Single-image-based georeferencing for unmanned aerial vehicles

Applications and practical considerations

Ralph L. Stoop [1], Markus Sax[2], Dejan Seatovic[3], and Thomas Anken[4]

Abstract: Georeferencing is important for many applications in precision farming, in particular those based on unmanned aerial vehicles (UAVs). In this context, georeferencing typically relates the optical features of UAV images to their actual position in the 3D world, creating a grid map of the area of interest. Although state-of-the-art georeferencing methods are very accurate, these methods rely on multiple-view geometry reconstruction, which requires largely overlapping images of high quality. Acquiring such images can be difficult in practice, given the low-cost requirements for precision farming. In this paper, we study the practical applications and challenges of a simple, computationally inexpensive and fast method for georeferencing that is solely based on single images. Our method only uses an affine transformation where the UAV's height is adjusted by a digital terrain model and does not require overlapping images. We find that our single-image-based method can be used for smart farming applications, where spatial accuracies of around 25 cm are sufficient.

Keywords: unmanned aerial vehicles, UAVs, georeferencing, structure-from-motion, precision agriculture

1 Introduction

Knowing the exact location of objects of interest, such as weeds, is crucial in precision agriculture. For example, weeds detected in an image captured by an unmanned aerial vehicle (UAV) need to be located accurately for subsequent spot spraying. This involves georeferencing, which attributes a geographic location, typically in longitude/latitude, to each point of an image. However, georeferencing an object based on a single image is generally not possible because image acquisition is a lossy process due to the projective transformation from three dimensions (3D) to two dimensions (2D). Hence, additional information, such as additional images from other viewpoints or complementary sensory data, is required. Traditionally, georeferencing is achieved with high accuracy using aerial

[1] Agroscope, Digitale Produktion, Tänikon 1, 8356 Ettenhausen, ralph.stoop@agroscope.admin.ch, https://orcid.org/0000-0002-5138-8020
[2] Agroscope, Digitale Produktion, Tänikon 1, 8356 Ettenhausen, markus.sax@agroscope.admin.ch
[3] Ostschweizer Fachhochschule OST, ISF Institut für intelligente Systeme und Smart Farming, Oberseestrasse 10, 8640 Rapperswil, dejan.seatovic@ost.ch
[4] Agroscope, Digitale Produktion, Tänikon 1, 8356 Ettenhausen, thomas.anken@agroscope.admin.ch

triangulation (AT) with the use of ground control points (GCPs) [Mi16]. GCPs are accurately located points of the scene of interest which exhibit clear visual features. This makes their later identification in the images easy and allows for image localization, alignment and distortion corrections. However, the use of GCPs is too costly, in particular for typical applications in precision agriculture. Direct georeferencing has emerged as an alternative that does not require GCPs. Instead, data from different sensors, such as from the global navigation satellite systems (GNSSs) or inertial measurement units (IMUs), are used to find a globally consistent reconstruction of the scene (e.g., [CRS17; Ga18; Sy20]). Despite their great accuracy, these methods are computationally demanding. Depending on the resolution and hardware, a planar view of the whole scene (orthomosaic) might take several hours for one hectare to compute.

In this paper, we discuss our simplified, single-image-based direct georeferencing method which serves as an alternative to a full photogrammetrical pipeline implemented in the Innosuisse "Rumex" project [IN24; Sa23]. Based on a simple linear transformation described by a 3x3 matrix, we achieve errors of Rumex locations in the order of 25 cm using our prototype system. While this accuracy might not be sufficient for direct treatment, it is suitable to estimate Rumex's location and to fuse it into a map, which could then be delivered directly to the farmer. The maps can be calculated in approximately real-time using standard laptop hardware, allowing the farmer to start treatment promptly.

2 Material and methods

2.1 Overview

The goal of the "Rumex" project is to develop a fully automated detection pipeline that outputs the accurate location of Rumex plants on meadows, given a UAV flight capturing high-resolution RGB color images. An overview of the envisioned workflow is shown in Figure 1. In the current implementation, image transfer from the drone (DJI Matrice 300 RTK with Zenmuse P1 camera system) to the server is still done manually using standard secure digital memory cards after the flight. However, proof of concept of 5G image transfer was also demonstrated using an alternative UAV system. Each image is input into a trained object detection model, and pixel locations (bounding boxes) of the detected plants are returned. To relate the image coordinates of the plants to their actual positions in the 3D world, two complementary pipelines were implemented: a fully photogrammetrical georeferencing method with accuracies of approximately 2cm (green arrows "alignment and inference", "surface model", "orthomosaic" and "digital map/mission file" in Figure 1) and the simplified georeferencing method discussed in this work (yellow arrows "direct georeferencing" and "simplified map"). While the photogrammetrical method first aligns all images using bundle adjustment, it creates a full, dense surface model from which the orthomosaic is obtained via projecting the dense structure to a plane. The orthomosaic not only provides a highly detailed, geometrically corrected bird's eye view of the meadow but also associates a location on the earth's surface to each pixel. Although very accurate, the applicability of the method is hampered by its computational demand and may take hours to complete.

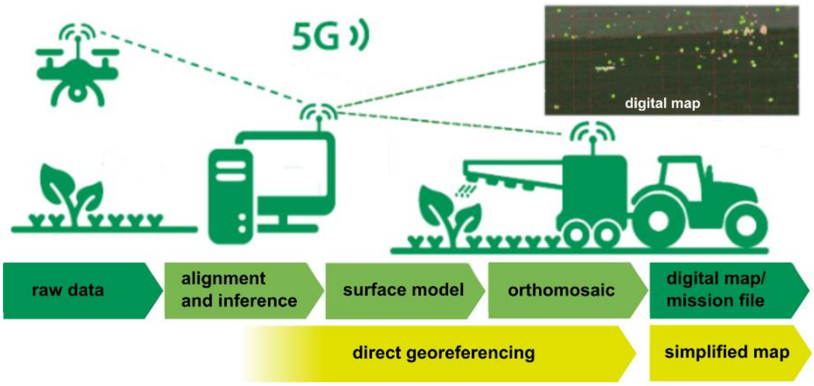

Fig. 1: Overview of the 'Rumex detection' pipeline. Image adapted from [Sa23]

2.2 Georeferencing method

To allow a fast, in-situ localization of the plants, an alternative, direct georeferencing method was implemented. Our approach is based on the ideal pinhole camera model, which is a widely employed approximation for direct georeferencing [HZ04]. The pinhole camera describes a simplified mathematical relation between the 2D pixel coordinates of an image point and its corresponding 3D location in the real world. Considerable simplification is achieved by neglecting the non-idealities of the lens, including, e.g., geometric distortions. The resulting mapping from 3D to 2D described by the pinhole camera is a perspective projection that can be formulated in homogenous coordinates using a linear transformation with the 3 x 4 camera matrix C. While the parameters of C can be derived from the camera intrinsically and extrinsically, our method further reduces the complexity of the georeferencing problem using two fundamental assumptions: First, the ground surface of the scene is assumed to be a perfect plane parallel to the camera's image plane, a condition often referred to as nadir. Second, at any instance of time, the distance between the two planes is assumed to be constant (constant flying height). These assumptions allow to describe the transformation of a pixel in the 2D image with coordinates $p_{im} = (x_{im}, y_{im})^T$ to the world coordinates $p_w = (x_w, y_w)^T$ (in longitude, latitude) using homogenous coordinates by the simplified 2D camera model:

$$\begin{pmatrix} x_w \\ y_w \\ 1 \end{pmatrix} = \begin{pmatrix} s_x \cos\theta & -s_y \sin\theta & t_x \\ s_x \sin\theta & s_y \cos\theta & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{im} \\ y_{im} \\ 1 \end{pmatrix} \tag{1}$$

In the above equation, the upper left (2 x 2) matrix involving sin(·) and cos(·) describes a 2D rotation of yaw angle θ while the upper right (2 x 1) vector translates the origin of the image center by $t_{x,y}$. The translation $t_{x,y}$ is measured by the GNSS RTK positioning system of the UAV at the time of image acquisition. The assumption of a constant flying height ensures a constant $s_{x,y}$ during a flight. These scaling factors are calculated from the ground sampling distance in x and y direction, given the camera's intrinsic parameters. Lastly, assuming the ideal nadir condition imposes roll and pitch angles to be zero, and hence, only the IMU readings of the yaw angle (θ) at each image are taken into account.

2.3 Experimental setup and error analysis

To validate the proposed georeferencing method, measurements were conducted at six different flat meadows in Tänikon (Aadorf, Switzerland). The ground truth positions ("target") of GCPs were measured using a Trimble R8 GNSS system (Trimble Inc, Westminster Colorado, USA). Flying height was $h = 12$ m and speed was $v = 1$ m/s. Overall, a total of 80 measurements was obtained and compared with their ground truth data. As figure of merit, the total error $\Delta d = \sqrt{\Delta Lat^2 + \Delta Long^2}$ is used, where $\Delta Lat = Lat_{pred} - Lat_{target}$ is the latitudinal error and $\Delta Long = Long_{pred} - Long_{target}$ is the longitudinal error.

3 Results and discussion

A scatter plot of the georeferencing errors Δd is shown in Figure 2. The achieved mean total error is approximately 0.20 cm. This finding has been recently reproduced (mean total error of 0.24m) in a complementary study. The errors obtained are in the order of typical Rumex plants. Note that we generally observe larger latitudinal errors compared to longitudinal errors. This finding could be reproduced in another flight at 30m flying height over an unrelated field with randomly placed GCPs. Although the actual reason for this behavior remains unknown, one explanation could be drift in the IMU. However, this hypothesis needs to be investigated in more detail in the future.

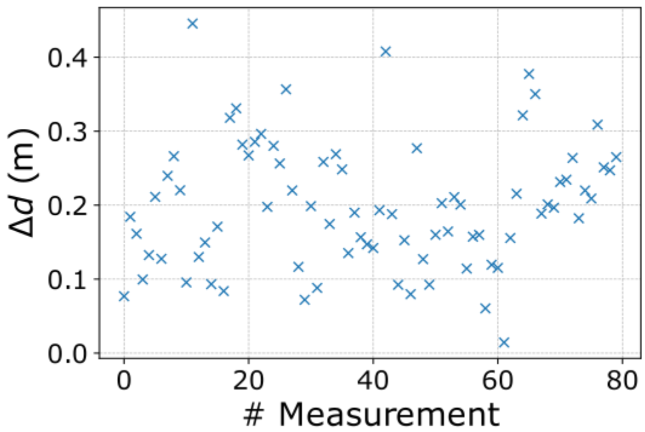

Fig. 2: Georeferencing errors in meters

Assuming a planar scene may not be justified for some meadows. As demonstrated in ongoing work, the presence of pronounced slopes requires an iterative correction of the flying height by the following algorithm: First, a prediction of the object of interest's position is calculated with model (1) using the value of the pre-assumed flying height (here 12m). Then, the actual flying height is refined by the difference between the UAV's altitude and the altitude value of a digital terrain model (DTM) at the location of the initial prediction. Then, a new prediction is obtained using the refined height. Although this procedure could be repeated to further refine the prediction, convergence of the error after one iteration is observed. In ongoing work, the improvements of the predictions are demonstrated for steep meadows, whereas the predictions do not improve for flat areas, as expected.

4 Conclusion and outlook

In this work, a simplified georeferencing method is described and validated for the case of Rumex detection on meadows with negligible slope. Based only on a linear transformation, mean errors of approximately 25 cm or less were achieved. Although it is still an open question whether the achieved accuracy is sufficient for an automated treatment, it seems to be sufficient for a map delivered to the farmer for fast manual treatment. Future work will aim at better quantifying the improvements gained from the iterative flying height adjustment in areas with notable topography. Also, a better understanding of the georeferencing errors, in particular the observed difference in magnitude between longitudinal and longitudinal errors, could provide possible improvements to the method.

Bibliography

[CRS17] Cucci, D. A.; Rehak, M.; Skaloud, J.: Bundle adjustment with raw inertial observations in UAV applications. ISPRS Journal of Photogrammetry and Remote Sensing 130, S. 1–12, 2017.

[Ga18] Gabrlik, P. et al.: Calibration and accuracy assessment in a direct georeferencing system for UAS photogrammetry. International Journal of Remote Sensing 39, 15–16, 4931–591999, 2018.

[HZ04] Hartley, R. I.; Zisserman, A.: Multiple View Geometry in Computer Vision, ISBN 0521540518, Cambridge University Press, 2004.

[IN24] Schweizerische Agentur für Innovationsförderung. „Verbesserung der Schweizer Agrarproduktion durch Digitalisierung“, https://www.innosuisse.admin.ch/de/verbesserung-der-schweizer-agrarproduktion-durch-digitalisierung, Stand: 15.10.2024.

[Mi16] Mian, O. et al.: Accuracy assessment of direct georeferencing for photogrammetric applications on small unmanned aerial platforms. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-3/W4, S. 77–83, 2016.

[Sy20] Syetiawan, A. et al.: Precise Topographic Mapping Using Direct Georeferencing in UAV. IOP Conference Series: Earth and Environmental Science 500, Nr. 1, 012029, 2020.

[Sa23] Sax, M. et al.: Broad-leaved dock control by unmanned aerial vehicles: What image quality is needed for a plant detection? LAND.TECHNIK AgEng, 2427:505–12, 2023.