

Research on the Visual Navigation Methods based on Scene Modeling and Environmental Understanding

Guo Binghua and Cen Zhisong

Department of Electronics and Information Engineering, Zhaoqing University,
Zhaoqing Guangdong 526061, China.

Keywords: Visual Navigation, Scene Modeling, Environmental Understanding, Mathematical.

Abstract. In this paper, we research on the visual navigation methods based on scene modeling and the general environmental understanding. Computer vision's ultimate goal is to fully understand the natural scene, it involves several tasks at the same time, target detection meaningful region labeling and 3D reconstruction and image understanding as one of basic work has been one of the research hotspots and difficulties. Image contains a large amount of the information, but also more redundant information, the key to extract essence can correctly interpret images information become one of the important research content features said. Under this need, we propose the scene modeling and general environmental understanding based method for visual navigation that is meaningful.

1. Introduction

Based on navigation line and road signs of visual navigation is one of the simplest applications, the robot working in a structured environment, there is a clear navigation line and road signs reference, the robot information extracted from the image navigation line and road signs, and then the real-time control. In this way, the requirements of the image processing are simple, fast processing speed, good real-time control while considering the characteristic of this style of navigation and the complexity of system building and cost-effective and many robots using monocular vision system for navigation.

Visual navigation according to the visual navigation sensor types can be divided into passive and the active vision navigation. Passive vision navigation is dependent on the visible and invisible light imaging technology, the method of CCD camera as a typical form of passive imaging sensors, widely used in various kinds of general visual navigation system. Visual navigation and positioning system according to the dependence of the map can be divided into the: map navigation system, navigation system and the no map navigation system that can be organized as the follows. (1) The initial robot navigation method is developed for the indoor robot, and the early development is navigation method based on the map. Early attempts are prominent features of basic environment using two-dimensional projection annotation in the map. (2) With the deepening of research, how to independent real-time detection in the unknown environment has become the research focus that needs to solve problem of autonomous navigation, map, and positioning. Positioning and drawing and drawing at the same time and at the same time positioning method to solve the unknown environment to complete the drawing real-time positioning, and the positioning tasks at the same time. (3) No map navigation method does not need information to conduct comprehensive description on the environments and the optical flow method, the navigation method based on feature tracking or appearance is not main research direction of visual navigation map method through image sequence to estimate optical flow method or feature point movement, in the process of vehicle movement with the field of view changes [1-2].

Under this background, in this paper, we will propose the novel algorithm on the visual navigation methods based on scene modeling and environmental understanding.

2. Our Proposed Methodology

2.1 Scene Modeling Methodology.

Image understanding is a research about the computer system design and experiment, through the adoption of one or more method features extracted from the image, and use a certain control structure model and algorithm of feature matching, and appropriate form of image analysis and interpretation. Image understanding of the semantic analysis in the field of application of the potential is huge. In the image of basic rich semantic knowledge can provide more accurate image search engines, to generate intelligent digital image photo albums and the virtual world of the visual scene.

Most current context method mainly analysis the image of the semantic context and spatial context as semantic context can be obtained from the other two kinds of context reasoning compared with the scene target symbiosis. The change of the scale and spatial context scope is larger, and the knowledge is more likely to get a symbiotic relationship, and calculates faster. Blend in image semantic context features formed the global and local two analysis strategy, based on the context scenario analysis and context analysis based on the target that satisfy the listed condition [3].

$$z = \arg \max_z P(z | d) \quad (1)$$

One obvious strategy is based on this model into the geometry or space relation, namely uses the topic at the same time the corresponding visual description of the semantic features and basic location information as the location information contained in the model explicitly semantic feature that reflect the relationship between the space constraints which has better effect analysis. The weight of heavy effect of classification and effect is poor as the integration of learning result is weighted combination of weak classifier which could be formulized as equation 2.

$$\sum_{t=1}^T \sum_{x \in D} a_t h_t(x) \quad (2)$$

The assumption is not restrictive conditions, because of any representation, contain a great group can always by redefining great group of potential function for it contains all the mass of the product of the basic potential function, which is transformed into only contains the great mass of representation and sometimes in order to calculate conveniently that often with the great mass of representation.

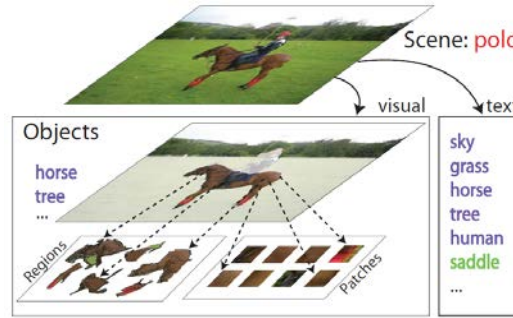


Fig. 1 The Illustration of the Scene Modelling Procedures

2.2 Image based Modeling.

Color image has become a common form of image that the color image recognition is the focus of research. Color image edge detection, threshold selection is very important. Currently widely used 3D scanning technology can accurately obtain geometric model of the real scene, but at the same time usually cannot obtain color information of the model. In recent years, primary image based modeling technology that has been widely concerned. In the model data acquisition phase, to get to model the appearance of the object data, such as the width dimension, depth data. In the field of the research and application of different, according to different requirements, the method to get the data types and also quite different, and use of equipment and the process of data modeling also have difference.

In the frame of the Euclidean reconstruction, we can be the shortest distance between two straight lines in the middle as their intersection point, but in the frame of the projective reconstruction, this conclusion is not established. To this end, we can turn triangulation steps to image space according to the listed features and expressions [4].

$$\min \left(D(m, P_1^{projection} M^{projection}) \right) \quad (3)$$

As mentioned above, more than just by corresponding relationship between the feature points on the image to get the feature points of the projective reconstruction, but the projective reconstruction results cannot reflect the real scene of space shape, also is unable to meet the requirements of people on the vision, so it is necessary to upgrade for the metric reconstruction as the follows.

$$m_{reconstructed} = P_k^{projection} Q Q^{-1} M \quad (4)$$

With the depth of the image data for study of human motion detection is usually a reference image that is commonly used in ordinary human motion detection algorithm, there are mainly based on the human body feature matching method based on background subtraction and division two ideas.

2.3 Environmental Perception.

Unmanned ground platform is a set of environmental perception, dynamic decision-making and the planning, behavior control and execution and other functions in a body's comprehensive system. Among them, the environmental awareness is unmanned platform system is the premise and basis for decision making and planning, in the study of the unmanned ground platform system, occupies very important position and platform is unknown, the work environment can be known, the fixed structure environment that also can be complicated and changeable unstructured environment.

- Unmanned ground platform environment perception system to complete the basic multimodal sensor collected a large number of redundant data fusion processing as the first condition is to calibration and the sensors for each sensor to the establishment of coordinate system, and the transformation between the coordinate system.
- In the process of the path detection and recognition, the major effect of the basic visual image processing on the path detection accuracy in addition to the image gradient edge blur, irregular shape, the main sheltered from the influence of shadow.
- Color and shape detection has to sign for a degree of the classification, but the same color and shape of road signs still contain a variety of signs, such as blue, circular road signs. Therefore, need to extract feature parameters to achieve further classification.

Of the unmanned ground platform based on visual information processing technology in military weapons and equipment, transportation, health care, and all aspects of the daily life that has a broad application prospect and development space with increase of application requirements, environment of the platform for the unmanned ground vehicle carrier is more complex, changing operation mode, operation speed is more quickly, giving environment perception of visual information more and more data, and in such the aspects as information processing efficiency and accuracy are also put forward higher requirements, as a result, problems to be solved are also more and more accordingly and the final version of the challenge could be organized as the formula 5.

$$G_{finalized} = \operatorname{argmin} \{ \|x_i - x\| \} \quad (5)$$

2.4 Visual Navigation Approaches.

At present, the way of visual navigation is generally can be divided into rail line guide, roadside line extraction and pavement area segmentation, etc. The so-called guide trajectory is planned in the advance the robot to walk line that make the robot followed the movement. Using the advantages of this method is simple and can adapt to various road environment that do not need other reference. The difficulty is how to accurately will guide line partition. In general, the brightness of the pavement is relatively large, the lower part and gathering in the middle of the image for road, so the road region segmentation method can set off from chain code tracing method is adopted to image segmentation into two parts, and the road and then let the robot movement in the middle of the road all the time, and this method use is the premise of the brightness of the road and the road differ significantly [5].

$$\delta r_{eck}^e = 0, \quad \Theta_{cke}^e = 0 \quad (6)$$

However many state constraint Kalman filter will feature points with basic geometric constraint relationship between the sequence images directly into geometric constraint relationship between the sequence images as the formula 6. This method avoids the large number of feature points location parameter estimates and effective use of the general feature points in consecutive images more frames observation information as geometric constraints between the sequence images represent the camera

position, posture variation over time, thus combination can effectively inhibit the position and pose error of inertia navigation system. Use camera model, calculated, according to the linear relationship will affect the accuracy of the calculation results and aiming at this problem, the author designed a simple correction algorithm, by calibration in advance some space in the world coordinate system, the coordinates of piecewise linearization measurements while this can compensate for the effects of the distortion in the very great degree, the algorithm of real-time and precision can satisfy the demand of the mobile robot vision navigation and positioning.

3. Simulation and Verification

In this section, we simulate the proposed method. WordNet network provides a standard way of the semantic description that should be included in the image semantic understanding conform to the prior knowledge of WordNet structure and grammar rules, the image contains the intrinsic semantic information. The current number of image library has reached the standard and conform to WordNet class hierarchy, has extensive target category information, and associated information. In the listed figure two, we illustrate the scene understanding simulation result.

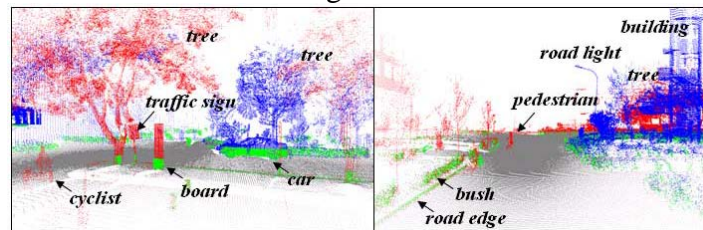


Fig. 2 The Demonstration of the Scene Understanding Simulation Result

4. Summary

In this paper, we research on the visual navigation methods based on scene modeling and general environmental understanding. This paper mainly summarizes the two-dimensional image of image semantic understanding, and more basic multi-dimensional image collection contains richer semantic relevance knowledge, to extract valuable information as well as the realization of the semantic image understanding which will expand the research field of image understanding, at present has become the development direction of image semantic analysis and the research hot spot. In the future research, we will combine more machine learning and data mining algorithms for systematic optimization.

Acknowledgement

This manuscript is financially supported by the special innovation fund projects in the Guangdong province (2014KTSCX191).

References

- [1]. Wang, Huayan, Stephen Gould, and Daphne Roller. "Discriminative learning with latent variables for cluttered indoor scene understanding." *Communications of the ACM* 56.4 (2013).
- [2]. Wojek, Christian, et al. "Monocular visual scene understanding: Understanding multi-object traffic scenes." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 35.4 (2013).
- [3]. Mottaghi, Roozbeh, et al. "Human-Machine CRFs for Identifying Bottlenecks in Scene Understanding." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 38.1 (2016).
- [4]. Wang, Haoxiang, and Jingbin Wang. "An effective image representation method using kernel classification." *Tools with Artificial Intelligence (ICTAI), 2014 IEEE 26th International Conference on. IEEE, 2014.*

- [5]. Li, Congcong, Devi Parikh, and Tsuhan Chen. "Automatic discovery of groups of objects for scene understanding." *Computer Vision and Pattern Recognition (CVPR)*.