論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	A Robust Visual-Feature-Extraction Method in Public Environment
著者(和文)	ｶｺﾞｾｰ
Author(English)	Gangchen Hua
出典(和文)	学位:博士（工学）, 学位授与機関:東京工業大学, 報告番号:甲第9871号, 授与年月日:2015年3月26日, 学位の種別:課程博士, 審査員:長谷川　修,新田　克己,三宅　美博,小野　功,長谷川　晶一
Citation(English)	Degree:, Conferring organization: Tokyo Institute of Technology, Report number:甲第9871号, Conferred date:2015/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

（博士課程）
Doctoral Program

論 文 要 旨

THESIS SUMMARY

専攻： Department of	知能システム科学　専攻	申請学位（専攻分野）： Academic Degree Requested	博士 Doctor of　（工学　）
学生氏名： Student's Name	Gangchen Hua	指導教員（主）： Academic Advisor(main)	長谷川　修
		指導教員（副）： Academic Advisor(sub)	

要旨（英文 800 語程度）

In this study we describe a new feature extracting method that can extract robust features from a sequence of images and also performs satisfactorily in a highly dynamic environment. This method is based on the geometric structure of matched local feature points. When compared with other previous methods, the proposed method is more accurate in appearance-only simultaneous localization and mapping (SLAM). When compared to position-invariant robust features, the proposed method is more suitable for low-cost, single conventional-lens cameras with narrow fields-of-view.

We tested our method in an outdoor environment at Shibuya station. We captured these images by using a conventional hand-held, single-lens video camera. These environments of experiments are public environments without any planned landmarks. The results show that the proposed method can accurately obtain matches for two visual-feature sets, and using the proposed method, stable and accurate appearance-only SLAM can be achieved in public dynamic environments.

In chapter 6, we show the proposed method based hybrid SLAM and visual odometry also work better than other previous methods in public dynamic environments.

Simultaneous localization and mapping (SLAM) is widely used to generate maps for localization or autonomous robotic navigation.

Appearance-only SLAM is a type of low-cost solution. Moreover, SLAM based on visual features has abundant information that can be used for matching and recognition.

There are two kinds of visual features, including local visual features and global visual features. Global visual features ([2]etc.) extract one feature for each image and are used for vision-based generic recognition (object classification, scene classification etc.) generally. Typical local visual features extract multiple feature points at each image's multiple interesting locations. Local visual features based solutions are widely used for vision-based recognition (robotic navigation, SLAM etc.). Because the vision-based SLAM is a kind of specific recognition, we have to use the local visual feature. Scale-invariant Feature Transform (SIFT)[3] and Speeded Up Robust Features(SURF)[4] are typical local visual features.

SIFT calculates scale-invariant features at interesting locations, these features are 128-dimension vectors. The feature matching and indexing process uses a modification of the k-d tree algorithm called the Best-bin-first search method that can identify the nearest neighbors with high probability using only a limited amount of computation.

SURF is similar to SIFT, however, for faster processing SURF's feature is 64-dimension. The proposed method uses SURF to track feature points.

M. Cummins et al. proposed a rapid method based on the probabilistic bail-out condition for appearance-only SLAM. It is called FAB-MAP. An o_ine dictionary need to be generated before running, so FAB-MAP is not a complete online incremental solution.

FAB-MAP uses SURF as local visual feature. F However, the appearance of objects in the actual world is always dynamic. Many appearance-only SLAM methods are based on the hypothesized static environment. These methods use SIFT or SURF. These local-visual features do not have a strong innocence on moving objects such as walking humans in cafeterias, stations, or shopping malls.

For appearance-only SLAM, the visual features' robustness and the effectiveness of the matching are important.

A. Kawewong et al. proposed a method that tracks robust features in a sequence of images, called position-invariant robust features (PIRF). PIRF extracts common features by referring to past images. So it is more robust than original SIFT or SURF in dynamic environments. In addition, A. Kawewong et al. proposed two online-incremental-

Appearance-only methods for SLAM PIRF-nav and PIRF-nav2.0 on the basis of PIRF. The methods in PIRF-nav and PIRF-nav2.0 perform better than in dynamic environments.

However, PIRF-nav is based on SIFT , FAB-MAP and PIRF-nav2.0 are based on SURF. Because they match and index features only on basis of n-dimensional nearest neighbors , they are called pure bag-of-words (BoW) methods.

To avoid this problem, we propose a method called Incremental Center of Gravity Matching (ICGM) that uses relative geometric structure of local feature points to track robust features in a sequence of images. Fig. 1.3 shows the ICGM's basic algorithm to distinguish incorrect matches and correct matches. Similar to PIRF, we propose a SLAM method that extracts robust features by referring to past image (I_{t-1}) on the basis of ICGM. It is called single-directional ICGM (Fig. 4.2). It works better than PIRF.

Moreover, because of the reasons described in 2.1.3 and 2.1.4. PIRF as well as single-directional ICGM always causes significant loss of features. So we also proposed a method called double-directional ICGM. Double-directional ICGM extracts robust features not only by referring to past image (I_{t-1}) but also by referring to future image (I_{t+1}).

Incremental Center of Gravity Matching (ICGM) is a novel method to extract robust visual-features from sequence of images in in crowded public environment. Using the proposed method, dynamic

objects (pedestrians, cars etc.) can be ignored.
Robust features extracted by the proposed method can improve performance of vision based localization.
This paper descried the basic algorithm of ICGM and its applications: appearance-only SLAM, hybrid SLAM and visual odometry.
Results of appearance-only SLAM, hybrid SLAM and visual odometry shows that the proposed method: Incremental Center of Gravity Matching (ICGM) works good in crowded public environment.
Experiments results proved that ICGM is an effective matching and feature extraction method. I believe that in the future, the proposed method should be more widely used.
Today, high-performance hand-held smart phones have become very popular. For the proposed method to possess high robustness while using hand-held devices, ICGM maybe applied to many types of platforms (including hand-held smart phones) for navigation by pedestrians. This is our future goal.
Deciding the threshold Thr automatically is one of our important future study.
When compared to PIRF-Nav 2.0, the processing speed of the proposed method is possibly relatively fast. However, we are currently considering replacing SURF with a type of corner detector (FAST, HARRIS etc.). Although SURF, SIFT possesses more information for visual recognition, FAST or HARRIS is faster.
Furthermore, ICGM can extract robust corner detectors on the basis of their geometric structure. Therefore, we intend to use FAST, HARRIS for faster processing.