BIOSYSTEMS ENGINEERING 166 (2018) 116–125

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Analysis of two visual odometry systems for use in an agricultural field environment

Stefan K. Ericson [a,*], *Björn S. Åstrand* [b]

[a] *School of Engineering Science, University of Skövde, Skövde, Sweden*
[b] *School of Information Science, Computer and Electrical Engineering, Halmstad University, Halmstad, Sweden*

ARTICLE INFO

Article history:
Received 26 August 2016
Received in revised form
12 November 2017
Accepted 21 November 2017

Keywords:
Visual odometry
Agricultural field robots
Visual navigation

This paper analyses two visual odometry systems for use in an agricultural field environment. The impact of various design parameters and camera setups are evaluated in a simulation environment. Four real field experiments were conducted using a mobile robot operating in an agricultural field. The robot was controlled to travel in a regular back-and-forth pattern with headland turns. The experimental runs were 1.8–3.1 km long and consisted of 32–63,000 frames. The results indicate that a camera angle of 75° gives the best results with the least error. An increased camera resolution only improves the result slightly. The algorithm must be able to reduce error accumulation by adapting the frame rate to minimise error. The results also illustrate the difficulties of estimating roll and pitch using a downward-facing camera. The best results for full 6-DOF position estimation were obtained on a 1.8-km run using 6680 frames captured from the forward-facing cameras. The translation error (x, y, z) is 3.76% and the rotational error (i.e., roll, pitch, and yaw) is 0.0482 deg m^{-1}. The main contributions of this paper are an analysis of design option impacts on visual odometry results and a comparison of two state-of-the-art visual odometry algorithms, applied to agricultural field data.

© 2017 IAgrE. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Visual odometry (VO) is a method for estimating the position of a camera from an image sequence. In VO, consecutive image frames in a sequence are matched for correspondence and the relative poses between the frames are accumulated. This estimates the travelled path with up to six degrees of freedom (DOF). This technique is applied to agricultural field robots to increase navigation precision compared with that of current GPS navigation systems and to make robots that can operate closer to crops than can current systems.

VO has been around for many years, one of the pioneering studies being by Nistér, Naroditsky, and Bergen (2004). Their work introduces an algorithm in which feature points are extracted from the images, matched to each other, and finally used for motion estimation. Outliers among the points are removed using random sample consensus (RANSAC) (Nistér, 2005). The algorithm is applied to a dataset acquired using a vehicle with a forward-facing stereo camera. Many state-of-the-art methods still use the same approach (Cvišić & Petrović, 2015; Geiger, Ziegler, & Stiller, 2011; Kitt, Geiger, & Lategahn, 2010) but differ slightly in how the features are

* Corresponding author. School of Engineering Science. Box 428, 54128 Skövde, Sweden.
E-mail address: stefan.ericson@his.se (S.K. Ericson).
https://doi.org/10.1016/j.biosystemseng.2017.11.009
1537-5110/© 2017 IAgrE. Published by Elsevier Ltd. All rights reserved.

Nomenclature

α	Field of view of the camera
R	Rotation matrix
T	Relative translation
ω_x	Angular velocity along x-axis
ω_y	Angular velocity along y-axis
ω_z	Angular velocity along z-axis
b	Camera baseline
C_k	Cumulative pose at frame k
d	Disparity between cameras
dTr	Relative pose change of the camera between consecutive frames
f	focal length of the camera
f_s	Frame rate (Hz)
h	Camera height above the ground (m)
k	frame index
T_x	Translational velocity along x-axis
T_y	Translational velocity along y-axis
T_z	Translational velocity along z-axis
v_x	Projected 2D flow field along x-axis
v_x	Projected 2D flow field along x-axis
v_y	Projected 2D flow field along y-axis
v_{max}	Maximum camera velocity (m/s)
x	image points x-coordinate
x_l	x-coordinate of point in left image
x_r	x-coordinate of point in right image
y	image points y-coordinate
Z	Distance to point (depth)
DOF	Degrees of freedom
DW	Downward facing camera
FW	Forward facing camera
ICP	Iterative closest point
IMU	Inertial measurement unit
RANSAC	Random sample consensus
RTK-GPS	Real-time kinematic Global Positioning System
VO	Visual odometry

extracted and matched and in how the outliers are removed. Position estimation can also be improved by making assumptions as to the environment or by adding a motion model of the vehicle, the goal being to reduce the cumulative error.

One reportedly successful method using downward-facing cameras in an agricultural application has been presented by Jiang et al. (2014). Their robot, Gantry, is in the form of a 3-m-high square-shaped table with a combined driving and steering wheel at each leg. Two downward-facing cameras mounted at the top of the robot are used for the VO. In one experiment conducted in a soybean field, the path follows a regular back-and-forth track with a total of 13 headland turns. This experiment has a track length of 2.5 km and consists of 11,700 frames. Their results indicate that the translation error (2-DOF) was under 5.12 m, which corresponds to 0.2% of the travelled distance. A shorter path (i.e., 386 m and 1300 frames) on a grass road was also evaluated and the reported result is 1.6%.

For urban environments there are publicly available datasets for VO evaluation (Geiger, Lenz, & Urtasun, 2012). One such dataset, the KITTI Vision Benchmark Suite, consists of several sequences of images captured from forward-facing cameras mounted on a car roof. Ground truths are available for some, but not all, sequences; the remaining sequences are used for the evaluation of algorithms and a ranking is published online.[1] Similar datasets are unavailable for the agricultural case, so a similar comparative ranking does not exist.

Two state-of-the-art methods are selected for evaluation in this paper, the method used with the Gantry robot (Jiang et al., 2014) and the C++ VO library Libviso (Geiger, 2015; Geiger et al., 2011). Gantry was specifically developed for use in agricultural fields and its reported error was under 0.2%. The Libviso method is also intended for use in the agriculture field environment (Markt & Technik, 2015). The highest-ranked VO algorithm on the KITTI benchmark (Cvišić & Petrović, 2015) is based on the Libviso method. Cvišić and Petrović (2015) use the same feature extractor but apply a more sophisticated outlier rejector, so only selected feature points are used in the motion estimator; they report translation error of 0.88% and rotational error of 0.0022 deg m^{-1}. The KITTI benchmark list reports 2.44% and 0.0114 deg m^{-1} translation and rotational errors, respectively, for the Libviso method with stereo cameras, as used in this paper.

To improve the accuracy of VO, this paper seeks new knowledge of cumulative error when VO is used in an open field environment typical for agricultural fields with low-height crops. The accuracy is evaluated by comparing two algorithms, the Gantry and Libviso methods, on both simulated data and on real data captured by a mobile robot. Using both simulated and real field data allows the impact of different design choices to be evaluated, improving our understanding of how various parameters and settings affect the VO results in an agricultural field environment.

The main contributions of this paper are an analysis of design option impacts on visual odometry results and a comparison of two state-of-the-art visual odometry algorithms, applied to agricultural field data.

2. Visual odometry theory

This section presents the theory related to VO. Consider a 6-DOF VO system in which the relative pose change of the camera between consecutive frames is modelled as a rigid motion transform. The transform can be written as shown in Equation 1 (Scaramuzza & Fraundorfer, 2011):

$$dTr = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \tag{1}$$

where $\mathbf{R} \in SO(3)$ is the rotation matrix and $\mathbf{T} \in \Re^{3\times1}$ is the relative translation. The cumulative pose C at frame k can be obtained using Equation (2):

$$C_k = C_{k-1} dTr \tag{2}$$

These are the basic equations of the odometry, and the goal is to find dTr and C_k for each frame. It should be mentioned that the transformation matrix is an overdetermined

[1] The KITTI Vision Benchmark Suite, http://www.cvlibs.net/datasets/KITTI/eval_odometry.php; accessed 20 June 2016.

Download English Version:

https://daneshyari.com/en/article/8054847

Download Persian Version:

https://daneshyari.com/article/8054847

Daneshyari.com