



ISSN: 2753-4154 (Online)

Type of the Paper (Article)

The methodology of the archival aerial image orientation based on the Structure-from-Motion method

A.Karol Karwel¹, Jakub Markiewicz¹

¹ Department of Photogrammetry, Remote Sensing and Spatial Information Systems, Faculty of Geodesy and Cartography, Warsaw University of Technology, pl. Politechniki 1, Warsaw 00-661, Poland; artur.karwel@pw.edu.pl (A.K.K), jakub.markiewicz@pw.edu.pl (J.M.)

Corresponding Author: artur.karwel@pw.edu.pl, Tel.: +48-22-234-5765

Received date: 10/11/2022; Accepted date:01/12/2022; Published date: 16/12/2022

Abstract: Nowadays, archival images find increasingly finding their way into geospatial applications, namely, among others, multi-temporal analysis, documentation reconstruction or change detection. It is, therefore, necessary to determine the images' external orientation elements that allow the images' position to be reconstructed in the assumed reference system. This paper aims to present a methodology for the extended evaluation of the automatic orientation process of archival images based on the commonly used Structure- from-Motion (SfM) approach. The work carried out presents: (1) the influence of parameter selection on the accuracy, number and distribution of tie points in the descriptor matching process at the pairwise image bundling stage using the descriptor matching approach together with the use of Random sample consensus filtered triangulation (RANSAC), (2) analyses of the reciprocal orientation quality of the images on detected points (control points) in the bundle adjustment process using simultaneous verification of the matching quality on check points, and (3) analysis of the external orientation accuracy. Points detected and matched using the SIFT algorithm on archival images of a fragment of Warsaw from 1986, 1994, and 2014 were used as reference data. A comparative analysis of the obtained results with the data obtained using the algorithms implemented in the Agisoft Metashape software (standard approach) shows that the relative orientation reprojection RMSE is about 4 times better, and detected points are even more robust.

Keywords: aerial images, archival images, bundle adjustment, RANSAC, reprojection error, Structure-from-Motion, SIFT detector

1. Introduction

Nowadays, archival aerial imagery is increasingly used, among others, in multi-time analyses [1], the reconstruction of the land development status [2], inventory [3] or the detection of objects that no longer exist [4]. To carry out spatial analyses, it is necessary to have information about the elements of the internal and external orientation of archival data [5,6]. Due to the poor quality of archival images (analogue images with poor radiometry and internal geometry) and the often lack of a Camera Calibration Certificate (without internal orientation parameters), the use of the FBM algorithm [7,8] in the orientations mentioned above is significantly difficult or impossible to perform. For this reason, the critical stage is to



perform the Structure-from-motion – bundle adjustment process, which allows you to obtain the parameters mentioned above. In most cases, the orientation of archival aerial images is associated with the classical approach, which is based on the following stages: (1) internal orientation consisting in reconstructing the camera parameters used while taking images based on the camera calibrate certificate; (2) automatic or semi-automatic measurement of tie points with relative orientation aimed at interconnecting images into one coherent block and giving them relative orientation parameters (i.e. relative to the reference image) and (3) exterior orientation (consisting in fitting previously oriented interior images) based on the control points (natural points) identifiable both in the present (digital) and archival images searched manually by an observer for the adopted reference system related to the national geodetic spatial reference system or world geodetic system [9,10]. In most cases, the orientation of the archival image is performed using the Feature Based Machining (FBM) and Area Based Machining (ABM) methods [11]. The accuracy of the orientation of archival images using the FBM and ABM methods allows obtaining the orientation quality up to 0.5 pixels [5,12–14].

The structure-from-motion (SfM) approach for archival images was proposed in this investigation. The SfM approach was divided into two main parts; (1) the full process in the Agisoft Metashape software (starting from keypoint detection to the final bundle adjustment) and (2) The test results refer to both the images orientation obtained in Agisoft Metashape (unknown detection algorithm) as well as the orientation based on the points detected using the SIFT algorithm. The result of the performed experiments shows that to obtain the correct result of the orientation of the photo block based on the Agisoft software, it is also necessary to use the point filtering module (the reprojection size criterion). It has been noticed that the orientation of images based on the points detected by the SIFT detector gives similar accuracy results as the orientation of images based on the detected points by the Agisoft Metashape detector (AM), with appropriate filtration. Suppose Agisoft Metashape does not use point filtering (which is a common practice of many users). In that case, the accuracy of the image block's orientation based on the methodology proposed by the authors is much higher.

The methodology of the orientation of aerial photographs by the SfM method using the SIFT detection algorithm presented by the authors is characterised by the uniqueness of the detected points in the characteristic local image changes (gradient changes); invariability of the detected points from radiometric and geometric distortion; stability of points independent of scale change or light conditions; the uniqueness of the detected points based on certain characteristics - which ensures the correct and reliable orientation of the images block, high rigid of their geometry and robustness to errors.

2. The Principles of the Structure-from-Motion approach

Modern software packages, applications, and solutions for automatic image orientation and 3D shape reconstruction utilise libraries and algorithms that combine methods applied in Computer Vision (CV) and photogrammetric approaches. This type of algorithm, especially a Structure-from-Motion [15], that allows for 3D scene reconstruction with images orientation is used in video games assets [15], virtual tours [16], virtual and augmented reality [17], navigation [18], spatial planning, and cultural heritage [19,20], among others.

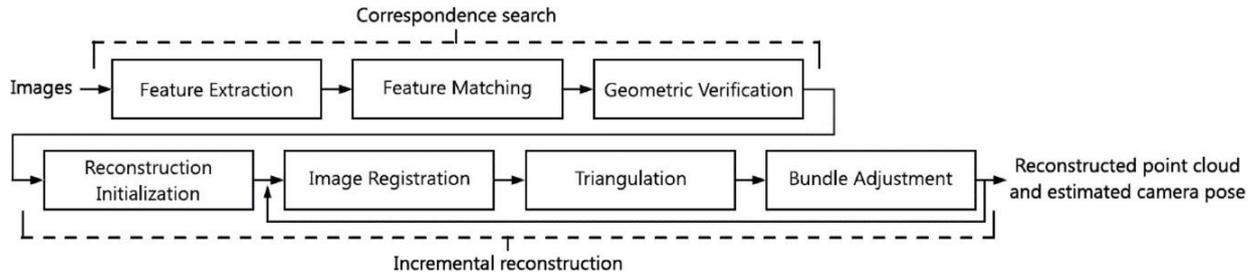


Figure 1. The Incremental SfM procedure [15].

The main idea of the SfM is to reconstruct the three-dimensional structure based on image series acquired from different positions. Figure 1 shows the overview of the incremental SfM process shown, which contains the following steps: (1) feature extraction; (2) feature matching; (3) geometric verification; (4) reconstruction initialisation; (5) image registration (also known as an image orientation); (6) triangulation with geometrical verification; and (7) bundle adjustment. To summarise, the SfM method can be divided into two main phases: the correspondence search phase (steps 1 to 3) and the iterative reconstruction phase (steps 4 to 6). This article describes the correspondence search phase and bundle adjustment step.

2.1. The Feature Extraction

The first step in the Structure-from-Motion method relies on the determination of the characteristic features, also named keypoints (detection part) and the assignment of the local characteristic of the image intensity (description part). The feature extraction step is performed on each image separately and based on the algorithms and methods which detect features invariant to image translation, scaling, and rotation, partially invariant to illumination changes, and robust to local geometric distortion [8]. Nowadays, two types of algorithms are used, namely corner/point detectors (such as FAST, BRISK, etc.) and blob detectors (i.e. SIFT and its modifications) [18,21–28]. Since many features detector exists, in this section, only SIFT [29] algorithm is described, which was used in this investigation.

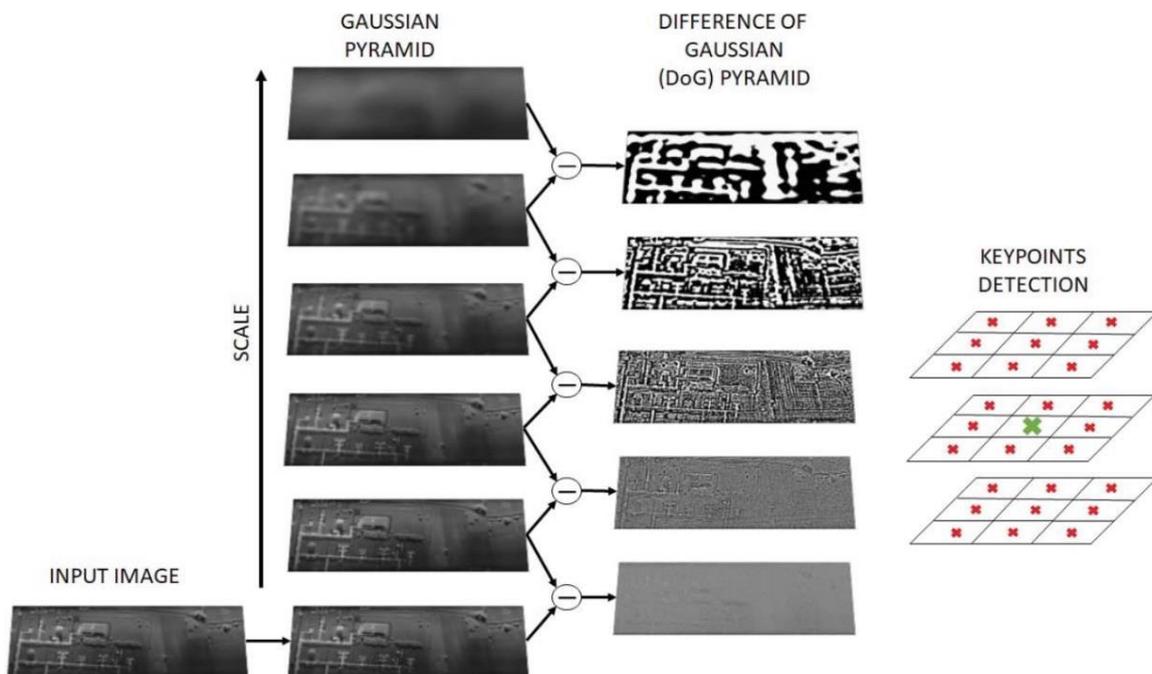


Figure 2. The diagram of SIFT keypoint detection algorithm [30]



The SIFT (Scale Invariant Feature Transform) algorithm was initially proposed by Lowe [29] for the registration of close-range images. The main idea of this algorithm is to process images in the following order (Fig. 2) (1) generate Different-of-Gaussian (DoG) images on different scale space layers (Fig. 3); (2) find minima and maxima of the DoG result; (3) low-contrast and edge response points are discarded; (4) assignment to keypoints dominant orientation. This makes it possible to detect stable for matching and recognition.

2.2. *The Feature Description*

After keypoint extraction, the next step in the SfM process is a feature description that considers the information about the gradient changes in the local neighbourhood. In the literature, several descriptors exist, such as SIFT, SURF, and Daisy [31]; in this investigation SIFT descriptor was used. The SIFT descriptor's main idea is to calculate local image gradients at a selected scale around the region's key point under study. The descriptor's work is based on analysing histograms of 4×4 pixel neighbourhood orientations with 8 bins each. The histograms are derived from magnitudes and orientations sampled in a 16×16 region around the keypoint, so that for each histogram, a 4×4 subregion of the original neighbourhood region is sampled. The magnitude and orientations of the image gradient are probed around the location of the keypoint, using the scale of the keypoint to select the image. To obtain orientation invariance, the descriptor coordinates and gradient orientations are rotated relative to the keypoint orientation [32].

2.3. *The Feature Matching, Geometrical Verification, and Bundle adjustment*

When the keypoints have assigned descriptor features, it is possible to determine which images have a correspondence and which overlap. In consequence, it is possible to determine the tie points. The determination of tie points is carried out in a two-stage process consisting of (1) an initial determination of potential pairs of points based on descriptor matching and (2) geometric verification based on an iterative approach based on the RANSAC method considering homography functions. Different strategies can be used for effectively computing matches between images (i.e.), but two usually used are approximate nearest-neighbour-based point matching and brute-force matching. Considering the descriptor matching, it is not guaranteed that the candidates for the tie points found to correspond to the 3D point in the scene and outliers could be included. Due to that fact, geometrical verification is needed that eliminate outliers and improve the quality of tie points and, consequently, the quality of the final image orientation.

Bundle adjustment plays an important role in geodesy and 3D reconstruction. Bundle adjustment constitutes a core component in most state-of-the-art multi-view geometry systems. It is typically invoked as a final refinement stage to approximate initial scene estimates and a means for removing drift in incremental reconstructions [33]. Bundle adjustment makes it possible to determine the orientation of all the images in a block while minimising reprojection and camera calibration parameters in the self-calibration process.

3. **Materials and Methods**

The proposed method of images orientation based on the extent Structure-from-Motion approach is a multi-stage process (Fig. 3); it consists of:

1) Detection of keypoint points in images using a SIFT

The selection of the function's input parameters, i.e. sigma value, edge threshold, octave parameter, and contrast parameter, was determined by empirical analyses (see paragraph 4.1). In the end, the following values were decided upon : (1) sigma =1.4, (2) nOctaveLayers = 3; (3) contrastThreshold = 0.06; (4) edgeThreshold = 10, which was a compromise



between the number of points, their distribution and the accuracy of connecting pairs of images.

2) Keypoint description

Keypoint description of the features of the detected points based on the 16 x 16 neighbourhood around the keypoints. As a result, the 128-dimensional vector is generated related to the Lowe theory (see paragraph 2.2).

3) Initial matching based on the descriptor matching

Matching on two images in longitudinal (images with a typical coverage of at least 60% taken along the direction of flight) and transverse (images with a coverage of at least 30% from two adjacent rows) coverage of the described points using the FLANN with the following parameters: (1) FLANN_DIST_EUCLIDEAN, (2) KDTreeIndexParams = 4, (3) SearchParams: checks = 32, eps = 0.01, sorted = True.

4) Clusterisation candidates of tie points

The clustering process was performed automatically, making it possible to determine groups of points to be selected for geometric verification using the RANSAC algorithm and homography (see section 5). In the first stage, for each point in the set of potential tie points, the 4 nearest neighbours were searched for which the distance was calculated. Then the median distance for all points was then determined, which was used as an input parameter in the DBSCAN algorithm. To determine the clusters in the DBSCAN function, 3 x median as the distance between points and a minimum number of points equal to 4 were used.

5) Geometric filtering of the matched potential points using the extended RANSAC function

The new geometric tie point verification approach was based on clusters detected using the DBSCAN and the RANSAC approaches. In contrast to the commonly used RANSAC method, in the proposed solution, the points to determine the homography were randomly chosen as follows: (1) only one point was independently selected from each cluster, (2) the geometry distribution of the points detected in each cluster was checked – if the distance between points were not shorter than 5 x median (computed in clusterisation step) new point was drawn, (3) the RANSAC method was used to determine the homography parameters – based on points taken from each cluster, (4) if the threshold condition was met in the dataset, the deviation value was calculated for all points and pairs of points with deviations greater than the threshold were iteratively removed, (5) if the condition was not met, further points were selected from the detected clusters at point 2. Determination of the 3D- coordinate of each filtered point using the homography function given the threshold window parameter of 0.5, 2, and 5 pixels. The filtered points were a geometrically consistent set.

6) Import tie points into Agisoft Metashape via API (Application Programming Interface)

7) Dividing tie points into control tie points (CTP) and check tie points (CHTP). CHTPs accounted for 10% of the total number of points

In the classic approach to the relative orientation of a group of images, an assessment of accuracy is made only based on the points used for bundle adjustment (BA). The disadvantage of this approach is the lack of complete controllability, as validation is performed on the points used to build the mathematical model. Therefore, in the proposed approach, points were divided into two groups: control tie points (CTPs) used for bundle adjustment and check tie points (CHTPs), which did not participate in the BA and served as an independent check. It was decided to use 10% of all tie points as a CHTP.



- 8) Initial alignment of the images on control tie points (CTP) (Agisoft Metashape). Verification of reprojection error. Filtering of points with a reprojection error greater than 1 pixel. Re-alignment

The final process of relative images orientation was a two-stage process. In the first step, the initial orientation elements of the images were determined based on CTP. In the next step, points for which the linear reprojection error value was higher than 1 pixel were filtered out, and the final bundle adjustment was performed.

- 9) Analysis of the obtained orientation results on control and check tie points (Agisoft Metashape) based on the values of x and y reprojection errors

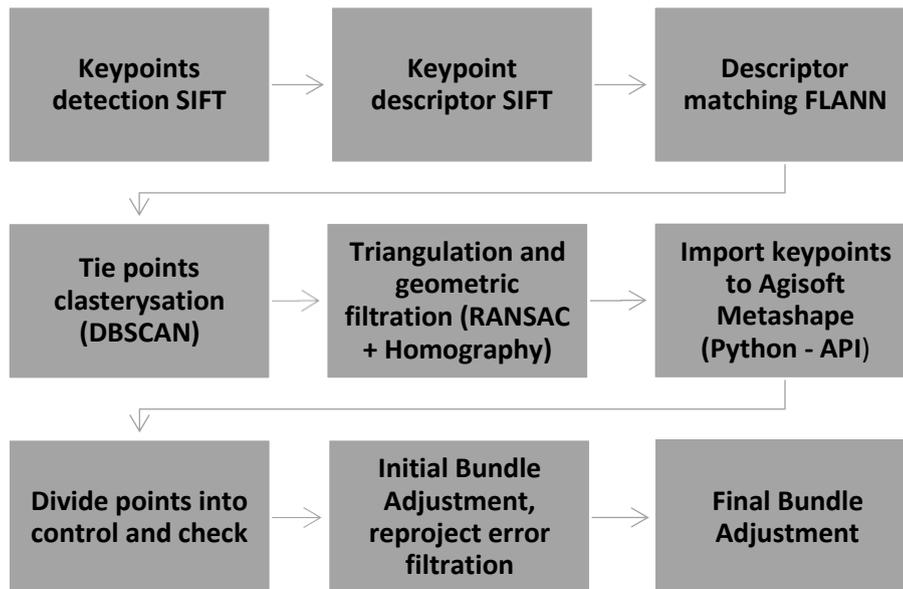
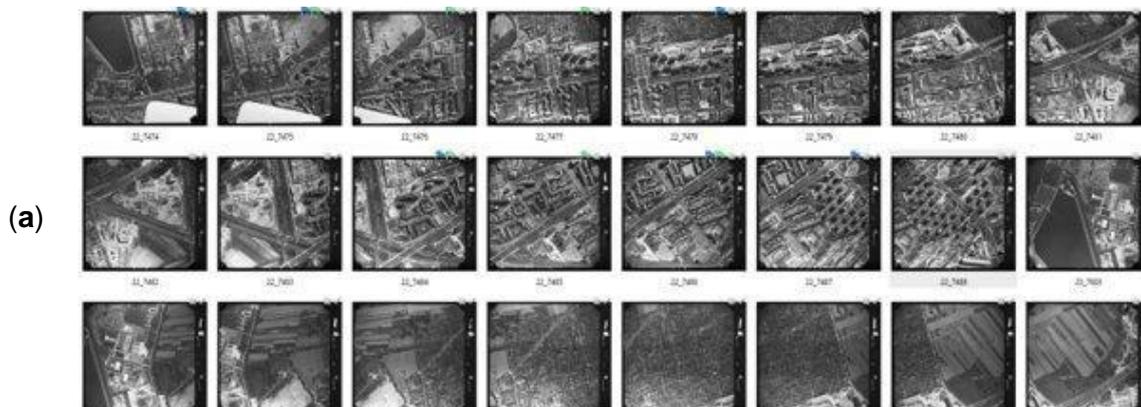


Figure 3. The diagram of the stages of the orientation of the aerial images based on the extended Structure-from-Motion approach

3.1. The test site description

The test data consisted of aerial images taken in 1986, 1994 and 2014, covering part of the area of Warsaw. Images from 1986 consisted of 2 series of 13 images taken in the 1: 3000 scale, scanned with a pixel of 14 μm . The images from 1994 consisted of 2 series of 9 images taken at a scale of 1: 5000, scanned with a pixel of 14 μm . Images from 2014 were 1 series consisting of 4 images with a field pixel dimension of 24 cm. Camera calibration metrics were not used to orient the images from all three periods. For the orientation of the images from 1986 and 1994, the camera's focal length and location of fiducial markers were used from the available documentation.



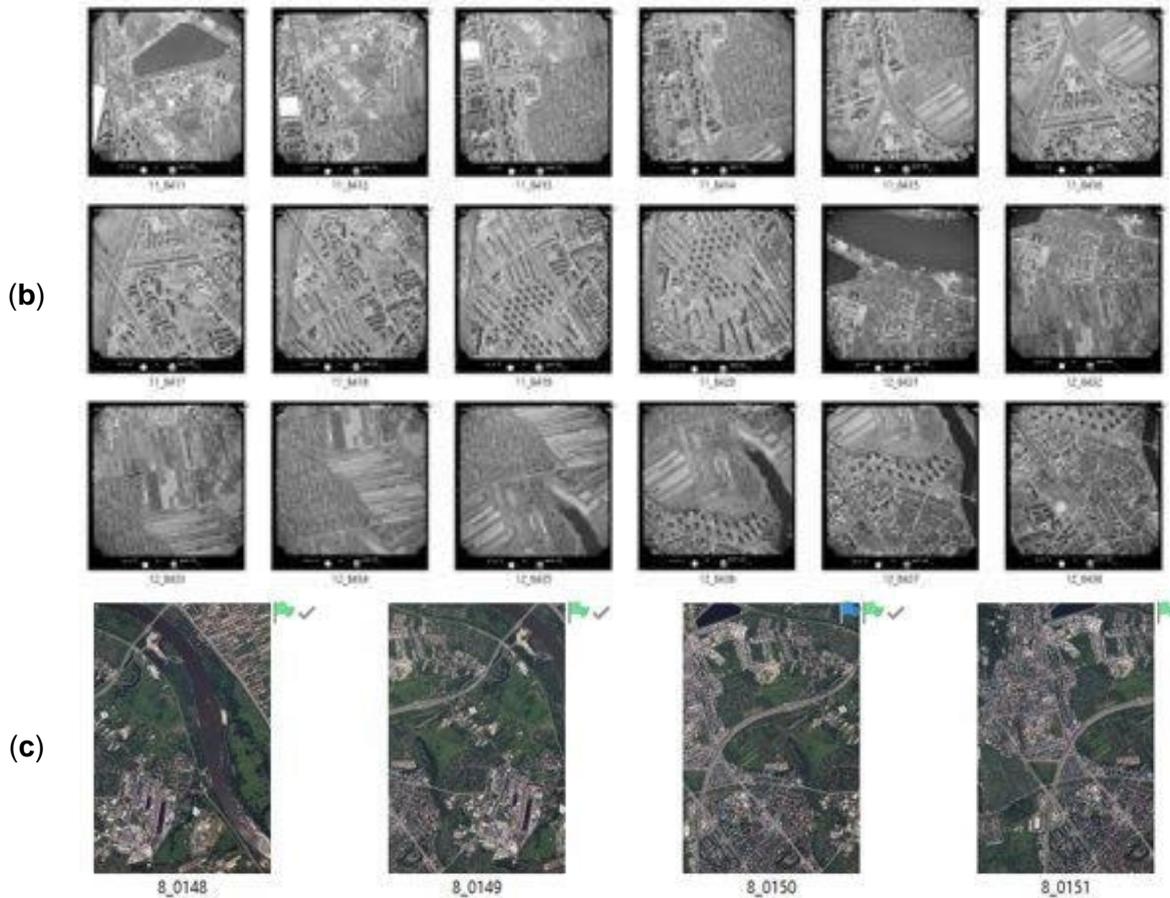
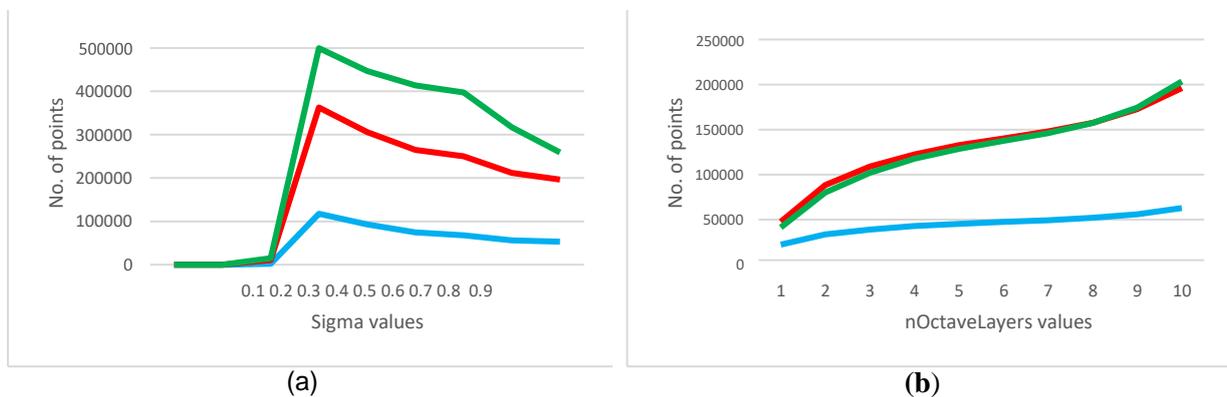


Figure 4. The image samples for (a) 1986, (b) 1994 and (c) 2014

4. Results

4.1. SIFT parameters determination

The first investigation is based on analysing the influence of the input parameters of the SIFT algorithm, namely sigma, nOctaveLayers, contrast and edge. For the study, the images were downsampled into the 4K resolution (3840 x 2160 pixels) and used for further processing. The result is shown in Figure 5.



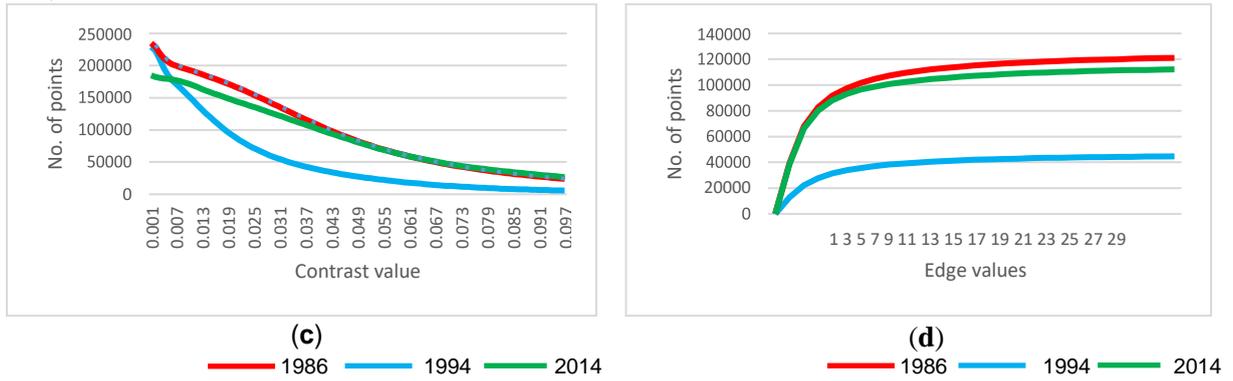


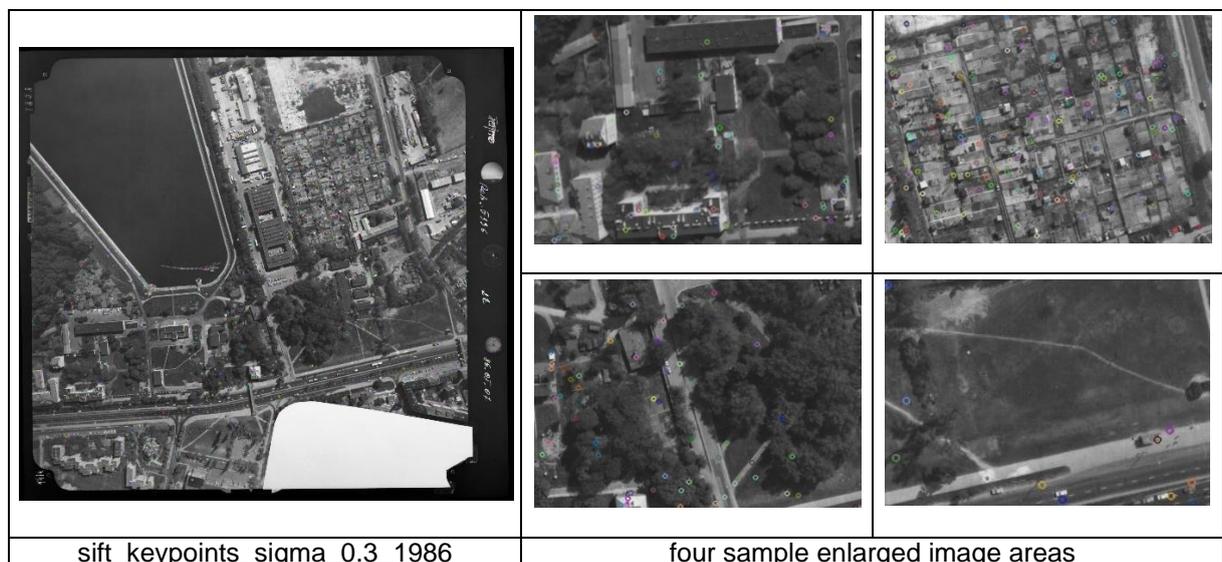
Figure 5. The influence of the (a) sigma parameter values, (b) the nOctaveLayers parameter value, (c) contrast parameter values, and (d) the edge parameter values on the number of detected

Figure 5(a) shows a graph relating to the influence of the sigma parameter on the number of detected points in the 1986, 1994 and 2014 images by the SIFT algorithm. The plot shows that the number of points for the sigma range between 0.1 and 0.3 significantly increases. The maximum values were obtained for the sigma equal 0.4 (for the image 1986, 1994, 2014); above it, the number of detected points decreased. The distribution and number of detected points depending on the sigma parameter are shown in Figure 6.

Figure 5(b) shows the graph relating to the influence of the nOctaveLayers parameter. The plots of the functions for the three years of data are approximately linear from nOctaveLayers ranging from 1 to 30. The shapes of the linear relationship between the number of points and nOctaveLayers value are similar for 1986 and 2014 years due to the similar scale of the images taken. It should be emphasised that the highest number of points was obtained for the nOctaveLayers number equal to 10.

Changing the contrast values reduces the number of points exponentially for the 2014 images and semi-logarithmically for the analogue images (Fig. 5c). No significant change in the number of points can be seen for a contrast parameter value of 0.037 for the image from the 2014 year and 0.061 for the archive images.

The increase in the number of points with a change in the edge parameter threshold progresses logarithmically (Fig. 5d). No significant increase in the number of detected points for all images is observable for an edge threshold value of 7.



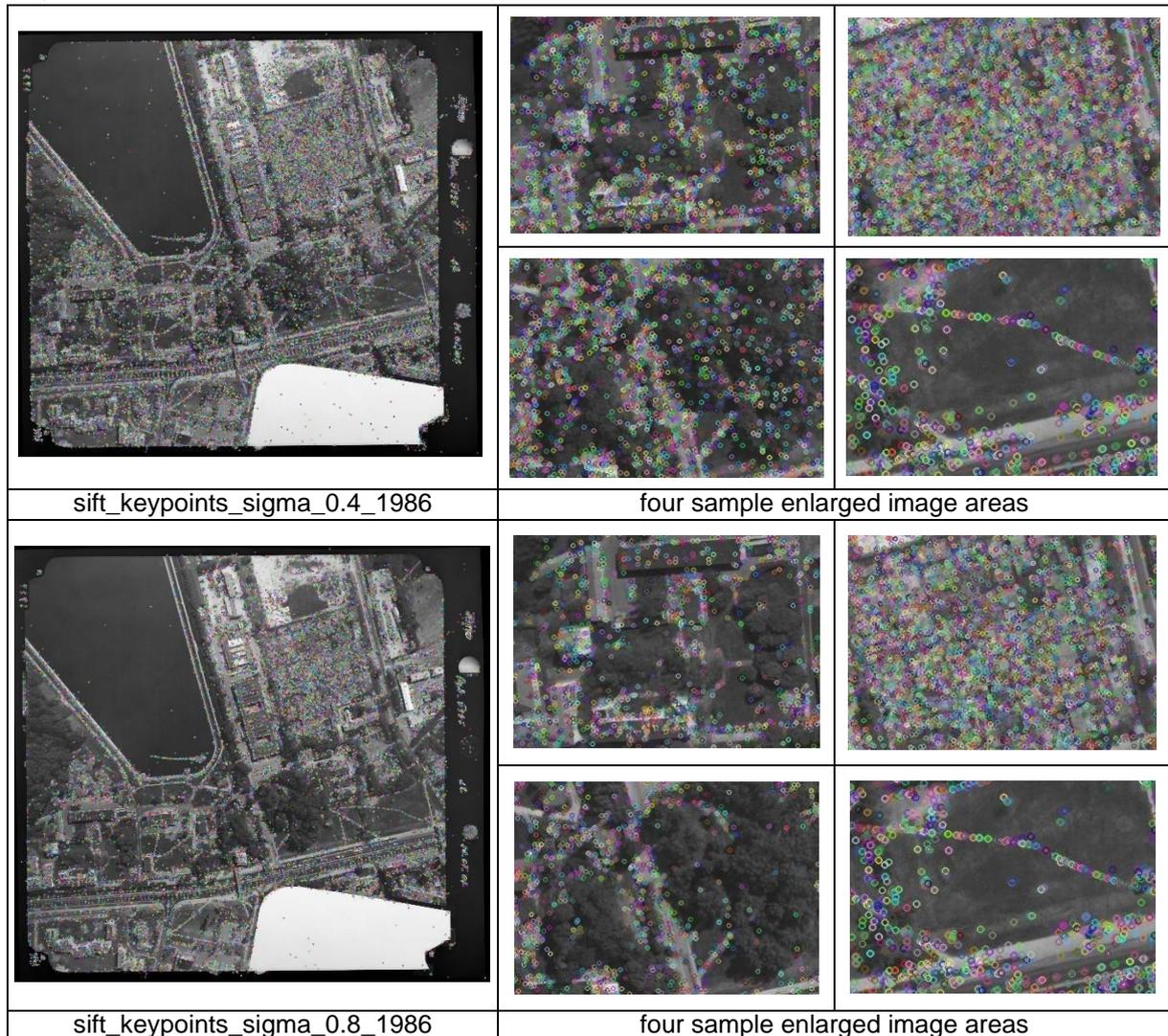


Figure 6. The distribution and number of points detected by the SIFT detector depending on the sigma parameter

The effect of changing the sigma parameter results in a 'blurring' of the image created during the DoG image generation. This relationship is reflected in the number of points, as shown in Fig. 5a. This is also visible in the image space (Fig. 6). Using a sigma value of 0.3 enabled the points to be detected evenly in the space of the entire image; the points were not detected on water, grass, etc. It appears that the detected points are robust according to the assumptions of Loew's theory [24]. The use of higher sigma values significantly impacts the number of points. However, some points appear not robust as they were detected on the water and the white massed area. Despite this, it should be noted that increasing this coefficient allows more points to be detected in wooded areas and areas with grass. This is a significant relationship, appended, for example, during the orientation of archival aerial photographs where agricultural or wooded and shrubby areas have been mapped.

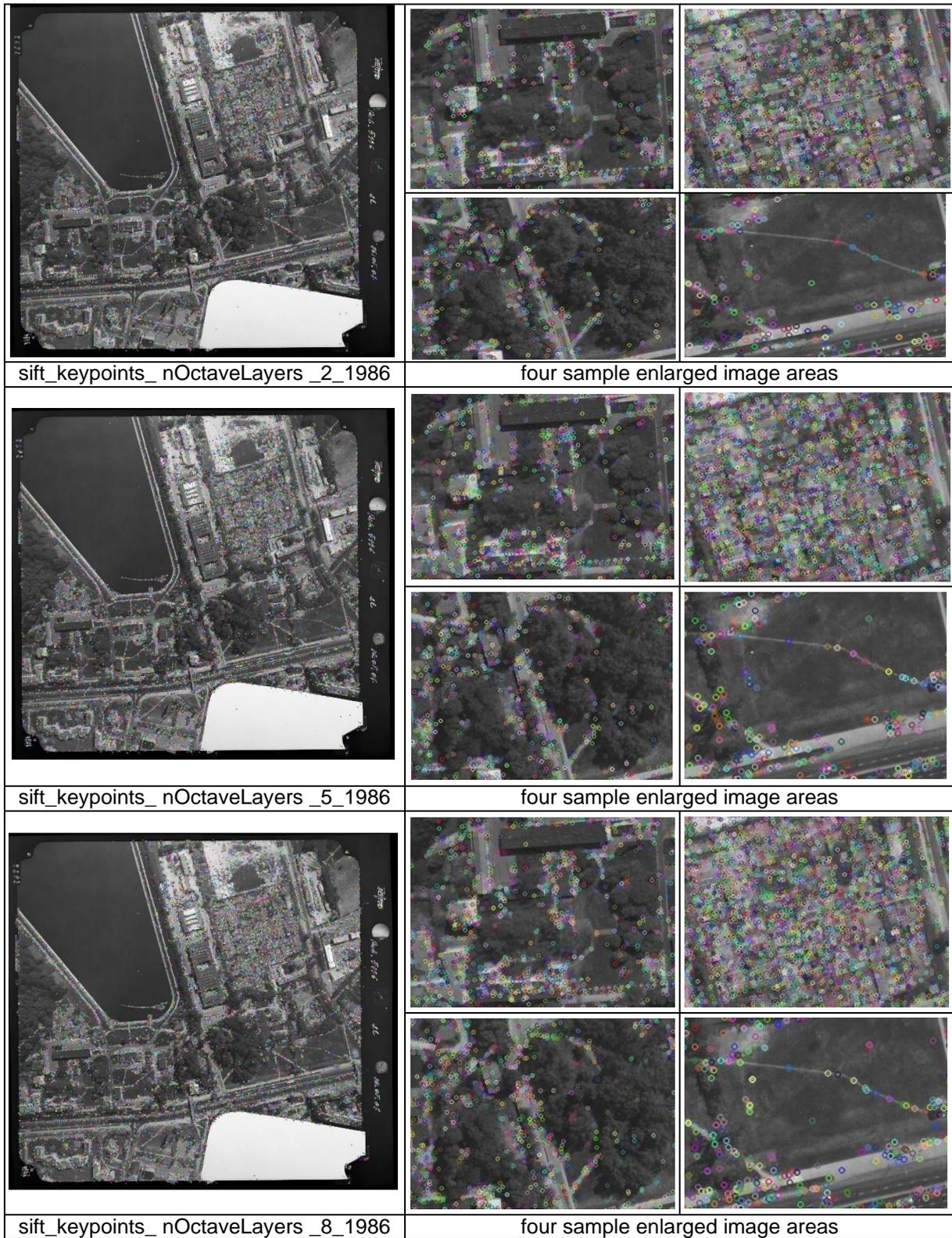


Figure 7. The distribution and number of points detected by the SIFT detector depend on the octave parameter

Increasing the number of nOctaveLayers to check increases the number of detected points in areas of varying contrast and significant changes in gradient values (Fig. 7). Therefore, changing this parameter only affects the detection of a larger number of robust points, according to Lowe's theory.

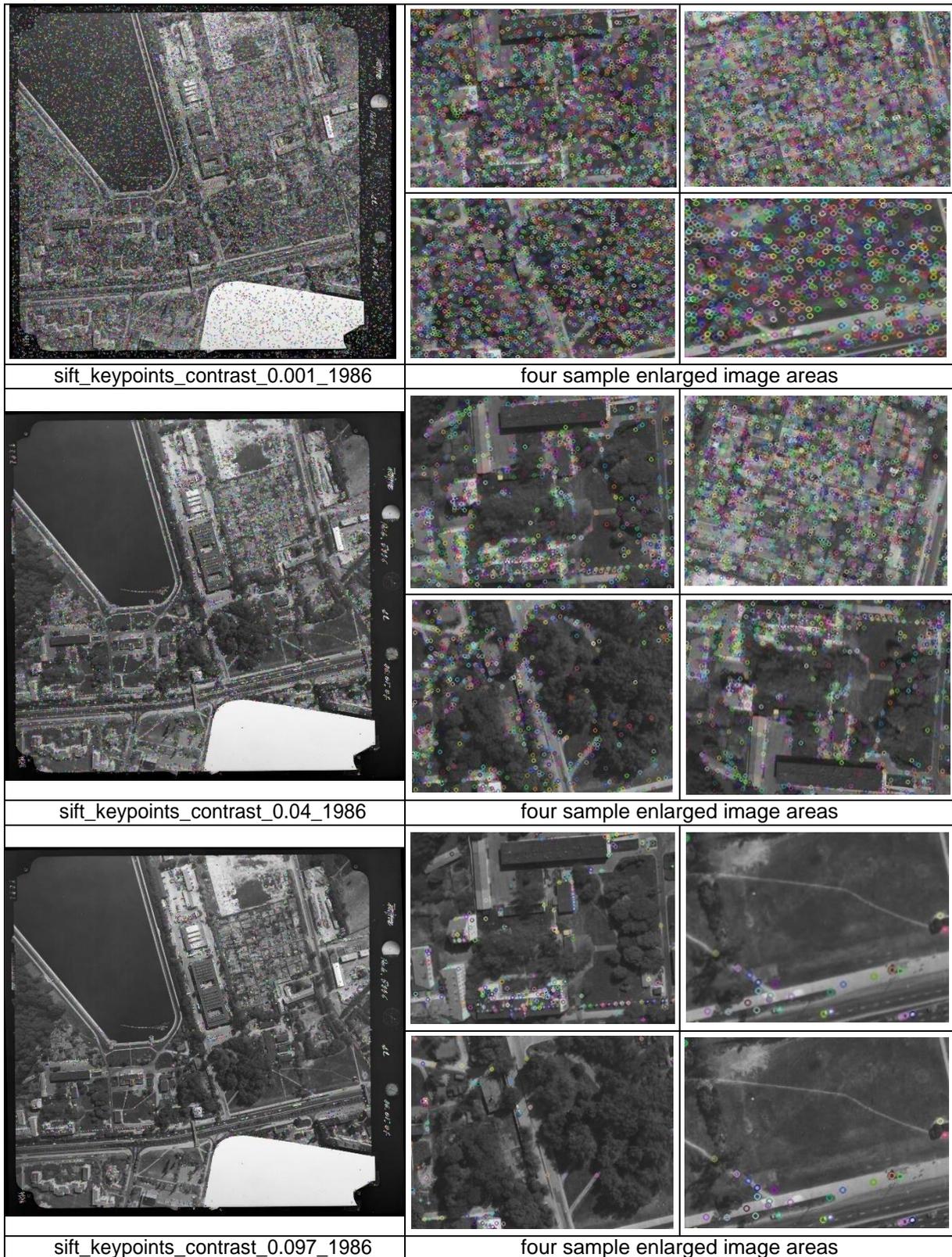


Figure 8. The distribution and number of points detected by the SIFT detector depending on the contrast parameter

The increase in the contrast parameter results in a decrease in the number of detected points. In contrast, a reduction in the contrast parameter results in a significant increase in the number of detected points (Fig. 8). A higher number of detected points does not affect its quality, as evidenced, for example, by the distribution of points on water, in a white area, or



their dense distribution in wooded areas. The selection of the parameter contrast equal to 0.097 resulted in no detection of keypoints on trees or grass.

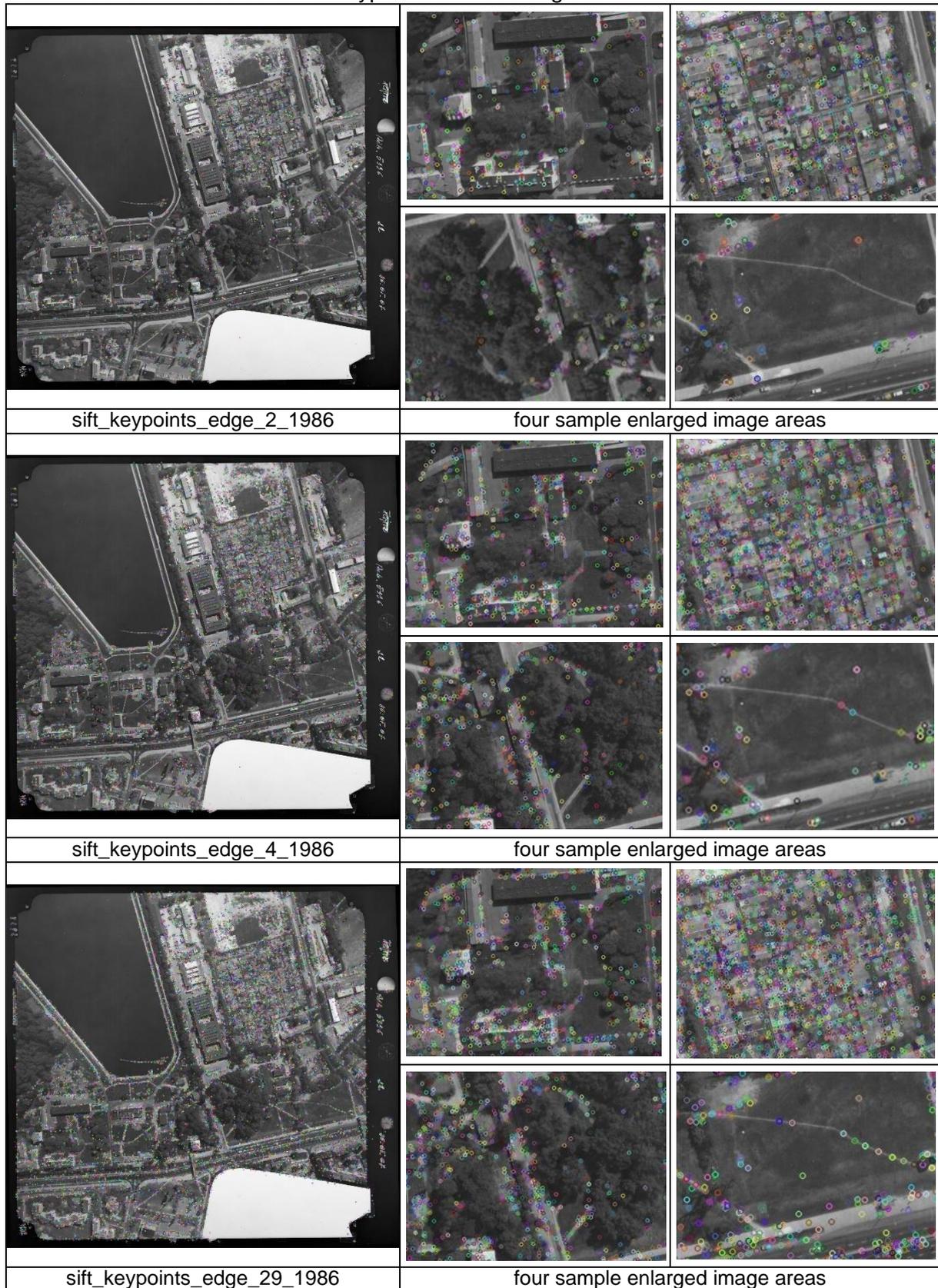


Figure 9. The distribution and number of points detected by the SIFT detector depend on the edge parameter



Figure 9 shows the effect of the edge parameter on the number of points detected in the images. As the size of the edge parameter increases, the number of points detected increases. This increase is logarithmic. Points are seen both on buildings and trees, grass and ground. For the water area, points are not detected. The number of points detected using the edge parameter is much smaller for the remaining parameters of the SIFT algorithm.

To summarise, it should be stated that the sufficient influence on the number and point distribution on images have sigma and contrast parameters. It also has a strong influence on the robustness of detected features.

4.2. *The influence of the selection input parameter in the RANSAC algorithm*

The next step of image orientation was based on geometric verification of the tie points extracted from the descriptor matching with the RANSAC approach. To select the input parameters of the RANSAC algorithm, it was decided to analyse the influence of the confidence value (Fig. 10) and the reprojection error threshold value (Fig. 11).

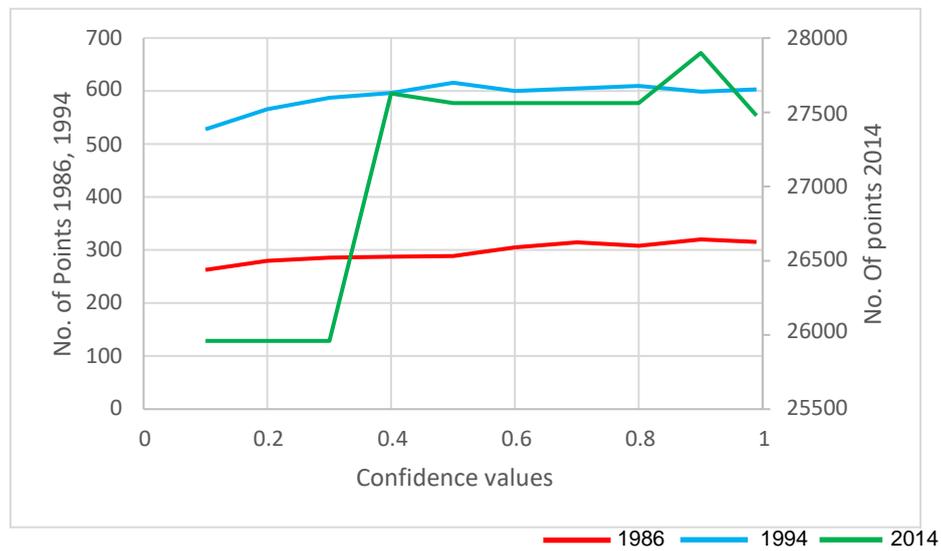


Figure 10. Influence of the confidence parameter values on the number of points detected

Figure 10 shows the influence of the confidence parameter on the number of points detected by the SIFT detector for three periods (1986, 1994, 2014). The relationship between the number of points and confidence values for analogous images is approximately linear and similar. The number of detected points for a given year slightly differed depending on the change in the confidence parameter. The plot analysis shows that the number of points for a confidence value of about 0.4 does not increase significantly. Therefore, a confidence value of 0.4 can be taken as the input parameter for the RANSAC algorithm.

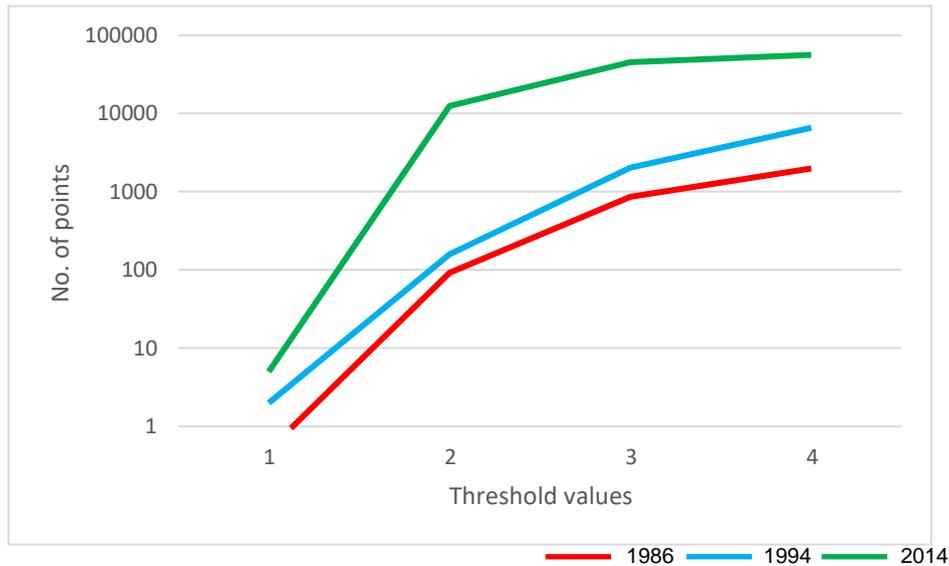


Figure 11. Influence of the threshold parameter values on the number of detected points

Figure 11 shows the relationship between the reprojection error threshold and the number of detected points. It might be stated that the number of detected points increases with the increase of the threshold values, and the shape of the graphs is similar for 2014, 1994 and 1986 periods. However, an increase in the value of the threshold parameter causes an increase in the number of detected points, with a higher reprojection error.

4.3. The tie point quality assessment of pair of images

To assess the quality of the pairwise photo matching, it was decided to analyse the reprojection error values for all pairs of images together. The following statistical indices were used: maximum deviation, minimum deviation, mean deviation, median deviation SMAD (Eq. 1) and RMSE deviation (Eq. 2).

$$SMAD = 1.4826 * median(|x_i - \delta|) \quad (1)$$

where: x_i is a value in the data set, and δ is the median of the data set.

$$RMSE = \sqrt{\frac{\sum(x_i - \mu)^2}{N}} \quad (2)$$

where: Σ means "sum of", x_i is a value in the data set, μ is the mean, and N is the number of data points.

Table 1 The tie point quality assessment for SIFT detector

Year	Coordinate	Max Deviation (pix)	Min Deviation (pix)	Mean deviation (pix)	Median deviation (pix)	Smad (pix)	RMSE Deviation (pix)	Number of points
1986	X	2.9	-2.9	1.1	0.0	1.5	1.4	1340
	Y	2.9	-2.9	1.1	0.0	1.5	1.3	
1994	X	3.0	-3.0	1.3	-0.1	1.7	1.5	3709
	Y	3.0	-3.0	1.3	0.0	1.8	1.5	
2014	X	3.0	-3.0	0.5	0.0	0.5	0.6	51993
	Y	3.0	-3.0	1.1	-0.1	1.3	1.3	

Table 1 contains the statistic of tie points reprojection errors on images blocks from 1986, 1994 and 2014. The maximum deviation of all years for the x and y coordinates was 3 pixels, and the minimum – 3 pixels. The mean deviation for all years' x and y coordinates



was equal 1 pixel, excluding the x coordinate for 2014, which was 0.5 pixels. The median value for both coordinates in all years fluctuates around 0 pixels. The smallest SMAD error value was obtained for 2014, and it was 0.5 pixels for the x coordinate, and for the y coordinate amounted to 1.3 pixels; for the remaining years, this value was 1.5 pixels for 1986 and about 2 pixels for 1994. The lowest RMSE Deviation value was for 2014 ($x = 0.6\text{pix}$, $y = 1.3\text{pix}$), and for 1986 and 1994, it was about 1.5 pixels (x , y coordinates). The last column of the Table 1 relates to the average number of points detected per one stereo pair (images with joint coverage).

5. Discussion - the final bundle adjustment assessment

The final and most important step in combining a block of aerial images is determining relative orientation elements in a bundle adjustment process. However, bundle adjustment allows for determining the orientation of all the images in a block while minimising reprojection and camera calibration parameters in the self-calibration procedure; in this investigation, the Authors decided to decline self-calibration. This was due to the use of metric archival aerial photographs, on which fiducial marks were measured, determining the internal orientation elements of the photographs. To assess the accuracy of the SfM process as a whole, and the bundle adjustment stage, in particular, it was decided to analyse the RMSE reprojection error values on the tie control points used for orientation and the tie check points for validation (Table 2). In addition, the distribution of reprojection errors on tie check points was analysed (Fig. 12). Ground truth data was prepared for comparison using the commercial software Agisoft Metashape. It was also decided to filter out the outlier points based on a threshold of 1 pix for the reprojection error values for the x and y axes.

Table 2. The results of the bundle adjustment process – relative orientation

Detector	RMSE reprojection [pix] 2014 (control points)	RMSE reprojection [pix] 2014 (check points)	RMSE reprojection [pix] 1994 (control points)	RMSE reprojection [pix] 1994 (check points)	RMSE reprojection [pix] 1986 (control points)	RMSE reprojection [pix] 1986 (check points)
SIFT	0.19	0.20	0.53	0.53	0.44	0.44
Agisoft	0.24	0.29	2.11	1.92	1.84	1.85
Agisoft after filtration	0.19	0.19	0.46	0.47	0.45	0.45

Table 2 shows the reprojection error values for the relative orientation of 1986, 1994 and 2014 photo blocks based on the tie points detected by the SIFT algorithm and the method implemented in the Agisoft Metashape(AM) with and without filtration. For all three periods, the reprojections error for the SIFT and AM detectors after filtration had approximately the same values, but differences are in hundredths. Due to the quality of images, especially digital images, the smallest RMSE values for control and check were obtained for 2014 and did not exceed 0.3 pixels. Significant differences are noticeable in the case of archive images. Comparing the results obtained using the extended Structure-from-Motion with those obtained using the approach commonly used in AM (without filtering), it can be seen that, using the method proposed by the authors, the accuracy of the BA process is about 4 times higher. The application of a non-standard approach to outlier filtering, based on the use of APIs and the analysis of reprojection error values, made it possible to obtain results similar to the proposed data processing method.

To complete the statistical analysis of the bundle adjustment process's quality, checking the distribution of reprojection error variances was necessary. For this purpose, it was



decided to present the results as boxplots (Fig. 12). Due to the geometric and radiometric quality, these distributions should be considered separately, i.e. the digital images from 2014 and the analogue archive images. For the block of digital images from 2014:

- the median value both for the X and Y coordinates for all methods was equal to 0;
- lower and upper quartile was ± 0.1 pix, and the lower and upper whisker ± 0.2 pix for all methods for X - coordinate,
- comparing the results of the extended SfM method with AM and AM with filtration, there is a significantly higher number of outliers,
- according to the bundle adjustment theory, the reprojection value for the y-coordinate should be approximately for aerial images. This condition is met by using a method based on the SIFT algorithm. For the AM results, a significant number of outliers can be seen.

For the analogue images of 1986 and 1994: a similar relationship can be observed:

- the median value both for the X and Y coordinates for all methods was equal to 0;
- lower and upper quartile was ± 0.3 pix, and the lower and upper whisker ± 1 pix for SIFT and AM filtered method for X - coordinate,
- lower and upper quartile was ± 0.8 pix and lower and upper whisker ± 2.5 pix for AM without filtration for X - coordinate,
- only for AM without filtration for X – coordinate outliers exist, and a significant number also exists for Y-coordinate (both Agisoft methods). This determined that the Agisoft Metashape method should not be treated as a robust solution.

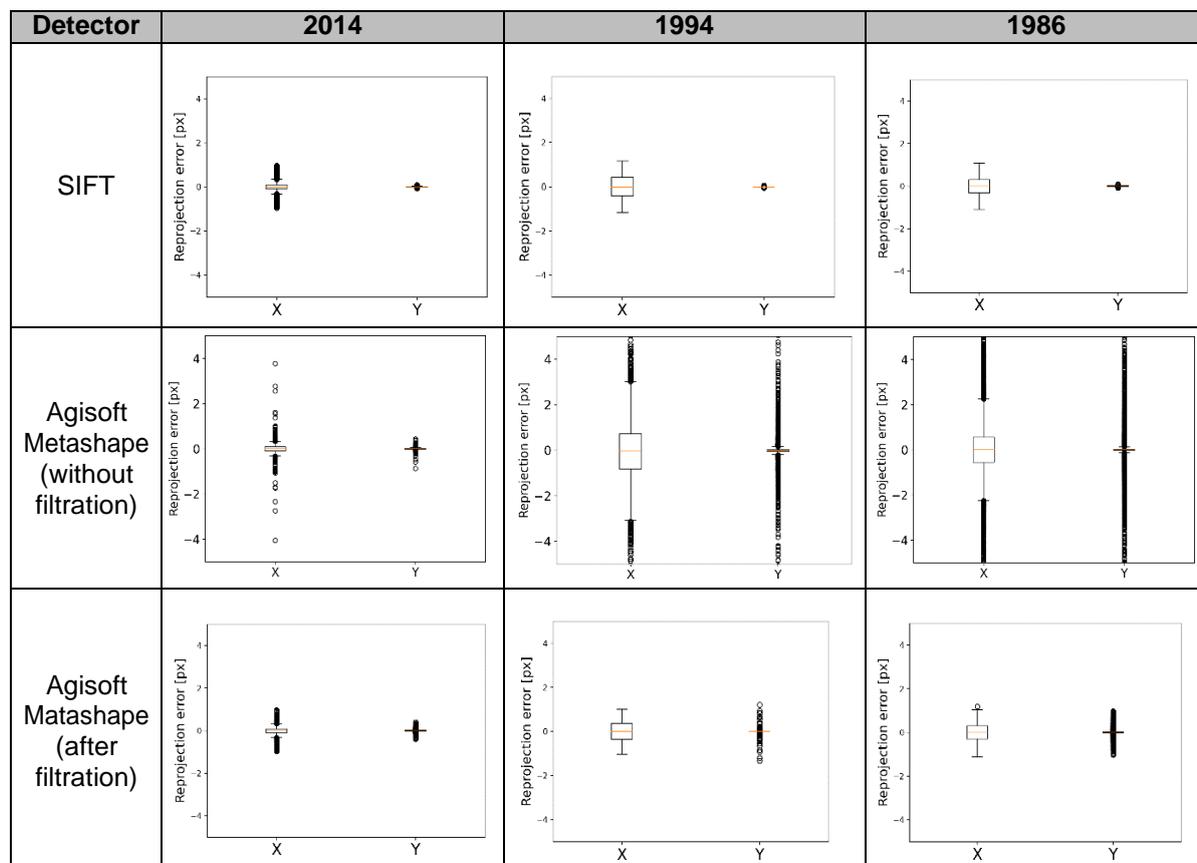


Figure 12. Assessment of the accuracy of the interior orientation of images blocks based on the mean errors on the tie check points



6. Conclusions

The proposed methodology for the orientation of archival aerial photographs based on the commonly used Structure-from-Motion approach met the expected objectives. The concept of the methodology using the SIFT machine vision algorithm allowed the determination and description of points, which, after the application of appropriate filtering techniques, constituted a set of unique, unambiguous, invariant and stable points of the photogrammetric matrix constituting the basis from the performance of interior orientation of blocks of images. A comparative analysis of the results based on SIFT algorithm with the data obtained using the algorithms implemented in the Agisoft Metashape software shows that the relative orientation reprojection RMSE error is 4 time better than the results obtained from the commercial software. The proposed method is more robust because the detected tie points allow for reduce to zero (without outliers) of the reprojection error for the y- coordinate, which is consistent with the theory of the bundle adjustment for aerial images.

Authors Contributions: J.M. and A.K.K. organised the idea's conceptualisation and methodology employed in this paper. Following that A.K.K carried out the experimental design. A.K.K worked on the data processing, J.M. and A.K.K carried out the original writing and draft preparation. J.M., A.K.K undertook the data analysis. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Farella, E.M.; Morelli, L.; Remondino, F.; Mills, J.P.; Haala, N.; Cromptvoets, J. The EurosdR Time Benchmark for Historical Aerial Images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* **2022**, *43*, 1175–1182, doi:10.5194/isprs-archives-XLIII-B2-2022-1175-2022.
2. Ratajczak, R.; Crispim-Junior, C.F.; Faure, E.; Fervers, B.; Tougne, L. Automatic Land Cover Reconstruction from Historical Aerial Images: An Evaluation of Features Extraction and Classification Algorithms. *IEEE Trans. Image Process.* **2019**, *28*, 3357–3371, doi:10.1109/TIP.2019.2896492.
3. Zawieska, D.; Markiewicz, J.; Kopiasz, J. Development of True Orthophotomaps of the Fortified Settlement at Biskupin, Site 4, Based on Archival Data. *Archaeol. Prospect.* **2019**, *26*, 333–360, doi:10.1002/arp.1748.
4. Nebiker, S.; Lack, N.; Deuber, M. Building Change Detection from Historical Aerial Photographs Using Dense Image Matching and Object-Based Image Analysis. *Remote Sens.* **2014**, *6*, 8310–8336, doi:10.3390/rs6098310.
5. Nocerino, E.; Menna, F.; Remondino, F. Multi-Temporal Analysis of Landscapes and Urban Areas. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *XXXIX-B4*, 85–90, doi:10.5194/isprsarchives-xxxix-b4-85-2012.
6. Redecker, A.P. Historical Aerial Photographs and Digital Photogrammetry for Impact Analyses on Derelict Land Sites in Human Settlement Areas. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. Vol. XXXVII. Part B8. Beijing 2008* **2008**, 5–10.
7. Poli, D.; Casarotto, C.; Strudl, M.; Bollmann, E.; Moe, K.; Legat, K. USE of HISTORICAL AERIAL IMAGES for 3D MODELLING of GLACIERS in the PROVINCE of TRENTO. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* **2020**, *43*, 1151–1158, doi:10.5194/isprs-archives-XLIII-B2-2020-1151-2020.
8. Tuytelaars, T.; Mikolajczyk, K. Local Invariant Feature Detectors: A Survey. *Found. Trends Comput. Graph. Vis.* **2007**, *3*, 177–280, doi:10.1561/06000000017.



9. Sevara, C.; Verhoeven, G.; Doneus, M.; Draganits, E. Surfaces from the Visual Past: Recovering High-Resolution Terrain Data from Historic Aerial Imagery for Multitemporal Landscape Analysis. *J. Archaeol. Method Theory* **2018**, *25*, 611–642, doi:10.1007/s10816-017-9348-9.
10. Fieber, K.D.; Mills, J.P.; Miller, P.E.; Clarke, L.; Ireland, L.; Fox, A.J. Rigorous 3D Change Determination in Antarctic Peninsula Glaciers from Stereo WorldView-2 and Archival Aerial Imagery. *Remote Sens. Environ.* **2018**, *205*, 18–31, doi:10.1016/j.rse.2017.10.042.
11. Jiayan Shen, Xiucheng Guo, Wenzong Zhou, Y.Z. and J.L. Feature Point Matching Method for Aerial Image Based on Recursive Diffusion Algorithm. *Symmetry Mdpi* **2021**, *13*(3), 407.
12. Kapnias, D.; Milenov, P.; Kay, S. *Guidelines for Best Practice and Quality Checking of Ortho Imagery*; 2008; ISBN 9789279109690.
13. Sauerbier, M.; Kunz, M.; Fluehler, M.; Remondino, F. PHOTOGRAMMETRIC RECONSTRUCTION OF ADOBE ARCHITECTURE AT TÚCUME , PERU Swiss Federal Institute of Technology. **2004**.
14. Micheletti, N.; Lane, S.N.; Chandler, J.H. Application of Archival Aerial Photogrammetry to Quantify Climate Forcing of Alpine Landscapes. *Photogramm. Rec.* **2015**, *30*, 143–165, doi:10.1111/phor.12099.
15. Bianco, S.; Ciocca, G.; Marelli, D. Evaluating the Performance of Structure from Motion Pipelines. *J. Imaging* **2018**, *4*, 1–18, doi:10.3390/jimaging4080098.
16. Saurer, O.; Fraundorfer, F.; Pollefeys, M. OmniTour: Semi-Automatic Generation of Interactive Virtual Tours from Omnidirectional Video. *Proc. 3DPVT* **2010**, 1–8.
17. Zakiah Noh, Mohd Shahrizal Sunar, Z.P. A Review on Augmented Reality for Virtual Heritage System. *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*; Springer **2009**.
18. Markiewicz, J.S. The Use of Computer Vision Algorithms for Automatic Orientation of Terrestrial Laser Scanning Data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* **2016**, *41*, 315–322, doi:10.5194/isprsarchives-XLI-B3-315-2016.
19. Hatzopoulos, J.N.; Stefanakis, D.; Georgopoulos, A.; Tapinaki, S.; Pantelis, V.; Liritzis, I. Use of Various Surveying Technologies to 3D Digital Mapping and Modelling of Cultural Heritage Structures for Maintenance and Restoration Purposes: The Tholos in Delphi, Greece. *Mediterr. Archaeol. Archaeom.* **2017**, *17*, 311–336, doi:10.5281/zenodo.1048937.
20. Tobiasz, A.; Markiewicz, J.; Lapinski, S.; Nickel, J.; Kot, P.; Muradov, M. Review of Methods for Documentation, Management, and Sustainability of Cultural Heritage. Case Study: Museum of King Jan III's Palace at Wilanów. *Sustain.* **2019**, *11*, doi:10.3390/su11247046.
21. Moussa, W.; Abdel-Wahab, M.; Fritsch, D. An Automatic Procedure for Combining Digital Images and Laser Scanner Data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, XXXIX-B5, 229–234, doi:10.5194/isprsarchives-xxxix-b5-229-2012.
22. Moussa, W. *Integration of Digital Photogrammetry and Terrestrial Laser Scanning for Cultural Heritage Data Recording*; **2014**; ISBN 9783769651379.
23. Stephens, C.H. & M. A COMBINED CORNER AND EDGE DETECTOR.
24. Low, D.G. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput.*



- Vis.* **2004**, 91–110.
25. Bay, H.; Ess, A.; Tuytelaars, T.; Van Gool, L. Speeded-Up Robust Features (SURF). *Comput. Vis. Image Underst.* **2008**, *110*, 346–359, doi:10.1016/j.cviu.2007.09.014.
 26. Bay, H.; Tuytelaars, T.; Van Gool, L. Speeded up Robust Features. *Lect. Comput. Vis. – ECCV 2006* **2006**.
 27. Nurminen, K.; Litkey, P.; Honkavaara, E.; Vastaranta, M.; Holopainen, M.; Lyytikäinen-Saarenmaa, P.; Kantola, T.; Lyytikäinen, M. Automation Aspects for the Georeferencing of Photogrammetric Aerial Image Archives in Forested Scenes. *Remote Sens.* **2015**, *7*, 1565–1593, doi:10.3390/rs70201565.
 28. Guoshen Yu, J.-M.M. ASIFT: An Algorithm for Fully Affine Invariant Comparison.
 29. Lowe, D.G. Object Recognition from Local Scale-Invariant Features. *Proc. Int. Conf. Comput. Vision, Corfu* **1999**.
 30. Abratkiewicz, K.; Gromek, D.; Samczynski, P.; Markiewicz, J.; Ostrowski, W. The Concept of Applying a SIFT Algorithm and Orthophotomaps in SAR-Based Augmented Integrity Navigation Systems. *Proc. Int. Radar Symp.* **2019**, *2019-June*, 1–12, doi:10.23919/IRS.2019.8768136.
 31. Tola, E.; Lepetit, V.; Fua, P. DAISY: An Efficient Dense Descriptor Applied to Wide-Baseline Stereo. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 815–830, doi:10.1109/TPAMI.2009.77.
 32. Brown, M.; Lowe, D. Invariant Features from Interest Point Groups. **2013**, 23.1-23.10, doi:10.5244/c.16.23.
 33. Chen, Y.; Chen, Y.; Wang, G. Bundle Adjustment Revisited. **2019**.