韓國 컴퓨터情報學會 2006하계학술발표논문집&학회지
第14卷 第1號, 2006. 6.

연속된 3차원 영상에서의 통계적 물체인식

장 대 식*, 이 양 원*, 성 국 서*

Probabilistic Object Recognition in a Sequence of 3D Images

Jang Dae Sik*, Yang Won Rhee*, Guo Rui Sheng*

요 약

냉장고나 에어컨 등과 같은 비교적 크고 자주 움직이지 않는 물체들에 대한 인식은 실내 환경에서의 SLAM (Simultaneous Localization and Map building) 문제에서 중요한 전역적 고정 특징으로 사용될 수 있다는 측면에서 그 필요성이 크다. 본 논문에서는 연속적으로 획득되는 3차원의 영상 장면들을 사용하여 이러한 큰 물체들을 안정적으로 인식할 수 있는 방법을 제안한다. 제안하는 방법에서는 파티클 필터(Particle Filter)를 기반으로 연속적인 3차원 영상에서 점진적으로 3차원의 물체를 인식하는 방법을 사용한다. 이를 위해 인식하고자 하는 하나의 물체를 표현하는 파티클(Particle) 들을 3차원의 장면에 뿌리고, 3차원 선들의 정합을 통해 각 파티클에 대한 정합 확률을 계산한다. 이 확률과 정합된 파티클의 비율을 기반으로 3차원 환경 속에 놓여진 물체를 인식할 수 있으면 물체의 자세 또한 함께 인식될 수 있다. 실험 결과를 통해 파티클 필터에 기반한 점진적이고 확률적인 물체인식의 가능성을 보이고 SLAM 문제에 응용한 결과도 함께 보여준다.

Abstract

The recognition of a relatively big and rarely movable object, such as refrigerator and air conditioner, etc. is necessary because these objects can be crucial global stable features of Simultaneous Localization and Map building(SLAM) in the indoor environment. In this paper, we propose a novel method to recognize these big objects using a sequence of 3D scenes. The particles representing an object to be recognized are scattered to the environment and then the probability of each particles is calculated by the matching test with 3D lines of the environment. Based on the probability and degree of convergence of particles, we can recognize the object in the environment and the pose of object is also estimated. The experimental results show the feasibility of incremental object recognition based on particle filtering and the application to SLAM

▸ Keyword : 물체인식(Object Recgnition), 파티클필터(Particle Filter), SLAM(SLAM)

• 제1저자 : 장대식
* 국립군산대학교 컴퓨터정보과학과

Ⅰ. 서 론

Object recognition has been one of the most challenging issues in computer vision and intensively investigated for several decades. In particular, the object recognition has played an important role for manipulation and SLAM in robotics fields.

Many researchers suggested various 3D object recognition approaches. Among them, the model-based approach mentioned in this paper is the most general one for recognizing shapes and objects. It recognizes objects by matching features extracted from an object on the scene with the stored features of the object in advance[1]. Some famous model-based recognition studies are as follows.

The method suggested by Fischler and Bolles [2] uses RANSAC for recognizing objects. It projects points of all models on the scene and decides if the projected points are similar to those of the captured scenes and recognizes the object based on the similarity. This method is not so efficient because the procedure including assumption and verification is repeated many times to get an accurate result

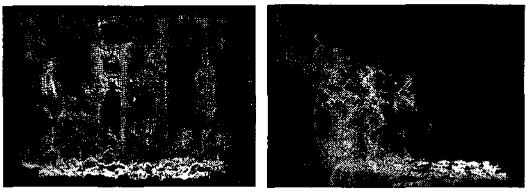
In addition, Johnson and Herbert [4] proposed a spin-image based recognition algorithm in cluttered 3D scenes and Andrea Frome et al. [3] compared the performance of 3D shape contexts with that of spin-image. Jean Ponce et al. [5] introduced a 3D object recognition approach using affine invariant patches. However, these methods are working well only when accurate 3D data or fully textured environments are provided, while our approach makes it possible to recognize objects when there are lots of noises and uncertainties in the captured scenes stemming from low-quality sensors.

In this paper we propose a new approach to recognize big and rarely movable objects in a sequence of images by applying a probabilistic model in noisy and texture-less environments.

Ⅱ. Extraction of 3D Line Features

We use 3D lines as key features for object recognition because 3D data can be obtained robustly on the boundaries of objects even in texture-less environments as shown in Fig. 1(a)

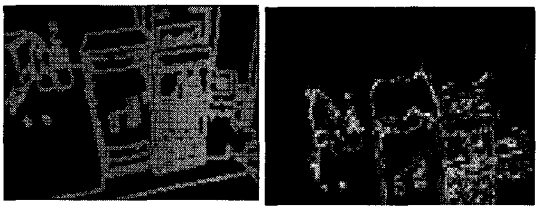
Due to the poor accuracy of 3D data as illustrated in Fig. 1(b), all lines are firstly extracted from 2D images and the 2D lines can be transformed to 3D lines by mapping corresponding 3D points (2D images and their corresponding 3D points are captured at the same time from a stereo camera).



(a) 3D data from stereo camera (b) Noisy 3D data

Fig. 1. 3D data

(a) Experimental environment (b) Edges extracted



(c) 2D lines (d) 3D lines

Fig. 2. Results of the line feature extraction in 2D and 3D

We made a simple algorithm to find 2D lines based on the edge following approach so that we could find most of lines in a scene efficiently. First of all, the edges are drawn by canny edge detection algorithm. Then, we categorize the edges as horizontal, vertical and diagonal line segments based on the connectedness of edges. 2D lines are found by connecting each line segments with adjoining line segments considering the aliasing problem. 3D lines can be obtained if there are corresponding 3D points at the pixels of 2D lines. The 2D lines are transformed into 3D lines by assigning the 3D positions of the corresponding 3D points. Fig. 2 shows the results of line extraction in 2D and 3D.

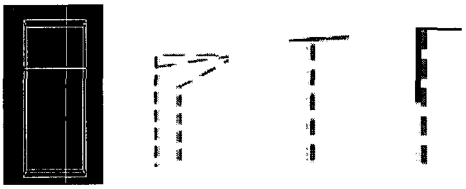
III. OBJECT RECOGNITION BASED ON PARTICLE FILTER

3.1 Initial Particle Generation

An object to be recognized is modeled with 3D lines and we consider this line set as a particle. Many particles which represent this object will be spread into possible positions of objects in a 3D scene based on the initial particle generation. Fig. 3 (a) illustrates the particle of an objects.

At every scene, the particles are initially generated to find other possible positions of an object, which could not be extracted in previous scenes.

Fig. 3 shows how particles are generated by using directional features of lines. In case of (b), many particles can be generated by rotating vertical lines based on core vertical line. All particles located apart from floor are eliminated because refrigerator does not stand up with certain gap from floor



(a) The model of a refrigerator (b) Vertical line (c) Horizontal line (d) Relationship of lines

Fig. 3. A particle of one object and the initial particle generation using lines (dot : particles generated, solid: 3D line in the scene)

3.2 Determination and updating of the probabilitiesof particle

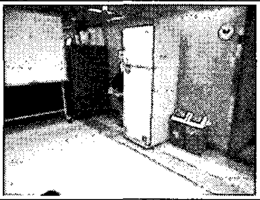
The probability is obtained after getting positive and negative similarity of each particle by using 3D line in space. The positive similarity shows how much the lines composed of particles are matched with lines in space after projecting them into space. It is decided by following two elements.

The first element is S1 and it is the degree that lines composed of particles are matched with lines inspace. S1 is determined as follows. It is tested if there is a 3D line around each line of a particle or not. And the similarities in length, orientation, and distance of the corresponding lines are verified. The second element is S2. It shows how many lines of a particle are matched with the lines in space. These two elements of similarity S1 and S2 are integrated by weighted sum.

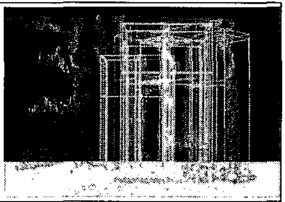
On the other hand, the negative similarity shows how lines in space are matched with those of particles. For example, if an air conditioner having the same shape and dimension with refrigerator is existing in space, the positive similarity of the air conditioner is identical to that of a refrigerator. But in this case, the number of lines included in the air conditioner is greater than that of lines comprised in a refrigerator due to the geometric shape difference. All particles at time t-1 should be propagated to scene at time t to update the probability of particles. Probabilities of particles newly founded at time t and existing particles are updated based on (1).

$$P_t(n) = \sum_{n=0}^{N_p}\sum_{m=0}^{N_p}\frac{1}{d}P_{t-1}(m)P_t(n) \qquad (3.1)$$

where d is the difference of the pose of $P_{t-1}(m)$ and $P_t(n)$.



(a) 2D image of environment



(b) The particles at first scene

(c) The particles at second scene

(d) The particles at third scene

Fig. 4. The distribution of particles in a sequence of images (The particles are represented by green boxes)

After that, particles are sampled again according to the probabilities of them. Particleswith high probability can generate more particles and particles, while those with low probability are disappeared.

Fig. 4 shows how particles are updated in continuous scenes.

IV. EXPERIMENTAL RESULT

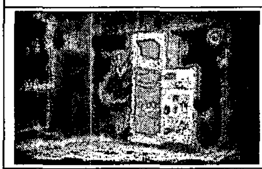
This paper aims to make a robot to know the location of large objects such as refrigerators when the robot is navigating. We assume that all objects to be applied for this approach should have majorstraight line features enough for recognition. In order to get 2D images and 3D data, a stereo camera named as Bumblebee is used. The stereo camera is mounted on the end effecter of an arm with an eye-on-hand configuration. Fig. 5 shows the stereo camera and the robot used forexperiment and an eye-on-hand configuration.

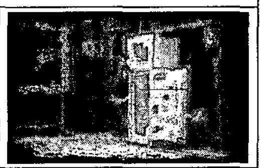
(a) Stereo camera (b) Robot (c) Eye on hand

Fig. 5. Equipments for the experiment

Fig. 6 shows experimental results with different viewpoints. Three sequences of images were used and the robot is looking at and approaching the object from different directions of 0, 30 and 60 degree respectively. Although the probability of each particle is getting lower if the angle between the robot and the object is increasing from 0 to 60 degree, these results are acceptable since the probabilities of particles are over a predefined threshold of 0.7. The blue box in the scene means the estimated model position after recognition.

(a) 0도





(b) 30도 (c) 60도

Fig. 6.The recognition results from different viewpoints

Fig. 7 shows the result in the presence of occlusions. A mannequin was put in front of the object to make occlusions. Even though the mannequin stands in front of the refrigerator, the position of the object was recognized successfully with 3 consecutive scenes.

(a) Occluded by mannequin	(b) Occlusion with another view angle (30도)

Fig. 7. The recognition results with static occlusions

V. Conclusion

In this paper, the method to recognize objects and estimate their poses by using sequential scenes in noisy environments is suggested.

Under the assumption that we already know the object to be recognized, particles of the object are scattered into the 3D scene and the probability of each particle is determined by matching 3D lines. And the probabilities of particles are updated in the same way after reading the next scene and then the object is measured and its pose is estimated.

This method can be applied to recognize large objects such as refrigerators, air conditioners and bookcases that have many line features. It is proved byexperiment that this method is robust to orientation changes and occlusions. Moreover, it can be used to perform SLAM more reliably by providing the positions and poses of the recognized objects as landmarks.

참고문헌

[1] Farias M. F. S., Carvalho J. M.: Multi-view Technique For 3D Polyhedral Object Rocognition Using Surface Representation. Revista Controle & Automacao pp. 107-117, 1999

[2] Fischler M. A., Bolles R. C.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Comm. Assoc. Comp. Mach, Vol. 24 No. 6 pp. 381-395, 1981

〔3〕 Frome A., Huber D., Kolluri R., Bulow T., Malik J.: Recognizing Objects in Range Data Using Regional Point Descriptors. European Conference on Computer Vision, Prague, Czech Republic, 2004
〔4〕 Johnson A. E., Hebert M.: Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes.IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21 No. 5 pp. 433-449, 1999
〔5〕 Fred R., Svetlana L., Cordelia S., Jean P.: 3D Object Modeling and Recognition Using Affine-Invariant Patches and Multi-View Spatial Constraints. CVPR pp. 272-280, 2003

저자소개

장 대 식
1999년 2월 : 숭실대학교 컴퓨터공학박사
1999년 ~ 20002년 : 줄라이네트(주) 연구팀장
2002년 ~ 2004년 : (주)인컴아이엔씨 연구팀장
2004년 ~ 2005년 : 성균관대학교 연구교수
2005년 ~ 현재 : 국립군산대학교 교수

관심분야 : 컴퓨터비전, 패턴인식, 3차원 물체인식, 로봇 지능