

스테레오 기법을 적용한 3차원 물체인식 시스템

허윤석* · 홍봉화**

요 약

3차원 영상의 깊이 정보는 3차원 물체를 2차원 스크린에 투사시킴에 의해 손실될 수 있다. 만약 깊이 정보가 재 저장되고 3차원 물체 인식을 위해 사용된다면 더 효율적인 인식시스템을 만들 수 있다. 이 정보는 재 저장하기 위하여 스테레오 알고리즘으로부터 형상을 이용한다.

본 논문에서는 3차원 Hough 변환 영역을 3차원 물체 표현에 채택하여 3차원 물체 인식시스템을 제안한다. 본 시스템은 정합 시간 감소를 위하여 물체의 이동 벡터와 미지의 입력 영상을 8진 트리 구조로 이루어진 기준 영상과 비교하여 정합 단계를 이용한다. 8진 트리 구조 코드는 3차원 물체의 형태 기반에 사용되었다. 모의실험 결과 제안된 3차원 물체 인식 시스템은 만족할 수 있는 성능을 보였다.

1. 서론

3차원 물체인식 분야는 통상의 2차원 세계를 대상으로 한 화상 처리 문제와 비교하면 다음의 특징이 있다. 첫째, 하나의 장면이 많은 제약 조건하에서 영상으로 되기 때문에 영상 자체만으로는 그 장면을 회복하기에 충분한 정보가 제공되지 않으며, 실세계의 3차원 장면이 2차원 영상으로 투영될 때 물체의 거리 정보가 붕괴된다. 따라서 영상 자료가 갖는 모호성(ambiguity)을 해결하기 위하여 적절한 가설과 영상 자료에서부터 얻어지는 측정값을 이용하는 연구가 필요하다. 둘째, 보는 방향에 따라서 물체의 형상이 다르게 보이게 되는 문제를 해결해야 한다. 셋째, 하나의 영상이 형성되는 과정에는 많은 요소들이 작용하기 때문에, 예를 들어 하나의 물체에 관한 영상은 그 물체의 표면 물질, 주변환경, 발광체의 각

도, 주변의 빛, 카메라의 각도와 특성들에 의해 많은 영향을 받게 된다. 이러한 모든 요소들에 대해 한 화소의 명암도가 어느 정도 영향을 미치는지 고려하여야 한다. 넷째, 3차원 영상을 정합에 이용하기까지는 물체의 면들을 분류해 내기 위해 영상을 영역화 하고, 영역화된 부분을 표현할 수 있는 인자(parameter)를 추출해 내는 등의 전처리 단계가 필요하다. 이 과정은 잡음에 민감할 뿐만 아니라 계산 시간 면에서 전체 3차원 인식 시스템의 가장 큰 병목 구간으로 작용하게 된다.

이러한 특징을 가지는 3차원 물체 인식분야에서 3차원 Hough변환 공간을 이용하여 3차원 물체를 다룰 경우 전처리 단계 없이 바로 3차원 Hough변환 공간상에서의 정합이 가능하며 물체의 깊이 정보를 포함한 X-Y-Z 공간상의 한 벡터를 크기와 방향의 공간으로 변환시켜 다루기 때문에 인식 대상이 가지고 있는 정보의 손실이 적어 더 좋은 인식률을 나타내는 알고리즘 적용이 가능할 것이다.

본 논문에서는 Hough변환에서 필요로 하는

* 충청대학 전자정보과 교수
** 경희사이버대학교 조교수

물체의 깊이 정보를 얻기 위한 방법으로 인간의 시각 시스템과 유사한 스테레오 기법(shape from stereo)을 사용할 것을 제안하며, 정합시 비교하여야 하는 상당히 많은 회전 이동의 경우수를 줄이기 위하여 우선 물체의 진행 방향을 이용하여 진행방향벡터를 구한 후 그 벡터를 축으로한 회전이동에 대해 물체를 정합 하도록 하는 방법을 고안하였다. 그리고 표준 물체를 8진 트리 구조로 저장해 놓아 Hough변환에 의해서 얻어진 회전 이동량에 따른 2차원 투영영상을 합성해 내어 입력영상과의 정합 유무를 최종적으로 판별해 낼 수 있는 수단으로 이용하도록 하는 시스템을 제안한다.

II. 스테레오 기법

스테레오 기법이란 수평적으로 위치한 서로 다른 2대의 카메라로부터 같은 물체의 이미지를 얻어내어, 3차원 영상이 2차원 스크린으로 투영될 때 소멸된 물체의 거리 정보(depth information)를 복원하는 방법이다.

2.1. 스테레오 기법의 개요

우선 스테레오 기법을 사용하기 위해서는 입력영상은 다음과 같은 조건을 만족해야만 한다. 첫째, 두 카메라의 초점사이에는 적절한 거리 차이가 있어야 한다. 둘째, 영상은 구별되는 특징들이 나타날 수 있도록 충분히 무늬화(textured)되어야 하고 균일한 명암을 가지면 아니 된다. 셋째, 거리 정보는 두 초점에서 동시에 보이는 영상점에서만 계산되어야 한다.

스테레오 기법은 일정한 거리를 두고 위치한

두 카메라에서 각각 얻은 영상에서의 동일한 한 점에 대해서 시각 차이(disparity)가 나타나게 되고 두 영상에서의 이 시각 차이는 물체의 거리 정보에 반비례하여 나타나게 된다. 이러한 과정을 설명하기 위해 그림1과 같은 좌표구조를 생각한다.

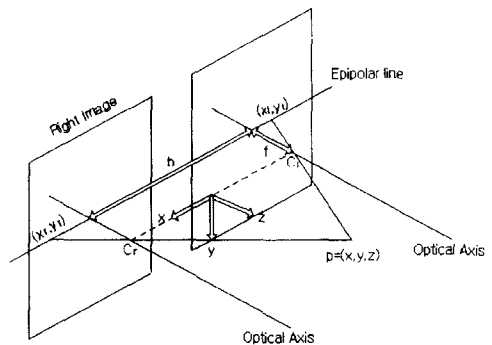
그림1에서 f 는 초점으로부터 이미지 평면까지의 거리이고, b 는 2대의 카메라 사이의 거리차이다. 위 그림으로부터 우리는 다음과 같이 쓸 수 있다.

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

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

$$z(x_l - x_r) = fb \quad (3)$$

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



(그림 1) 스테레오 기하 구조

이 식에서 $x_l - x_r$ 은 두 카메라에서 검출된 시각 차이(disparity)값으로써 물체의 거리 정보 z 에 반비례함을 알 수 있다. 따라서 이 식을 이용해 스테레오 영상의 거리 정보(depth map)를 복원할 수 있다. 이를 위하여 대부분의 스테레오 기법은 다음의 단계를 포함한다.

- (1) 왼쪽과 오른쪽 영상에서 특징점들을 찾는다.
- (2) 두 영상의 특징점들간의 대응관계를 찾는다.
- (3) 대응되는 특징점들로 부터의 시각 차이 (disparity)를 측정한다.
- (4) 변위로부터 거리 지도(depth map)를 얻어 3차원 표면을 재구성한다.

력영상과의 정합 유무를 판별해 내도록 하는 시스템을 연구하였다.

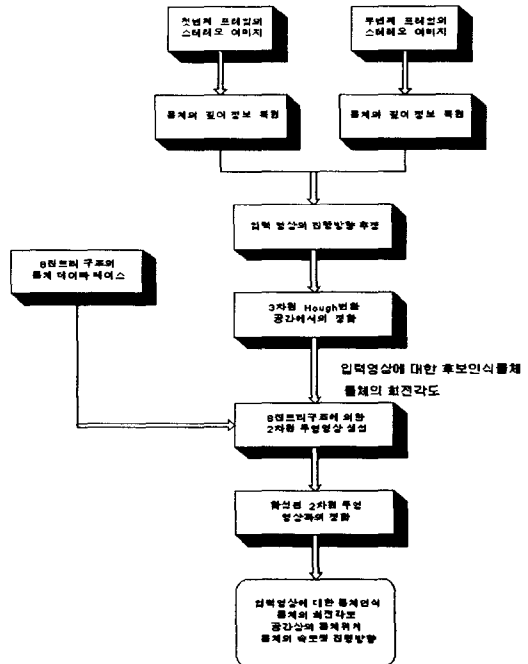
고안한 스테레오 기법을 이용한 3차원 물체인식 시스템의 전체 구성도는 그림 2와 같으며 이 시스템은 전체적으로 3부분으로 구성되어 있다.

- (1) 스테레오 기법에 의한 물체의 거리 정보 복원 부분
- (2) 3차원 Hough변환에 의한 영상 분류 및 정합 부분
- (3) 8진 트리 구조에 의하여 모델링 된 2차원 투영영상과의 정합 부분

III. 스테레오 기법을 이용한 3차원 물체인식 시스템

3.1. 개요

3차원 Hough 변환을 이용하여 3차원 물체인식을 구현할 때 필요한 물체의 깊이 정보를 복원하기 위하여 레인지 파인더를 사용한다. 그러나 레인지 파인더를 사용하는 방법은 먼 거리에 놓인 물체와 빠르게 움직이는 물체에 대해서는 적용하기가 쉽지 않다는 단점을 가지고 있다. 그러나 Hough변환에서 필요로 하는 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. 그러나 3차원 Hough변환에서 필요한 물체의 거리 정보 복원 문제가 해결되었다 할지라도 3차원 Hough변환에서는 물체 정합시 비교하여야 하는 많은 회전 이동의 경우 수에 대한 정합이 문제가 된다. 이와 같은 단점을 극복하기 위해 우선 물체의 진행 방향을 이용하여 진행방향 벡터를 구한 후 그 벡터를 축으로 한 회전이동에 대해 물체를 정합 하도록 하는 방법을 제안한다. 그리고 최종적으로 표준 물체를 8진 트리 구조로 저장해 놓아 Hough변환에 의해서 얻어진 회전 이동량에 따른 2차원 투영영상을 합성해 내어 입



(그림 2) 3차원 물체 인식 시스템 블록도

3.2. 물체의 진행 방향 추정

스테레오 이미지에 의해서 물체의 거리 정보가 복원되었으면 시간차를 갖는 두 frame간의 영상차를 이용해서 물체의 속도 및 진행 방향벡터를 결정 할 수 있으며, 이 벡터를 물체의

Hough변환시 시점벡터로 설정해 이 시점벡터를 회전축으로 한 회전과의 정합만을 계산하여 계산량과 메모리량을 대폭 줄일 수 있다. 예를 들어 Hough변환 공간상의 시점벡터에 대한 샘플링을 약 6도로 가정한 경우 Hough변환시 단위구 상에 존재하는 회전축의 수는 1102개가 되고 이에 대해 60개의 회전이 가능하므로 비교대상은 약 66000개가 된다. 반면에 회전축이 정해지면 비교대상을 회전축에 대한 회전만을 고려하면 되기 때문에 비교대상을 60개로 줄일 수 있다.

진행방향 벡터는 다음과 같은 공식에 의해서 구해진다.(그림 3)

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

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

$$\Delta z = z(t+1) - z(t) \quad (7)$$

● 등간격 $\Delta x = \frac{2\pi R}{60\text{개}}$ (8)

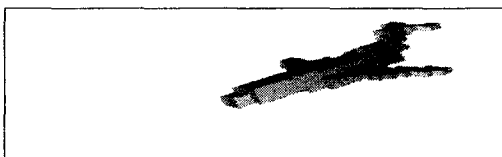
● 한 라인당 샘플링수 (9)

$$n = \frac{2\pi R'}{\Delta x} = \frac{2\pi R'}{\frac{2\pi R}{60}}$$

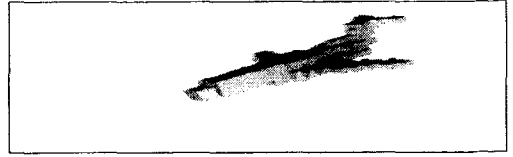
$$= \frac{R'}{R} 60 = \frac{R \times \cos\theta}{R} \times 60 = 60 \cos\theta$$

● 전체 샘플링수

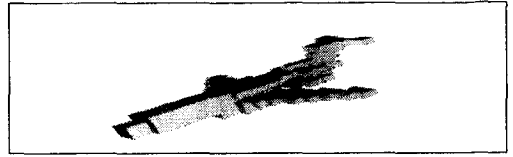
$$= 2 \times \sum_{\theta=0}^{90} 60 \cos\theta - 60 \quad (10)$$



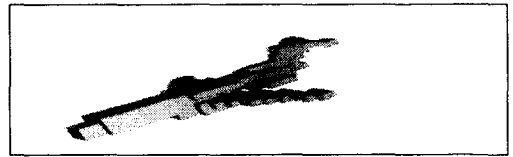
(a) n번째 프레임의 왼쪽 이미지



(b) n번째 프레임의 오른쪽 이미지



(c) n+1프레임의 왼쪽 이미지



(d) n+1프레임의 오른쪽 이미지

(그림 3) n, n+1 프레임에서의 스테레오 영상

IV. 모의실험 및 결과 고찰

본 논문에서는 Hough변환 과정에서 필요한 스테레오 이미지에 의한 깊이 정보 복원 과정과 정합대상을 줄이기 위한 회전축 설정에 관한 부분에 대해서 모의 실험하였다. 모의실험에 사용

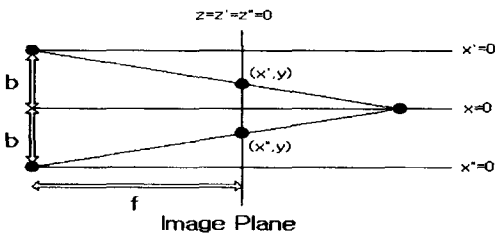
된 영상은 3차원 물체의 모델링에 의해 합성된 이미지를 사용하였고 실험에 사용된 이미지의 합성방법은 앞에서 소개했던 8진 트리구조의 자료구조를 이용하여 합성하였다.

모의실험에 사용된 프레임당 스테레오 이미지는 다음의 공식을 이용하여 만들었다.

$$x' = \frac{(x+h)f}{z} \tag{11}$$

$$x'' = \frac{(x-h)f}{z} \tag{12}$$

여기서 x' 은 스테레오 이미지의 왼쪽 이미지이고, x'' 은 스테레오 이미지의 오른쪽 이미지이다. $2h$ 는 카메라 사이의 거리이고 z 는 이미지의 x 점에서의 깊이 정보이다. 이것은 그림 4에서 볼 수 있다.



(그림 4) 스테레오 카메라 영상 시스템

이렇게 합성된 이미지로부터 스테레오 기법(shape from stereo)을 이용해 물체의 깊이 정보를 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 정합을 시도하는데 이용할 수 있었다.

모의실험의 오차계산 및 타당성의 검증을 위해 난수 발생기에 의한 X,Y,Z축의 난수를 발생

시켜 20개의 임의의 공간상에 이동물체를 위치시키고 표준 시각방향에서 바라보았을 때의 투영영상을 얻어내어 위 알고리즘을 적용시켜 모의 실험한 결과를 표 1에 나타내었다.

이 표에서 알 수 있듯이 평균 오차가 샘플링 오차보다 작으므로 대체로 Hough변환을 이용한 정합시 비교대상을 회전축을 기준으로 6도 내에 있는 5개의 회전축에 관한 회전에 대해서만 정합을 시도해도 될 것이다.

<표 1> 3 방향 벡터에서의 비행 물체 결과

모델	진행방향	계산된 진행방향	각도 오차	계산된 속도
	158, 191, -203	176, 197, -251	4.532	364.398
Plane1	227, -119, -214	254, -134, -256	1.856	384.718
	212, 62, 273	207, 71, 350	7.124	412.783

<표 2> 각 방향 벡터에서의 평균

모델	평균 각도 오차	최대 각도 오차	최소 각도 오차
Plane1	3.606	11.694	0.931



(a) 평면도



(b) 측면도



(c) 정면도



(d) (158, 191, -203)

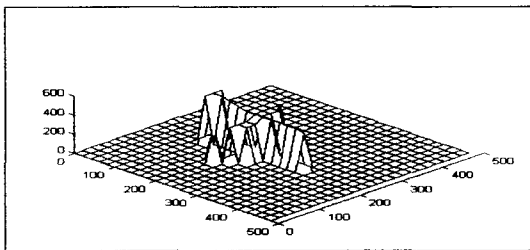


(e) (227, -119, -214)

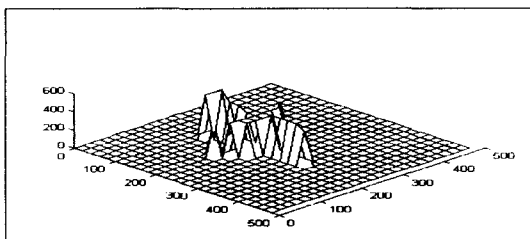


(f) (212, 62, 273)

(그림 5) 8진 트리에 의한 영상 모델링의 예



(a) 입력 영상의 깊이 정보



(b) 스테레오기법에 의해 복원된 물체의 깊이 정보

(그림 6) 비행 물체 1의 깊이 정보

V. 결론

본 논문에서는 Hough 변환에서 필요로 하는 물체의 깊이 정보를 얻기 위한 방법으로 인간의 시각 시스템과 유사한 스테레오 기법(shape from stereo)을 사용할 것을 제안하며 정합시 비교하여야 하는 상당히 많은 회전 이동의 가짓수를 줄이기 위한 방법으로 물체의 진행 방향을 이용하여 시점 벡터를 구한 후 이 시점벡터를 회전축으로 삼아 그 축에 관한 회전이동에 대해서만 물체를 정합 하도록 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 방법을 사용할 수 있음을 확인하였다.

참고문헌

- [1] Tzay. Y. Young, *Handbook of Pattern recognition and image processing: Computer vision, Vol.2*, 1994, 240-319.
- [2] R. Krishnapuram and D. Casasent. Determination of three-dimensional object location and orientation from range images, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(11), Nov. 1989, 1158-1167.
- [3] Takeo Kanade, A stereo matching algorithm with an adaptive window: Theory and experiment, *IEEE Transactions on pattern analysis and machine intelligence*,

- 16, 1994. 9.
- [4] D. C. W.Pao and H.F.Li, Shapes recognition using the straight line hough transform: Theory and generalization, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14(11), 1992. 11, 1076-1089.
- [5] R. Hoffman and A. K. Jain, Segmentation and classification of range images, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 9(5), 1987. 9, 608-620.
- [6] C. H. Chien and J. K. Aggarwal, Volume/Surface octrees for the representation of 3-D object, *CGIP*, 36, 1986, 100-113.
- [7] T. M. Silberberg, L. Davis and D. Harwood, An iterative Hough procedure for three dimensional object recognition, *Pattern Recognition*, 17(6), 1984, 621-629.
- [8] Pollard et al, PMF: A stereo correspondence algorithm using a disparity gradient limit, *Perception*, 1985, 14, 490-470.

Three-Dimensional Object Recognition System Using Shape from Stereo Algorithm

Yun-Seok Heo* · Bong-Hwa Hong**

Abstract

The depth information of 3D image lost by projecting 3D-object to 2D-screen for earning image. If depth information is restored and is used to recognize 3D-object, we can make the more effective recognition system. We often use shape from stereo algorithm in order to restore this information.

In this paper, we suggest 3-D object recognition system in which the 3-D Hough transform domain is employed to represent the 3-D objects. In this system, we use the moving vector of object to reduce matching time and In second matching step, the unknown input image is compared with the reference images, which is made with octree codes. Octree codes are used in volume-based representation of a three dimensional object.

The result of simulation show that the proposed 3-D object recognition system provides satisfactory performance.

Key words : stereo algorithm, recognition system, 3-D objects

* Dept. of Electronic information, Chung Cheong University

** Dept. of Information Communication Eng., Kyung Hee Cyber University