EVALUATION OF VISUAL ODOMETRY ALGORITHMS FOR TREE NURSERY INSPECTION MACHINES

Vitalijs Osadcuks[1], Kristaps Muska[1], Andrejs Zujevs[2], Arturs Ardavs[2]
[1]Latvia University of Life Sciences and Technologies, Latvia; [2]Riga Technical University, Latvia
vtl@tvnet.lv

Abstract. Tree nurseries are responsible for providing strong and healthy seedlings to ensure their development into trees and increase overall efficiency of forest regrowing process. Seedlings during the first stages of development are usually grown in large greenhouses or dedicated open fields. Despite growing indoors in a greenhouse, a lot of weed seeds may get into the soil in different ways: ventilation, mobile machines, workers, birds etc. Currently weeding is performed by periodically removing the trays affected and working manually. This is time consuming and costly process. The ultimate solution would be full automation of inspection and weeding process on-site using some mobile robotized equipment. Such equipment will have to localize in the greenhouse to be able to operate on plants at different areas. Aim of the current study is to evaluate possibility to use main camera primarily intended for plant inspection also for localisation purposes in tree nursery greenhouses using visual odometry (VO). Four different VO methods were compared on a set of 111 experimental images. Results show, that phase cross correlation method with its 1.056 s execution time for 111 frame images and 0.7 mm error over one metre is the best option to use for linear movement of a sensor bundle in a greenhouse environment.

Keywords: tree nursery, camera, visual odometry, algorithms.

Introduction

For successful operation of industrial and recreational forestry as well as reforestation of territories in general quality planting material is essential. According to Official statistics of Latvia annual reforestation is more than 40 thousand ha during last 5 years and requires more than 60 million of seedlings per year [1]. Tree nurseries are responsible for providing strong and healthy seedlings to ensure their development into trees and increase overall efficiency of forest regrowing process. Seedlings during the first stages of development are usually grown in large greenhouses or dedicated open fields. To increase the area use effectiveness, seeds are planted into trays, which are placed as tight as possible on the available space. A typical installation of this kind for *Pinus sylvestris* L. is shown in Fig. 1. At the far end of the greenhouse irrigation installation can be seen. It moves on a rail under the top of the roof and performs watering and fertilizing of the plants.

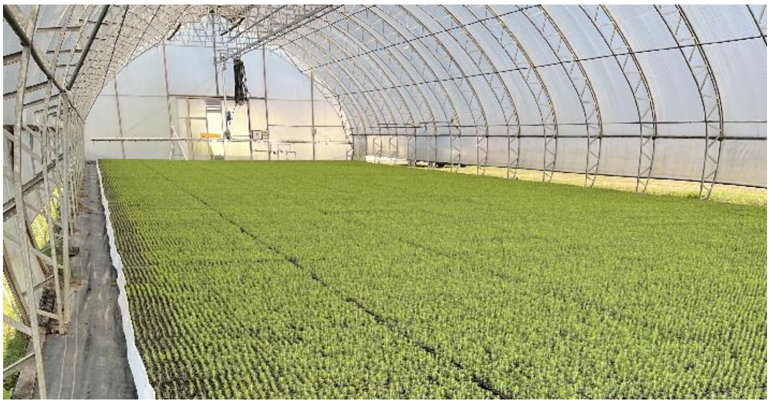

Fig. 1. **Seedlings in a greenhouse of "Norupes" tree nursery, Ogre, Latvia**

Despite growing indoors in a greenhouse, a lot of weed seeds may get into the soil in different ways: ventilation, mobile machines, workers, birds etc. Currently weeding is performed by periodically removing the trays affected and working manually. This is time consuming and costly process. Pathways between trays are also no-solution as they will decrease total seedling output and will have more negative economic effect than labour cost for relocation of trays during weeding. Also, timely recognition of diseases is essential to prevent loss of planting material. These problems grow only increase when growing outdoors.

The ultimate solution would be full automation of inspection and weeding process on-site without relocation of trays. This could be achieved by a robotized unit moving on linear rails in the same way as irrigation equipment or these two could be organized into a single machine. First step of this ultimate solution is to introduce monitoring system capable of detection weeds and diseases as well as monitor total health of seedlings. This study is part of a research project aiming to develop a vision sensor bundle attachable to irrigation system. The main idea is to use existing irrigation infrastructure to move sensor bundle in the longer dimension of greenhouse and add additional rail on irrigation tube itself for the shorter dimension. Working laboratory prototype is shown on Fig. 2.

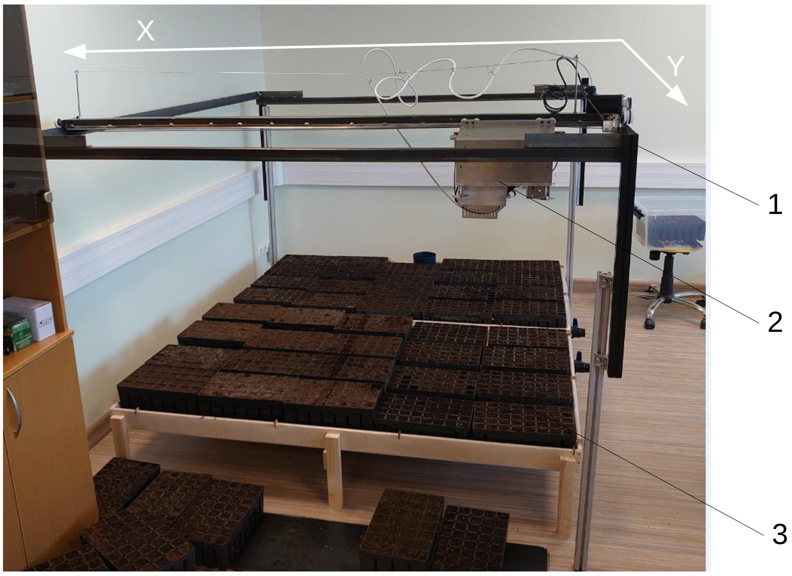

Fig. 2. **Laboratory prototype of tree nursery greenhouse monitoring system:** 1 – Axis with driving stepper motors; 2 – sensor bundle with cameras; 3 – trays with seedlings

As any autonomous mobile equipment such sensor bundle should be able to localize itself in the coordinate frame of the greenhouse. The most obvious way would be to use encoders or range sensors. On irrigation systems available on market usually inductive type sensors are used, but their precision may not be enough: there are 4 cogs on 200 mm pulley giving 157 mm resolution. In any case additional encoder for the second axis is necessary. In current laboratory prototype positioning is performed using rubber wheel encoders Sick DBV50 for both axes.

Use of dedicated localization sensors may come with a number of problems:

- in general greenhouse environment is considered quite harsh for electronics (moist, fertilizers, plant protections chemicals etc), fine mechanical encoders may not be reliable in such environment [2];
- there may be also system integration issues to read the position information from different irrigation equipment;
- relatively long distances and moving parts require quality connection cables.

Aim of the current study is to evaluate possibility to use main camera primarily intended for plant inspection also for localisation purposes in tree nursery greenhouses using visual odometry (VO). VO is the process of estimating the pose and motion of the camera from an image sequence; using consumer-grade cameras rather than expensive sensors or systems, VO is an inexpensive and simple approach to estimate the location of robots and vehicles [3]. Visual odometry is successfully finding its place in autonomous vehicles for applications in agriculture. Examples like unmanned aerial vehicles (UAV) for survey and spraying tasks as well as ground vehicles for various operations directly on plants [4]. There are a number of different VO algorithms with their strengths and weaknesses. In essence VO algorithm is detecting pixel displacement in two consequent images from cameras thus estimating camera movement in terms of pixels. To obtain movement in real life units, pixels should be converted into length units.

Current case, when camera is moving along two axes resulting only in translation movement makes the task easier from computation point of view as rotation may not be considered. Results of this research may also be applicable for similar precision irrigation/fertilizing systems based on moving frames. Also, they may be relevant for precision plant-level agricultural equipment that moves only along rows like robotized or tractor-mounted laser weeders, where rotation movement may be neglected.

Materials and methods

The evaluation consists of three steps.

1. Evaluation of height changes.
2. Running test sequence of images and comparison of real pixel displacement and one obtained by visual odometry method.
3. Evaluation of time for each method to process images (mean for image pair and total time for the sequence).

Sensor bundle incorporates Alvium G5-508 RGB camera with Edmund optics TECHSPEC® HR Series 16 mm Fixed Focal Length lens, pixel size 3.45 µm. This was used as data source of image sequence for visual odometry. Calibration of the camera (obtaining intrinsic camera matrix) was performed and images were undistorted using open-source computer vision library OpenCV [5]. Images were taken on laboratory prototype running for one meter every 10 mm, 111 images in total. A typical image (raw and undistorted) is shown on Fig. 3. All images were taken on constant height 500 mm.

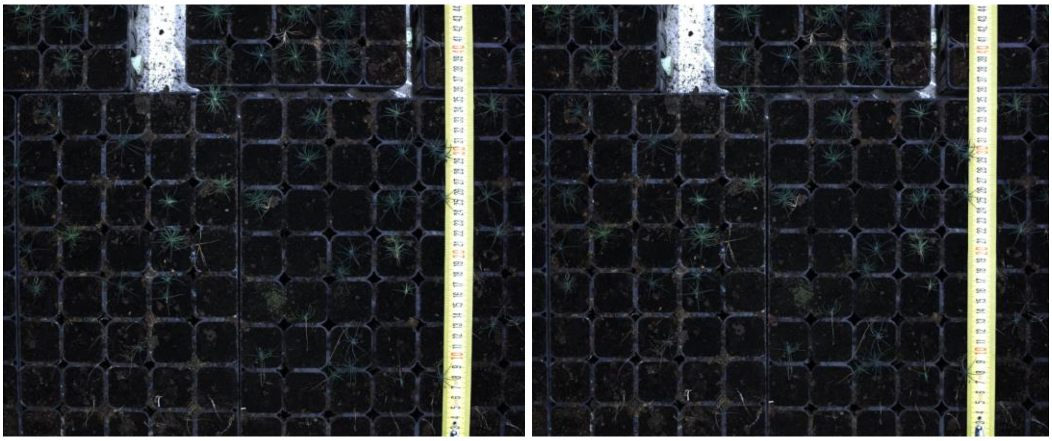

Fig. 3. **A typical image from obtained sequence: raw distorted image on the left and undistorted after processing in OpenCV on the right**

"Ground truth" or real pixel movement for 100 image sequence was determined manually by counting pixel displacement between undistorted frames using ruler ticks as reference.

To evaluate performance of VO in terms of real-world units pixel size in millimetres is necessary. Pixel size in millimetres for both camera sensor axis can be easily calculated using well-known imaging equation:

$$K = \frac{f}{H_c \cdot s_p} = \frac{R \cdot f}{S \cdot H_c} \tag{1}$$

where K – pixel to mm ratio for camera sensor, pix·mm^{-1};
f – camera focal length, mm;
S – sensor size, mm;
s_p – pixel size, mm;
H_c – distance from lens focal point to surface;
R – camera resolution, pixels.

Maximum error for ground truth is one pixel, which results in ±(500*0.003454/16 = 0.1 mm·pix^{-1}) for above-given camera parameters and H_c = 500 mm.

Four visual odometry methods were used for evaluation. Each method detects pixel movement between two consecutive images using different approach.

- Optical flow (Lucas-Kanade) – computes motion between sparse key points using changes in pixel intensity [6].

- Dense optical flow (Farneback) – computes motion for every pixel using changes in pixel intensity [7].
- 2D image correlation – measures similarity by sliding one over the other to calculate a correlation score to find best alignment [8].
- Phase cross correlation – measures phase differences in the frequency domain to estimate motion [9].

Python implementations of all methods was used for evaluation. Multiple libraries were utilized, including OpenCv (opencv-python) for optical flow, signal processing (scipy.signal) for 2D image correlation, image registration (skimage.registration) for phase cross correlation, scientific computing (numpy) for calculations and time measurement (time). To exclude possible effect oh height changes and lens sensor-system, pixel-mm ratio K was obtained directly from test images by counting pixels in the size of ruler. VO was performed on cropped images to reduce the effect from lens distortion and to reduce computing resources and processing time. The original image had a resolution of 2432x2018 pix, the cropped image sizes for every method are given in Table 1.

Table 1

Image sizes for different methods

Method	Cropped Area (%)	Size, pix
Optical flow	30	605x729
Dense Optical Flow	30	605x729
Correlate 2D	10	200x242
Phase cross correlation	10	201x243

For correlate 2D and phase cross correlation the cropped area was 10% of the middle. For optical flow a larger cropped middle area was required – 30%, as smaller image regions led to difficulties in accurately estimating motion between the pictures.

Execution speed of all methods was compared on three different computers (see Table 2). Execution time was measure was performed by Python “time” module, image reading from disk was not included in time measurement.

Table 2

Computers used in execution speed comparison

<table>
<tr><th colspan="2">Machine</th><th>C1 (Google colab)</th><th>C2 (personal)</th><th>C3 (notebook)</th></tr>
<tr><td rowspan="5">CPU</td><td>Model</td><td>Intel(R) Xeon(R) CPU @ 2.20GHz</td><td>13th Gen Intel(R) Core(TM) i7-13700KF</td><td>Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz</td></tr>
<tr><td>Clock speed</td><td>2.20GHz</td><td>3.4 GHz</td><td>1.6 GHz</td></tr>
<tr><td>Cache Size</td><td>56.32 MB</td><td>30 MB</td><td>6 MB</td></tr>
<tr><td>Cores</td><td>1</td><td>16</td><td>4</td></tr>
<tr><td>Threads</td><td>2</td><td>24</td><td>8</td></tr>
<tr><td>RAM</td><td>Total RAM</td><td>12.69 GB</td><td>64 GB</td><td>12 GB</td></tr>
</table>

Results and discussion

Evaluation of height change effect

Ratio K is strongly dependent on camera distance to surface. In case of UAV applications distance H_c is usually measured using dedicated sensor, e.g. AFBR-S50LV85D Time-of-Flight laser rangefinder or similar. In case of ground vehicles this distance is much smaller and changes insignificantly and often is taken as constant [4]. In our application height changing happens manly due to swinging horizontally positioned irrigation tube and unevenness in seedling tray placement. Height changes are estimated to be H_c = 500 ± 20 mm. For f = 16 mm, s_p = 3.45 μm this will result in pixel ratio partial error

$$K = 9.275\ (+0.386, -0.357)\ \text{pix}\cdot\text{mm}^{-1} \text{ or } (+4.2, -3.9)\ \%.$$

In displacement terms for 1 m, we will get error about 4 cm. In order to decrease this error additional height measurement would be necessary. However, if comparing with resolution of existing inductive sensor encoder with 4 cogs (157 mm), VO without dedicated H_c measurement sensor is possible to use for short movement detection with better accuracy and clear accumulated error at each pulse from the encoder.

Another option available for greenhouse application is fact, that objects of known size – seedling trays – are always in vicinity of the camera. Advantage of it can be taken, by calculating ratio K for visible sizes in pixel of tray cells detected by some computer vision algorithm.

Evaluation of displacement

The comparison between the measured displacement in each frame and the values calculated using the OV methods are presented in Fig. 4.

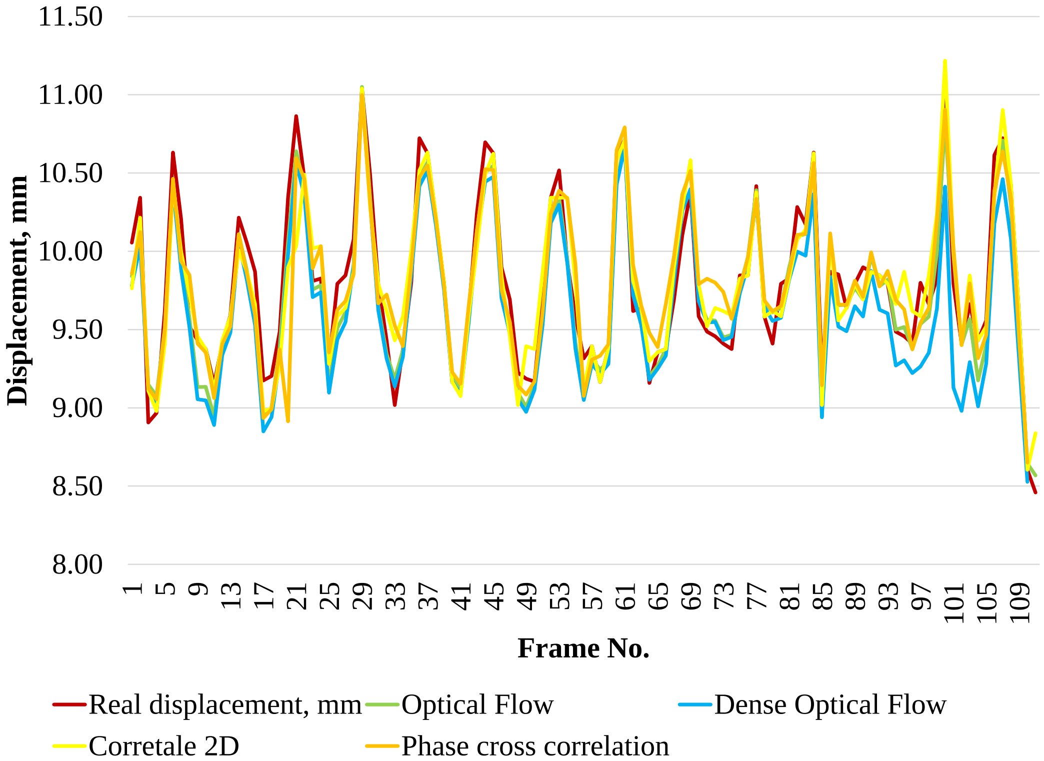

Fig. 4. **Frame-wise motion detection comparison of different VO methods and real measurement**

Comparing the real displacement in each individual frame (represented by the red line) to the calculated displacements obtained by the OV methods, the following results were observed: optical flow (error range: -3.17% to 2.72%), dense optical flow (error range: -6.78% to 2.30%), correlate2d (error range: -7.65% to 4.59%), phase cross corelation (error range: -13.81% to 6.36%). However, for total distance (summarized in Table 3) individual frame errors average and the final measurement is close to obtained by other authors. Comparing the VO methods calculated displacements to the actual displacement, the error margin ranges from -0.04% to 1.63%. The results, listed in order of increasing error, are as follows: phase cross correlation (error: -0.04%), optical flow (error: -0.65%), dense optical flow (error: -0.94%), correlate 2D (error: 1.63%). For example, for translation movement and VO using the most precise (and resource consuming) method Correlate 2D error was only 0.2 mm which falls into range obtained in [4].

Evaluation of execution time

The total calculated distance and execution time of the OV methods are given in Table 3. The execution times of the methods across all three computers described in Table 2 (C1-C3), listed from fastest to slowest, are as follows: phase cross correlation, optical flow, dense optical flow, correlate 2D. The fastest execution times were observed with phase cross correlation and optical flow, with the highest recorded execution time remaining under 4 seconds in all cases. Dense optical flow and correlate 2D demonstrated significantly slower performance, with the longest execution time exceeding 30 seconds.

According to the obtained results Phase cross correlation method with its best execution time and reasonable error is the best option to use for linear movement of a sensor bundle in a greenhouse environment.

Table 3

Summary of displacement and execution time measurements for each of the methods

Method	Total distance, mm	Com-parison, %	Error, %	Execution time total, s			Execution time average per frame, s		
				C1	C2	C3	C1	C2	C3
Real movement	1086.2	100	0	–	–	–	–	–	–
Optical Flow	1079.1	99.35	-0.65	3.142	1.502	3.574	0.028	0.014	0.032
Dense Optical Flow	1068.9	98.41	-1.58	47.083	33.868	48.555	0.424	0.305	0.437
Correlate 2D	1086.4	100.01	0.01	52.625	31.541	47.138	0.474	0.284	0.425
Phase cross correlation	1085.9	99.97	-0.03	2.13	1.056	2.845	0.019	0.01	0.026

Conclusions

1. Phase cross correlation method with 1.056 s execution time for 111 frame images and 0.7 mm error over one metre is the best option to use for linear movement of a sensor bundle in a greenhouse environment.
2. VO camera to surface distance changes for 500 ± 20 mm height considered in this study theoretically may result in VO errors in the range of –3.9% to +4.2%.
3. To increase overall accuracy, it is recommended to add height measurement sensor for more accurate pixel to mm conversion; or to use object of known size in the view of VO camera. In the case of tree nursery seedling trays with cells of regular size could be used for that purpose.

Acknowledgements

The study was supported by funding from LAD project No. 23-00-A01612-000002.

Author contributions

Conceptualization, V.O.; methodology, VO. and K.M.; software, K.M.; validation, A.Z. and A.A; formal analysis, V.O.; experiments, A.Z. and A.A.; writing – original draft preparation, V.O.; writing – review and editing, A.Z. and A.A.; visualization, V.O., K.M.; project administration, A.Z.

All authors have read and agreed to the published version of the manuscript.

References

[1] Official statistics of Latvia. Forest regeneration and afforestation. [online] [06.05.2025] Available at: https://stat.gov.lv/en/statistics-themes/business-sectors/forestry/8918-forest-regeneration-and-afforestation?themeCode=ME

[2] Osadcuks V., Pecka A., Lojans A., Kakitis A. Experimental research of proximity sensors for application in mobile robotics in greenhouse environment, Agron. Res., vol. 12, no. 3, 2014, pp. 955-966, Accessed: May 05, 2025. [online] [06.05.2025] Available: https://www.researchgate.net/publication/286171148_Experimental_research_of_proximity_sensors_for_application_in_mobile_robotics_in_greenhouse_environment

[3] Bai Y., Zhang B., Xu N., Zhou J., Shi J., Diao Z. Vision-based navigation and guidance for agricultural autonomous vehicles and robots: A review, Comput. Electron. Agric., vol. 205, 2023, p. 107584, DOI: 10.1016/J.COMPAG.2022.107584

[4] Zaman S., Comba L., Biglia A., Ricauda Aimonino D., Barge P., Gay P. Cost-effective visual odometry system for vehicle motion control in agricultural environments, Comput. Electron. Agric., vol. 162, no. March, 2019, pp. 82-94, DOI: 10.1016/j.compag.2019.03.037

[5] OpenCV: Camera Calibration and 3D Reconstruction. [online] [06.05.2025] Available at: https://docs.opencv.org/4.x/dc/dbb/tutorial_py_calibration.html
[6] Mitiche A., Aggarwal J. K. Optical flow estimation. Springer Topics in Signal Processing, 10, 2014, pp. 41-93. DOI: 10.1007/978-3-319-00711-3_3
[7] Farnebäck G. Two-Frame Motion Estimation Based on Polynomial Expansion. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2749, 2023, pp. 363-370. DOI: 10.1007/3-540-45103-X_50
[8] Wu H., Sun C., Liu Y., Cui Z., Qi W., Liu Y. A Fast Image Template Matching Algorithm Based on Normalized Cross Correlation. Journal of Physics: Conference Series, 2020. DOI: 10.1088/1742-6596/1693/1/012163
[9] Yan H., BMVC J. L. Robust Phase Correlation based Motion Estimation and Its Applications, 2008. DOI: 10.5244/C.22.104
[10] Ericson S. K., Åstrand B. S. Analysis of two visual odometry systems for use in an agricultural field environment, Biosyst. Eng., vol. 166, 2018, pp. 116-125, DOI: 10.1016/j.biosystemseng.2017.11.009