Cost-effective geocoding with exterior orientation for airborne and terrestrial archaeological photography – possibilities and limitations

Martin Wieser[a], Geert Verhoeven[b,c], Christian Briese[a,c], Michael Doneus[b,d,c], Wilfried Karel[a,b] and Norbert Pfeifer[a]

[a]Department of Geodesy and Geoinformation, Vienna University of Technology, Gusshausstrasse 27-29, 1040 Vienna, Austria, (martin.wieser; christian.briese; norbert.pfeifer)@geo.tuwien.ac.at

[b]VIAS – Vienna Institute for Archaeological Science, University of Vienna, Franz-Klein-Gasse 1, 1190 Vienna, Austria, (geert.verhoeven; michael.doneus; wilfried.karel)@univie.ac.at

[c]LBI for Archaeological Prospection and Virtual Archaeology, Franz-Klein-Gasse 1, 1190 Vienna, Austria

[d]Department for Prehistoric and Historical Archaeology, University of Vienna, Franz-Klein-Gasse 1, 1190 Vienna, Austria

Cost-effective geocoding with exterior orientation for airborne and terrestrial archaeological photography – possibilities and limitations

Martin Wieser, Geert Verhoeven, Christian Briese, Michael Doneus, Wilfried Karel and Norbert Pfeifer

Abstract:

Taking a photograph is often considered to be an indispensable procedural step in many archaeological fields (e.g. excavating), whereas some sub-disciplines (e.g. aerial archaeology) often consider photographs to be the prime data source. Whether they were acquired on the ground or from the air, digital cameras save with each photograph the exact date and time of acquisition and additionally enable to store the camera's geographical location in specific metadata fields. This location is typically obtained from GNSS (Global Navigation Satellite System) receivers, either operating in continuous mode to record the path of the camera platform, or the position is observed for each exposure individually. Although such positional information has huge advantages in archiving the imagery, this approach has several limits as it does not record the complete exterior orientation of the camera. More specifically, the essential roll, pitch and yaw camera angles are missing, thus the viewing direction and the camera rotation around it. Besides enabling to define the exact portion of the scene that was photographed (essential for proper archiving), these parameters can also aid the subsequent orthophoto production workflows and even guide photo acquisition.

This paper proposes a cost-effective hard- and software solution (camera position: 2.5 m and orientation in static conditions: maximally 2°, both at 1σ) to record all indispensable exterior orientation

parameters during image acquisition. After the introduction of the utilized hardware components, the software that allows recording and estimating these parameters as well as embedding them into the image metadata is introduced. Afterwards, the obtainable accuracy in both static (i.e. terrestrial) and dynamic (i.e. airborne) conditions are calculated and assessed. Finally, the good use of this solution for different archaeological purposes will be detailed and commented where needed, while an outlook on future developments finalizes this article.

1. Introduction

1.1. Photography in archaeology

For at least a century, photography has been integrated in archaeological practice, primarily because it allows for a straightforward pictorial record. While far from being the sole source of data in most archaeological fields, taking a photograph is often considered to be an indispensable procedural step. Apart from its more mundane function as illustration and memory aid, excavation photography was often promoted as a means of registering soil features and sections since the 1950s. To that end, different types of photo platforms were developed, allowing for an elevated view on the excavation, and subsequent photogrammetric processing (e.g. [1], [2]). Besides this ground-based imaging, photographs are not only considered a cherished commodity but even an indispensable basis for the sub-discipline of aerial archaeology ([3], [4]).

The advent of digital photography in the 1990s opened up a completely new world for a lot of people, including archaeologists. There has never been a moment in history when so many people actually owned and used a photo camera, either embedded in a mobile phone or as a device designed merely for taking photographs. Since the exact time and specific camera parameters related to the image acquisition are now directly and unambiguously stored inside the Exif (Exchangeable image file format) metadata fields of the image, also the documentational and archival value of the photographs improved. However, to be truly of archival significance, the images' metadata should also hold information on the exact position the image was taken from and – even better – the exact

location of the scene or object depicted in the photograph. In the next section, it will be shown why this archaeological desideratum was mainly pushed in the discipline of airborne photography.

1.2. Archaeological airborne remote sensing

Remote sensing can be broadly defined as the collection of information about a scene or object without having direct physical contact with it [5]. In archaeology, remote sensing mainly relies on passive air- and spaceborne imaging or the active sounding technique known as Airborne Laser Scanning (ALS), although the specific selection of appropriate techniques is entirely due to the nature of most (hidden) archaeological remains. The latter can show up on the earth's surface in a number of ways. Often, the remains of buried archaeology change the chemical and physical properties of the surrounding soil matrix and these might result in colour or height differences of vegetation on top of the remains (i.e. crop or vegetation marks) or distinct tonal variances in the ploughed soil (soil marks) ([6], [7]). Mapping the resulting variances in soil or plant reflectance is generally performed by air- or spaceborne imaging devices. Although archaeology often relies on satellite sensors to record these reflectance differences [8], passive airborne imaging such as archaeological aerial photography is by default better suited for the discovery and detailed recording of site-level visibility marks due to its inherent higher spatial resolution.

To date, the common practice of archaeological aerial photographic reconnaissance is quite straightforward and seems not to have significantly changed over the past century. In general, photographs are acquired from the cabin of a low-flying aircraft (preferably a high-wing airplane) using a small- or medium-format hand-held photographic/still frame camera [9]. Once airborne, the archaeologist flies over targeted areas and tries to detect possible archaeologically-induced crop and soil marks. As soon as an archaeological feature is detected, it is orbited and commonly documented from various oblique viewpoints. This type of aerial photographic reconnaissance has been the workhorse of all archaeological remote sensing techniques since it is one of the most cost-effective methods for site discovery and the non-invasive approach yields easily interpretable imagery with abundant spatial detail.

Due to the fact that flying paths and photo locations are never predefined in this oblique reconnaissance approach and accurate mapping and photo interpretation necessitates knowledge about the part of the earth's surface covered by the aerial image, this information should ideally be recorded during photo acquisition. If not, the subsequent image management and interpretation workflow becomes very time-consuming and certain questions are difficult to

answer (e.g. "Where was this photograph taken?" or "Which pictures cover that area?"). In the worst case scenario, retrieving the exposure location of a specific photograph might even prove impossible.

1.3. Objectives

The aim of this research was to link a digital camera with a cost-effective hardware solution to observe all exterior orientation parameters (see 2.2 for a detailed description) at the moment of image acquisition, so that the depicted object area can be computed and stored. The hardware of the system should consist of off-the-shelve components that enable a straightforward logging of all essential parameters for further handling in a specifically developed MATLAB-based processing chain. The resulting exterior orientation should be saved inside the Exif metadata tags of the image as well as an additional sidecar metadata file. The solution should be flexible enough to work both in an airborne and terrestrial environment, while the accuracy and precision of the estimated exterior orientation must be sufficient for image archiving and – if possible – as initial values for an orthophoto production workflow. More specifically, this means that the discrepancy between the observed and the real exterior orientation parameters should be maximally 3° for the rotation values and 3 meters for the positional values. Several ground-based and airborne tests will enable to assess the usability of the proposed solution with respect to these objectives.

2. Archaeological State-of-the Art

2.1. Geocoding in (aerial) archaeology

Generally, embedding geographic coordinates into (aerial) imagery can be executed using three possible approaches: a software, hardware, and hybrid approach. In its simplest form (i.e. the software approach), the user has to manually or semi-automatically input coordinates extracted from any other georeferenced spatial dataset. This approach takes, however, place after the flight, maybe supported by a flight protocol, but is not advised for the previously mentioned reasons.

More handy is the hybrid soft- and hardware solution, which tags the photographs with the locations stored in the continuous track log of any external, handheld Gobal Navigation Satellite System (GNSS) receiver or more dedicated GNSS data loggers such as Qstarz's BT-Q1000XT Travel Recorder or the GiSTEQPhotoTrackr Mini. After the aerial sortie, many commercial or freely available software packages can synchronize both data sources by comparing the time-stamped GNSS track with the time of image acquisition stored in the Exif metadata fields of the aerial image. Subsequently, the coordinates of the corresponding GNSS point are written as new location data into

the image file or in a separate *.xmp sidecar file, which features the same name as the image file and stores the metadata using Adobe's eXtensible Metadata Platform (XMP) data model [10]. This hybrid workflow is still very popular in archaeological aerial photography, as the GNSS logger can easily be left unattended in the front of the cockpit (see Figure 1). In the case of terrestrial imaging, a separate GNSS logger is often too cumbersome. Excavation photography therefore uses mainly the third approach: hardware-based geocoding.

Hardware-based geocoding is very straightforward, since the camera's firmware (i.e. hardware embedded software) takes care of all the rest. Several compacts (e.g. Sony Cyber-shot DSC-HX30V and Canon PowerShot S100), bridge (e.g. Nikon Coolpix P510 and Sony Cyber-shot DSC-HX100V) or Single-Lens Reflex (SLR) cameras (e.g. Sony SLT-A99, Nikon D5300 and Canon EOS 6D) already feature a built-in GNSS receiver. More common is the option to physically link a separate GNSS receiver to a digital camera. Until a few years ago, only the high-end Nikon digital SLR cameras such as the D2X(s), D2Hs, D1X, D1H and D200 together with the Fuji S5 Pro supported this flexible way of geocoding. Currently, several Canon and Pentax SLR models also offer this as an option, while Samsung, Canon, Leica and Nikon even included the option to attach a manufacturer-specific GNSS receiver to one or more of their mirrorless cameras.

Irrespective of the workflow (hardware, software or hybrid), the end result is a so-called geocoded image: an image that was assigned a geographic identifier in its metadata (a geocode) to pinpoint its location somewhere on the earth. Since this is generally done by writing geographical coordinates into some pre-defined Exif metadata

Figure 1. The aerial photographer inside a Cessna 172 holding a two-camera rig while the GNSS receiver is located in the front of the cockpit (photograph by F. Vermeulen).

tags of that particular photograph, location stamping or geotagging are often used synonyms for this type of image geocoding.

2.2. Exterior orientation

When applying any of the aforementioned geocoding methods, the Exif tags will only represent the position of the camera at the moment of exposure. This is by no means an accurate way of describing the scene that is captured in the (aerial) image. To achieve this, additional information is needed. A camera is always placed at a certain spatial location, but it is also pointed into a specific direction. The camera location is defined by the perspective centre O whose coordinates (X_O, Y_O, Z_O) are expressed in a global Coordinate Reference System (CRS) XYZ (Figure 2). Of those three coordinates, Z_O is often equal to the flying height H above the reference ellipsoid for an aerial image [11].

The angular orientation of the camera is defined by three rotation/Euler angles roll (ω), pitch (φ) and yaw (κ) which define the angular relationships between the axis of the camera CRS (xyz in Figure 2) and the global CRS (XYZ in Figure 2). The first rotation "roll" (ω) is about the X axis, the second rotation "pitch" (φ) about the Y axis and the last rotation "yaw" (κ) is about the Z axis (see Figure 2). Together, these six parameters establish the so-called exterior orientation [12] that describe the spatial position and orientation of the camera CRS with respect to the ground CRS. Synonyms, often used in the field of computer vision, are camera extrinsics, camera pose or extrinsic orientation.

The rotation angles of the camera can be obtained by a so-called Inertial Measurement Unit (IMU) or Inertial Reference Unit (IRU), which

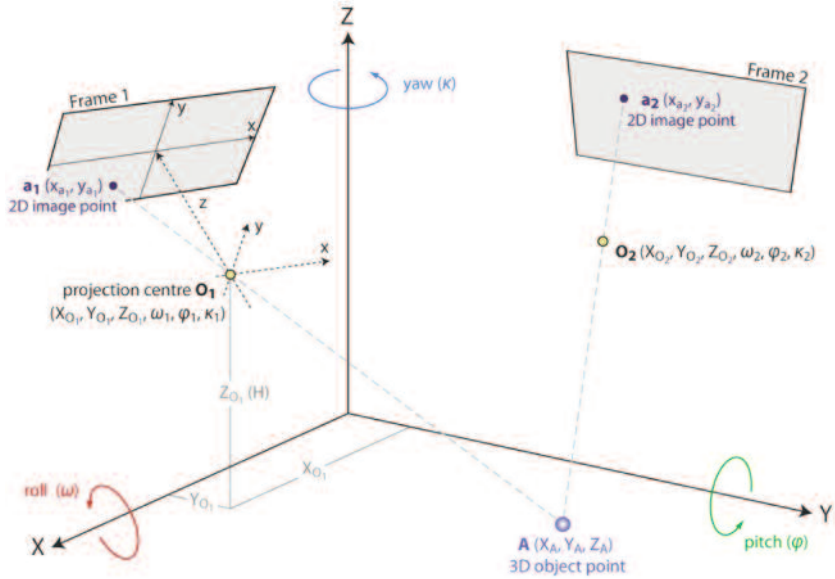

Figure 2. The exterior orientation of a photograph in relation to a global CRS (XYZ) (illustration by G. Verhoeven).

consists of accelerometers and gyroscopes which are rigidly mounted on a common base to maintain the same relative orientation. Accelerometers measure acceleration in m/s² or G-force (*g*), which can be static (e.g. gravity) and dynamic (i.e. caused by the deviation from a linear movement). Although these accelerometers cannot decide whether this amount of acceleration is static or not, if one knows that the accelerometer is not moving (or more correctly: not deviating from a linear movement), it can be used to measure the amount of static acceleration due to gravity. As a result, its orientation toward the earth's surface can be computed. Hence, accelerometers are often used for tilt-sensing [13]. This fact is exploited by all modern digital photo cameras to inform the user if the image was shot in portrait or landscape mode. Gyroscopes measure angular velocity (i.e. the speed by which something is spinning around its axis) in rotations per minute (rpm) or degrees per second (°/s). Since gyros are not affected by gravity, they complement accelerometers. The combination of accelerometers and gyroscopes (both inertial sensors) with supporting electronics and one or more navigational computers is denoted an Inertial Navigation System (INS). To know the exact direction with respect to magnetic North, the IMU's inertial sensors are often complemented by a magnetometer.

2.3. Archaeological benefits of geocoding with exterior orientation

Combining all exterior orientation parameters from the GNSS/IMU solution with the interior/inner orientation of the camera (also called camera intrinsics and comprising all parameters that describe the internal geometry of the camera as it was during the time of exposure [13]) unequivocally defines the position and orientation of the aerial image. Finally, the complete Field of View (FoV: the angle in object space over which objects are recorded in a camera)can be calculated from the combined play between both the physical size of the camera's sensor and the principal distance (i.e. the distance measured along the optical axis from the perspective centre of the lens to the image plane) of the lens attached [14].

With all this information, mono-plotting or mono-photogrammetry allows to compute the footprint of the image by projecting the image corners onto an existing Digital Surface or Terrain Model (DSM/DTM). This is accomplished by constructing a ray that originates at the projection centre and which goes through the image corner (or any image point for that matter). Where this ray intersects the DSM/DTM, the object point corresponding to the image point can be found. The exterior orientation and the DSM/DTM have to be specified in the same CRS and in this system also the object points are provided. Although the term "image footprint" is generally used in air- or spaceborne remote sensing, it is defined here as "the object area covered by the image"

so that it can be applied to terrestrial imaging too. When the image footprint can be complemented with the object point coordinates of the principal image point, all essential variables needed for fast and accurate archiving of (aerial) archaeological images are in place. In the special case of a vertical aerial photograph, where the image is considered as horizontally aligned with the photographed object, the projection centre of the image can be projected along the vertical direction.

Moreover, this information can be used in speeding up the orthophoto and DSM creation in a computer vision-based workflow. Since its introduction in archaeological research about fifteen years ago (e.g. [15], [16]), the computer vision techniques known as Structure from Motion (SfM) and dense Multi-View Stereo (MVS) have become very popular in archaeology. Nowadays, an SfM and MVS pipeline can almost be considered a standard tool in many aspects of both aerial and terrestrial archaeological research (e.g. [17–31]). Although the camera pose and the parameters of inner orientation are computed during the SfM stage ([32], [33]), the internal camera parameters can be accurately determined beforehand by a geometric camera calibration procedure [34]. This leaves only the exterior parameters to be computed during the SfM step. However, the SfM algorithm can rely on initial values that are provided by the IMU and GNSS log files, which means that processing time and gross relative orientation errors of the photograph can be minimized.

Finally, information about the exact coverage of each photograph could guide image acquisition. In cases where one wants to document an archaeological object or scene and extract an orthophoto and three-dimensional (3D) model from the imagery, it is of importance to have a decent camera network (e.g. [35]). This means that images should have sufficient overlap, while the number of images should be kept to a minimum. Both conditions can be verified already during image acquisition by projecting the FoVs onto an existing dataset such as a 3D model or an orthophotograph.

3. Method: Hard- and Software Developments

3.1. Hardware

3.1.1. Digital still camera

So far, only (semi-) professional Nikon digital SLR cameras have been used. Although this choice was determined by the availability of the Nikon cameras within the research team, they also offer several other advantages. Nikon was the first to support GNSS connections with their digital SLR cameras. As a result, many commercial GNSS solutions for hardware-based geotagging can be incorporated in comparison tests (see section 4). Second, to the author's knowledge, only Nikon's semi-

pro and pro level digital SLRs store the sub-second timing as metadata tags. Most cameras use a temporal resolution of one second since the date/time fields in the original Exif 2.3 specification are defined this way [36]. Although there are Exif fields that provide sub-second information (i.e. SubSecTime, SubSecTimeOriginal, SubSecTimeDigitized), they are often 00 or always have identical values. Also, the GPSTimeStamp Exif field only has one second resolution [36]. Although appropriate in most cases, it can be crippling for scientific photography that requires exact synchronisation with GNSS and IMU data.

3.1.2. GNSS/IMU

A cost-effective GNSS/IMU solution is provided by the ArduPilot Mega 2.0 (APM 2.0) [37], an open source autopilot system featuring an integrated MediaTek MT3329 GNSS chipset [38], a three-axis magnetometer and the InvenSense's MPU-6000: a six-axis gyro and accelerometer device [39]. In a first stage, the synchronisation between the APM 2.0 and the camera had to be established using a hardware-based solution. To this end, the APM 2.0 was directly connected to the Nikon by the ten-pin remote terminal. Using a Nikon N10 cable, the APM 2.0 board could be powered with the camera battery (hence avoiding the need to rely on additional batteries, which is, however, still possible). Moreover, the camera cable transfers a signal which indicates whether the camera button is pressed or not.

Besides the Nikon ten-pin cable, a standard flash sync cord with a coaxial PC (Prontor/Compur) 3.5 mm connector is also implemented for synchronisation. Similar to the ten-pin connector, this PC sync card features a locking thread for a reliable and sturdy connection. Every time a photograph is taken,the PC sync terminal sets the signal of the PC Sync cable to ground for the complete duration of the exposure and this signal is used by an interrupt handle of the microcontroller (APM 2.0). Since this sync terminal provides a highly accurate time stamp and the generated pulse is very clear, it allows to distinguish every individual photograph. The PC cord functions thus as the primary connection for data synchronisation, while the ten-pin cable is used to power the APM 2.0 and additionally serves as a synchronisation back-up.

However, all this would be useless if it remained impossible to log the GNSS/IMU data that are needed for the estimation of the external orientation of the acquired images. To this end, the standard APM 2.0 firmware was replaced and just a part of the software modules of the ArduPilot [37] are used to log all parameters of interest. These are the moment of photo acquisition as well as the GNSS (5 Hz) and IMU (200 Hz) values over the entire time span of the image acquisition, all with accurate time relations. Saving the entire data stream is enabled by a small serial data logger – called OpenLog – which holds up to

16 GB microSD cards [40]. As a result, there is ample of space to log all necessary data for hours. Moreover, the data access is straightforward (only a simple MicroSD card reader is needed). This whole sensor package is housed in a black plastic box and mounted on the hot shoe on top of the camera. To establish the accurate position and orientation of this box and its contained GNSS and IMU components, a mounting calibration can be performed (see 3.2.3).

3.2. Software

3.2.1. GNSS/IMU post-processing

Although the hardware solution was at this stage more or less fixed, some further software issues had to be solved before a working solution was achieved that acquired the correct positional and orientation values. The time dependent position is directly obtained from the GNSS receiver. To this end, the small displacement of around 10 cm between the perspective centre of the lens and the GNSS receiver is neglected since the observed precision of the MT3329 GNSS chipset is approximately 2.5 meter at 1σ when using a satellite-based augmentation system such as WAAS (Wide Area Augmentation System) or EGNOS (European Geostationary Navigation Overlay Service) [38].

The actual orientation parameters are calculated from the IMU data stream. The InvenSense's MPU-6000 is built with three accelerometers and three gyroscopes which are placed orthogonal on three axes. Both sensor types are based on MEMS (Micro Electro Mechanical Systems) technology [39]. To get the correct orientation values, a strap-down calculation is performed as described by Wendel [41]. First, just the gyroscopes' data are used. Due to the high bias-drift of MEMS-IMUs, the orientation values provided by the gyroscopes have to be updated with pitch and roll angle values estimated from the accelerometers and the yaw angle given by the magnetometer. These updates are just allowed under certain circumstances. Accelerometers, for example, can only be used to update pitch and roll angle in conditions without acceleration (e.g. static or with a constant movement). In such a condition, the earth gravity vector is the only remaining acceleration and therefore can be used to calculate roll and pitch angle of the IMU ([42], [43]).

3.2.2. Combining the data streams

Two different workflows were developed to link the actual image file with the computed orientation and position. The first method uses Phil Harvey's ExifTool [44] to write the complete exterior orientation information directly into the image's metadata. Because the Exif 2.3 specification supports GPSImageDirection (i.e. yaw), the values for pitch and roll are also written in custom-made metadata tags of the GNSS

attribute section, although they are provided by the IMU (APM 2.0) or a magnetic-based compass and have nothing to do with the GNSS signal. The second method creates an additional XMP sidecar file with the same name as the image file and the .xmp extension. Both methods have pros and cons (e.g. the first method does not create additional files but only a small number of software packages can read all embedded and non-standard metadata tags). Since both approaches are implemented in the presented MATLAB-based post-processing software, different image processing workflows can be accommodated.

3.2.3. Mounting calibration

Due to the fact that the APM 2.0 is mounted on the camera's hot shoe, the exact position and orientation of its sensors is not the same as for the camera. Additionally, the attitude relationship between the APM 2.0 and the camera will most likely slightly change every time the sensors are mounted on top of the camera. A camera mounting calibration (also called boresight calibration) mathematically describes the translation and rotation between the camera's coordinate reference system (CRS) and the APM 2.0 CRS (often also called misalignment), hereby enabling a reliable coordinate transformation between both systems. In other words: a mounting calibration is essential if one wants to transfer the APM 2.0 observed exterior orientation values to the aerial image. Since the GNSS positional precision is many times lower than the displacement between the APM 2.0 and the camera, the translation component is negligible. Being the only remaining parameter, the rotation between APM 2.0 and camera can be computed when both their exterior orientations are known (not all six parameters have to be known, but only the three rotation angles). The camera's rotation angles can be extracted by means of control points measured in the image, while the APM 2.0's exterior orientation (again, limited to only the three rotation angles) is given by its IMU and magnetometer measurements. Using both sets of orientation values, the mounting values are computed and applied to the APM 2.0 observed exterior orientation so that they are expressed in the camera's CRS. See [45] for more details on this procedure.

4. Results

Since it was the aim to validate the solution presented here in terms of archaeological usability for both terrestrial (i.e. static to low dynamic) and airborne (i.e. highly dynamic) applications, the next section will detail both tests. Previously, the terrestrial capabilities of the APM 2.0 solution have already been assessed in a low dynamic application and compared to an equally expensive commercial solution [45]. This comparison with Solmeta's Geotagger Pro 2 indicated the high

precision of the proposed solution in static image acquisition conditions. More specifically, standard deviations between the reference orientation angles and those logged with the APM 2.0 were reported to equal 1.3° for pitch and 0.8° for roll angle, with a 0° mean difference between both measurements. These values were significantly lower than those obtained by the Solmeta (standard deviations of 7.4° for pitch and 12.5° for the roll angle), which is an expected result given the fact that the latter solution does not provide access to its raw sensor data, while the update rate is also limited to 1 Hz. In both solutions, the provided yaw angles were obviously much less precise (standard deviations up to 12° were observed), although those of the APM 2.0 solution could be post-processed to achieve better results (with a standard deviation around 5°).

Since the previous tests were performed using a building façade, an archaeologically more relevant and less two-dimensional test scene was chosen. In the next sections, the archaeological test area will be described after which the new terrestrial and airborne tests are detailed.

4.1. Archaeological case study

The Roman town of *Carnuntum*, capital of the former Roman province *Pannonia superior* and located approximately 40 km south-east of Vienna (Austria) on the southern bank of the Danube river (N 48° 6' 41", E 16° 51' 57" – WGS84), was chosen as case study area. Since the Roman legionary camp with attached *canabae* and a civil town covers about 650 hectares, it constitutes a very large "archaeological site" ([46], [47]). Because a large part of its surface is covered by agricultural land, *Carnuntum* is ideal to cover with aerial photography (e.g. [46]). Besides, *Carnuntum* is also characterized by a few upstanding monuments such as two amphitheaters and the Heidentor. This latter monument (Figure 3), whose archaeological dating and meaning are still debated, measures approximately 15 m by 15 m and has a height of circa 14 m ([47], [48]). Given its dimensions, archaeological value and the fact that it is free-standing, the Heidentor is often chosen as an archaeological bench-marking target (e.g. [49]).

4.2. Terrestrial test

To test the APM 2.0 hard- and software solution in a low dynamic archaeological setting, a network of 24 well-identifiable targets was established around the Heidentor (Figure 3A). The position of these targets was surveyed and expressed in a European CRS (ETRS89 / UTM zone 33N; EPSG 25833). Afterwards, 69 photographs were taken all

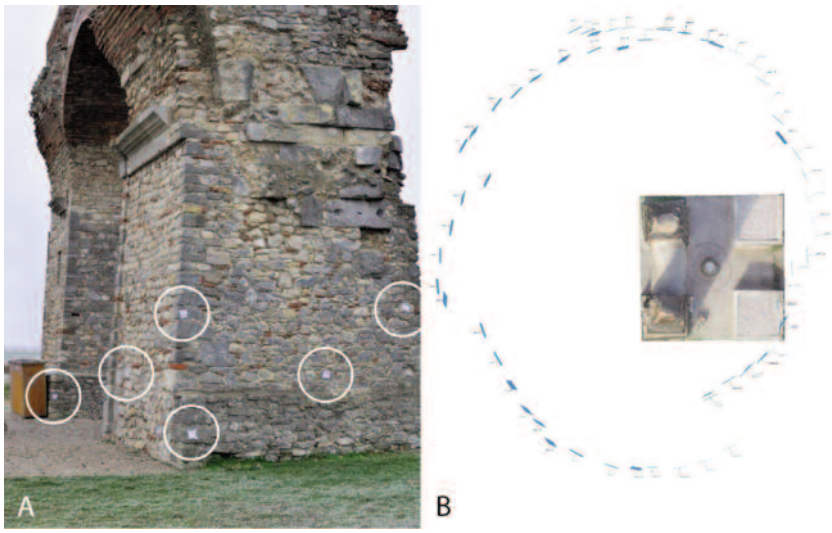

Figure 3. A. The Heidentor and some of the attached targets; B. The positions and orientation of all 69 terrestrial images around the Heidentor, visualized as an orthographic nadir view (photograph and illustration by G. Verhoeven).

around this test field using a Nikon D300 equipped with a prime 35mm lens (AF-S DX Nikkor 35mm f/1.8G) (Figure 3B). During image acquisition, care was taken to acquire the images with a wide variety of orientation angles. The APM 2.0 was mounted on top of the digital camera as described in [45]. Since this configuration is identical to the initial terrestrial tests, it is possible to apply the same workflow and directly compare the obtained accuracy and precision.

In a first step, the SfM approach embedded into Agisoft's PhotoScan Professional edition (version 1.0.1 build 1812) was used. By treating the previously mentioned targets as ground control points (GCPs), PhotoScan could compute the reference exterior orientation of all images in an indirect way (Figure 3B). Afterwards, these results were compared to the orientation angles that were acquired by the APM 2.0 after post-processing. The position of the camera was provided by the GNSS receiver and not subjected to further processing.

Using the MATLAB-based workflow, the complete exterior orientation estimated by the APM 2.0 (i.e. direct georeferencing information) could be written into every image and compared to the reference values obtained in the GCP-based workflow with PhotoScan (i.e. indirect georeferencing information). Figure 4 shows the differences between both observations for every photograph. From this figure, it can already be seen that the amount of outliers is very small and the accuracy of the pitch and roll angles is better than for the yaw angle. The slightly larger differences of the yaw angle are due to the magnetometer itself as well as magnetic disturbances by external influences. In summary, the standard deviation σ equals 0.7° for the pitch differences, while the differences in roll and yaw amount

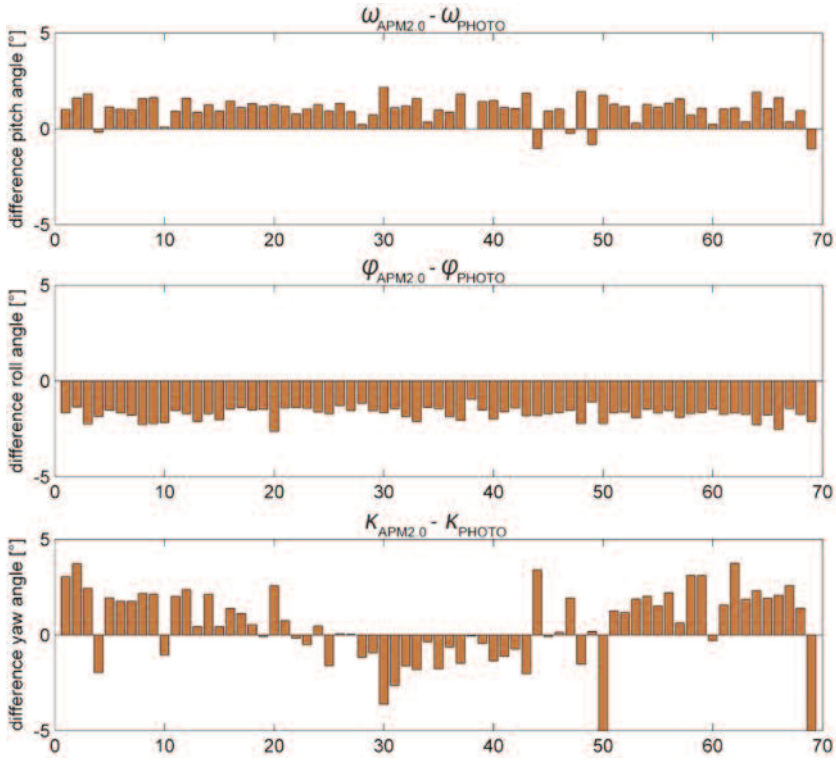

Figure 4. Differences in pitch, roll and yaw angles from direct (APM 2.0) and indirect observations (GCPs + PhotoScan) (illustration by M. Wieser).

to 0.4° and 2° respectively. Compared to the initially reported standard deviation values that resulted from the building façade test [45], the higher precision of the yaw angle is noticeable. This is entirely due to an optimized magnetometer calibration. With the improved yaw angle estimation, all three orientation angles have a precision ≤ 2° at 1σ.

As can be clearly seen in Figure 5, the average angle differences deviate from zero. In quantitative terms, the median of the differences equal -1.1° for pitch, 1.7° for roll and 0.6° for the yaw angle. This arises from the mounting error of the APM and the camera as described in 3.2.3. Since a mounting calibration is easy to calculate using the established GCP network, the values were recomputed and showed no longer any bias (i.e. the differences were spread around zero). This lack of systematic errors indicates the high accuracy of the estimated rotation angles.

4.3. Airborne test

On the 13th of June 2013, an archaeological reconnaissance flight was executed above the Roman town of *Carnuntum*. During the flight, the APM 2.0-based solution logged all exterior parameters while images were acquired with a Nikon D300 equipped with a 50 mm lens

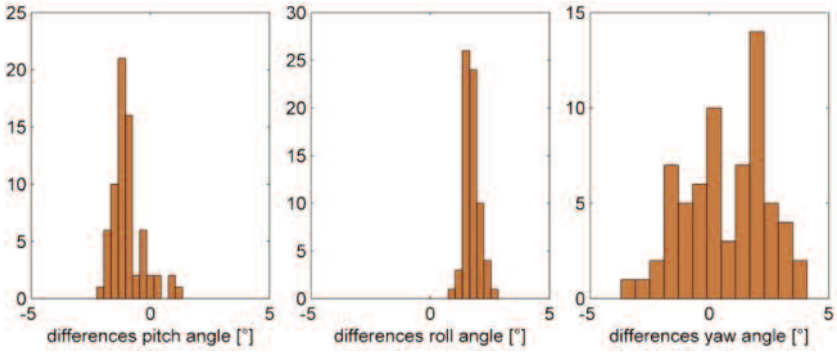

Figure 5. Histograms of the differences seen in Figure 4 (illustration by M. Wieser).

(AF Nikkor 50 mm f/1.8D). Although no GCPs were specifically materialized for this aerial survey, several of them could be identified in existing orthophotographs.

For this test, a set of twelve photographs acquired around the civil amphitheater were selected. Using GCPs and SfM (in PhotoScan), the reference exterior orientation of these images could again be determined. However, since the differences with the logged APM 2.0 GNSS data seemed so unusually large (Figure 6), a few images were selected for manual processing in ORIENT/ORPHEUS (a universal photogrammetric adjustment system developed at the Vienna University of Technology; [50]). As the exterior orientation estimates from ORPHEUS were almost identical to the values computed by PhotoScan, PhotoScan's results were proven to be reliable. Furthermore, another GNSS track (recorded with a Garmin eTrax) (Figure 8) of the flight was provided which also validates the results of the SfM approach.

As can be seen in Figure 6 and Figure 7, both the positional and rotational components delivered by the APM 2.0 solution are very different from the reference values. Figure 6 shows the APM 2.0 recorded flight track as white dots. This track features many data gaps, while its shape is characterized by very strange and often sharp bends. Both "artefacts" are most likely the result of the airplane's shielding effect, which is apparently a very difficult hurdle to overcome by the built-in GNSS receiver and antenna. Along the track, fourteen exposure stations are indicated with a red number. The corresponding reference photo location is pointed out with a green photo number. One can see that the positional differences are not constant. Since the GNSS logging frequency is 5 Hz, the horizontal travelling distance between two logging events can maximally be 11.1 m (given an airplane speed of 200 km/h or 55.6 m/s), which does also not explain the positional differences of hundreds of meters.

However, even if the GNSS receiver would accurately estimate the camera's position, the proposed solution would still not be suitable for

Figure 6. The reference footprints and camera locations (green) of respectively four and twelve aerial images compared with the APM 2.0-derived footprints and camera locations (red). Details can be found in the text (illustration by M. Wieser and G. Verhoeven).

aerial documentation due to the bad orientation angles. This can be seen in Figure 7, which depicts a bar chart of the differences in pitch, roll and yaw angles from the direct (APM 2.0) and indirect observations (GCPs + PhotoScan). In summary, the standard deviation equals 4.1° for the pitch differences, while the differences in roll and yaw amount to 7.0° and 23.5° respectively. Moreover, the median of the differences equals -8.8° for pitch, -1.6° for roll and -1.8° for the yaw angle. These figures were not optimized by a mounting calibration, since this procedure does not make any sense in this case because the adjustments would still be very small compared to the errors. When using the complete exterior orientation in a mono-plotting, the reference and estimated footprints can also be compared. The four footprints that are visualized in Figure 6 (the numbers and colours correspond to the photographs from which they were derived) clearly indicate that the obtained results cannot be considered sufficient for documentation and archiving purposes (e.g. compare the footprints of photo 155 which had overall the most accurate exterior parameters).

To make sure that these inferior results were not only occurring when making turns with the aircraft, four different photographs were selected for manual processing in ORPHEUS. The latter were acquired while the aircraft was not maneuvering but only flying in a straight line.

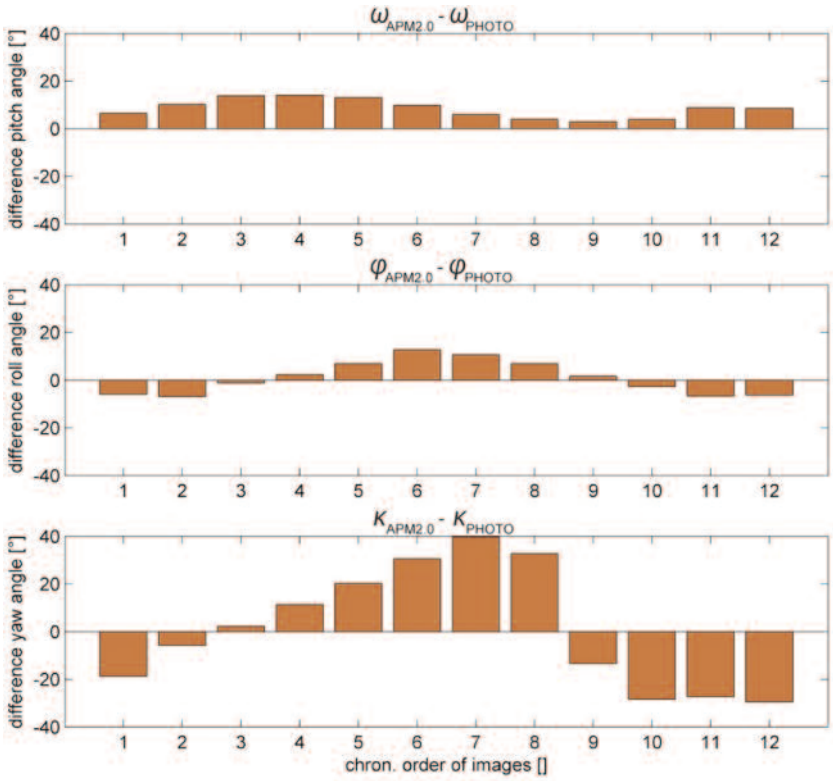

Figure 7. Differences in pitch, roll and yaw angles from direct (APM 2.0) and indirect observations (GCPs + PhotoScan) (illustration by M. Wieser).

The differences between the reference and estimated positional values and orientation angles was in the exact same range as the other twelve photographs.

As mentioned in the previous paragraphs the recorded flight track with the APM 2.0 is very poor (displacement of hundreds of meters) due to the unreliable APM 2.0's GNSS receiver. Therefore, the mono-plotting result of the image footprints is strongly affected by positional errors (see Figure 6). With the help of the second GNSS receiver (Garmin eTrax) in the aircraft (used similar to the one in Figure 1) significantly higher accurate positions could be determined. Figure 8 presents the mono-plotting results with this improved flight trajectory (the orientation angles were used in the same manner as before). It can be clearly seen in Figure 8 that the positional information provided by the second GNSS receiver (white track) has an improved accuracy of a few meters (compared to the reference photo location) and therefore, the image footprints are now located significantly closer to the reference footprints. However, due to the bad orientation angles, big displacements remain. For the point coordinates of the projected (see 2.3) principal image point in the reference coordinate frame, displacements from 140 m to 300 m of the four selected images can

Figure 8. The projected coordinates of the principal image point (middle point) and the footprints of four reference aerial images (green) compared to the results of the APM 2.0. Instead of the GNSS data recorded by the APM 2.0 the additional GNSS receiver was used (red). The flight track (white) of the additional GNSS receiver has an accuracy of a few meters in respect to the reference image location. (illustration by M. Wieser and G. Verhoeven).

be found. However, compared to Figure 6,the influence on the displacements could be reduced extensively.

Since no commercial solution was simultaneously taken aloft, it remains impossible to compare these results with those from an off-the-shelve low-cost product. However, since most low-end geotaggers feature a similar GNSS receiver and an IMU output that is even more limiting (only yaw is provided given that the pitch angle does not surpass certain limits, there is not raw data access and only 1 Hz logging), it is expected that the proposed solution will still be more accurate. Notwithstanding, these test results clearly indicate that the APM 2.0-based solution is too unreliable for systematic and archaeologically relevant documentation of the image footprint in the dynamic, airborne case.

5. Discussion and Outlook

In terms of positioning and orientation hardware, several new technologies and devices have been developed in the past decades. In the last years, both the cost and dimensions of many of these solutions have decreased. GNSS sensors are nowadays found in many electronic devices and their integration with digital cameras have become a common approach. Furthermore, the developments in the design of IMUs currently allow a stable, quite accurate and high frequent (several hundred hertz) estimation of the 3D-orientation of the sensing platform.

However, the IMU that has been applied here is certainly not the most accurate and stable one (in terms of drift rate). While the results in the terrestrial application were quite promising, the device turned out not to be apt when dealing with the highly dynamic conditions encountered in archaeological aerial photography. Although it was previously thought [45] that further developments in the post-processing workflow could partly remedy this, initial tests with Kalman filtering (e.g. [41]) revealed that the quality of the recorded low-cost IMU and GNSS data is simply insufficient to estimate a better exterior orientation for imagery acquired from an agile platform. Although the obtained positional and rotational information is still thought to be more accurate and precise than any off-the-shelve low-cost solution currently available, these results still clearly indicate that the proposed hard- and software solution is not suitable for documenting the footprint of aerial images yet. As a result, it is advised to apply the proposed solution for image management only in static or low dynamic conditions such as those described here.

With the help of an advanced GNSS receiver an improved result could be determined (see Figure 8). The improved coordinates of the projected principal image point deviate between 140 m and 300 m from the locations obtained from photogrammetry. This is still not accurate enough to be used as exterior orientation. However, it could be good enough for simple archiving purposes, where image locations are often documented as dots on small-scaled maps (typically 1: 50000). In that way, it could be used for a general indication of the image footprint location. Further field-tests will therefore integrate a better GNSS system in the APM 2.0 hardware solution. Although it was not tackled in this article, the authors also already obtained a much better hardware solution: the XsensMTi-G-700 GPS/INS. While the cost of this component is several times higher than the APM 2.0 solution, it should provide very reliable results in both static and dynamic conditions (which has already been confirmed by some initial tests). As a result, future airborne tests will only incorporate this more accurate INS. A thorough assessment of this Xsens MTi-G-700 GPS/INS in a real aerial survey environment, its integration with a digital camera and its potential for a direct georeferencing workflow (i.e. image restitution for which the image's exterior orientation values are directly determined using GNSS/INS and not indirectly using GCPs) will be reported in a future paper.

One more interesting issue still needs to be mentioned. Although the yaw angles that were estimated in terrestrial application of the APM 2.0 were quite accurate and precise, they sometimes seemed to be wrong when plotted at the position estimated by the GNSS receiver. This was due to the low precision of the latter device: approximately 2.5 meter at 1σ when using EGNOS or WAAS [38]. This

issue is not trivial, as it will result in a wrong image footprint (which can, even in the case of terrestrial applications with their characteristic short object distance, be completely off). For very accurate image footprints, the only solution is therefore to apply a differential or RTK (Real Time Kinematic) GNSS solution. The same, however, also holds for the more accurate Xsens INS solution. Although its IMU is much more precise than the one of the APM 2.0, its GNSS receiver is still non-differential. When the directly observed exterior orientation will be used to compute an accurate footprint of the terrestrial imagery (e.g. for documentation purposes or to establish a proper camera network for SfM-based documentation), a differential GNSS solution will thus be essential. Luckily, more and more compact and affordable solutions are also introduced into the market (e.g. Piks; [51])

Finally, the development of a small tool that calculates the footprint of the photograph out of the acquired exterior orientation values and the given inner orientation is also in progress. So far, a few examples have been shown in Figure 6 and Figure 8, but the tool is still in its infancy and further improvements are essential. The final aim is to store this computed footprint with the object point coordinates of the principal image point afterwards in a GIS system for an improved spatial management of aerial archaeological photographs.

6. Acknowledgements

This research is being carried out with the financial support of the Austrian Science Fund (FWF): P24116-N23.

The Ludwig Boltzmann Institute for Archaeological Prospection and Virtual Archaeology (archpro.lbg.ac.at) is based on an international cooperation of the Ludwig Boltzmann Gesellschaft (A), the University of Vienna (A), the Vienna University of Technology (A), the Austrian Central Institute for Meteorology and Geodynamic (A), the office of the provincial government of Lower Austria (A), Airborne Technologies GmbH (A), RGZM-Romano-Germanic Central Museum in Mainz (D), RAÄ-Swedish National Heritage Board (S), IBM VISTA-University of Birmingham (GB) and NIKU-Norwegian Institute for Cultural Heritage Research (N).

7. References

[1] Hampl, F., Archäologische Feldphotographie. *Archaeologia Austriaca*, 1957, 22, 54–64.

[2] Whittlesey, J.H., Photogrammetry for the Excavator. *Archaeology*, 1966, 19 (4), 273–276.

[3] Crawford, O.G.S., *Air Survey and Archaeology*, Ordnance Survey, Southampton, 1924.

[4] Brophy, K. and Cowley, D.C., eds., *From the air: Understanding aerial archaeology*, Stroud, Tempus, 2005.

[5] Lillesand, T.M., Kiefer, R.W., and Chipman, J.W., *Remote sensing and image interpretation*, 5th ed., Wiley, New York, 2004.

[6] Scollar, I., *et al.*, *Archaeological Prospecting and Remote Sensing*, Cambridge University Press, Cambridge, 1990.

[7] Wilson, D.R., *Air photo interpretation for archaeologists*, 2nd ed., Tempus, Stroud, 2000.

[8] Lasaponara, R. and Masini, N., eds., *Satellite remote sensing: A new tool for archaeology*, Dordrecht, Springer, 2012.

[9] Wilson, D.R., ed., *Aerial reconnaissance for archaeology*, London, The Council for British Archaeology, 1975.

[10] Adobe Systems Incorporated, 2013, *Extensible Metadata Platform (XMP)* [online], Adobe Systems Incorporated. Available from: http://www.adobe.com/products/xmp/ [Accessed 14 Apr 2013].

[11] Wolf, P.R. and Dewitt, B.A., *Elements of photogrammetry with applications in GIS*, 3rd ed., McGraw-Hill, Boston, 2000.

[12] Kraus, K., *Photogrammetry: Geometry from images and laser scans*, 2nd ed., Walter de Gruyter, Berlin - New York, 2007.

[13] Mikhail, E.M., Bethel, J.S., and McGlone, J.C., *Introduction to modern photogrammetry*, Wiley, New York, 2001.

[14] Verhoeven, G., Digitally Cropping the Aerial View. On the Interplay between Focal Length and Sensor Size. *Archeologia Aerea. Studi di Aerotopografia Archeologica*, 2008, 3, 195–210.

[15] Pollefeys, M., Koch, R., Vergauwen, M., and van Gool, L., Virtualizing Archaeological Sites, *in:Proceedings of the 4th International Conference on Virtual Systems and Multimedia (VSMM) 98*, 1998.

[16] Pollefeys, M., Koch, R., Vergauwen, M., and van Gool, L., Automated reconstruction of 3D scenes from sequences of images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2000, 55 (4), 251–267.

[17] Doneus, M., Verhoeven, G., Fera, M., Briese, C., Kucera, M., and Neubauer, W., From deposit to point cloud – A study of low-cost computer vision approaches for the straightforward documentation of archaeological excavations. *Geoinformatics*, 2011 [online], 6 (XXIIIrd International CIPA Symposium), 81–88. Available from: http://geoinformatics.fsv.cvut.cz/pdf/geoinformatics-fce-ctu-2011-06.pdf [Accessed 30 Sep 2011].

[18] Lerma, J.L., Navarro, S., Cabrelles, M., Seguí, A.e., Haddad, N., and Akasheh, T., 2011, Integration of Laser Scanning and Imagery for Photorealistic 3D Architectural Documentation, *in*: Wang, C.-C., ed., *Laser Scanning, Theory and Applications*, 413–430.

[19] Remondino, F., Barazzetti, L., Nex, F., Scaioni, M., and Sarazzi, D., UAV photogrammetry for mapping and 3d modeling – current status and future perspectives –, *in*: Eisenbeiss, H., Kunz, M., and Ingensand, H., eds., *Proceedings of the International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g)*, Zürich, 2011.

[20] Verhoeven, G., Taking computer vision aloft - Archaeological three-dimensional reconstructions from aerial photographs with PhotoScan. *Archaeological Prospection*, 2011 [online], 18 (1), 67–73. Available from: http://onlinelibrary.wiley.com/doi/10.1002/arp.399/abstract [Accessed 24 Feb 2011].

[21] Appetecchia, A., *et al.*, *New methods for documentation and analysis in building archaeology - prestudy: A project funded by the Swedish National Heritage Board, R & D funds*, Lund, Riksantikvarieämbetet, 2012.

[22] Bezzi, L., 2012, *3D documentation of small archaeological finds* [online]. Available from: http://arc-team-open-research.blogspot.com.br/2012/08/3d-documentation-of-small.html [Accessed 11 Oct 2012].

[23] Forte, M., Dell'unto, N., Issavi, J., Onsurez, L., and Lercari, N., 3D Archaeology at Çatalhöyük. *International Journal of Heritage in the Digital Era*, 2012, 1 (3), 352–378.

[24] Kersten, T.P. and Lindstaedt, M., Potential of Automatic 3D Object Reconstruction from Multiple Images for Applications in Architecture, Cultural Heritage and Archaeology. *International Journal of Heritage in the Digital Era*, 2012, 1 (3), 399–420.

[25] Lo Brutto, M., Borruso, A., and D'Argenio, A., UAV Systems for Photogrammetric Data Acquisition of Archaeological Sites. *International Journal of Heritage in the Digital Era*, 2012 [online], 1 (supplement 1 (EUROMED 2012)), 7–14. Available from: http://multi-science.metapress.com/content/7t71837015467552/fulltext.pdf [Accessed 3 Jan 2013].

[26] Lo Brutto, M. and Meli, P., Computer Vision Tools for 3D Modelling in Archaeology. *International Journal of Heritage in the Digital Era*, 2012 [online], 1 (supplement 1 (EUROMED 2012)), 1–6. Available from: http://multi-science.metapress.com/content/n865nn8573586777/fulltext.pdf [Accessed 3 Jan 2012].

[27] Opitz, R.S. and Nowlin, J., Photogrammetric Modeling + GIS: Better methods for working with mesh data. *ArcUser*, 2012 [online], Spring 2012, 46–49. Available from: http://www.esri.com/news/arcuser/0312/files/archaeology-inventory.pdf [Accessed 4 Feb 2013].

[28] Reinhard, J., Things on strings and complex computer algorithms: Kite Aerial Photography and Structure from Motion Photogrammetry at the Tulul adh-Dhahab, Jordan. *AARGnews*, 2012, 45, 37–41.

[29] Scollar, I. and Giradeau-Montaut, D., Georeferenced Orthophotos and DTMs from Multiple Oblique Images. *AARGnews*, 2012, 44, 12–17.

[30] Verhoeven, G., Doneus, M., Briese, C., and Vermeulen, F., Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. *Journal of Archaeological Science*, 2012, 39 (7), 2060–2070.

[31] Verhoeven, G., Taelman, D., and Vermeulen, F., Computer vision-based orthophoto mapping of complex archaeological sites: the ancient quarry of Pitaranha (Portugal-Spain). *Archaeometry*, 2012, 54 (6), 1114–1129.

[32] Hartley, R. and Zisserman, A., *Multiple view geometry in computer vision*, 2nd ed., Cambridge University Press, Cambridge, 2003.

[33] Szeliski, R., *Computer vision: Algorithms and applications*, Springer, New York, 2011.

[34] Sewell, E.D., Livingston, R.G., Quick, J.R., Norton, C.L., Case, J.B., Sanders, R.G., Goldhammer, J.S., and Aschenbrenner, B., 1966, Aerial cameras, *in*: Thompson, M.M., Eller, R.C., Radlinski, W.A., and Speert, J.L., eds., *Manual of Photogrammetry: Volume I*. 3rd ed., Falls Church, American Society of Photogrammetry, 133–194.

[35] Alsadik, B., Gerke, M., and Vosselman, G., Automated camera network design for 3D modeling of cultural heritage objects. *Journal of Cultural Heritage*, 2013, 14 (6), 515–526.

[36] Camera & Imaging Products Association, 2010–2012, *Exchangeable image file format for digital still cameras: Exif Version* 2.3, Tokyo, CIPA-JEITA (CIPA DC-008-2012 / JEITA CP-3451C). Available from: http://www.cipa.jp/english/hyoujunka/kikaku/pdf/DC-008-2012_E.pdf [Accessed 6 Apr 2013].

[37] Creative Commons 3.0, 2012, *ardupilot-mega*[online]. Available from: https://code.google.com/p/ardupilot-mega/ [Accessed 6 Apr 2013].

[38] MediaTek Incorporated, 2010, *MEDIATEK-3329 Datasheet. Rev.A03: 66-channel GPS Engine Board Antenna Module with MTK Chipset* [online], Hsinchu City, MediaTek Incorporated. Available from: http://inmotion.pt/documentation/diydrones/MediaTek_MT3329/mediatek_3329.pdf [Accessed 14 Apr 2013].

[39] InvenSense, *MPU-6000 and MPU-6050 Product Specification Revision 3.3*, Sunnyvale, 2012.

[40] SparkFun Electronics, 2012, *OpenLog* [online], Colorado, SparkFun Electronics. Available from: https://www.sparkfun.com/products/9530 [Accessed 14 Apr 2013].

[41] Wendel, J., *Integrierte Navigationssysteme: Sensordatenfusion, GPS und Inertiale Navigation*, Oldenbourg, München, 2007.

[42] Pfeifer, N., Glira, P., and Briese, C., Direct georeferencing with on board navigation components of light weight UAV platforms, *in*: Shortis, M.R., Wagner, W., and Hyyppä, J., eds., *Proceedings of the XXII ISPRS Congress: Technical Commission VII*, ISPRS, 2012, 487–492.

[43] Glira, P., Briese, C., and Pfeifer, N., Direkte Georeferenzierung von Bildern eines unbemannten Luftfahrzeuges mit LowCost-Sensoren. *VGI - Österreichische Zeitschrift für Vermessung und Geoinformation*, 2013, 2/13.

[44] Harvey, P., 2013, *ExifTool - Read, Write and Edit Meta Information!* [online]. Available from: http://www.sno.phy.queensu.ca/~phil/exiftool/ [Accessed 6 Apr 2013].

[45] Verhoeven, G., Wieser, M., Briese, C., and Doneus, M., Positioning in time and space – Cost-effective exterior orientation for airborne archaeological photographs. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2013 [online], II-5/W1, 313–318. Available from: http://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/II-5-W1/313/2013/isprsannals-II-5-W1-313-2013.pdf [Accessed 2 Sep 2013].

[46] Doneus, M., Eder-Hinterleitner, A., and Neubauer, W., Roman Carnuntum – Prospecting the largest archaeological landscape in Austria, *in*: Doneus, M., Eder-Hinterleitner, A., and Neubauer, W., eds., *Archaeological prospection: Fourth International Conference on Archaeological Prospection, Vienna, 19-23 September 2001*, Austrian Academy of Sciences, Vienna, 2001, 47–59.

[47] Jobst, W., *Das Heidentor von Carnuntum: Ein spätantikes Triumphalmonument am Donaulimes*, Verlag der Österreichische Akademie der Wissenschaften, Wien, 2001.

[48] Doneus, M., Gugl, C., and Doneus, N., *Die Canabae von Carnuntum: Eine Modellstudie der Erforschung römischer Lagervorstädte: von der Luftbildprospektion zur siedlungsarchäologischen Synthese*, Österreichische Akademie der Wissenschaften, Wien, 2013.

[49] Studnicka, N., Briese, C., Verhoeven, G., Kucera, M., Zach, G., and Ressl, C., The Roman Heidentor as study object to compare mobile laser scanning data and multi-view image reconstruction, *in*: Neubauer, W., Trinks, I., and Salisbury, Roderick B., Einwögerer, Christina, eds., *Archaeological prospection: Proceedings of the 10th International Conference on Archaeological Prospection*, Austrian Academy of Sciences, Vienna, 2013, 25–28.

[50] Department of Geodesy and Geoinformation, 2014, *ORIENT - A universal photogrammetric adjustment system* [online], Department of Geodesy and Geoinformation. Available from: http://photo.geo.tuwien.ac.at/software/orient-orpheus/ [Accessed 3 Feb 2014].

[51] Swift Navigation, 2014, *Piksi* [online], Swift Navigation. Available from: http://swift-nav.com/piksi.html [Accessed 3 Feb 2014].