The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-2/W8-2024
8th International ISPRS Workshop LowCost 3D - Sensors, Algorithms, Applications, 12–13 December 2024, Brescia, Italy

Exterior Orientation in a Box: Cost-Effective RTK/IMU-Based Photo Geotagging

Martin Wieser [1], Geert Verhoeven [2], Benjamin Wild [1], Norbert Pfeifer [1]

[1] Department of Geodesy and Geoinformation, TU Wien, Wiedner Hauptstraße 8/E120, 1040 Vienna, Austria
(martin.wieser, benjamin.wild, norbert.pfeifer)@geo.tuwien.ac.at

[2] Department of Prehistoric and Historical Archaeology, University of Vienna, Franz-Klein-Gasse 1, 1190 Vienna, Austria
geert.verhoeven@univie.ac.at

Keywords: Exterior orientation, Georeferencing, Geotagging, GNSS-RTK, IMU, Low-cost

Abstract

This paper presents a small, cableless and cost-effective (circa € 500 in 2022) device that embeds an RTK-enabled GNSS receiver and IMU in a 3D printed case to record the exterior orientation of photos via the camera's hot shoe. This hardware solution was developed within the academic graffiti project INDIGO to enable a faster and more robust exterior orientation (via incremental SfM) of the hundreds of weekly acquired graffiti photographs. The device relies solely on commercially available open-source development components that require only minimal knowledge of electronics. Moreover, it is small enough not to disturb the photographer but still provide constant visual feedback about the device's overall and positional solution status. To realistically test the achievable positional and rotational accuracy, we acquired a dense network of graffiti photos in INDIGO's test area. This test showed that the device provides centimetre-accurate photo positions and one-degree accurate pitch and roll values. The heading or yaw is probably only usable to some extent since it primarily relies on magnetometer measurements, which get quickly disturbed in an urban environment. In addition, the test revealed that the camera's internal IMU can provide roll and pitch values with sub-degree accuracy.

1. Introduction

1.1 Project INDIGO

Project INDIGO (IN-ventory and DI-sseminate G-raffiti along the d-O-naukanal; projectindigo.eu) was a two-year academic project launched in September 2021 through funding from the Heritage Science Austria programme of the Austrian Academy of Sciences (ÖAW). This project wanted to push the status quo boundaries in inventorying and understanding extensive graffiti-scapes (Verhoeven et al., 2022).

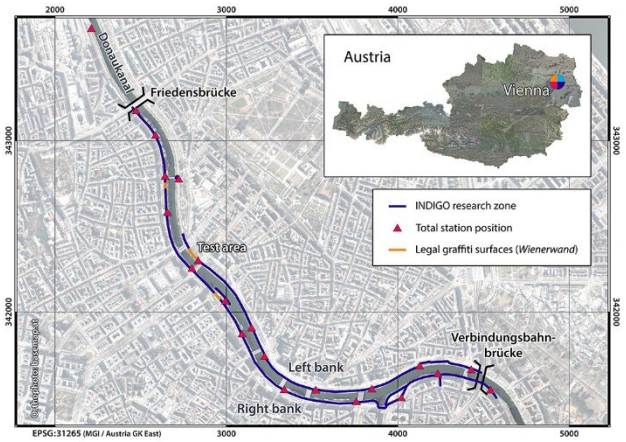

Figure 1. An overview of INDIGO's project area.

INDIGO aimed to ensure the digital survival of a significant portion of Vienna's graffiti-scape while also seeking new insights into socio-political and cultural aspects. Given Vienna's size, it would be impractical to consider the entire city as the research area. For this reason, INDIGO concentrated on one of Vienna's most popular tourist attractions and graffiti areas: the *Donaukanal* (Eng. Danube Canal), a river channel originating from the Danube River in the northwestern part of Vienna. To be more specific, the INDIGO project concentrated on all public surfaces surrounding this central waterway between two bridges: from the *Friedensbrücke* in the northwest to the *Verbindungsbahnbrücke* in the southeast (see Figure 1).

This section of the Danube Canal spans approximately 3.3 km when measured in the middle of the waterway. However, this number inadequately represents the extent of graffiti-covered surfaces examined by INDIGO. Graffiti appears on the left and right bank, above and below the walking path (Figure 2). When considering all graffiti-covered surfaces—walls, bridge pillars, and staircases—above the walking level, the total length reaches 8.5 km. Including the 4.4 km of graffiti below the walking surface results in nearly 13 km of continuous urban graffiti-ed surfaces that INDIGO had to monitor. Within this entire graffiti-scape, graffiti are only legal in three small areas that comprise less than 300 m (see Figure 1). These three legal stretches are part of Vienna's *Wienerwand* (Eng. Viennese wall), a joint label given to the city's legal graffiti zones (see www.wienerwand.at).

Figure 2. Surfaces with graffiti are located above and below the walking and biking area flanking the Danube Canal.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-2-W8-2024-463-2024

1.2 Improving Graffiti Photography

During the entire project period, one of the two photographers documented new graffiti at least once per week. INDIGO relied on two Nikon NIKKOR Z 20mm f/1.8 S lenses paired with a full-frame mirrorless Nikon Z7 II camera (see the lower inset of Figure 3) generating 45-megapixel photos. Both cameras always featured the same settings (Verhoeven et al., 2023). This enforced identical results (from a technical point of view) across imaging systems and ensured that the camera-related photo properties were appropriate for INDIGO's geometric and colourimetric processing pipelines (Molada-Tebar and Verhoeven, 2023; Molada-Tebar et al., 2024).

Initially, a Solmeta Geotagger GMAX was mounted on each camera. This unit uses the American GPS and Chinese Beidou Global Navigation Satellite System (GNSS) to compute the camera's location with a precision of about 2.5 metres (at one standard deviation). This precision can be reached in ideal scenarios because the unit uses the correction signals broadcasted by the satellite-based augmentation system EGNOS (European Geostationary Navigation Overlay Service). The estimated geographical latitude, longitude, and altitude values were directly written into the Exif metadata of the *.NEF and *.JPG photo files. These values were leveraged in the Structure from Motion (SfM) and image orthorectification workflows (Wild et al., 2022; Wild et al., 2023) for computational speed improvements. However, acquiring more accurate coordinates for every camera station was deemed helpful in decreasing processing time even further.

The INDIGO team has, therefore, developed an RTK/IMU (Real-Time Kinematic/Inertial Measurement Unit) device to record the camera's exterior orientation (i.e., its position and angular rotation) when a photo is acquired. Built from commercially available but cost-effective components housed in a 3D printed case, this device connects to the hot shoe on top of the camera. The device can use RTK GNSS correction services, for which the settings can be wirelessly controlled from a tablet or smartphone. Because the device provides accurate time-synchronisation (within tens of milliseconds) with the camera shutter, the positional and rotational values can be correctly logged for each acquired photo.

This paper will first detail this hardware device and the software developed to process the device's data. Afterwards, the accuracy of the obtained positional and angular data values will be quantified, and the main (dis)advantages of this low-cost hard- and software combination will be assessed.

2. Hard-, Firm-, and Software

2.1 Hardware

In recent years, low-cost and dual-frequency GNSS receivers have emerged, offering accuracies close to those of geodetic receivers. In addition, building electronic systems with pre-built boards and tons of open-source libraries for microcontrollers has become easier than ever. Mainly thanks to the birth of the Arduino platform in 2005, only basic electronic knowledge is needed to access a wide range of sensors and platforms, often used for home automation and robotics.

Today, companies specialising in selling electronic parts in the spirit of open-source have developed hundreds of electronic sensor boards and platforms for integration. Most prominent are the USA-based companies Sparkfun (www.sparkfun.com) and Adafruit (www.adafruit.com), which build easy-to-use electronic boards and provide libraries for all of them. The device presented in this paper is solely a combination of existing open-source development boards that need a minimum of soldering, all packed into a 3D-printed case (see Figure 3).

Figure 3. The RTK-enabled GNSS/IMU logging device. The lower inset shows the device on top of the Nikon Z7 II.

The main electronic components of the device are:

- Sparkfun's GPS-RTK2 board (www.sparkfun.com/products/15136), a breakout board for a u-blox ZED-F9P GNSS receiver (www.u-blox.com/en/product/zed-f9p-module), that allows to straightforwardly communicate with the GNSS receiver and access the Pulse-Per-Second (PPS) signal for reliable time synchronisation. The u-blox ZED-F9P module is a multi-constellation and multi-frequency GNSS receiver with a built-in RTK-processor. It comes with an open-source software library (SparkFun Electronics, 2024) to deal with all the ZED-F9P communication. The scientific community has already used this GNSS receiver recently (Hohensinn et al., 2022; Robustelli et al., 2023).
- Sparkfun's EPS32-S2 WROOM board (www.sparkfun.com/products/17743), a development board for Espressif's EPS32-S2 (www.espressif.com/en/products/socs/esp32-s2) module which features a built-in USB-C connector and a lithium accumulator charger. The ESP32-S2 is a low-cost and low-power system-on-a-chip microcontroller with a single-core RISC-V microprocessor and integrated Wi-Fi plus Bluetooth.
- Adafruit's 9 DoF (Degree of Freedom) IMU breakout board (www.adafruit.com/product/4646) featuring Bosch's BNO055 sensor (www.bosch-

sensortec.com/products/smart-sensor-systems/bno055). This solution integrates a 3-axis 14-bit accelerometer, a 3-axis close-loop 16-bit gyroscope, and a microcontroller running a sensor fusion software. The Adafruit unified sensor library (Adafruit, 2023) enables sensor communication.
- A Beitian BT-560 multi-band GNSS helical antenna (www.sparkfun.com/products/retired/17383).
- A single 2000 mAh Lithium Polymer battery (LiPo). This power source, which allows an operation time of about seven hours, can be charged via a USB-C plug on the side of the device.

In 2022, the total cost of one device was around € 500, with the GNSS board and helical antenna being the most expensive parts.

2.2 Firmware

The firmware developed for this device runs on the ESP32 module and is written in C/C++ using the Arduino environment. Open-source libraries exist for all breakout boards used. These libraries provide an abstraction of all sensor communication, allowing easy access to data and configurations. The main functional components of the firmware are:

- A **time-critical synchronisator** handles interrupts and time synchronisation. Using a hardware interrupt event on the GNSS's PPS signal and the received GNSS time, a highly accurate timing reference is constructed to which the camera's hot shoe exposure event is linked by another hardware interrupt, as such defining when the photos are acquired.
- An **exterior orientation processor** handles the communication with the GNSS receiver and the IMU. It also deals with the latter's internal calibration procedure. The display on top of the RTK/IMU device tells the user if calibration is needed and indicates when the IMU operates within its specified accuracy.
- An **NTRIP processor** handles all the communication with the NTRIP (Networked Transport of RTCM via Internet Protocol) server (authenticating, positional updates for the GNSS solution, receiving correction messages) via the RTCM (Radio Technical Commission for Maritime Services) protocol and forwards correction data to the GNSS receiver. Internet is provided over Wi-Fi by a smartphone or tablet.
- A **user interface** allows the user to access webpages, configure the device (as seen in Figure 4), and download its recorded data. This is all made possible via the built-in Wi-Fi of the ESP32-S2, so that the latter can run a basic web server.

The development of this hard- and firmware prototype focused on a smooth user experience. While existing small-form-factor RTK solutions are typically boxes with external cables, often designed for the attachment to drones and usually lacking detailed status information, this integrated solution is a single, 183 g compact device (see the lower inset of Figure 3 for scale) that provides much operational information and lacks any external cables (which would only interfere with the camera operator). For example, operational information is provided by four LEDs (indicating battery low, RTK fixed solution available, camera trigger captured, Wi-Fi connected) mounted on the front (see Figure 3 in the middle). Hence, they are visible to the camera operator at all times. A display on top of the device provides more information on the battery status, the RTK solution accuracy, and the number of images (see the upper inset of Figure 3).

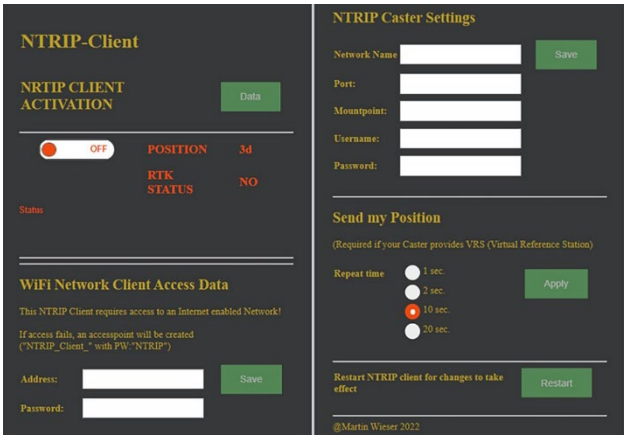

Figure 4. Webpage for the configuration of Wi-Fi and NTRIP settings, which is hosted on a webserver at the ESP32.

2.3 Software

We developed software to handle all post-processing of the RTK/IMU data (Section 3.3). Whereas modern geodetic GNSS receivers can perform on-device coordinate transformations and leverage RTCM-delivered distortion grids, this post-processing software deals with that. It can apply coordinate transformations to systems other than the Coordinate Reference System (CRS) of the RTK solution, including NTv2 (National Transformation version 2) rasters and time-dependent transformations. ExifTool (exiftool.org) is used to geotag the photos by incorporating positional and rotational info into their metadata.

3. Test Data Set

3.1 Data Acquisition

To assess the accuracy and useability of our RTK/IMU device, test data were collected on the 12th of October, 2022. A 45-megapixel full-frame mirrorless Nikon Z 7 II with a Nikon NIKKOR Z 20mm f/1.8 S lens were used to acquire a redundant network of 777 highly overlapping photos. The photographed scene was project INDIGO's dedicated test area (see Verhoeven et al. 2022), partly consisting of a legal graffiti zone. Besides the typically graffiti-covered walls, this test area also features a ramp and a metal bridge with graffiti on all pillars (see Figure 5).

Since it was planned to orient all photos with a Structure from Motion (SfM) approach, the image acquisition was executed according to well-known rules that enable a self-calibrating SfM algorithm to accurately estimate the interior and exterior camera orientations. Although the camera's optical axis was typically perpendicular to the colourful surfaces, a large set of convergent photos (i.e. with a non-orthogonal and inclined optical axis) was also acquired to achieve sufficient intra-image scale variation. The camera was usually held in landscape orientation, but care was taken to acquire enough photos with the camera rotated 180°, 90° clockwise, or 90° anti-clockwise around its optical axis (see Figures 5 and 7).

Due to the surface geometry of the scene, the entire photo set naturally also featured various object distances. Finally, the built-in camera vibration reduction was deactivated, and the focusing ring was immobilised with cellophane tape to ensure invariant interior orientation parameters throughout the image acquisition. The camera stored minimally compressed JPEGs, with *f*/5.6 dialled in for all exposures. This aperture provided sufficient depth of field and led to a high and uniform lens resolving power.

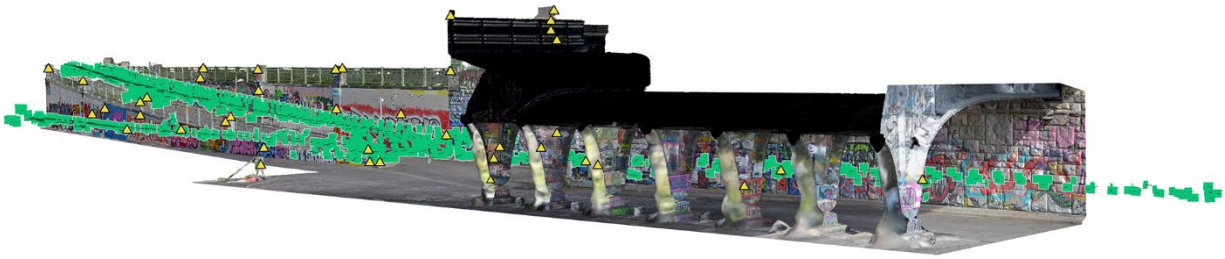

Figure 5. Coloured test area surface mesh. Camera stations and coded targets are symbolised by rectangles and triangles, respectively.

The RTK-enabled GNSS/IMU logging device was attached to the camera during the entire survey. RTK corrections were received via the Austrian EPOSA service (*Echtzeit-Positionierung-Austria*, Eng. Real-time positioning Austria) (www.eposa.at), for which the device was configured with the NTRIP Mounting Point RTK-32-4G (using GPS, GLONASS, Galileo, BDS) in VRS (Virtual Reference Station) Mode.

41 coded targets printed on white A4 paper were attached across the scene. The 3D coordinates of their centres were acquired by measuring distances and bearings from the known position of the Leica Viva TS16 total station, which was placed at four different locations to ensure visibility to all targets. The coordinates of these four total station locations were determined using free stationing. To that end, distances and bearings between the total station and a Leica GPR121 circular prism were measured to an average of nine visible Control Network Points (CNPs), each part of an extensive network established by the City of Vienna. The coordinates of these CNPs are expressed in the MGI/Austria GK East (EPSG:31256) CRS using the Vienna height (Wiener Null) vertical CRS (EPSG:8881). The latter is a simple offset to the GHA height (heights in use, EPSG:5778). The coordinates of all Viennese CNPs are available at www.wien.gv.at/ma41datenviewer/public/start.aspx.

3.2 Processing of the Reference Data

The 3D coordinates of the four TS16 centres were calculated using free stationing, which involves a 2D Helmert transformation to obtain the *x*- and *y*-coordinates, followed by a trigonometric determination of height to determine the *z*-coordinate. The parameters of the Helmert transformation were calculated from the CNP coordinates, which are known in both the target and the source CRS (a local coordinate system with the total station's centre as the origin). The transformation parameters were obtained after minimising the sum of the squares of the residuals. Residuals over 4 cm were considered outliers, and these CNPs were removed. In the end, the average standard deviation of the residuals equalled 9 mm in planimetry and 5 mm for the *z*-coordinates. Based on the total station's known positions and the reflectorless measured distances and bearings to the printed coded target centres, their 3D coordinates were determined in EPSG:31256/8881. All the abovementioned geodetic calculations were conducted in IDC EDV's Geosi Verm v. 21 (idc-edv.at/geosi/geosi-verm).

The 777 photographs were oriented in Agisoft Metashape Professional 2.1.3 using a maximum of 40k interest points and 4k tie points per image. The self-calibrating bundle adjustment of Metashape's SfM algorithm used the 3D coordinates of the 41 coded targets (indicated 1085 times) as constraints (with a 5 mm precision to weigh these external observations). After tie point cleaning and rerunning the bundle adjustment, georeferencing accuracies were all below 1 cm (see Table 1). That result is now called the "**reference block**" and used for comparison purposes.

Georeferencing accuracy metric	Value
$RMSE_x$	5.3 mm
$RMSE_y$	4.0 mm
$RMSE_{xy}$ (total planimetric accuracy)	6.6 mm
$RMSE_z$	5.8 mm
$RMSE_{xyz}$ (total 3D positional accuracy)	8.8 mm

Table 1. Accuracy metrics for the oriented image network.

3.3 Post-Processing of the RTK/IMU Data

The RTK/IMU device records several measurements per photo trigger, mainly the device's GNSS position in ellipsoid coordinates (in the reference station network's coordinate reference frame), the RTK solution flag, the RTK solution accuracy and the IMU values of the device as quaternions. However, the software performs two major transformation steps to bring all the recorded data into the CRS we use for analysis.

Coordinate Transformations: Since EPOSA was operating in coordinate reference frame ITRF2014 Ep. 2010.0 at the time of the survey, transformations need to be applied (see Figure 6) to get planimetric MGI/Austria GK East (EPSG:31256) coordinates and Vienna height *z*-coordinates (EPSG:8881). First, the recorded coordinates are transformed to ETRF2000 Ep. 2002.56. Second, a transformation (EPSG:9910) is applied in the form of an NTv2 raster ("GIS GRID 2021-09-2", provided by state authorities) to get into the national survey system MGI (*MilitärGeographisches Institut*). In addition, a height correction grid (EPSG:9499) applies a geoid undulation correction and a local height correction ("*Hoehengrid plus Geoid V2*") to get from ETRS89/GRS80 ellipsoid heights directly to MGI/GHA heights (EPSG:5778). Both grids are available at Proj.org (cdn.proj.org). Finally, we projected the planimetric coordinates to MGI/Austria GK East (EPSG:31256) and applied a height offset to end up with Vienna heights (EPSG:8881).

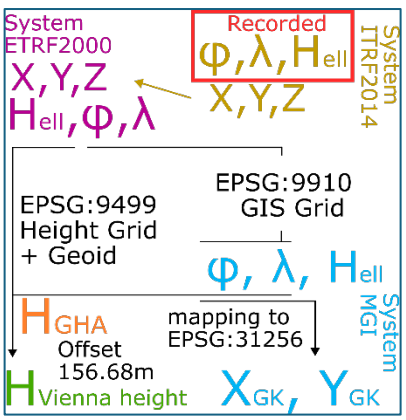

Figure 6. Coordinate transformation steps for our study area.

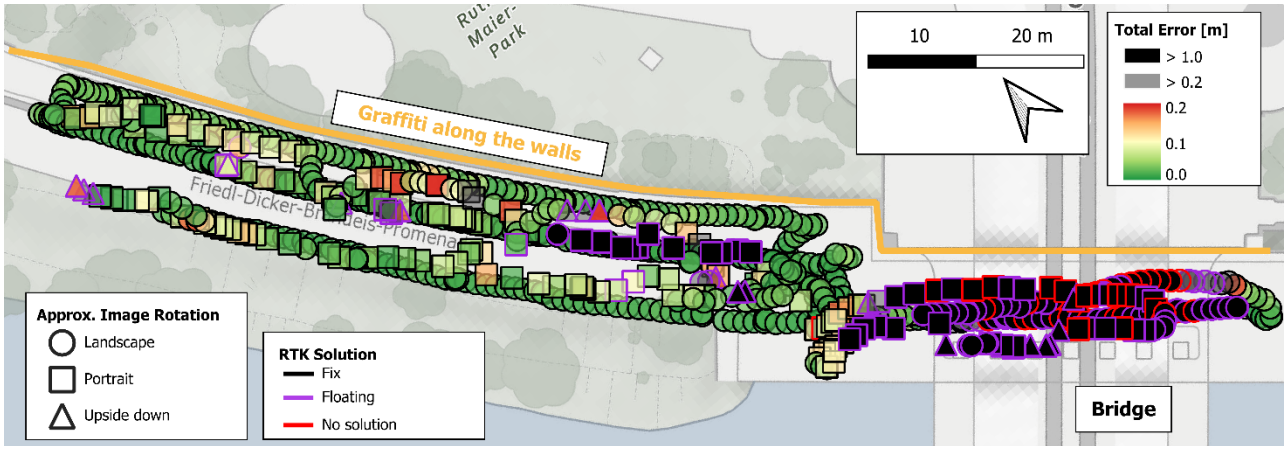

Figure 7. Camera positions with colour-coded total positional error. The right side of the map (south-east) shows the photos acquired under the steel bridge. The camera symbol indicates its rotation, while the symbol's border colour encodes the RTK solution status. The wall of interest with the graffiti is on the northern side indicated by the orange line.

Rotation Transformation: The CRS for the RTK/IMU device was defined with its origin at the antenna's approximate electrical phase centre; its X-axis points to the front of the device, with the Y-axis to its left and the Z-axis upwards. Based on a North-East-Down convention, the IMU provides the rotation of its CRS with respect to a local tangent plane as quaternions. These rotations are transformed to get them for the RTK/IMU CRS in an East-North-Up convention. Finally, position-dependent corrections for magnetic declination and grid convergence were applied. A sequence was used to get common roll, pitch, and yaw angles, where roll and pitch define the tilt to the local tangent plane. No full boresight/mounting calibration with respect to the camera CRS was performed for reasons described in Section 4.

4. Accuracy Analysis

The analysis consists of three parts. First, we analyse the exterior orientation derived from the RTK/IMU device and compare it with the exterior orientations of the "reference block", where solely the coded target's 3D coordinates were used in the bundle block adjustment. Second, we only use the positions derived from the RTK/IMU device to perform a bundle block adjustment to analyse the device's feasibility for INDIGO's intended workflow. Third, we investigate the rotation angles provided by the Nikon camera's internal IMU.

4.1 RTK/IMU Device Accuracy Assessment

As described above, the device records the RTK solution flag (0: No solution, 1: Floating, and 2: Fix) and position accuracies. The distribution of the quality of RTK positioning is shown in Figure 7. As expected, the positional error is mostly below 0.1 m in the open area. However, the RTK solution and overall GNSS positioning are not working under or close to the bridge, with positional errors of up to 20 m, even for a Floating solution (see Table 2). However, this area was explicitly chosen to reveal such issues and check if we can use information from the GNSS receiver to filter out points with degraded accuracy. In addition, Figure 7 reveals that for photos without an upside-up landscape rotation (i.e., those symbolised with square and triangular markers), the device tends to have greater positional errors and is quicker to lose its Fix solution. We suspect that if the antenna is not pointing to the sky, there are two primary error sources: GNSS shadowing by the user's body and likely the device's case on the one hand, and the antenna's electrical characteristics and phase centre position on the other. Although these potential factors were not analysed in detail, it is clear that the RTK/IMU device might not provide a Fix solution (resulting in positional errors beyond 1 m) when the camera is rotated 180°, 90° clockwise, or 90° anti-clockwise around its optical axis, even if these camera stations are located in areas that feature otherwise accurate positional measurements. Acquiring enough images with the camera in upside-up landscape mode is thus crucial.

Figure 7 and Table 2 also reveal that a Fix solution does not guarantee highly accurate results. More specifically, Table 2 clarifies that the GNSS receiver overestimates the accuracy of the recorded positions. By combining the RTK solution flag and the solution accuracy, we classified all recorded camera positions into two classes: "first-order quality" RTK-GNSS positions (with a Fix solution and a reported accuracy of 0.01 m) and all others. Of these 586 first-order positions, only 14 (or 2.4 %) have a total positional error of more than 15 cm compared to the reference block, with a maximal total positional deviation of 46 cm.

Solution Flag	Images	Recorded Position Accuracy	Total Error to Reference Block
No	46	0.12 m – 1.00 m	0.33 m – 13.00 m
Floating	133	0.01 m – 1.00 m	0.05 m – 21.00 m
Fix	598	0.01 m – 0.03 m	0.02 m – 0.53 m
Class	**Images**	**Recorded Position Accuracy**	**Total Error to Reference Block**
1st order quality	**586**	**0.01 m**	**0.02 – 0.46 m only 14 >15 cm**
All others	191	0.01 m – 1.00 m	0.05 m – 21.00 m

Table 2. Summary of the RTK solutions.

To compute these positional errors, one must first compute the lever arm (also known as the mounting or displacement vector) between the origin of the RTK/IMU CRS and the camera's CRS, which originates at the lens' perspective centre. Using all first-order quality positions, Metashape quantified the lever arm in the camera CRS as **x = -0.019 m, y = 0.089 m** and **z = 0.105 m**. Since these values were close to manually measured values, they were assumed to be accurate. The positions logged by the RTK/IMU device were then adjusted for this lever arm.

Figure 7 shows the differences between the 586 first-order quality image positions and the reference block. The overall RMSE of the total positional error (Euclidean distance) is 7.6 cm, with a small systematic offset for each coordinate component (see Table 3). Worth mentioning is that for 129 of those positions

(i.e., 22.0 %), the roll or pitch angle (compared to the local tangent plane) exceeded 30°, which goes to show that the RTK/IMU device can also produce accurate coordinates if they are indicated as first-order quality points.

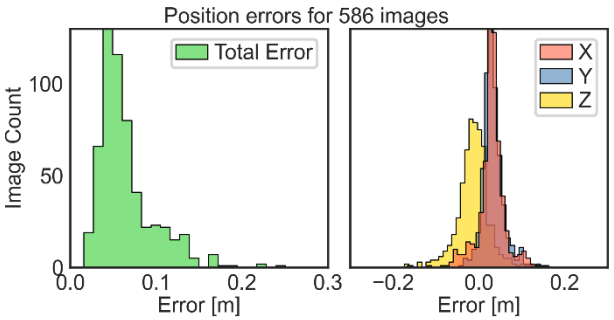

Figure 8. Histograms of positional errors. Left: total error; Right: error per coordinate component.

Although one could examine many possible explanations for the small systematic offsets (e.g., local distortions in the CNP network, the accuracy of the RTK correction), we refrained from doing so as these offsets were within our expectations.

Error	X [cm]	Y [cm]	Z [cm]	Total [cm]	Roll [°]	Pitch [°]	Yaw [°]
Mean	2.9	2.8	-1.9	6.5	0.17	0.73	2.93
Median	3.0	2.5	-1.8	5.4	0.01	0.59	1.79
St. Dev.	3.1	2.7	4.4	3.9	1.22	0.81	15.3
RMSE	4.3	3.9	4.8	7.6	1.23	1.09	15.6

Table 3. Statistics for positional and angular errors. For positional errors, only first-order camera stations were used.

The errors of the device's rotation angles are shown in Figure 9 and listed in Table 3. The RMSE of roll and pitch errors equals 1.23° and 1.09°, respectively (computed over all 777 photos). Such values are to be expected for that kind of IMU. Minimal offsets (see Table 3) are also discernable here. However, these are normal because we did not perform a boresight calibration to quantify the exact rotation between the camera's CRS and the CRS of the RTK/IMU device (see also Section 3.3) for two reasons: the orientation relationship between both CRSs will slightly change every time the RTK/IMU device is mounted on top of the camera, and the yaw values provided by the IMU are too inaccurate for a decent boresight calibration.

The yaw values suffer from deviations up to 15° and more (see Table 3 plus Figures 9 and 10). The main problem lies with the built-in magnetometer. It is the primary source for determining the yaw angle in quasi-static trajectories, but it suffers considerably from magnetic disturbances, like the steel bridge in our test area. Still, even on sections further away from the bridge, the yaw angle was sometimes entirely off. We have no valid explanation at the moment, but the metal fences or lantern posts along the wall might be the cause of these magnetic disturbances.

To use the provided orientation angles to the full extent (e.g., for direct georeferencing), one could try to better calibrate the IMU's magnetometer (which would reduce magnetic disturbances) or, more importantly, investigate if the chip provides accuracy values to remove erroneous measurements in post-processing. However, we currently cannot tell which yaw angles to discard.

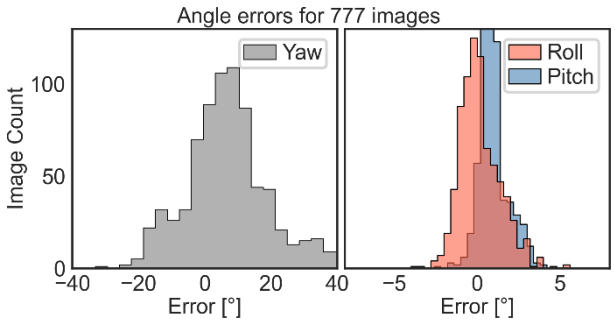

Figure 9. Histograms of angle errors. Yaw is on the left; Roll and pitch are on the right.

4.2 Bundle Block Adjustment with RTK-GNSS Positions

To check the device's feasibility in our workflow, we performed a self-calibrating SfM solely using the RTK-GNSS positions of first-order quality as constraints for the bundle adjustment. The SfM processing parameters were identical to the reference block, except that Metashape accounted for a lever arm. No further block optimisation was undertaken. Again, all 777 photos could be oriented.

As expected, the differences between these interior orientations and those of the reference block were negligible. Only the antenna lever arm was slightly different, with **x = -0.016 m**, **y = 0.132 m** and **z = 0.120 m**. A comparison of all exterior orientations with those of the reference block leads to the positional errors graphed in Figure 11. None of the 777 positions deviates by more than 0.2 m.

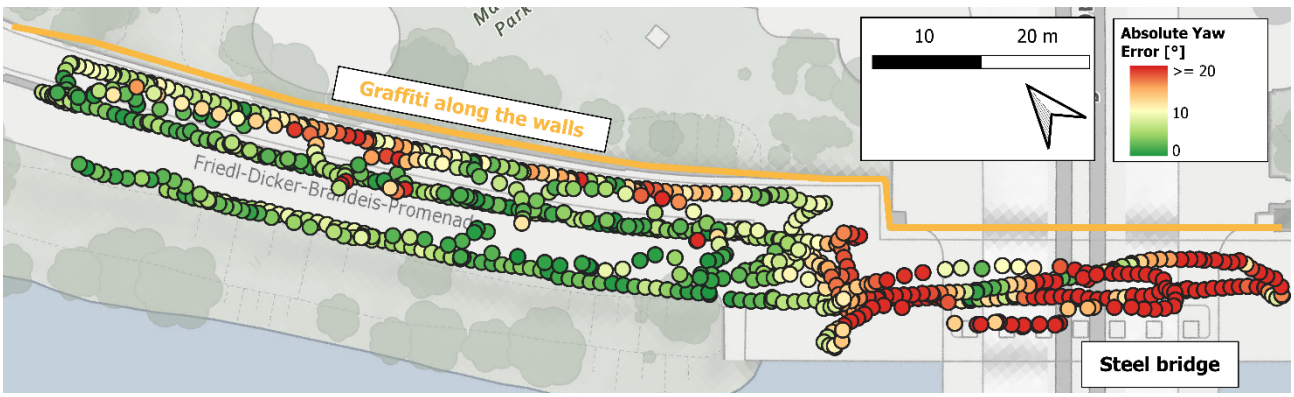

Figure 10. Camera positions with colour-coded yaw error. The right side of the map shows photos acquired under the steel bridge.

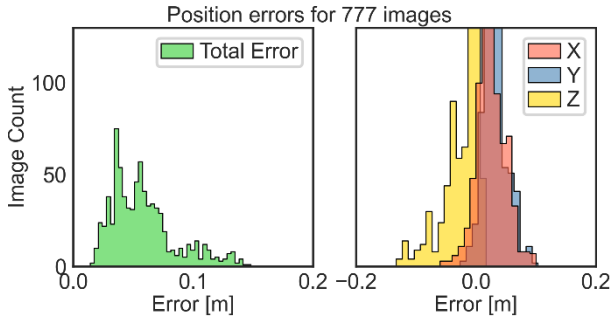

Figure 11. Histograms of positional errors between the reference block and one constrained by first-order positions. Left: total error; Right: error per coordinate component.

The bias in the X and Y coordinates is logically also present here. However, the errors for all 777 camera stations (see Table 4) are now slightly better than what the RTK/IMU device achieves for the first-order ones (see Table 3) because the positional and rotational values have been optimised in a bundle adjustment.

Error	X	Y	Z	Total	Roll	Pitch	Yaw
	[cm]			[cm]	[millidegree]		
Mean	2.4	2.9	-2.9	5.8	3	0	5
Median	2.3	2.6	-1.6	5.3	3	0	4
St. Dev.	2.4	1.9	3.0	2.7	47	67	35
RMSE	3.4	3.5	4.2	6.4	47	67	35

Table 4. Statistics of exterior orientation differences between the reference block and one constrained by first-order positions.

Figure 12 visualises that a bundle adjustment constrained with RTK-GNSS position yields accurate orientation angles. For most photos, all three rotation angles are within 0.1° of the reference block. This indicates that no large local deformations are present in the image block and opens up new possibilities for incremental SfM (see Section 5).

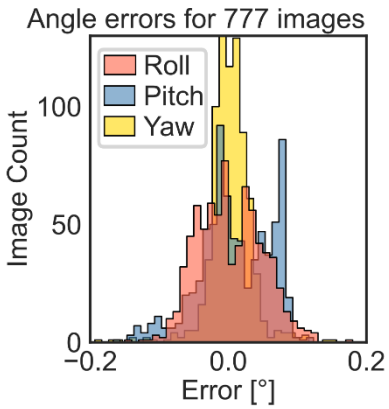

Figure 12. Histogram of roll, pitch and yaw errors between the reference block and one constrained by first-order positions.

However, the accuracy of these results may not be overrated. The image acquisition ensured that a highly redundant photo network was created, so block deformations will always be minimal if enough valid RTK-GNSS positions are available. Further tests would be needed to define the absolute minimum camera network configuration that yields angle errors in that range.

4.3 Nikon's Internal IMU

Many digital cameras, certainly newer ones like the Nikon Z7 II, rely on internal IMU data for image stabilisation, to present a virtual horizon in the viewfinder or on the back display, and to detect the camera's (upside-up or -down) landscape or portrait rotation. These IMU data are often stored as image metadata into manufacturer-specific Exif Makernote tags. However, for the Nikon Z7 II, these angles can not be used directly. Depending on the general rotation of the camera (stored in the Exif Orientation tag, ID 274), the recorded values need different post-processing. Rotated camera poses require a swap of the yaw and pitch values, sign changes or 180° subtractions. Since no rotation notation was found for this behaviour, values were altered programmatically.

Nikon's roll and pitch angles are pretty accurate (see Figure 13), with RMSE angle errors of 0.9° and 0.7°, respectively (see Table 5). These values are better than the current implementation of our IMU. In addition, only a few outliers are present in the roll and pitch values, showing the potential to use these angles to rotate a bundle block to the local tangent plane. However, the substantial deviations of the yaw angles (compared to the reference block) render them unusable.

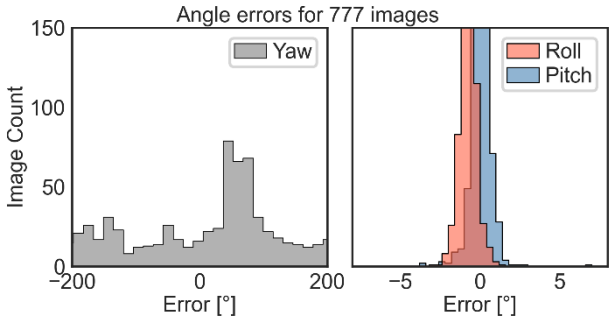

Figure 13. Histogram of the errors characterising the Nikon's recorded angles compared to the reference block.

Error [°]	Mean	Median	St.dev	RMSE	Min	Max
Roll	-0.7	-0.7	0.5	0.9	-2.8	1.2
Pitch	-0.2	-0.1	0.7	0.7	-3.8	6.5
Yaw	-15.6	-33.5	123.5	124.5	-259.6	350.5

Table 5. Statistics of the errors characterising the Nikon's recorded angles compared to the reference block.

5. Discussion

5.1 Exterior Orientation Accuracy

The low-cost device assessed here provides a camera's complete exterior orientation. However, the latter suffers from unreliable magnetometer measurements. Luckily, both the device's and the camera's IMU deliver reliable angles with respect to a local tangent plane. These angular data could thus be used to orient SfM-derived photo blocks at least vertically (see also Nocerino and Menna, 2023) if one can not rely on GNSS positions (e.g., inaccurate positional data or a sparse photo network).

Even though the device sometimes suffered from RTK solution problems, we could use the GNSS receiver's recorded data to remove erroneous position measurements. Given the challenging but realistic test area, the device's data can be used to accurately georeference a photo network or constrain a bundle adjustment.

5.2 Comparison with Similar Systems

Various GNSS-based systems exist to assist the exterior orientation of photo cameras. Oniga et al. (2024) have been using hardware components of the Hungary-based company Emlid (emlid.com) to obtain positional camera data via two different setups. However, these setups were bulkier than our solution or the recent systems from Geobsys and REDcatch.

The French company Geobsys developed the GEOSTIX, a small cylindrical case to house a single- or multiple-frequency GNSS receiver (www.geobsys.com/geostix). However, the system lacks an IMU and seems to come without a hot shoe synchronisation interface (although that is likely not so hard to solve). Another interesting option is offered by REDcatch. This Austrian company updated their bulkier 3D ImageVector (i.e., an RTK-GNSS receiver attached to the camera's hot shoe) to a more compact, cableless but sadly also IMU-less HotShoe RTK solution (www.redcatch.at/hotshoe-rtk). The authors tested none of these devices, but REDcatch certainly seems to offer a suitable ready-to-use device for centimetre-accurate image positioning. The GEOSTIX and HotShoe RTK solutions were also unavailable when we started developing our RTK/IMU solution.

5.3 INDIGO's Incremental SfM Workflow

Project INDIGO relied on two so-called "total coverage" surveys during which the entire research zone was photographed. Orienting these photos yielded a base photo network to which new graffiti photos were added via an incremental SfM step that limited the search space via the photo coordinates provided by the Solmeta Geotagger GMAX (see Section 1.2). The RTK/IMU device data now allows for optimising this workflow. One could first compute a temporary SfM solution for the new graffiti photos – with the positional data of the RTK/IMU device as constraints – to get improved angular orientation values (certainly for yaw). Combining these computed attitude angles with the logged RTK-GNSS positions could tightly constrain the object space when looking for tie point subsets of previously established image networks, or even finding older photos.

5.4 Disadvantages

Using our RTK/IMU device in project INDIGO revealed two disadvantages. First, additional mass on the camera – even when just a few hundred grams – makes long-duration image surveys slightly more exhausting. Second, our device does not get a sync signal when the camera operates in silent mode or when the interval priority timer mode is used. However, these are Nikon Z7 II limitations; they are not inherent to our RTK/IMU device.

6. Conclusion

This paper demonstrated that with modern open-source and cost-effective components, it is possible to build a relatively accurate RTK/IMU device to record the exterior orientation of camera stations during photo acquisition. An assessment of the device's provided parameters revealed a total positional accuracy of 6.4 cm (using only valid RTK measurements) with a 2.7 cm precision, while the rotational accuracy was circa 1° with respect to the local tangent plane. The yaw angle could not be filtered sufficiently. Afterwards, the text discussed the device's pros and cons concerning similar solutions and its intended workflow.

7. Acknowledgements

INDIGO was funded by the Heritage Science Austria programme of the Austrian Academy of Sciences (ÖAW).

References

Adafruit, 2023. Common sensor library. Version 1.1.14. https://github.com/adafruit/Adafruit_Sensor (28 October 2024).

Hohensinn, R., Stauffer, R., Glaner, M.F., Herrera Pinzón, I.D., et al., 2022. Low-Cost GNSS and Real-Time PPP: Assessing the Precision of the u-blox ZED-F9P for Kinematic Monitoring Applications. *Remote Sens*., 14, 5100, doi:10.3390/rs14205100.

Molada-Tebar, A., Verhoeven, G., Hernández-López, D., González-Aguilera, D., 2024. Practical RGB-to-XYZ Color Transformation Matrix Estimation under Different Lighting Conditions for Graffiti Documentation. *Sensors*, 24, 1743, doi:10.3390/s24061743.

Molada-Tebar, A., Verhoeven, G., 2023. Towards colour-accurate documentation of anonymous expressions. In *document | archive | disseminate graffiti-scapes. Proceedings of the goINDIGO 2022 international graffiti symposium,* Vienna, 11-13 May 2022, Urban Creativity: Lisbon, 86–130.

Nocerino, E., Menna, F., 2023. In-camera IMU angular data for orthophoto projection in underwater photogrammetry. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, 7, 100027, doi:10.1016/j.ophoto.2022.100027.

Oniga, E., Boroianu, B., Morelli, L., Remondino, F., Macovei, M., 2024. Beyond ground control points: cost-effective 3D building reconstruction through GNSS-integrated photogrammetry. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-2/W4-2024, 333–339, doi:10.5194/isprs-archives-XLVIII-2-W4-2024-333-2024.

Robustelli, U., Cutugno, M., Pugliano, G., 2023: Low-Cost GNSS and PPP-RTK: Investigating the Capabilities of the u-blox ZED-F9P Module. *Sensors*, 23, doi:10.3390/s23136074.

SparkFun Electronics, 2024. An Arduino library which allows you to communicate seamlessly with u-blox GNSS modules using the Configuration Interface. Version 3.1.7. https://github.com/sparkfun/SparkFun_u-blox_GNSS_v3 (28 October 2024).

Verhoeven, G., Wild, B., Schlegel, J., Wieser, M., Pfeifer, N., Wogrin, S., Eysn, L., et al., 2022. Project INDIGO – document, disseminate & analyse a graffiti-scape. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVI-2/W1-2022, 513–520, doi:10.5194/isprs-archives-XLVI-2-W1-2022-513-2022.

Verhoeven, G., Wogrin, S., Schlegel, J., Wieser, M., Wild, B., 2023. Facing a chameleon—How project INDIGO discovers and records new graffiti. In *document | archive | disseminate graffiti-scapes. Proceedings of the goINDIGO 2022 international graffiti symposium,* Vienna, 11-13 May 2022, Urban Creativity, 63–85.

Wild, B., Verhoeven, G., Wieser, M., Ressl, C., Schlegel, J., Wogrin, S., Otepka-Schremmer, J., Pfeifer, N., 2022: AUTOGRAF—AUTomated Orthorectification of GRAFfiti Photos. *Heritage*, 5, 2987–3009, doi:10.3390/heritage5040155.

Wild, B., Verhoeven, G., Wogrin, S., Wieser, M., Ressl, C., Otepka-Schremmer, J., Pfeifer, N., 2023. Urban creativity meets engineering. In *document | archive | disseminate graffiti-scapes. Proceedings of the goINDIGO 2022 international graffiti symposium,* Vienna, 11-13 May 2022, Urban Creativity: Lisbon, 131–145.