

AUTOMATIC GEO-REFERENCING OF ARCHAEOLOGICAL UAV PHOTOGRAPHS BY CORRELATION WITH PRE-EXISTING ORTHO-PHOTOS

W. Karel^{φ,ψ*}, M. Doneus^{φ,χ}, C. Briese^{χ,ψ}, G. Verhoeven^{χ,φ}, N. Pfeifer^ψ

^φ Vienna Institute for Archaeological Science, University of Vienna
Franz-Klein-Gasse 1, 1190 Vienna, Austria

^χ Ludwig Boltzmann Institute for Archaeological Prospection & Virtual Archaeology
Hohe Warte 38, 1190 Vienna, Austria

^ψ Department of Geodesy and Geoinformation, Vienna University of Technology
Gusshausstrasse 27-29, 1040 Vienna, Austria

KEY WORDS: geo-referencing, unmanned aerial vehicle, archaeology, automation, correlation, image matching

ABSTRACT:

We present a method for the automatic geo-referencing of archaeological photographs captured aboard unmanned aerial vehicles (UAVs), termed UPs. We do so by help of pre-existing ortho-photo maps (OPMs) and digital surface models (DSMs). Typically, these pre-existing data sets are based on data that were captured at a widely different point in time. This renders the detection (and hence the matching) of homologous feature points in the UPs and OPMs infeasible mainly due to temporal variations of vegetation and illumination. Facing this difficulty, we opt for the normalized cross correlation coefficient of perspectively transformed image patches as the measure of image similarity. Applying a threshold to this measure, we detect candidates for homologous image points, resulting in a distinctive, but computationally intensive method. In order to lower computation times, we reduce the dimensionality and extents of the search space by making use of a priori knowledge of the data sets. By assigning terrain heights interpolated in the DSM to the image points found in the OPM, we generate control points. We introduce respective observations into a bundle block, from which gross errors i.e. false matches are eliminated during its robust adjustment. A test of our approach on a UAV image data set demonstrates its potential and raises hope to successfully process large image archives.

1. INTRODUCTION

The aerial archaeology unit at the Institute of Prehistoric and Historical Archaeology of the University of Vienna stores more than 110,000 vertical and oblique aerial photographs (APs) of archaeological features in its image archive (Doneus et al., 2001). While its oldest records are analogue APs taken several decades ago with professional aerial large-format cameras, the archive has been augmented with photographs captured aboard UAVs in recent years.

Geo-referencing, which is the determination of the absolute i.e. earth-related orientation of these UPs, is a precondition for the mapping of the imaged archaeological features and for the production of high-resolution ortho-photos that can hence be overlaid with other spatial data. However, this work is time-consuming when done manually, and off-the-shelf automatic methods are unavailable. Unfortunately, most UPs have thus remained with only a roughly known geo-referencing that is based on the flight records. To find remedies for this dissatisfying situation is the main goal of this research.

1.1 OrientAL

The newly developed software OrientAL (Karel et al., 2013) has shown to be capable of automatically computing the relative orientation of these challenging archaeological UPs, which mostly show vegetation, are thus poorly textured, and picture few distinctive objects.

Relative orientation is carried out similar as in (Barazzetti et al., 2010) for sets of images that have been captured within a short

*Corresponding author. wilfried.karel@geo.tuwien.ac.at



Figure 1: Relatively oriented UPs with reconstructed sparse point cloud in model space, computed by matching the statistical descriptions of local neighbourhoods of scale-invariant image feature points.

period of time, and hence, the same scale-invariant feature points (Lowe, 2004, Bay et al., 2008) can be extracted from different views of low vegetation and be matched by statistical descriptions of their surroundings. In the subsequent computation of the relative orientations of image pairs (Nistér, 2003), outliers are eliminated by RANSAC (Fischler and Bolles, 1981), before resp. image observations are introduced into a bundle block for incremental reconstruction of the whole scene (see fig. 1).

For the relative orientation of imagery with inhomogeneous characteristics (e.g. from different cameras, of different scales), graph-

based image feature matching (Delponte et al., 2006, Leordeanu and Hebert, 2005) helps to cope with the increased outlier ratio in the results of feature descriptor matching.

However, existing graph-based methods are computationally too expensive for being applied to large data sets. Hence, OrientAL makes use of a semi-local algorithm (Liu and Marlet, 2012) that imposes geometric constraints in image space on pairs of potential matches, based on the scales and orientations of resp. scale-invariant feature points. Additionally, the statistical descriptions of texture along line segments that connect such pairs are tested on similarity in order to iteratively eliminate outliers.

1.2 Related work

UAVs may carry global navigation satellite system (GNSS) receivers and inertial measurement units (IMUs), which are typically combined with a barometer, as additional sensors that allow for a direct geo-referencing of UPs i.e. independent of the object and thus without usage of the image content.

We classify existing approaches for automatic geo-referencing into those making use of only such additional sensors (see subsection 1.2.1), and those using only the image content for that purpose (indirect geo-referencing). Combined approaches exist, which thus determine an integrated geo-referencing, and which we mention in the same subsection 1.2.2.

1.2.1 Direct geo-referencing Despite the typically strong accelerations on board that affect the usually light-weight, low-quality IMUs negatively, accuracies of better than a few metres can be achieved with direct geo-referencing (Eisenbeiss and Sauerbier, 2011, Verhoeven et al., 2013) under good conditions. This is usually sufficient for applications like archiving the imagery in a spatial data base and its subsequent retrieval, but may not be so for the purpose of the combination with other spatial data.

Ground control is indispensable when aiming at reliability and accuracy, which is only used by indirect geo-referencing. Nonetheless, the outcome of direct geo-referencing can largely aid that process.

1.2.2 Indirect and integrated geo-referencing (Conte and Doherty, 2008) integrate a visual navigation system into the otherwise direct geo-referencing of a UAV for the purpose of autonomous UAV navigation. On the one hand, the nadir-looking on-board video camera serves as a visual odometer that can resume navigation in case of GNSS outage and that corrects the drift of IMU readouts with its high update rate. However, the odometer gradually accumulates positional errors.

On the other hand, the authors run an edge detector both on selected still images i.e. UPs from the video camera, and on existing geo-referenced ortho-photos, arguing that edges are robust against illumination changes, and large edges are temporally stable. The resulting binary images are scaled to each other according to the barometer measurements and the known terrain height, and rotated about the optical axis to account for the heading, as determined by the overall navigation system. They determine the location of maximum edge overlap by shifting those images w.r.t. each other, from which the earth-related UAV position is derived.

The authors conclude that the performance of the visual navigation system generally decreases with decreasing flying height, because the structures in the field of view become smaller and hence temporally less stable. Also, the rate of correct matches decreases in rural areas due to fewer visible man-made structures.



Figure 2: Exemplary UP, rectified w.r.t. an adjusting plane of the sparse point cloud reconstructed along with relative orientations (see fig. 1).

Instead of the object's texture, (Eugster and Nebiker, 2009) use its shape for indirect geo-referencing, in the form of existing boundary representations of buildings. Based on the results of direct geo-referencing, they project the model boundaries into a virtual image plane. Besides that, they extract edges from the UP and match them with the projected image. The indirect geo-referencing is derived from that by spatial resection.

For an experiment at a custom test site, they report positional accuracies of better than 1m for their integrated approach, being superior to direct geo-referencing by a factor of over 3. However, they criticise that the outcome of their integrated approach heavily depends on the parameters for model matching, the quality of the building models, and that depending on the orientation of the building edges in the field of view, the spatial resection may be unreliable.

2. METHOD

As previously mentioned, the main purpose of this work is the development of viable solutions for the automatic geo-referencing of UPs found in large archaeological image archives like the one mentioned in the beginning of section 1. Fig. 2 shows an exemplary UP. Besides the otherwise often faint archaeological features of interest, such imagery mainly pictures rural land that is potentially flat and is mostly covered by meadows, cropland, and forests. It may have been captured at low flying altitudes, at bright sunlight, and at an arbitrary time, with or without additional sensors.

Considering the potential absence of additional sensors and the possibly flat terrain, we aim at indirect geo-referencing of UPs, and we rely on object texture instead of shape. The automatic computation of the geo-referencing shall be carried out with existing OPMs and DSMs serving as reference data, as these are country-wide available even for rural areas.

Looking at fig. 3, which shows the detail of an OPM that contains the foot-print of fig. 2, it becomes evident that the data that OPMs are based on are typically captured not only at a different time of day, but also in another season of a different year.

Hence, not only the direction, length, and strength of cast shadows are different from the ones found in the UPs, but also, veg-



Figure 3: Detail of a pre-existing OPM of the civil amphitheatre of Roman Carnuntum, Austria, containing the foot-print of fig. 2 in the west, with vegetation in a different phenological state.

etation is generally in a different phenological state and of a different size, and cropland is in another phase of cultivation, with different marks of ploughing.

Also, during the possibly long period between data captures, field boundaries may have changed, and building measures may have altered the few man-made objects in the imaged area.

Resulting from those conditions, we state the following requirements on our method. It shall:

- master large temporal variations of illumination and vegetation;
- work without observations of additional sensors, but benefit from them if available;
- not depend on buildings in the field of view, but take advantage of features found in rural areas;
- only depend on widely available external spatial data products;
- cope with a wide range of flying altitudes above ground, especially low ones with resp. small image foot-prints;
- not depend on large variations of terrain height;
- work with both vertical and oblique UPs;
- clearly indicate failure, while being successful often enough to be helpful.

Considering these requirements, we aim to extract homologous points from the OPMs and UPs, where terrain heights interpolated in the DSM are assigned to the points extracted from the OPMs, resulting in control points in object space. Based on those correspondences, we determine the spatial similarity transformation from model space.

Finally, we insert according image observations and observations of control point positions into the transformed bundle block constructed during relative orientation, which ends in a simultaneous, robust, and weighted adjustment of all observations, where erroneous matches are detected and eliminated, and results from direct geo-referencing can be incorporated beneficially, if available.

Unfortunately, the indirect geo-referencing of UPs by matching scale-invariant image feature points extracted from the two data sets, analogous to section 1.1, has shown to be infeasible. We attribute that to the large differences in shape, texture, and illumination.

Potentially small image foot-prints are the consequence of the requirement that our approach shall be able to also cope with low flying heights and normal lenses. This, however, results in the larger, temporally generally more stable feature points found in the OPM being too large to be fully covered by single UPs.

Attempts to match feature regions (Matas et al., 2004) detected in the two data sets have not been successful either. Figs. 4 and 5 detail the strong differences in illumination and indicate that image feature matching may be condemned to fail, because the strongest brightness changes result from cast shadows, and not from texture.

3. MATCHING BY CORRELATION

With feature matching being infeasible and image edges being dominated by cast shadows of different orientation, we opt for area-based matching.

Lacking any initial values, we execute a brute-force search for the maximum cross-correlation coefficients between planar patches of the two data sets as the most robust, but also computationally most expensive approach to determine homologous points.

In order to reduce the computational cost, we use image pyramids. Also, we reduce the dimensionality and extents of the search space by using as much a priori knowledge as possible, which originates from the relative orientation, experience, and e.g. the flight records:

- the approximate planar earth-related position of UP projection centres, known up to a few tens of metres;
- the approximate flying height, known up to about 20m;
- the relative orientation of UPs;
- the overall surface structure, known from the DSM and the sparse point cloud reconstructed along with the relative orientation of UPs;
- the typical extents of temporally stable objects.

Orientation parameters known from direct geo-referencing considerably reduce the extents of the initial search space further, if available.

In order to eliminate the roll and pitch angles, we compute an approximating plane through the sparse point cloud, w.r.t. which the UPs are rectified (see fig. 2), similar as in (Yang et al., 2010). Based on the approximate flying height and the UP's focal length, we additionally limit the range of relative image scales to be considered.

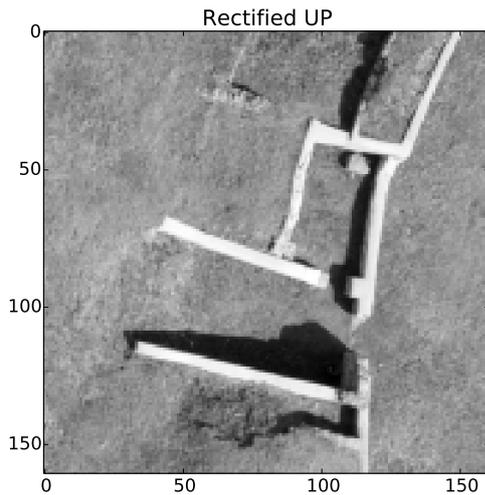


Figure 4: Detail of grey-scale version of fig. 2, rotated and scaled to match the imaged area of fig. 5, with cast shadows.

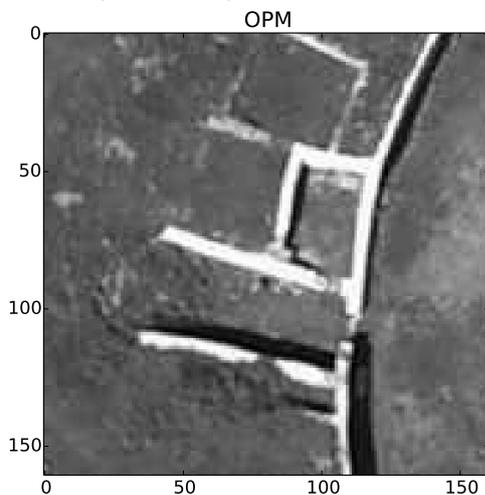


Figure 5: Detail of grey-scale version of fig. 3, with cast shadows, serving as template for image matching.

Thus, at the start of the quest for homologous patches, we vary

- the planar displacement,
- the relative image scale, and
- the azimuth.

Based on our observation that temporally stable objects are generally larger than a few metres, we choose the patch size as 20m. For each selection of relative image scale and azimuth, a map of correlation coefficients is generated, which is inspected for promisingly high values.

We finally apply adaptive least-squares matching (Gruen, 1985) on the resp. image regions, and accept patches as correct matches, if the resulting affine transformation shows plausible values and its application results in a high enough correlation coefficient.

For accepted matches, we introduce resp. image observations and observations of control points into the bundle block defined during the computation of the relative orientation.

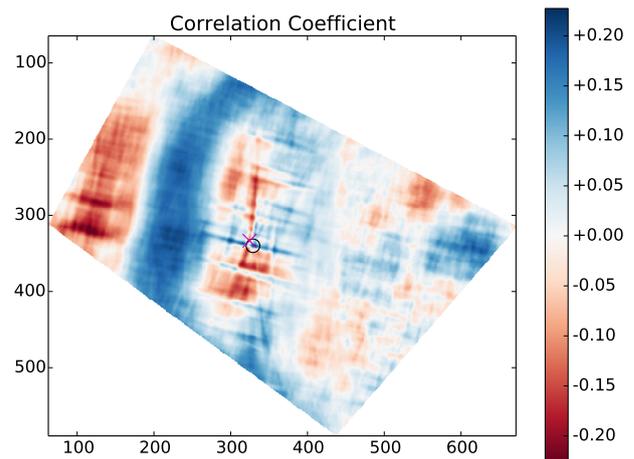


Figure 6: Correlation coefficients for all possible displacements with full overlap of the OPM detail (see fig. 5) w.r.t. the UP (see fig. 2), with the maximum value being marked by a black circle, close to the correct position (magenta cross).

Thus, the initial search space is fully four-dimensional in the beginning, and it needs to be sampled densely in order to account for small and elongated structures. However, with the first matching patches found, the search space can be narrowed dramatically. If computations are done in parallel, this results in a viable, robust approach.

4. APPLICATION

The archaeological feature we test our approach on is the 2nd century AD amphitheatre of Carnuntum. The Roman town of Carnuntum (today Petronell-Carnuntum in Austria, located 40 km south-east of Vienna on the southern bank of the Danube river), was home to some 50,000 inhabitants and consisted of both a Roman legionary camp with associated civilian settlement (canabae) and a civil town. The photographed amphitheatre is located outside the gates of the civil town.

The 83 UPs used in this test (see figs. 1, 2, 4) were acquired at the end of March 2011 around 9.30h using a radio-controlled microdrone md4-1000 quadcopter. In this case, a common Olympus PEN E-P2 (a 12.3 megapixel mirrorless Micro Four Thirds camera) equipped with an Olympus M.Zuiko Digital 17 mm f/2.8 lens was taken aloft by the UAV. The quadcopter was manually steered in four approximately parallel strips at a height of about 70m above the amphitheatre. During the flight, photographs were taken at specific spots in such a way that they have a sufficient amount of overlap with the neighbouring photographs.

The OPM used here (see figs. 3, 5) was captured at the beginning of August 2008 and features a ground sample distance of 12.5cm.

We implement our approach in OrientAL. As a first result, fig. 6 shows the map of normalized cross-correlation coefficients for all possible displacements with full overlap of the OPM detail (see fig. 5) w.r.t. the UP (see fig. 2). The maximum value of 0.23 is at the correct position within 1m. Application of adaptive least-squares matching raises this coefficient to 0.43, and reduces the positional error to 35cm. The transformation parameters result as plausible, with shift precisions of about one tenth of a pixel. Fig. 7 shows the same for another exemplary UP, yielding a positional error of 47cm after least-squares matching.

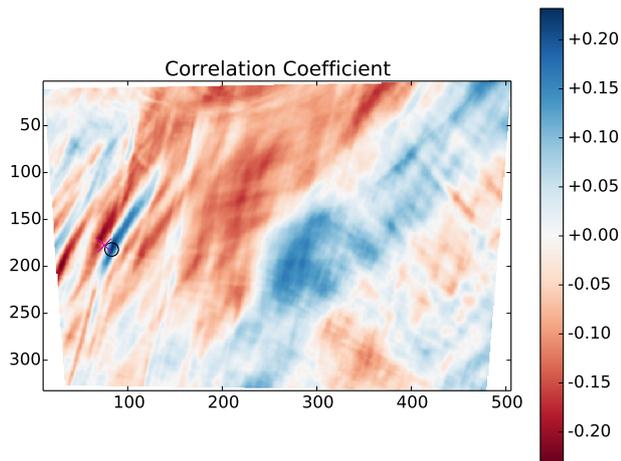


Figure 7: Correlation coefficients as in fig. 6, but for another exemplary UP, again with the maximum value marked by a black circle, being close to the correct position (magenta cross).

5. CONCLUSIONS & OUTLOOK

With image feature matching being infeasible for our application, we have presented a method that employs area-based matching of UPs and OPMs for the geo-referencing of UAV image data sets. Lacking any match candidates, we propose usage of a brute-force search using the correlation coefficient of image patches as similarity measure. In order to reduce the computational cost of this search, we opt for using as much a priori knowledge as possible, thereby limiting the search space.

For terrain that can sufficiently well be modelled by a plane, only two correct matches are already sufficient to determine the remaining unknowns of azimuth, origin, and scale. Our method has shown to be capable of detecting plenty of correct and accurate matches for structures on the ground, and hence we consider most of the requirements on our method stated in section 2 as fulfilled: it

- matches data from different years, captured in different seasons and at different times of day;
- works without observations of additional sensors, but may benefit from them;
- does not depend on buildings in the field of view, but takes advantage of features found in rural areas;
- only depends on OPMs and DSMs, two widely available external spatial data products;
- copes with low flying altitudes above ground with resp. small image foot-prints;
- does not depend on variations of terrain height;
- works with both vertical and oblique UPs;
- clearly indicates failure during the robust bundle block adjustment.

We largely attribute these promising results to the matching method, which operates directly in image space, instead of using abstract representations thereof. However, ways must be found to lower the ratio of spurious matches, which may reduce the quality of geo-referencing, or even prevent a successful result.

ACKNOWLEDGEMENTS

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

The Ludwig Boltzmann Institute for Archaeological Prospection and Virtual Archaeology (<http://www.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 Geodynamics (A), the office of the provincial government of Lower Austria (A), Airborne Technologies GmbH (A), RGZM-Roman-Germanic Central Museum Mainz (D), RA-Swedish National Heritage Board (S), IBM VISTA-University of Birmingham (GB) and NIKU-Norwegian Institute for Cultural Heritage Research (N).

REFERENCES

- Barazzetti, L., Remondino, F., Scaioni, M. and Brumana, R., 2010. Fully automatic UAV image-based sensor orientation. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII (Part 1/ICWG I/V), Calgary, Canada.
- Bay, H., Ess, A., Tuytelaars, T. and Gool, L. V., 2008. Speeded-up robust features (SURF). *Computer Vision and Image Understanding* 110(3), pp. 346 – 359. *Similarity Matching in Computer Vision and Multimedia*.
- Conte, G. and Doherty, P., 2008. An integrated UAV navigation system based on aerial image matching. In: *Aerospace Conference, 2008 IEEE*, pp. 1–10.
- Delponte, E., Isgrò, F., Odone, F. and Verri, A., 2006. SVD-matching using SIFT features. *Graphical Models* 68(56), pp. 415 – 431. *Special Issue on the Vision, Video and Graphics Conference 2005*.
- Doneus, M., Eder-Hinterleitner, A. and Neubauer, W., 2001. Archaeological prospection in Austria. In: M. Doneus, A. Eder-Hinterleitner and W. Neubauer (eds), *4th International Conference on Archaeological Prospection*, Austrian Academy of Sciences Press, Vienna, Austria, pp. 11–34.
- Eisenbeiss, H. and Sauerbier, M., 2011. Investigation of UAV systems and flight modes for photogrammetric applications. *The Photogrammetric Record* 26(136), pp. 400–421.
- Eugster, H. and Nebiker, S., 2009. Real-time georegistration of video streams from mini or micro UAS using digital 3d city models. In: *Proceedings of 6th International Symposium on Mobile Mapping Technology*, Presidente Prudente, Brazil.
- Fischler, M. A. and Bolles, R. C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24(6), pp. 381–395.
- Gruen, A., 1985. Adaptive least squares correlation: a powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography* 14(3), pp. 175–187.
- Karel, W., Doneus, M., Verhoeven, G., Briese, C., Ressler, C. and Pfeifer, N., 2013. OrientAL - automatic geo-referencing and ortho-rectification of archeological aerial photographs. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial*

- Information Sciences, II-5/W1. poster presentation: International Symposium of CIPA, Strasbourg, France; 2013-09-02 – 2013-09-06.
- Leordeanu, M. and Hebert, M., 2005. A spectral technique for correspondence problems using pairwise constraints. In: Computer Vision, 2005. ICCV 2005. 10th IEEE International Conference on, Vol. 2, pp. 1482–1489.
- Liu, Z. and Marlet, R., 2012. Virtual line descriptor and semi-local graph matching method for reliable feature correspondence. In: Proceedings of the British Machine Vision Conference, BMVA Press, pp. 16.1–16.11.
- Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. International journal of computer vision 60(2), pp. 91–110.
- Matas, J., Chum, O., Urban, M. and Pajdla, T., 2004. Robust wide-baseline stereo from maximally stable extremal regions. Image and Vision Computing 22(10), pp. 761 – 767. British Machine Vision Computing 2002.
- Nistér, D., 2003. An efficient solution to the five-point relative pose problem. In: Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on, Vol. 2, IEEE, pp. II–195.
- Verhoeven, G., Wieser, M., Briese, C. and Doneus, M., 2013. Positioning in time and space - cost-effective exterior orientation for airborne archeological photographs. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume II-5/W1. talk: International Symposium of CIPA, Strasbourg, France; 2013-09-02 – 2013-09-06.
- Yang, M. Y., Cao, Y., Förstner, W. and McDonald, J., 2010. Robust wide baseline scene alignment based on 3d viewpoint normalization. In: Advances in Visual Computing, Springer, pp. 654–665.