Robotics and Autonomous Systems 69 (2015) 40–51

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

Metric-based detection of robot kidnapping with an SVM classifier

Dylan Campbell [a,b,*], Mark Whitty [c]

[a] National ICT Australia (NICTA), Locked Bag 8001, Canberra ACT 2601, Australia
[b] Research School of Engineering, Australian National University, Canberra ACT 0200, Australia
[c] School of Mechanical and Manufacturing Engineering, UNSW Australia, Sydney NSW 2052, Australia

HIGHLIGHTS

- Two methods for kidnap detection using local pose estimation techniques are proposed.
- At least two independent ways of estimating relative pose are required.
- Metrics assessing the quality of a pose estimate are developed and evaluated.
- For applications with limited training data, a joint classifier performs well.
- If a large training dataset is available, an SVM classifier is more accurate.

ARTICLE INFO

Article history:
Available online 26 August 2014

Keywords:
Mobile robots
Failure detection
Robot programming
Support Vector Machine

ABSTRACT

Kidnapping occurs when a robot is unaware that it has not correctly ascertained its position, potentially causing severe map deformation and reducing the robot's functionality. This paper presents metric-based techniques for real-time kidnap detection, utilising either linear or SVM classifiers to identify all kidnapping events during the autonomous operation of a mobile robot. In contrast, existing techniques either solve specific cases of kidnapping, such as elevator motion, without addressing the general case or remove dependence on local pose estimation entirely, an inefficient and computationally expensive approach. Three metrics that measured the quality of a pose estimate were evaluated and a joint classifier was constructed by combining the most discriminative quality metric with a fourth metric that measured the discrepancy between two independent pose estimates. A multi-class Support Vector Machine classifier was also trained using all four metrics and produced better classification results than the simpler joint classifier, at the cost of requiring a larger training dataset. While metrics specific to 3D point clouds were used, the approach can be generalised to other forms of data, including visual, provided that two independent ways of estimating pose are available.

1. Introduction

Kidnapping occurs when a robot fails to correctly ascertain its position through dead-reckoning or other relative localisation techniques [1]. If not detected and resolved, kidnapping causes many existing localisation and mapping algorithms to malfunction. This results in incorrect maps and position estimates, which may render the robot unable to perform its function. In addition, a kidnapped robot may unwittingly perform dangerous actions, such as driving towards previously mapped hazards. While kidnapping may not occur regularly, the severity of the consequences is such that it is incumbent upon the robot designer to implement a system for kidnap detection.

When kidnapping occurs, the robot is unaware that its pose estimate is incorrect. It would be inefficient to run global localisation on every occasion, because a reasonable pose estimate is only unavailable when kidnapped and methods that incorporate this estimate are much quicker and more suited to real-time applications [2]. Hence, a means to detect kidnapping events is mandatory for an efficient mapping and localisation system that is robust to kidnapping events.

Extending the classification of Engelson and McDermott [1], three types of kidnapping can be differentiated. Type 1 occurs when the robot's position is changed significantly, but its position estimate does not change accordingly. This happens when the robot's position-tracking sensors fail to detect motion, such as

* Corresponding author at: Research School of Engineering, Australian National University, Canberra ACT 0200, Australia. Tel.: +61 401 867 186.
E-mail addresses: djcampbell.01@gmail.com, dylan.campbell@nicta.com.au (D. Campbell), m.whitty@unsw.edu.au (M. Whitty).

http://dx.doi.org/10.1016/j.robot.2014.08.004
0921-8890/

when the robot is carried, enters an elevator or encounters fog, dust or a blackout, which circumvent the wheel encoders, visual odometry system or both. In addition, sequential scan-matching algorithms, such as the Iterative Closest Point (ICP) [3] and the Normal Distributions Transform (NDT) [4] algorithms, can misalign scans and thereby fail to estimate the robot's location correctly.

Type 2 kidnapping occurs when the robot's internal position estimate changes significantly, but its actual location does not, such as when the position estimate is reset upon shutting down. The robot will need to relocalise within the global map when it is turned on, a circumstance known as the wake-up robot problem [5]. As well as system errors and instabilities, this form of kidnapping can also be caused by hardware issues, such as knocked cables or faulty connections, that may otherwise go undetected. A robust solution should be able to handle such failures.

Finally, Type 3 kidnapping is a short range variant of Type 1, occurring when the robot moves a short distance without sensing that it is moving. This primarily happens when the robot slips and the sensors, such as wheel encoders, incorrectly measure the displacement.

The action taken depends on the kidnap type: global localisation is required for Types 1 and 2, whereas local techniques, such as scan-matching or optical flow, are sufficient to correct the pose after Type 3 kidnapping. While it may seem unnecessary to detect Type 3 kidnapping, it is advantageous for an autonomous system to know when slip has occurred in order to change the driving strategy. In particular, slip may indicate that the robot has encountered an obstacle that it cannot detect, such as a rock abutting one of the wheels. Also, scans taken when slip was detected could be excluded from the global map to reduce local deformation.

In our previous work [6], we investigated a variety of metrics for use in a kidnap detection system. Furthermore, we proposed a joint classifier that combined a quality metric with a discrepancy metric to robustly identify when kidnapping occurred. While only detection metrics specific to 3D point clouds were evaluated, the approach was generalisable to 2D and other modalities, such as visual. The critical requirements were that two independent ways of estimating relative pose were available and that the quality of at least one of the pose estimates could be assessed. To detect Type 3 kidnapping, such as slip, another requirement was that the pose must be estimated in a way that is relatively immune to that form of kidnapping and in a way that is not immune.

In this work, we extend the previous detection formulation to a Support Vector Machine (SVM) approach [7,8]. Suitable when sufficient training data is available, this approach can easily integrate additional quality and discrepancy metrics to improve classification accuracy. Using an SVM with a non-linear kernel, all kidnap types can be distinguished, including Types 1 and 2 that were previously classified as identical. An SVM was trained and tested on the datasets used in our previous work [6], as well as the large, publicly-available Hannover2 dataset [9].

The remainder of this paper is organised as follows: Section 2 briefly reviews related work, Section 3 provides an overview of algorithms and concepts used in the method and Section 4 details how the proposed kidnap detection system functions. The experiments and results are presented in Sections 5 and 6 and are discussed in Section 7. Finally, the most important conclusions are summarised in Section 8.

2. Related work

The two main approaches to kidnap detection are to either incorporate new sensors, an application-specific solution, or remove dependence on local pose estimation entirely. The first approach only partially solves the problem, but can be useful in situations where one form of kidnapping predominates. For example, Lee et al. [10] use a wheel drop switch to detect kidnapping when a robotic vacuum cleaner is picked up and taken to another room, but cannot detect other forms of kidnapping.

Kidnapping due to elevator motion has been partially solved by using a barometer [11] or an accelerometer [12] to detect floor transitions. However, both sensors were unable to accurately assess the relative altitude of the robot and therefore additional information about the structure of the building was required to construct a full 3D map.

To address the problem of kidnapping when a vision-based robot enters a dark environment, Henry et al. [13] incorporated depth data into their visual SLAM approach. By using an RGB-Depth camera that can function in light deficient areas, they improved the robustness of their localisation system to kidnapping caused by sensor failure.

The second approach obviates the need for kidnap detection in the first place by performing global localisation regularly, regardless of whether the robot has been kidnapped or not. Such an approach was taken by Thrun et al. [14] and Milstein et al. [15], both of whom used a variant of the Monte Carlo Localisation algorithm. After a kidnapping event, the approaches gave increasing credence to kidnapping hypotheses as the robot progressed.

One problem with this approach is that kidnapping events are not detected immediately. As a result, the robot will operate with an incorrect internal map until sufficient evidence is attained to suggest that kidnapping has occurred. At best, this is inefficient, rendering the robot unable to perform its function. At worst, the kidnapped robot could unwittingly perform dangerous actions, such as driving towards previously identified hazards. In addition, the computational complexity of this approach scales linearly with the area of the mapped environment. It would be preferable for kidnapping to be detected in constant time, as our approach achieves.

For localisation systems that use visual odometry, the failure to track a sufficient number of features can be indicative of kidnapping, as asserted by Se et al. [16]. However, this approach to kidnap detection does not extend to other localisation methods, like scan-matching, and cannot detect Type 3 kidnapping, such as slip. Nonetheless, for visual odometry systems, a metric based on the number of tracked features could be incorporated into our method, allowing the detection of slip.

Another metric for kidnap detection was proposed by Choi et al. [17], based on the entropy of node probabilities in a topological map. However, this approach can only be used when a global map of the environment is available and cannot detect wheel slip.

3. Fundamentals

3.1. The Normal Distributions Transform algorithm

Two of the quality metrics evaluated in this work (Q_s and Q_h) were derived from the optimisation function and Hessian of the 3D Normal Distributions Transform (NDT) scan-matching algorithm [4], although the metrics are calculated without running the algorithm. It is particularly useful for kidnap detection because it provides an estimate of the variances of each pose degree of freedom [18], which can be used as a measure of registration quality.

The NDT algorithm subdivides a point cloud into a 3D grid of cells and computes a Probability Density Function (PDF) for each cell. A mixed normal and uniform distribution is used, since a pure Gaussian is not robust to outliers [19]. The algorithm then finds the transformation, corresponding to a change in pose, that maximises the likelihood that the points of another point cloud were sampled from this PDF surface. It does this by minimising an approximation of the negative log-likelihood of Ψ, as given in Eq. (1), where Ψ is the likelihood function, p is the PDF with mean and covariance