THE EUROSDR TIME BENCHMARK FOR HISTORICAL AERIAL IMAGES

E.M. Farella [1] *, L. Morelli [1,2], F. Remondino [1], J. P. Mills [3], N. Haala [4], J. Crompvoets [5]

[1] 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy
Web: http://3dom.fbk.eu - Email: <elifarella><lmorelli><remondino>@fbk.eu

[2] Dept. of Civil, Environmental and Mechanical Engineering (DICAM), University of Trento, Italy

[3] School of Engineering, Newcastle University, Newcastle upon Tyne, NE1 7RU, United Kingdom - jon.mills@newcastle.ac.uk

[4] Institute for Photogrammetry, University of Stuttgart, Germany - norbert.haala@ifp.uni-stuttgart.de

[5] Public Governance Institute, KU Leuven, Belgium - joep.crompvoets@kuleuven.be

Commission II / EuroSDR Theme Session

KEYWORDS: benchmark, historical aerial images, aerial triangulation, DSM, DTM, orthophoto, multi-temporal, EuroSDR

ABSTRACT

Automatic photogrammetric processing of historical (or archival) aerial photos is still a challenging task, particularly in cases of missing ancillary information, low radiometric and image quality, limited stereo coverage or large temporal span. However, with recent advances in photogrammetry and Artificial Intelligence (AI) algorithms for image processing and interpretation, an increasing number of applications are now feasible. The article presents the TIME (hisTorical aerIal iMagEs) benchmark (https://time.fbk.eu/), promoted by EuroSDR to explore the potential of historical aerial images. Realized in collaboration with various European NMCAs, the benchmark has garnered aerial image blocks and time series imagery captured since the 1950s. To support the photogrammetric processing of the digitized photos, ancillary data are supplied with available information about flight missions, taking cameras, and ground control points (GCPs). Several diverse investigations have been undertaken with the benchmark datasets, all captured over historical urban areas or landscapes. The paper describes the benchmark datasets and some potential research topics, presenting several tests and analyses realized with the collated and shared data.

Figure 1. Examples of historical aerial images collected for the TIME benchmark. Finland – 1951 (left); Estonia – 1962 (middle); Slovenia – 1979 (right).

1. INTRODUCTION

Historical images, defined here as photos dating from the origin of photography to the beginning of the digital era (c. 2005), are becoming an increasingly valuable information source, yet many automated processing issues remain. Whether aerial or terrestrial, nadir or oblique, many photo collections are under digitization worldwide, providing a significant amount of data for multiple potential applications (Schulz et al., 2021). Archives often offer long-term time series, relatively dense temporal resolution, relatively high geometric resolution and, quite frequently, stereo coverage. Historical aerial images often feature different image formats (e.g. 23x23 cm, 30x30 cm, microfilms, etc.) and were collected for various purposes (topographic map generation and updating, visual interpretation, reconnaissance, etc.).

EuroSDR, in collaboration with various European National Mapping and Cadastre Agencies (NMCAs), launched a research activity entitled hisTorical aerIal iMagEs (TIME - https://time.fbk.eu/) to investigate the potential of historical (often referred to as archival) aerial images and the actual bottlenecks in processing archival aerial datasets for mapping and other activities. Datasets such as those illustrated in Figure 1 are highly suited to map long-term land cover changes, often presenting stereo coverage, long time series, relatively dense temporal span (ca 3-5 years), and good geometric resolutions (ca 1-2m). However, despite the undoubted value of such sources of information for performing, e.g., multi-temporal urban analyses, the digital processing of historical imagery still has to address numerous open research issues. The photogrammetric processing of digitized historical photos is an attractive research task,

* Corresponding author

offering mapping benefits but also numerous investigation opportunities and challenges: missing or incomplete metadata, low radiometry and geometric resolutions, poor image quality arising from incorrect scanning activities or inappropriate conservation measures, large time intervals in multi-temporal datasets, reduced stereo coverage, etc. (Redecker, 2008; Nocerino et al., 2012, Cowley et al., 2013; Giordano and Mallet, 2019).

In the case of very old datasets, camera certificates are rarely available (also considering that rigorous calibration standards were only introduced in the 1960s), and this translates into missing Interior Orientation (IO) parameters and fiducial marks coordinates, notwithstanding the lack of flight plan information (for coarse image localization). Furthermore, poor image quality can limit the performance of automatic algorithms for salient feature extraction, complicating further 2D/3D image processes. Finally, the rare availability and visibility of multi-temporal GCPs in historical and modern urban scenarios limits automated georeferencing processes.

This paper presents the EuroSDR TIME benchmark, established to collect and harmonize a series of historical aerial images and share them with the research community to investigate open issues in various photogrammetric processing steps. The shared aerial photos (and associated ancillary data) aim to stimulate research activities to offer new procedures that can fully exploit the invaluable content of such historical data. TIME is believed to be the first photogrammetric benchmark dedicated to the processing of historical aerial photographs. Previous studies on archival aerial image processing have highlighted the complexity of handling these data, which are increasingly being digitized by European NMCAs (Giordano and Mallet, 2019).

2. RELATED WORK

As a tool for scientists, a benchmark is generally defined as a reference against which entities may be compared. A review of benchmarking initiatives for evaluating and comparing the performance of sensors, algorithms and methods for geospatial investigations was presented by Bakula et al. (2019). Many benchmark initiatives have been supported by ISPRS and/or EuroSDR, in line with their ambitions to support the geospatial community with research activities and datasets.

Although different in terms of test data and investigation aims, all benchmarks share the common focus on developing and testing automatic processing procedures (Haala, 2014), while simultaneously boosting robust data quality assessments (Nex et al. 2015). Several benchmarks on aerial imagery have been released, primarily focusing on image triangulation and DSM generation (Haala, 2014; Nex et al., 2015), object detection (Ding et al., 2021), semantic segmentation of urban 3D point clouds (Kölle et al., 2021; Hu et al., 2021) and image interpretation (Long et al., 2021). Benchmarks related to historical aerial imagery have not yet been made available, although Maiwald (2019) proposed terrestrial scenario datasets to evaluate the performance of several feature matching methods.

The literature shows that the most common investigations into historical aerial images focus on examining the standard pipeline for image triangulation, DSMs and orthoimagery production, including also the most recent SfM and MVS methods, especially for landscape change detection, 3D reconstruction of heritage structures and monitoring analyses (Wiedemann et al., 2000; Sauerbier et al., 2004; Walstra et al., 2004; Sonnemann et al., 2006; Verhoeven et al., 2013; Nebiker et al., 2014; Risbøl et al., 2014; Jao et al., 2014; Micheletti et al., 2015; Popelková and Mulková, 2016; Fieber et al., 2018; Lydersen et al., 2018; Sevara et al., 2018; Peppa et al., 2018; Kupidura et al., 2019; Osińska-

Skotak et al., 2019; Pinto et al., 2019; Grottoli et al., 2020; Poli et al., 2020).

Missing or incomplete camera and flight mission details are the primary reported problem for photogrammetric processing and the image triangulation of analogue photos. When camera certificates and fiducial marks coordinates are unavailable, an initial bottleneck relates to the derivation of camera interior orientation (IO) parameters, and the establishment of a mathematical relationship between the digital image and the camera's coordinate system. Procedures for approximately deriving interior orientation information and for fiducial transformation employing affine models were described in Nocerino et al. (2012), Nurminen et al. (2015), Salach (2017) and Poli et al., (2020).

Different studies have focused on the recovery of camera exterior orientation (EO) parameters and on automatic geo-referencing procedures of historical imagery, using photogrammetric methods (Kim et al., 2010; Redweik et al., 2010; Verhoeven et al., 2012; Karel et al., 2013; Nurminen et al., 2015; Giordano et al., 2018) or empirical transformations (Zambanini and Sablatnig, 2017). Multi-date aerial datasets can be co-registered using handcrafted operators (e.g. SIFT), contours or lines (Clery et al., 2014). These approaches normally fail to extract reliable correspondences when large time differences exist and/or a change of sensor occurs between epochs. For these reasons, recent works were based on learning-based approaches to extract reliable correspondences among multi-temporal images and improve image matching (Ressl et al., 2020; Zhang et al., 2021). Historical aerial images have frequently been used to interpret the territory and discover lost information (Sevara, 2013), applying monoplotting (Bozzini et al., 2012; Bayr, 2021) but also using simulation processes for supporting data interpretation (Siok and Ewiak, 2020). Further attractive research activities have focussed on exploiting historical aerial photographs to identify war-related bomb craters (Meixner and Eckstein, 2016; Valjavec et al., 2018; Clermont et al., 2019; Dolejš et al., 2020). Other studies have deepened the benefits of employing archival sources for producing land-cover and land-use historical maps (LULC) (Ratajczak et al., 2019; Mboga et al., 2020), also exploiting Artificial Intelligence (AI) algorithms.

AI-based techniques have also been adopted for tackling grayscale image colourization. Colourizing archival aerial data can be crucial for improving scene understanding and supporting other processing tasks, such as data classification. Several methodologies have been developed in this field for terrestrial scenarios, whereas few works have focussed on aerial-scale historical images (Seo et al., 2018; Dias et al., 2020; Poterek et al., 2020).

3. DATASETS DESCRIPTION

The TIME benchmark collected image blocks and ancillary information span from the 1940s until the 2000s. The contributing countries (Austria, Norway, Finland, Cyprus, Slovenia, Estonia, Poland, Italy) generally provided multi-temporal image blocks with somehow diverse ancillary data consisting of camera calibration, approximate exterior orientation (EO), ground control points (GCPs), ground truth, etc. In particular, the shared data (Table 1) includes:

- *Austria*: it comprises three multi-temporal image series from three different locations (Vienna, Steyr and Schladming), dating from 1950 to the early 2000s. Image blocks contain a variable number of images, ranging from 10 to 28, and more recent data are in RGB. Camera certificates, fiducial marks coordinates and approximated exterior orientation information are available. Some modern Ground Control Points (GCP) for

geo-referencing historical data are also available for the Vienna location.

- *Norway:* it consists of four series capturing the area of Austre-Moland and Arendal with a temporal duration of approximately ten years, starting from 1950. The 1963 dataset contains the largest number of images of the entire benchmark (185 images). The full aerial triangulation (AT) and some GCPs are provided.

- *Finland:* it is an aerial block of 25 panchromatic images captured over Olkkola in 1951. No camera certificate or approximate EO information are available but only some GCPs and fiducial marks coordinates are included.

- *Cyprus:* it includes five image-series of Kouflia Pafou, acquired between 1957 and 1993. Except for approximate camera centres and a few actual GCPs (partly) visible in the historical images, no further information is available.

- *Slovenia:* three time-series (from 1967 to 1998) depicting the southern area of Ljubljana are available. Nor camera certificates or fiducial marks coordinates are available. A further small dataset of digital images (2018) and LiDAR data (2015) are included for retrieving GCPs and geo-referencing historical images.

- *Estonia:* it features two image blocks over Voka in 1962 and Asuka in 1955. The Voka dataset includes 158 images, but unfortunately, no further information about the camera or GCPs is available. The same is true for the 43 images from 1955 of the Asuka block, where images are partially cropped, and fiducial marks are not always visible.

- *Poland:* two sets of the same aerial block over Warsaw in 1986 are shared. Analogue images were scanned using both a photogrammetric and a regular scanner. The available calibration certificate provides camera information and fiducial marks coordinates. A 2020 orthophoto can be used for retrieving GCPs and geo-referencing aerial data.

- Italy: they provided some 20 miscellaneous images acquired during the Second World War (1942 and 1945) depicting bomb damages over urban and rural areas. Radiometric and geometric resolution vary and no ground-truth data are available.

The benchmark datasets and ancillary information could stimulate several investigations, including:
- radiometric equalization and image enhancement;
- automatic aerial triangulation and geo-referencing;
- automatic DSM/DTM generation;
- bomb crater detection;
- multi-temporal analyses and change detection;
- land-use land-cover (LULC) map generation;
- automatic colourization of grayscale images.

In the next section, some investigations on the aforementioned R&D issues are presented.

COUNTRY	LOCATION	YEAR	CAMERA	IMAGES	IMAGE SIZE [cm]	SCAN RES. [mm]	CALIBRATION CERTIFICATE	FLIGHT HEIGHT [m]
Austria	Wien	1958	Wild RC5	15	18x18	14	yes	3700
		1976	Wild RC5/RC8	21	18x18	11,5	yes	2460
		2000	RC20	12	23x23	21	yes	3600
	Steyr	1953	Wild RC5	28	18x18	15	yes	4910
		1985	RC10	10	23x23	14.5	yes	4750
		2001	RC20	24	23x23	21	yes	3600
	Schladming	1953	Wild RC5	18	18x18	14	yes	4900
		1966	RC5/RC8	26	18x18	14	yes	4420
		2006	LC0030	17	23x23	20	yes	6750
Norway	Austre-Moland	1950	RMK21	110	23x23	15	no	3200
	Arendal-Moland-Tromøya	1963	RMK21	185	23x23	15	no	1240
	Arendal	1975	Wild-UAg473	43	23x23	12	no	830
	Øyestad - Arendal	1984	RC10_UAg	63	23x23	12	no	2240
Finland	Olkkola	1951	Mh-2	25	18x18	15	no	2000
Cyprus	Kouklia Pafou	1957	303	6	23x23	21	no	1800
		1963	RC50	14	23x23	14	no	1800
		1970	F49 MK4 & F49 MAG	12	23x23	21	no	2500
		1973	RC8-903	14	23x23	21	no	2200
		1993	RMK TOP15	24	23x23	14	no	1350
Slovenia	South of Ljubljana	1967	RC5	4	18x18	12,5	no	2700
		1979	RMK_15	8	23x23	12,5	yes	2140
		1998	LMK 1021	4	23x23	12,5	yes	4050
Estonia	Voka	1962	-	158	-	42	no	-
	Asuka	1955	-	42	-	-	no	-
Poland	Warsaw	1986 (phot. scan)	Wild RC10	38	23x23	14	yes	690
		1986 (regular scan)	Wild RC10	38	23x23	-	yes	690
Italy	Various	1942-1945	Various	20	23x23	Various	no	Various

Table 1. Some specifications of the EuroSDR TIME benchmark aerial datasets.

4. PERFORMED TESTS

4.1 Image pre-processing

Image enhancement includes techniques employed in image pre-processing to highlight salient image features by transforming pixel intensity values. Poor contrast and brightness levels, over- or underexposed images, are normally the result of erroneous acquisitions or scanning (digitization) processes (Baltsavias, 1999). Several solutions have been proposed for image correction and for balancing improvements and enhancements in brightness and contrast (Maurya et al., 2021). Some of the shared aerial blocks from the TIME benchmark present evident radiometric problems, which need to be resolved to enable and facilitate photogrammetric processing tasks. This is especially true for the different Cyprus series (Figure 2). Standard brightness and contrast adjustment algorithms, such as histogram equalisation, were employed before running further processing. Besides investigating more robust and automatic procedures for handling this processing task, further analyses should deepen the radiometric distortion and noise level introduced with these pixels' manipulation, especially in 3D reconstruction processes.
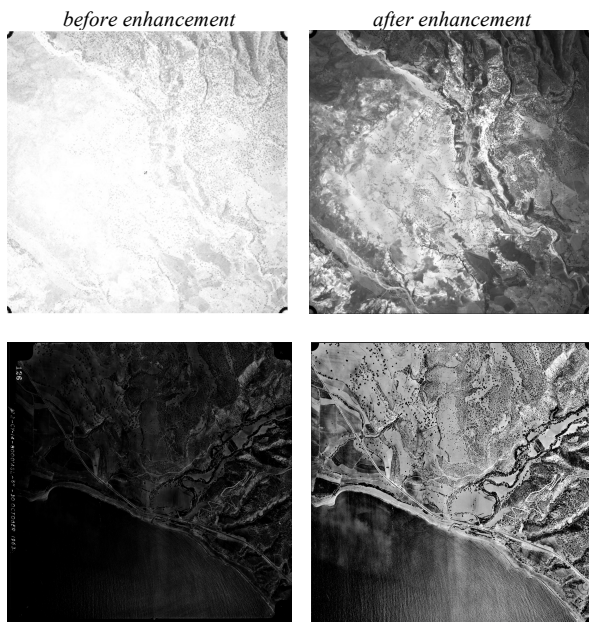

before enhancement   *after enhancement*

Figure 2. Exposure, contrast and brightness correction for the 1957 (upper images) and 1963 (lower images) Cyprus blocks.

4.2 Aerial triangulation (AT) and geo-referencing

In order to derive metric spatial information from aerial digitised photographs, an initial transformation between the pixel and the image coordinate system needs to be computed. This is defined by the fiducial marks (Figure 3) visible on the analogue photos, varying in shape and location (in the middle or in the corners of each side of the image or in all eight locations).

The transformation between the pixel and image coordinate systems can be rigorously resolved only when (i) the marks are clearly visible and accurately detected in the digitised images and (ii) the image coordinate system for the transformation is known. As in many datasets, and especially the oldest ones, no camera calibration certificates and fiducial marks coordinates are available, some initial approximations for the photogrammetric processing are needed (Nocerino et al., 2012). A coarse transformation between the pixel and image coordinate systems can be performed by adopting a virtual reference system and

assuming that the principal point is at the centre of the fiducial mark intersection. In our tests, pixel coordinates of fiducial marks with respect to the principal point were manually measured in some images, and their average value used for computing an affine transformation for all of the block.
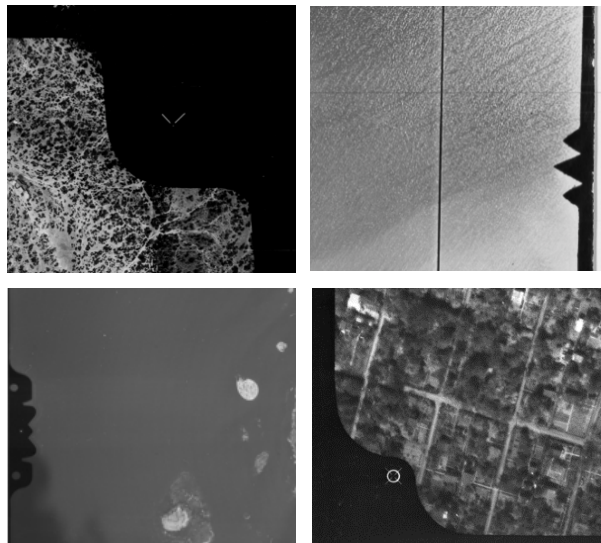

Figure 3. Examples of different types of fiducial marks visible, or partially visible, in the TIME archival images.

Nowadays, some photogrammetric software offers semi-automatic solutions for handling this process (Giordano and Mallet, 2018), mainly employing pattern matching methods for the automatic detection and measure of fiducial marks coordinates, starting from a single archival image. In our tests, the principal point position was used for cropping the single frames and ensuring the same sensor format for all the blocks. Then, camera parameters can be estimated in a bundle block adjustment, eventually including also distortion parameters (Poli et al., 2020).

Several blocks or subsets of the TIME benchmark were also processed with different photogrammetric software to assess the capability of automatic SfM-based processes with analogue grayscale images. Tie points were extracted using hand-crafted (RootSIFT), hybrid (HP) and learning-based (LoFTR, KeyNet, D2Net, RoRD) approaches. Results highlight the potential of newly developed solutions in finding tie points among multi-temporal datasets (Figure 4).

Reference GCPs are necessary for scaling and geo-referencing purposes. This stage is typically difficult and time-consuming, considering that stable-over-time GCPs are rarely available or clearly visible in historical imagery (Figure 5).

When some actual/modern GCPs are available (e.g. for the Vienna dataset) through maps or orthophotos, the manual identification and measurement of these points in the historical images followed by a Helmert transformation is the most common procedure for geo-referencing image datasets. However, this process can be quite complex and laborious, and conspicuous errors can be introduced in this phase. As an example, two temporal series of Vienna, from 1958 and 1976 (15 and 21 images, respectively), were geo-referenced using this procedure, applying a standard rigid transformation.

Using 10 manually identified GCPs, the RMSE is about 4 meters for the 1958 images (GSD≃0.5 m) and 3.5 meters for the 1976 images (GSD≃0.2 m). From this example and further tests conducted on several TIME benchmark datasets, the automatic AT of archival aerial images remains to be still a tricky and insidious processing task.
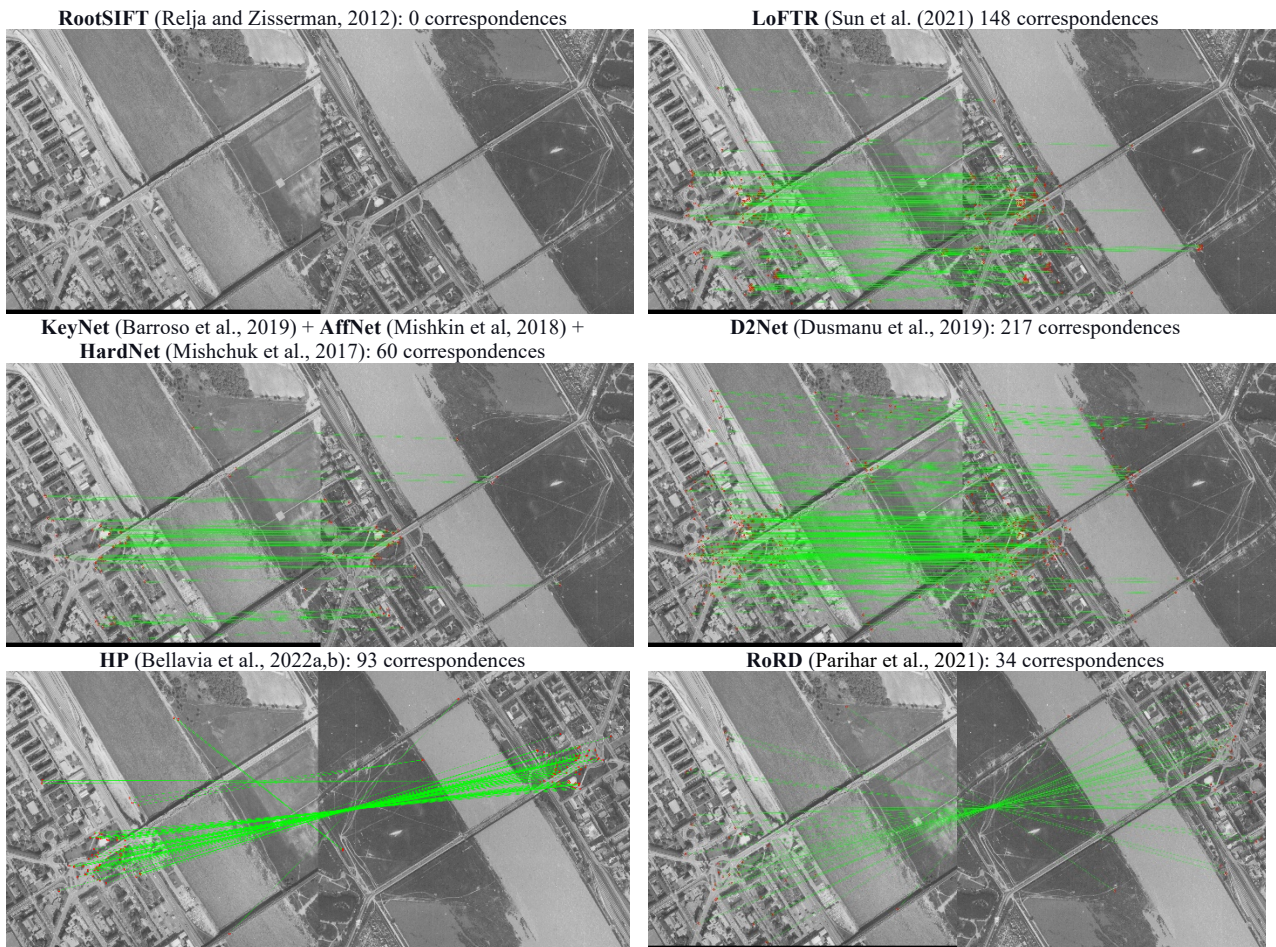
RootSIFT (Relja and Zisserman, 2012): 0 correspondences

LoFTR (Sun et al. (2021) 148 correspondences

KeyNet (Barroso et al., 2019) + **AffNet** (Mishkin et al, 2018) +
HardNet (Mishchuk et al., 2017): 60 correspondences

D2Net (Dusmanu et al., 2019): 217 correspondences

HP (Bellavia et al., 2022a,b): 93 correspondences

RoRD (Parihar et al., 2021): 34 correspondences

Figure 4. Image pairs over Vienna from the 1958 and 1976 blocks. Extracted (and geometrically verified) correspondences with hand-crafted, hybrid and deep learning approaches. Only HP and RoRD methods could handle rotated images.

Figure 5. Example of barely visible GCP (left images) in historical images (Cyprus, 1993).

Increasingly robust procedures need to be developed, especially when dealing with large temporal spas, missing ancillary and ground reference information.

4.3 DSM and orthoimages production

DSMs and orthophotos are the most common product of the photogrammetric processing of archival aerial images. Once the EO parameters are recovered, DSMs can be produced by means of dense image matching (DIM) algorithms (Figure 6), and then

orthorectification and mosaicking can be performed. The correctness of the AT obviously conditions the quality of these products and the extraction of reliable and precise information from historical data.
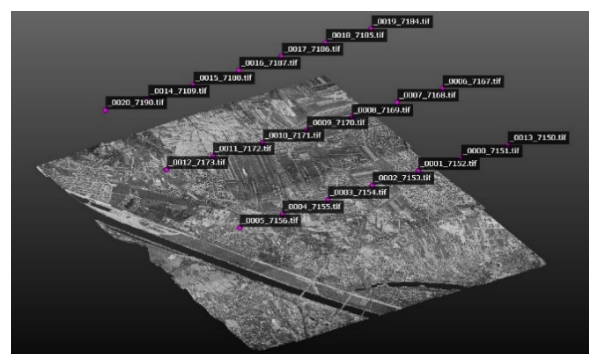

Figure 6. AT and DIM results for the 1976 aerial block (21 images) from the Vienna dataset.

In order to assess the capability of automatic algorithms for the DSMs production, tests were conducted on the same Austrian temporal series (1958 and 1976) presented in the previous section, and the final orthomosaics were visually evaluated and compared with some actual data (Figures 7 and 8).
A different example of AT and DSM production from a subset (31 images) of the oldest Norway block (110 images) is shown in Figure 9.
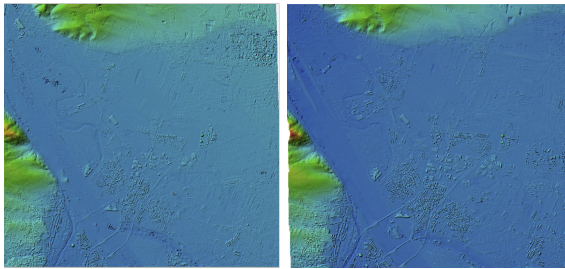
Figure 7. Multi-temporal (1958 - left and 1976 - right) DSM (5 m resolution) over Vienna (Austria).

Figure 8. Georeferenced and orthorectified images over Vienna (Austria) produced from the 1958 dataset, overlaid in Google Earth for visual quality control.

Norwegian blocks are among the most complete in the TIME benchmark, in terms of both number of images and available ancillary information. Provided data were used to re-process the image block and generate a DSM of the area of interest.
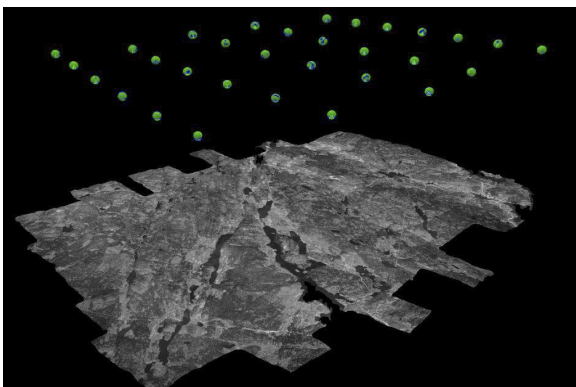

Figure 9. Automatic aerial triangulation and dense image matching results (1 m resolution) for a subset of the Austre Moland dataset (Norway), acquired in 1950.

The numerous lakes present in this dataset (black areas in the historical images) produced various errors in the DSM. The improvement of DSMs quality is a further open research task, considering that matching results are conditioned by the radiometric saturation, scanning errors, dark regions and shadows, typically present in historical imageries.

4.4 Automatic colorization of historical images

Automatic grayscale image colorization is an emerging topic following the latest results of CNN (Liu et al., 2018) or GAN-based (Poterek et al., 2020) deep learning methods. Colorized historical photos can effectively aid further image processing

tasks (data classification, object recognition, etc.). Several methods have been already been developed for tackling this image processing task in terrestrial scenarios, poorly performing on aerial-scale historical data. A few methods have been recently presented for this specific application, however, source codes are unavailable for tests and comparisons.

Therefore, a new deep learning method was implemented by the authors for handling the automatic colorization of several TIME benchmark datasets. The Italian imageries, including miscellaneous data captured between 1942 and 1945, were selected for testing the colorization performance of the proposed technique (Figure 10). The implemented network architecture is a combination of U-NET (Ronneberger et al. 2015) and the hypercolumn techniques (Hariharan et al., 2015), called "Hyper-U-NET". Some examples of the colorization outputs with the proposed techniques are presented in Figure 10. Further tests and quantitative analyses are needed.
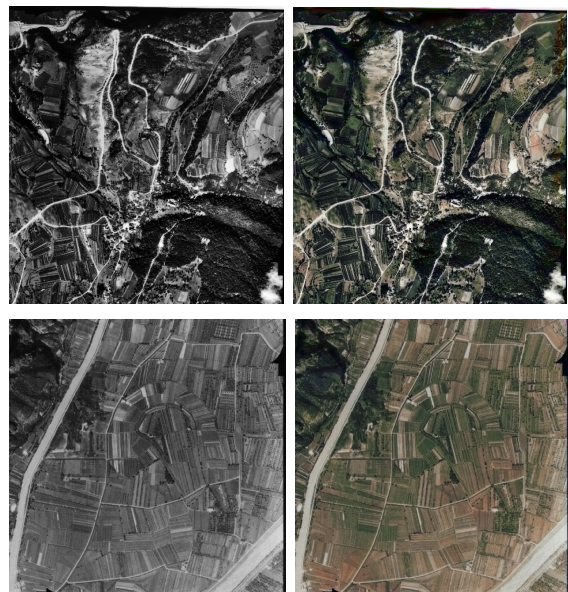

Figure 10. Example of colorization results.

5. CONCLUSIONS

The TIME benchmark is an ongoing EuroSDR research activity, offering access to many historical aerial image blocks to test and validate 2D and 3D image processing algorithms. The benchmark follows an increasing number of scientific publications related to historical airborne images as well as an increasing effort of local and national mapping authorities (NMCAs) to create digital archives and valorize the contents. The aerial blocks shared by several NMCAs are heterogeneous in terms of number of images, available ancillary information, GSD, image quality, overlap, etc. This paper has presented the main characteristics of the collected data, some possible investigation topics, and processing tests with several automatic solutions.

Issues and bottlenecks encountered in the presented experiments proved that archival aerial image processing is a wide and challenging research field. Many gaps in the currently available solutions can preclude the full exploitation of this incredible source of information on past urban scenarios.

EuroSDR and authors believe in the enormous potential of the TIME benchmark and in the value of historical aerial imagery. It is anticipated that the research community will benefit from the various datasets and that new innovative methods will be created to boost the use of these sources.

ACKNOWLEDGEMENTS

REFERENCES

Bakuła, K., Mills, J. P., and Remondino, F., 2019. A review of benchmarking in photogrammetry and Remote Sensing. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* XLII-1/W2, pp. 1-8.

Baltsavias, E., 1999. On the performance of photogrammetric scanners. In: Fritsch, D., Spiller, R. (Eds.), *Photogrammetric Week '99*, pp. 155-173.

Barroso-Laguna, A., Riba, E., Ponsa, D., Mikolajczyk, K., 2019. Key.Net: Keypoint detection by handcrafted and learned CNN filters. Proc. *ICCV*.

Bayr, U., 2021. Quantifying historical landscape change with repeat photography: an accuracy assessment of geospatial data obtained through monoplotting. Int. Journal of Geographical Information Science, Vol. 35, pp. 2026-2046.

Bellavia, F., Morelli, L., Menna, F., Remondino, F., 2022a. Image orientation with a hybrid pipeline robust to rotations and wide-baselines. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* Vol. XLVI-2/W1-2022, pp. 73-80

Bellavia, F., Colombo, C., Morelli, L., Remondino, F., 2022b. Challenges in image matching for cultural heritage: an overview and perspective. Proc. *FAPER2022, Springer LNCS*, in press.

Bozzini, C., Conedera, M., and Krebs, P., 2012. A new monoplotting tool to extract georeferenced vector data and orthorectified raster data from oblique non-metric photographs. *Int. Journal of Heritage in the Digital Era*, 1 (3), pp. 499-518.

Cléry, I., Pierrot-Deseilligny, M., Vallet, B., 2014. Automatic georeferencing of a heritage of old analog aerial photographs. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, 2(3).

Clermont, D., Kruse, C., Rottensteiner, F. and Heipke, C., 2019. Supervised detection of bomb craters in historical aerial images using convolutional neural networks. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* 42(2/W16), pp. 67-74.

Cowley, D. C., Ferguson, L., & Williams, A., 2013. The aerial reconnaissance archives: a global aerial photographic collection. In W. S. Hanson & I. Oltean (Eds.), *Archaeology from historical aerial and satellite archives*, pp. 13-30.

Dias, M., Monteiro, J., Estima, J., Silva, J. and Martins, B., 2020. Semantic segmentation and colorization of grayscale aerial imagery with W-Net models. *Expert Systems*, 37(6), p.e12622.

Ding, J., Xue, N., Xia, G.S., Bai, X., Yang, W., Yang, M.Y., elongie, S., Luo, J., Datcu, M., Pelillo, M. and Zhang, L., 2021. Object detection in aerial images: A large-scale benchmark and challenges. *arXiv preprint arXiv:2102.12219*.

Dolejš, M., Pacina, J., Veselý, M. and Brétt, D., 2020. Aerial bombing crater identification: exploitation of precise digital terrain models. *ISPRS Int. Journal of Geo-Information*, 9(12), p.713.

Dusmanu, M., Rocco, I., Pajdla, T., Pollefeys, M., Sivic, J., Torii, A., Sattler, T., 2019. D2-net: A trainable CNN for joint detection and description of local features. Proc. *CVPR*.

Fieber, K.D., Mills, J.P., Miller, P.E., Clarke, L., Ireland, L., and Fox, A.J., 2018. Rigorous 3D change determination in Antarctic Peninsula glaciers from stereo WorldView-2 and archival aerial

imagery. *Remote Sensing of Environment,* 205, pp. 18-31.

Giordano, S., Mallet, C., 2019. Archiving and geoprocessing of historical aerial images: Current status in Europe. *EuroSDR Official Publication No. 70.*

Giordano, S., Bris, A. L., Mallet, C., 2018. Toward automatic georeferencing of archival aerial photogrammetric surveys. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. 4(2).

Grottoli, E., Biausque, M., Rogers, D., Jackson, D.W. Cooper, J.A.G., 2020. Structure-from-motion-derived digital surface models from historical aerial photographs: A new 3D application for coastal dune monitoring. *Remote Sensing*, 13(1), p.95.

Haala, N., 2014. Dense Image Matching Final Report. *EuroSDR Publication Series*, Official Publication No. 64, pp. 115-145.

Hariharan, B., Arbeláez, P., Girshick, R. and Malik, J., 2015. Hypercolumns for object segmentation and fine-grained localization. Proc. *IEEE CVPR*, pp. 447-456.

Hu, Q., Yang, B., Khalid, S., Xiao, W., Trigoni, N., Markham, A., 2021. Towards semantic segmentation of urban-scale 3d point clouds: a dataset, benchmarks and challenges. *arXiv:2009.03137.*

Karel, W., Doneus, M., Verhoeven, G., Briese, C., Ressl, C. and Pfeifer, N., 2013. Oriental - automatic geo-referencing and ortho-rectification of archaeological aerial photographs. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-5/W1, pp. 175-180.

Kim, J. S., Miller, C. C. and Bethel, J., 2010. Automated georeferencing of historic aerial photography. *Journal of Terrestrial Observation*, 2(1), pp. 6.

Kölle, M., Laupheimer, D., Schmohl, S., Haala, N., Rottensteiner, F., Wegner, J.D., Ledoux, H., 2021. The Hessigheim 3D (H3D) benchmark on semantic segmentation of high-resolution 3D point clouds and textured meshes from UAV LiDAR and Multi-View-Stereo. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, Vol. 1.

Kupidura, P., Osińska-Skotak, K., Lesisz, K., Podkowa, A,. 2019. The efficacy analysis of determining the wooded and shrubbed area based on archival aerial imagery using texture analysis. *ISPRS International Journal of Geo-Information*, Vol. 8(10):450.

Jao, F.-J., Chu, H.-J., Tseng, Y.-H., 2014. Historical image registration and land-use land-cover change analysis. *Environments*, 1(2), pp. 181-189.

Liu, H., Fu, Z., Han, J., Shao, L. and Liu, H., 2018. Single satellite imagery simultaneous super-resolution and colorization using multi-task deep neural networks. *J. Visual Communication and Image Representation*, 53, pp. 20-30.

Lydersen, J.M., Collins, B.M., 2018. Change in vegetation patterns over a lLarge forested landscape based on historical and contemporary aerial photography. *Ecosystems*, 21, 1348-1363.

Long, Y., Xia, G.S., Li, S., Yang, W., Yang, M.Y., Zhu, X.X., Zhang, L. and Li, D., 2021. On creating benchmark dataset for aerial image interpretation: reviews, guidances, and million-AID. *IEEE J. Selected Topics in Applied Earth Observations and Remote Sensing*, 14, pp. 4205-4230.

Maiwald, F., 2019. Generation of a benchmark dataset using historical photographs for an automated evaluation of different feature matching methods. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* XLII-2/W13, pp. 87-94.

Maurya, L., Lohchab, V., Mahapatra, P.K. and Abonyi, J., 2021. Contrast and brightness balance in image enhancement using Cuckoo search-optimized image fusion. *Journal of King Saud University-Computer and Information Sciences*.

Mboga, N., Grippa, T., Georganos, S., Vanhuysse, S., Smets, B., Dewitte, O., Wolff, E. and Lennert, M., 2020. Fully convolutional networks for land cover classification from historical panchromatic aerial photographs. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167, pp. 385-395.

Meixner, P. and Eckstein, M., 2016. Multi-temporal analysis of WWII reconnaissance photos. *Int. Arch. Photogramm. Remote*

Sens. Spatial Inf. Sci., *41*.

Micheletti, N., Lane, S. and Chandler, J., 2015. Application of archival aerial photogrammetry to quantify climate forcing of alpine landscapes. *The Photogrammetric Record,* 30(150).

Mishchuk, A., Mishkin, D., Radenovic, F., Matas, J., 2017. Working Hard to Know Your Neighbor's Margins: Local Descriptor Learning Loss. Proc. *NIPS*.

Mishkin, D., Radenovic, F., Matas, J., 2018. Repeatability is not enough: Learning affine regions via discriminability. Proc. *ECCV*.

Nebiker, S., Lack, N., Deuber, M., 2014. Building change detection from historical aerial photographs using dense image matching and object-based image analysis. *Remote Sensing*, 6, pp. 8310-8336.

Nex, F., Gerke, M., Remondino, F., Przybilla H.-J., Bäumker, M., Zurhorst, A., 2015. ISPRS benchmark for multi-platform photogrammetry. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.,* Vol. II-3/W4, pp. 135-142.

Nocerino, E., Menna, F., Remondino, F., 2012. Multi-temporal analysis of landscapes and urban areas. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. 39(4), pp. 85-90.

Osińska-Skotak, K., Jełowicki, Ł., Bakuła, K., Michalska-Hejduk, D., Wylazłowska, J., & Kopeć, D., 2019. Analysis of using dense image matching techniques to study the process of secondary succession in non-forest Natura 2000 habitats. *Remote Sensing*, Vol. 11(8):893.

Parihar, U.S., Gujarathi, A., Mehta, K., Tourani, S., Garg, S., Milford, M., Krishna, K.M., 2021. RoRD: Rotation-robust descriptors and orthographic views for local feature matching. Proc. *IEEE CVPR*

Peppa, M.V., Mills, J.P., Fieber, K.D., Haynes, I., Turner, S., Turner, A., Douglas, M. and Bryan, P.G., 2018. Archaeological feature detection from archive aerial photography with a SfM-MVS and image enhancement pipeline. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* XLII-2, pp. 869–875.

Pinto, A.T., Gonçalves, J.A., Beja, P. and Pradinho Honrado, J., 2019. From archived historical aerial imagery to informative orthophotos: A framework for retrieving the past in long-term socioecological research. *Remote Sensing*, 11(11), p.1388.

Poli, D., Casarotto, C., Strudl, M., Bollmann, E., Moe, K., and Legat, K., 2020: Use of historical aerial images for 3D modelling of glaciers in the Province of Trento. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 43(B2-2020), pp. 1151–1158.

Popelková, R. and Mulková, M., 2016. Multitemporal aerial image analysis for the monitoring of the processes in the landscape affected by deep coal mining. *European J. Remote Sensing*, 49(1), pp. 973-1009.

Poterek, Q., Herrault, P.A., Skupinski, G. and Sheeren, D., 2020. Deep learning for automatic colorization of legacy grayscale aerial photographs. *IEEE J. Selected Topics in Applied Earth Observations and Remote Sensing*, 13, pp. 2899-2915.

Ratajczak, R., Crispim-Junior, C.F., Faure, É., Fervers, B. and Tougne, L., 2019. Automatic land cover reconstruction from historical aerial images: An evaluation of features extraction and classification algorithms. *IEEE Transactions on Image Processing*, 28(7), pp. 3357-3371.

Redecker, A.P., 2008. Historical aerial photographs and digital photogrammetry for impact analyses on derelict land sites in human settlement areas. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* Vol. 37(B8), pp. 5-10.

Redweik, P., Roque, D., Marques, A., Matildes, R. and Marques, F. 2010. Triangulating the past-recovering Portugal's aerial images repository. *Photogrammetric Engineering and Remote Sensing*, 76, pp. 1007-1018.

Relja, A., Zisserman, A., 2012. Three things everyone should know to improve object retrieval. Proc. *IEEE CVPR*.

Ressl, C., Karel, W., Piermattei, L., Puercher, G., Hollaus, M. and Pfeifer, N., 2020. Multi-epoch bundle block adjustment for processing large dataset of historical aerial images. In *EGU General Assembly Conference Abstracts* (p. 22544).

Risbøl, O., Briese, C., Doneus, M., & Nesbakken, A., 2014. Monitoring cultural heritage by comparing DEMs derived from historical aerial photographs and airborne laser scanning. *Journal of Cultural Heritage*, 16(2), pp. 202-209.

Ronneberger, O., Fischer, P. and Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. *Inter. Conference on Medical image computing and computer-assisted intervention*, pp. 234-241. Springer, Cham.

Salach, A. (2017). SAPC–application for adapting scanned analogue photographs to use them in structure from motion technology. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 42, 197.

Sauerbier, M., Kunz, M., Fluehler, M., Remondino, F., 2004: Photogrammetric reconstruction of adobe architecture at Tucume, Peru. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. 36(5/W1).

Schulz, J., Cramer, M., Herbst, T., 2021. Evaluation of Phase One Scan Station for Analogue Aerial Image Digitisation. *Journal of Photogrammetry, Remote Sensing and Geoinformation Science* (PFG), Vol. 89(5), pp. 461-473.

Seo, D.K., Kim, Y.H., Eo, Y.D. and Park, W.Y., 2018. Learning-based colorization of grayscale aerial images using random forest regression. *Applied Sciences*, 8(8), p.1269.

Sevara, C. 2013. Top secret topographies: recovering two and three-dimensional archaeological information from historic reconnaissance datasets using image-based modelling techniques. *Int. J. Heritage in the Digital Era*, Vol. 2, pp. 395-418.

Sevara, C., Verhoeven, G., Doneus, M., Draganits, E., 2018. Surfaces from the visual past: recovering high-resolution terrain data from historic aerial imagery for multitemporal landscape analysis. J*ournal of Archaeological Method and Theory*, Vol. 25, pp. 611-642.

Siok, K. and Ewiak, I., 2020. The simulation approach to the interpretation of archival aerial photographs. *Open Geosciences*, 12(1), pp.1-10.

Sonnemann, T., Saurbier, M., Remondino, F., Schoretter, G., 2006. Reality-based 3D modeling of the Angkorian temples using aerial images. Proc. of *2nd International Conference on "Remote Sensing in Archaeology"*, pp. 573-579.

Sun, J., Shen, Z., Wang, Y., Bao, Y., Zhou, X., 2021. LoFTR: Detector-free local feature matching with Transformers. Proc. *CVPR*.

Valjavec, M.B., Zorn, M. and Ribeiro, D., 2018. Mapping war geoheritage: Recognising geomorphological traces of war. *Open Geosciences*, 10(1), pp. 385-394.

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

Verhoeven, G., Sevara, C., Karel, W., Ressl, C., Doneus, M., & Briese, C., 2013. Undistorting the past: new techniques for orthorectification of archaeological aerial frame imagery. *Good practice in archaeological diagnostics: Natural Science in Archaeology*, pp. 31-67.

Walstra, J., Chandler, J., Dixon, N. and Dijkstra, T., 2004. Time for change-quantifying landslide evolution using historical aerial photographs and modern photogrammetric methods. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 35(B4), pp. 475–480.

Wiedemann, A., Hemmleb, M., Albertz, J., 2000. Reconstruction of historical buildings based on images from the Meydenbauer archives. *Int. Arch. Photogramm. Remote Sens.*, Vol. 33(B5/2), p. 887-893.

Zambanini, S. and Sablatnig, R., 2017. A Hough voting strategy for registering historical aerial images to present-day satellite imagery. Proc. *ICIAP*, pp. 595-605, Springer.

Zhang, L., Rupnik, E. and Pierrot-Deseilligny, M., 2021. Feature matching for multi-epoch historical aerial images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 182, pp.176-189.