 발생원인의 대부분이 스테레오 영상에 의한 잘못된 깊이 정보에서 비롯되었으므로 좀더 정확한 깊이 정보를 구할 수 있는 알고리즘을 적용한다면 정합 시 필요한 비교대상을 더욱 줄일 수 있을 것으로 생각된다.



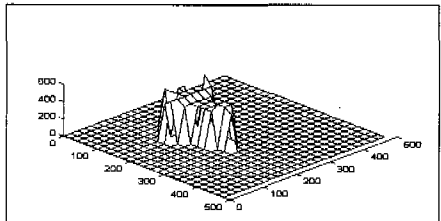
(a) 깊이정보(I)



(b) 깊이정보(II)



(c) 복원된 물체의 깊이정보(I)



(d) 복원된 물체의 깊이정보(II)

그림 8. 실험영상(I)과 (II)의 깊이정보

IV. 결론

본 논문에서는 Hough변환에서 물체의 깊이 정보를 얻기 위한 방법으로 인간의 시각 시스템과 유사한 스테레오 기법을 사용할 것을 제안하며 정합 시,

많은 회전 이동의 가짓수를 줄이기 위한 방법으로 물체의 진행 방향을 이용하여 시점 벡터를 구한 후 이 시점벡터를 회전축으로 삼아 그 축에 관한 회전 이동에 대해서만 물체를 정합하도록 하는 방법을 제안하였다. 3차원 물체를 계층적으로 묘사할 수 있는 기법인 8진 트리 구조 모델을 이용해서 물체를 합성해 내어 모의 실험한 결과 약 3.6도의 평균 오차만이 발생되어 제안한 방법의 효용성을 확인할 수 있었다.

참 고 문 헌

[1] D. Marr and T. Poggio, "A computational theory of human stereo vision" pp.301-328, 1979

[2] D. C. W. Pao and H.F.Li, "Shapes recognition using the straight line hough transform : theory and generalization", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.14, No.11, pp.1076-1089, 1992.

[3] T. M. Silberberg, L. Davis and D. Harwood, "An iterative Hough procedure for three dimensional object recognition", Pattern Recognition, Vol.17, No.6, pp.621-629, 1984

[4] Tzay. Y. Young, "Handbook of Pattern Recognition and Image Processing : Computer Vision", Vol.2, pp. 240-319, 1994

[5] Takeo Kanade, "A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment", IEEE Trans on Pattern Analysis and Machine Intelligence, Vol 16, 1994

[6] R. Krishnapuram and D. Casasent. "Determination of three-D object location and orientation from range images." IEEE Trans. on Pattern Analysis and Machine Intelligence. Vol.11, No.11, pp.1158-1167, Nov. 1989

[7] C. L. Jackins and S. L. Tanimoto, "Oct-tree and their use in representing three - dimensional object", CGIP Vol.14, pp.249 - 270, 1980

[8] 박영태, "Octree 모델에 근거한 고속 3차원 물체 인식", 1996

[9] 구자성, 이상욱 "Hough 변환을 이용한 캐드 기반 삼차원 물체 인식", 전자공학회논문집, pp. 35-44, 1995

[10] 박순영, 천성일 "RP코드 합성을 기반으로 한 세 방향 영상에서의 3차원 모델의 복원", 전자공학회지, pp. 106-114, 1994

박 성 만(Sung-Man Park) 학생회원
경희대학교 전자공학과졸

곽 윤 식(Yoon-Sik Kwak) 정회원
현재:충주대학교 컴퓨터공학과 조교수
제23권 6호 참조

이 대 영(Dae-Young Lee) 정회원
현재:경희대학교 교수
제 23권 6호 참조