Measurement Science and
Technology

PAPER • **OPEN ACCESS**

Autonomous image background removal for accurate and efficient close-range photogrammetry

To cite this article: J Eastwood *et al* 2023 *Meas. Sci. Technol.* **34** 035404

OPEN ACCESS
IOP Publishing

Measurement Science and Technology

Meas. Sci. Technol. **34** (2023) 035404 (10pp)

https://doi.org/10.1088/1361-6501/aca497

Autonomous image background removal for accurate and efficient close-range photogrammetry

J Eastwood[1,*], R K Leach[1,2] and S Piano[1]

[1] Manufacturing Metrology Team, Faculty of Engineering, University of Nottingham, Nottingham, United Kingdom
[2] Taraz Metrology Ltd, Nottingham, United Kingdom

E-mail: JoeE@stwood.co.uk

Received 27 July 2022, revised 7 November 2022
Accepted for publication 21 November 2022
Published 7 December 2022

Abstract

Close-range photogrammetry can be used to reconstruct dense point clouds of an object with very high surface coverage, making it useful for manufacturing metrology tasks such as part inspection and validation. However, compared to competing techniques, data processing times can be slow. In this paper we present a method to autonomously remove the background from the images within a photogrammetric dataset. We show that using masked images directly in the reconstruction results in much lower data processing times, with lower memory utilisation. Furthermore, we show that the point density on the object surface is increased while the number of superfluous background points is reduced. Finally, a set of reconstruction results are compared to a set of tactile coordinate measurements. Reconstructions with the background removed are shown to have a standard deviation in the point to mesh distance of up to 30 µm lower than if the background is not removed. This improvement in standard deviation is likely due to the static background, relative to the object on the rotation stage, causing triangulation errors when points are detected and matched on this background data. The proposed approach is shown to be robust over several example artefacts and can, therefore, be implemented to improve the measurement efficiency and measurement results of photogrammetry coordinate measurement systems.

Keywords: photogrammetry, optical coordinate metrology, 3D imaging, image processing, automation

(Some figures may appear in colour only in the online journal)

1. Introduction

Close-range photogrammetry, referred to here as photogrammetry, is a method of optical coordinate metrology which reconstructs an object's surface by detecting, matching and triangulating features across a set of overlapping photographic images [1]. Optical methods of coordinate metrology are growing in popularity due to their non-contact probing methods and high surface coverage. Photogrammetry has advantages over competing optical methods, such as laser scanning [2] and digital fringe projection [3], in that it requires relatively inexpensive hardware and allows for inherently multi-view measurement. However, these benefits come at the cost of slow reconstruction times.

In manufacturing metrology applications, we are usually only concerned with measuring points on a given part,

[*] Author to whom any correspondence should be addressed.

so any background points reconstructed are not useful to the measurement task and can be removed. In this paper we present a method for improving the efficiency of photogrammetric reconstructions by removing superfluous background pixels from the captured images used in reconstruction. We show not only that this improves the speed of reconstruction and reduces the number of background points, but also improves the measurement results' agreement with traceable tactile coordinate measurement machine (CMM) data.

There has been some research effort into decreasing the processing time of photogrammetric reconstructions. Eastwood *et al* [4] proposed a view-planning optimisation approach to minimise the number of imaging positions while maintaining reconstruction quality. Because each feature in an image is a potential match with each feature in every other image, feature matching algorithms scale poorly with the number of images in the measurement data; therefore, by minimising the number of images taken computational expense can be reduced. However, reducing the number of images in the measurement data will eventually impact the measurement result even if captured from the most optimal positions. For example, computer vision tasks often only use two images to reconstruct a scene at high speeds, but the accuracy requirements of this task are far lower than needed for precise measurement [5, 6]. Therefore, it is desirable to increase the per-image processing efficiency.

Removing the background from images has the potential to improve computational time as it reduces the number of features present in the image which will be extracted and then matched. Most current approaches to background removal rely on manual masking of images by the user [7, 8]. If the background is static relative to the camera, such as in measurement systems using a rotation stage, this can be exploited to remove the background in an automated way (see [9]). Furthermore, as static background feature matches can cause the reconstruction algorithms to fail, the removal of these features has the additional benefit of making the reconstruction more stable. Because of these benefits, commercial photogrammetry software packages can accept masks as part of their reconstruction algorithms. OpenMVG [10], an open-source structure-from-motion library, can use binary masks to determine which features are included in the reconstruction. However, generating these masks is left entirely up to the user. Agisoft Metashape [11], a commercial photogrammetry package, can generate image masks but requires the user to manually outline the object in a sub-set of the images used for reconstruction.

In this work we present a method for automated masking of the object from the background of the measurement system and propose an algorithm that performs well across a range of object geometries and materials. We show that passing these masked images to the photogrammetric reconstruction algorithms directly decreases processing time, memory usage and the number of reconstructed background points. We further show that, when background masking is applied, the number of object surface points reconstructed increases and that the measurement result agrees more closely with a measurement from a calibrated tactile CMM over repeated measurements.

2. Background removal technique

For the background removal algorithm to be general across any object which can be placed in the measurement volume, we make one major assumption about the measurement system: the proposed approach assumes that the background of the scene never contains any closed contours regardless of the measurement head position. Figure 1 shows how the Taraz P2 [12] measurement system in this paper meets this requirement.

As can be seen in figure 1, all the background contours are open and thus the largest closed contour in the image must represent the boundary of the object. As such, the problem of background masking can be reduced to the extraction of the largest closed contour in the image.

2.1. Algorithm detail

Python bindings for the OpenCV image processing library [13] was used to perform all image processing operations and file input/output (IO) in the implementation presented here. The steps used to robustly extract the largest closed contour from an image can be split into three stages; preprocessing, edge extraction, and contour selection. The details of each stage of the background removal pipeline are summarised in the diagram shown in figure 2.

First, the image is converted to grayscale as the required contours can be extracted from image intensity information alone. Next, the mean pixel intensity is calculated across the entire image; individual pixel values are then scaled linearly so that the average intensity across the image is equal to 52. This process is to correct any changes due to material differences between different objects. Fifty-two was used as it was the mean intensity recorded across a range of artefact measurements.

During edge detection, high spatial frequency information such as the rough surface texture of an additively manufactured (AM) part can negatively impact edge detection and contour extraction from the image. To prevent this effect, a denoising scheme is applied to the image. Recent publications have suggested multiple approaches for image noise reduction. Popular approaches include wavelet transforms [14–16], non-local methods [17, 18], and deep learning (DL) [19]. Due to our desire for simplicity, computational efficiency, generalisation and robustness an edge preserving smoothing filter was selected as the best approach. To this end, a bilateral filter is applied to the image [20], the bilateral filter was chosen as it can smooth out high spatial frequency information while preserving edges. A bilateral filter is composed of two Gaussian convolutions, one spatial and one intensity filter (referred to as the range filter). The spatial filter f acts as a standard Gaussian blur parameterised by the bilateral kernel size k and the spatial

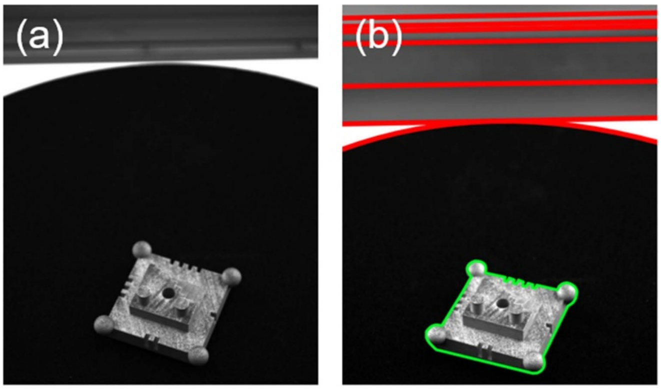

Figure 1. An example image (a) from the Taraz P2 measurement system, (b) with extracted contours highlighted, open contours shown in red the largest closed contour shown in green.

standard deviation σ_s. The range filter g acts over the space of pixel intensities and is parameterised by k and the intensity standard deviation σ_r. The result of a bilateral filter on image $\mathbf{I}$ is calculated by,

$$\mathbf{I}' = \mathbf{I} * (f(k,\sigma_s) \times g(k,\sigma_r)), \quad (1)$$

where $*$ is the convolution operator. For small values of σ_r this filter results in pixels which are spatially close to the current pixel but remote in intensity contributing much less to the output than pixels which are close both spatially and in intensity. Therefore, pixels which lie on opposite sides of a boundary do not contribute highly to the smoothing operation compared to pixels on the same side of this boundary. As $\sigma_r \to 255$ for 8-bit images, the bilateral filter acts like a Gaussian blur. In this case, through experimentation, the filter values were set to $k = 25$ pixels, $\sigma_r = 25$, $\sigma_s = 150$. These values result in large Gaussian blurring on faces but strong edge preservation. Figure 3 shows the effect of this filter on an example image and the impact on the performance of Canny edge detection [21].

Once the image has been filtered, Canny edge detection is applied (for full details see [15]). In brief, image gradients are extracted, areas of high gradient are taken to be edges, these edges are then thinned using non-maximum suppression in the direction of the gradient at that location, and edges are further refined using a hysteresis pruning algorithm. Image gradients can be found efficiently by decomposing the Sobel operator into four 1D convolutions given by,

$$\mathbf{G}_x = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} * \left(\begin{bmatrix} 1 & 0 & -1 \end{bmatrix} * \mathbf{I}\right), \quad (2)$$

$$\mathbf{G}_y = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} * \left(\begin{bmatrix} 1 & 2 & 1 \end{bmatrix} * \mathbf{I}\right), \quad (3)$$

where $\mathbf{G}_x$ and $\mathbf{G}_y$ represent the horizontal and vertical components of the gradient respectively. From these components, the gradient magnitude $\mathbf{G}$ and direction Θ can be calculated from,

$$\mathbf{G} = \sqrt{\mathbf{G}_x^2 + \mathbf{G}_y^2}, \quad (4)$$

and,

$$\Theta = \mathrm{atan2}\left(\mathbf{G}_y, \mathbf{G}_\mathrm{x}\right). \quad (5)$$

Each pixel is set to the value of the local image gradient at that image coordinate. Then, every pixel is compared to its two neighbours in the direction of the local image gradient. If the pixel is not a maximum compared to these neighbours, it is set to zero. This process can be iterated until only thin edges remain. Finally, hysteresis pruning is applied to the remaining gradient values to produce the final detected edges. Two threshold gradient values are set, one high and one low. If the image gradient at a given pixel is larger than the high threshold, it is considered an edge pixel and is left untouched. If the image gradient at a given pixel is lower than the low threshold, it is not considered an edge pixel and is set to zero. If an image pixel lies between the two thresholds, it is considered an edge pixel only if at least one of its eight neighbours is also considered an edge pixel. Setting the low and high thresholds is normally done by the user—in this case, because we desire automation as well as the algorithm to operate on any object, a slightly modified version of the Canny edge detector is used called AutoCanny [22]. In this case, the high and low thresholds are set based on the median image intensity $\widetilde{\mathbf{I}}$ as,

$$t_\mathrm{high} = \min\left(\left[255, (1+0.33)\cdot\widetilde{\mathbf{I}}\right]\right), \quad (6)$$

$$t_\mathrm{low} = \max\left(\left[0, (1-0.33)\cdot\widetilde{\mathbf{I}}\right]\right). \quad (7)$$

AutoCanny was found to perform well over a set of artefacts of many shapes and materials. The detected edges are dilated with a 25 square pixel kernel and then blurred. The blurring helps connect any discontinuities in the extracted edges, which can then be eroded to re-thin the new connected edges. The final edge image is passed to cv2::findContours which implements a pixel following algorithm to extract continuous contours from the detected edges. Finally, the contours are sorted by the area they inscribe, and the contour of maximum area is selected. This boundary is then used to mask the background from the image by setting each pixel outside its area to zero, and each pixel within its area equal to its value in the original colour image.

Figure 4 summarises each stage of the background removal pipeline and figure 5 shows the results when applying this method on a range of artefacts.

2.2. Experimental procedure

In each experimental test, images were collected using the Taraz Metrology P2 system [12]. Every scan was comprised of

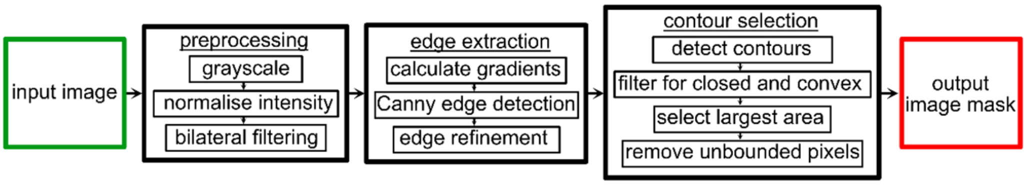

Figure 2. The proposed background removal pipeline.

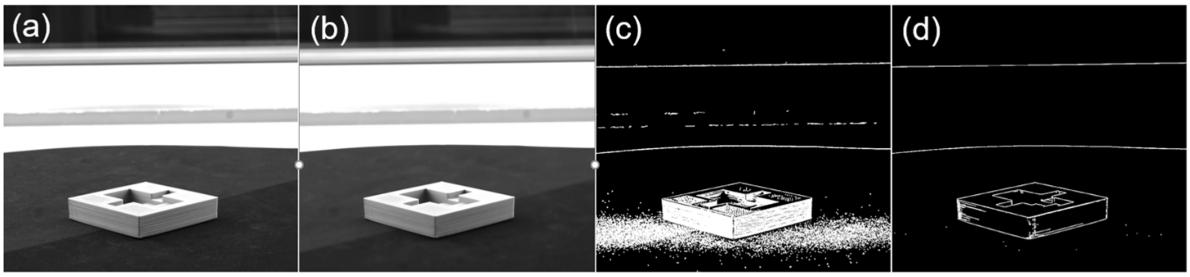

Figure 3. Showing part of (a) an input image, (b) a filtered image, (c) edge detection on (a), (d) edge detection on (b).

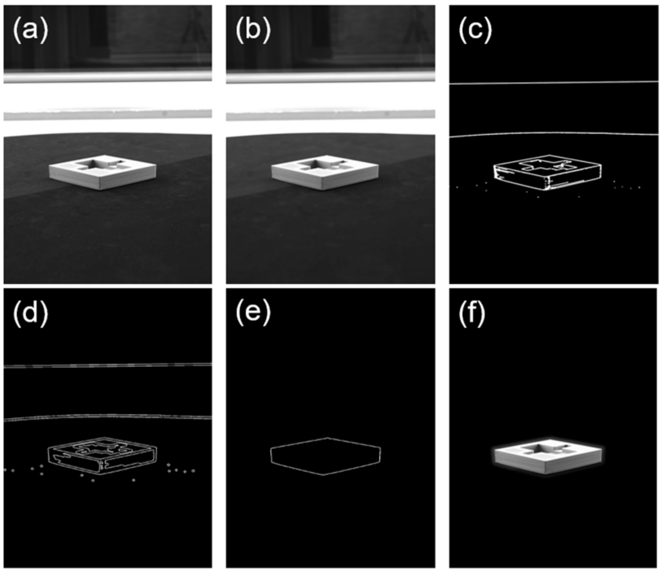

Figure 4. Background removal pipeline (a) captured image, (b) filtered image, (c) detected edges, (d) extracted contours, (e) selected maximum contour, (f) masked image.

60 pairs of stereo images captured in two equally spaced rings of 30 positions. Figure 6 shows an example of a reconstructed scene showing the 120 individual imaging positions.

Two aretefacts, shown in figure 7, were used to test the impact of applying background removal on photogrammetric reconstruction. Artefact 1 shown in figure 7(a) was measured once, while the measurement of artefact 2, shown in figure 7(b), was repeated three times to assess the variance and repeatability of the method. During each reconstruction, processing time and memory utilisation were recorded. Finally, each reconstruction was compared to a set of CMM measurements.

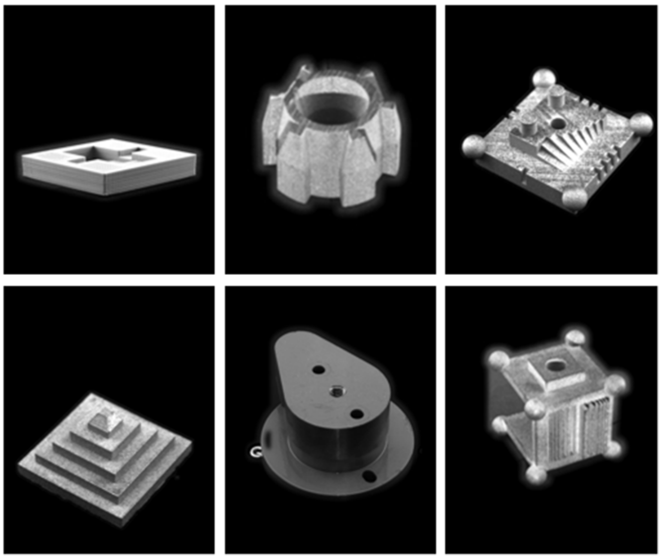

Figure 5. Example results of background removal across a range of artefacts.

3. Results

Reconstruction was performed using Agisoft Metashape using 'high' camera alignment settings and 'medium' dense reconstruction settings [11]. Every measured point cloud was then scaled using the stereo baseline distance between the optical centres of the cameras in the measurement head. Any background points were then manually removed to assess the ratio of object to background points in the scene.

3.1. Impact on reconstruction efficiency and point density

Figure 8 shows the reconstructed dense point clouds of artefact 1, both with and without image masking applied.

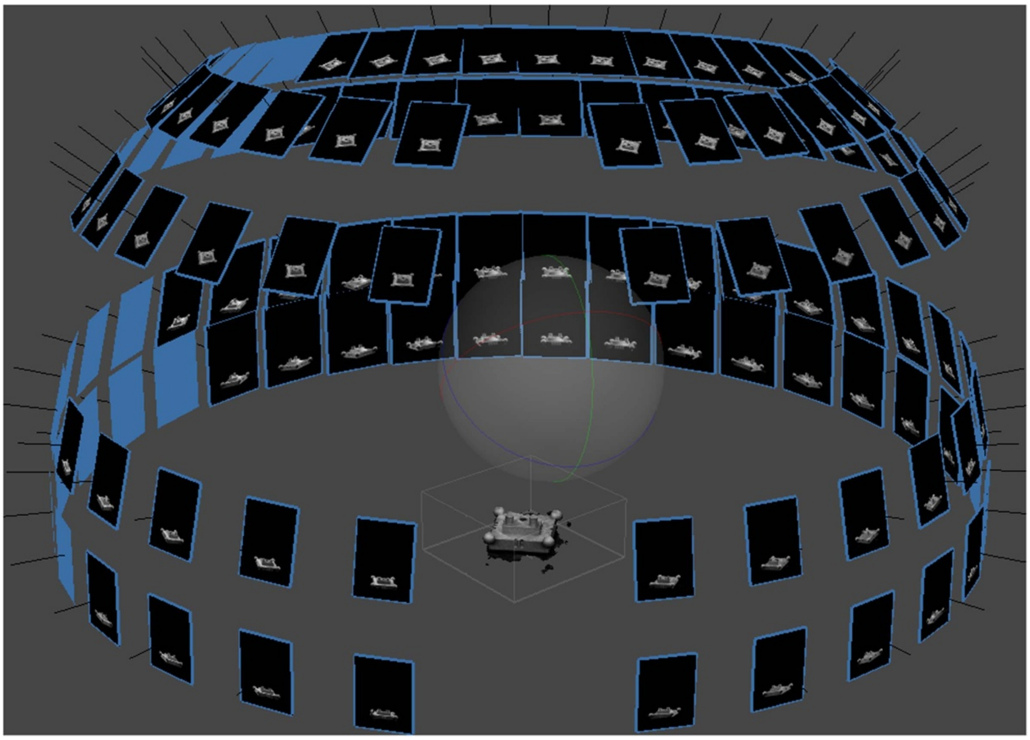

Figure 6. Imaging positions used for every scan.

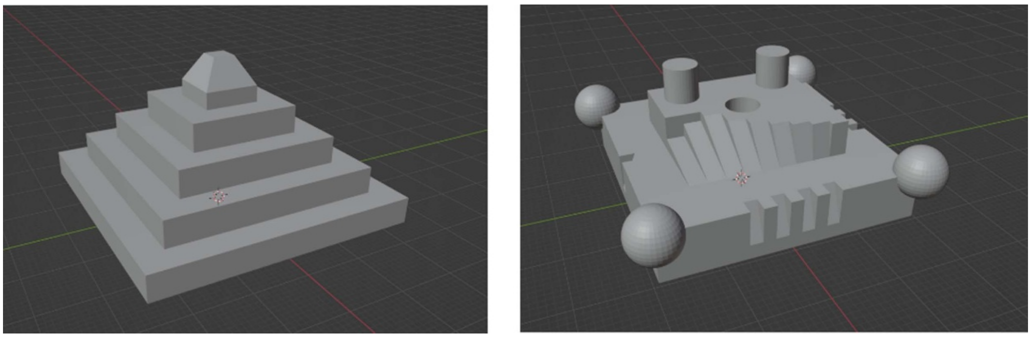

Figure 7. Two test artefacts used (a) artefact 1 (b) artefact 2.

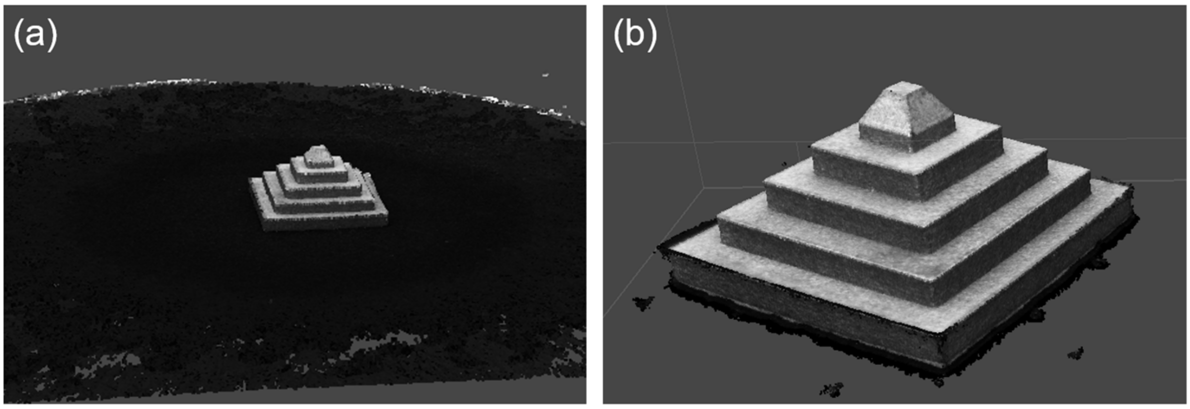

Figure 8. Reconstructions of the artefact 1 (a) without background masking (b) with background masking.

Table 1. Comparison of reconstruction metrics for artefact 1.

	Without background masking	With background masking	Difference
Densification time (s)	1398	903	−495
Object points	215 467	1140 950	+925 483
Background points	4302 679	71 500	−4231 179

It is clear that a large number of background points, shown in black, were produced in figure 8(a) but that this number was significantly reduced by using the background masking, as shown in figure 8(b). Table 1 summarises the impact on performance of the proposed method when compared with traditional reconstruction.

As can be seen in table 1, applying the background masks reduced overall processing time by approximately eight minutes. The reason for this speed increase is clear as when the masks are applied around four million fewer overall points are reconstructed. Of these four million missing points, the majority are from the background, which would be removed in further data analysis steps anyway. The number of points on the surface of the pyramid object itself has increased by almost a million points.

Figure 9 shows a comparison of reconstructions of artefact 2 with and without background masking.

Again, figure 9(a) shows a large number of superfluous background points were reconstructed compared to when masking was applied in figure 9(b). Tables 2–4 show a detailed breakdown of the processing time, memory usage and points reconstructed, averaged over the three repeated measurements of the Tomas artefact.

The results on artefact 2 agree with what was shown for artefact 1. The overall processing time, the number of background points reconstructed and memory usage are reduced, but the number of points reconstructed on the object surface are increased.

3.2. Comparison to tactile CMM

The reconstructed dense point clouds were then triangulated into a mesh. When the background points had been removed, an iterative closest point algorithm was employed to register the meshes to data taken from a tactile CMM. Specifically, a Mitutoyo Crysta Apex S 7106 [23] with a 1 mm probe tip diameter was used in scanning mode. The CMM maximum permissible error ($E_{0,\mathrm{MPE}}$) is manufacturer calibrated as ($1.7 + 3.0\ l\ \mathrm{M}^{-1}$) µm. Each CMM artefact measurement was repeated three times. The point to mesh (PTM) distances could then be calculated between the CMM and photogrammetry data. The results for artefact 1 are shown in figure 10.

As can be seen in figure 10, the unmasked reconstruction contains a higher number of outlying points, shown in red. Figure 11 shows a comparison between the histograms of PTM distances across the two measurement comparisons using 400 bins.

As can be seen in figure 11, when background masking is applied the PTM distance spread is reduced. Fitting a Gaussian to the distribution in figure 11(a) yields a standard deviation of 85 µm, while fitting a Gaussian to the distribution in figure 11(b) yields a standard deviation of 70 µm. In addition to the lower deviation in the PTM distances, there are also many fewer outliers when masking is applied; this can be seen in the spike on the far left of the distribution in figure 11(a) which represents PTM distances larger than 500 µm.

Figure 12 shows one of the three repeat measurements of artefact 2. Figure 13 shows the combined histograms over all three repeat measurements for both masked and unmasked reconstructions.

In figure 13(a) the combined standard deviation of the PTM distances over three repeat measurements was 93 µm. When background masking was applied to the measurement data in figure 13(b) the standard deviation reduced to 70 µm with masking, representing a decrease of 23 µm. Additionally, the number of outlying points with PTM distances greater than 500 µm also, was again, reduced by applying background masking. Figure 12(a) shows that many of the outlying points are concentrated around the more complex features such as the cylinders, spheres and recesses which were reconstructed more faithfully in figure 12(b), as shown by the reduced number of red points.

4. Discussion

As can be seen in figures 9–12 the agreement with CMM is improved when background masking is applied. This is likely mainly due to the removal of the static background from the image. As was discussed in section 1, the static portion of the background, present due to the use of a rotation stage, creates a set of points which are static relative to the camera while the rest of the points have undergone some relative motion. This means that when the bundle adjustment algorithm attempts to globally optimise the camera positions and point locations, the triangulation of the object points can be degraded. Figure 14 shows a comparison of the cylindrical features on artefact 2. It can be seen that the reconstruction quality of these features improves when background masking is used. Because the cylinders on artefact 2 are prominent features and intersect with the static portion of the background in many views, this reinforces the hypothesis that it is the removal of the static background that leads to improved measurement results.

A further possible reason for the improved measurement result when masking is applied are the patch-based algorithms used for densification. These dense reconstruction algorithms operate by growing and refining rectangular patches of points. In the case of the masked data, these patches can only be produced on the object surface, whereas in the unmasked case many are created in the background data.

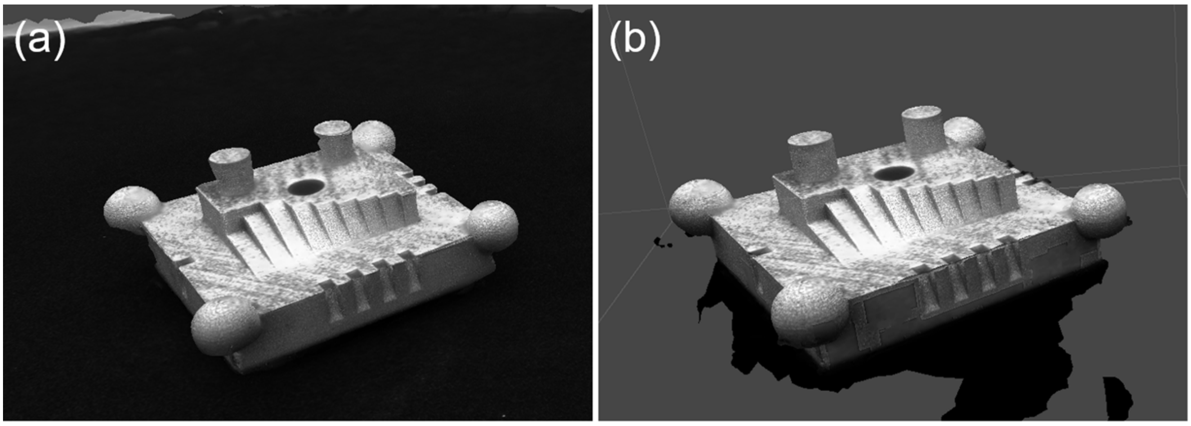

Figure 9. Dense reconstructions of artefact 2 (a) without background masking (b) with background masking.

Table 2. Comparison of time expended at each reconstruction step, averaged across three reconstructions of artefact 2.

	Time (s)				
	Image processing	Feature matching	Camera alignment	Densification	Total
No Mask	0.0	157.7	81.0	2459.0	2697.7
Mask	180.2	223.7	52.7	1143.3	1599.7
Difference	180.2	66.0	−28.3	−1315.7	−1098.0

Table 3. Comparison of memory usage at each reconstruction step, averaged across three reconstructions of artefact 2.

	Memory usage (GB)				
	Image processing	Feature matching	Camera alignment	Densification	Total
No Mask	0.000	0.365	0.067	2.999	3.430
Mask	0.471	0.536	0.099	0.923	2.028
Difference	0.471	0.171	0.032	−2.075	−1.402

Table 4. Comparison of points reconstructed, averaged across three reconstructions of artefact 2.

	Points			
	Sparse points	Dense points	Object points	Background points
No Mask	58 373	3600 644	282 721	3317 923
Mask	71 436	377 862	312 560	65 301
Difference	13 062	−3222 783	29 839	−3252 622

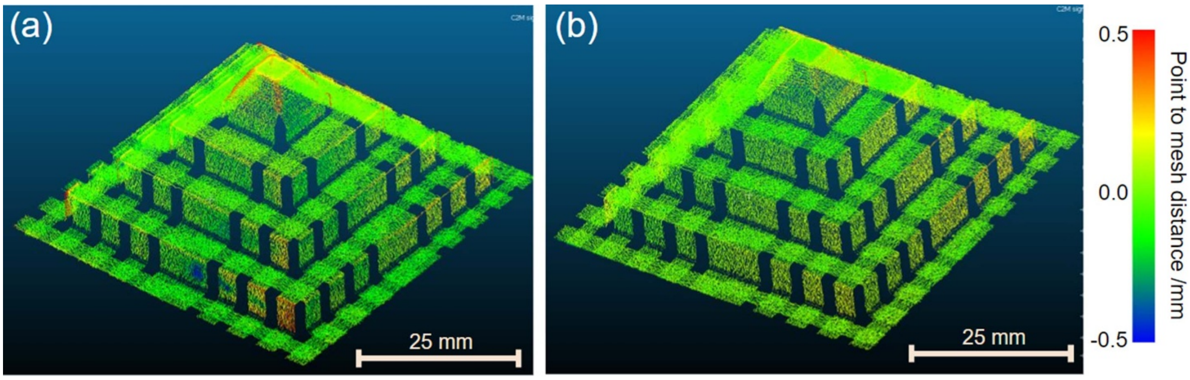

Figure 10. Comparison of PTM distances for artefact 1 (a) without background masking (b) with background masking.

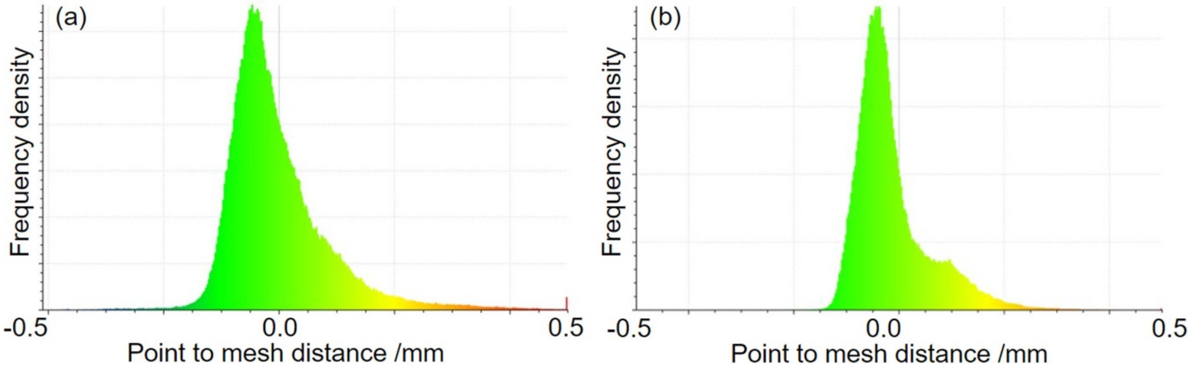

Figure 11. Comparison of the distribution of PTM distances for artefact 1 (a) without background masking (b) with background masking.

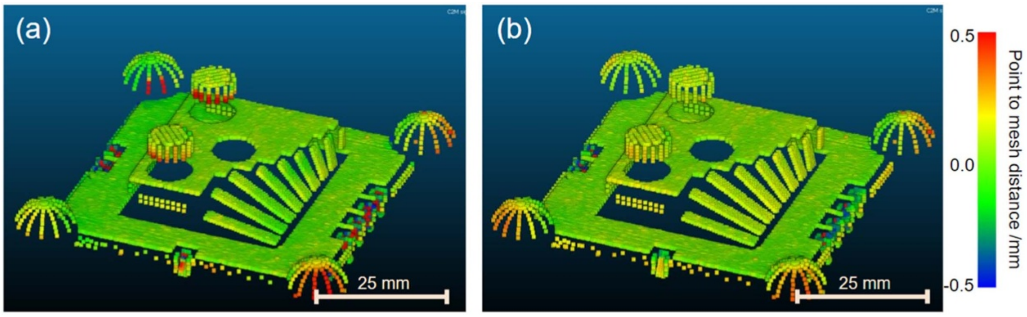

Figure 12. Comparison of PTM distances for artefact 2 (a) without background masking (b) with background masking.

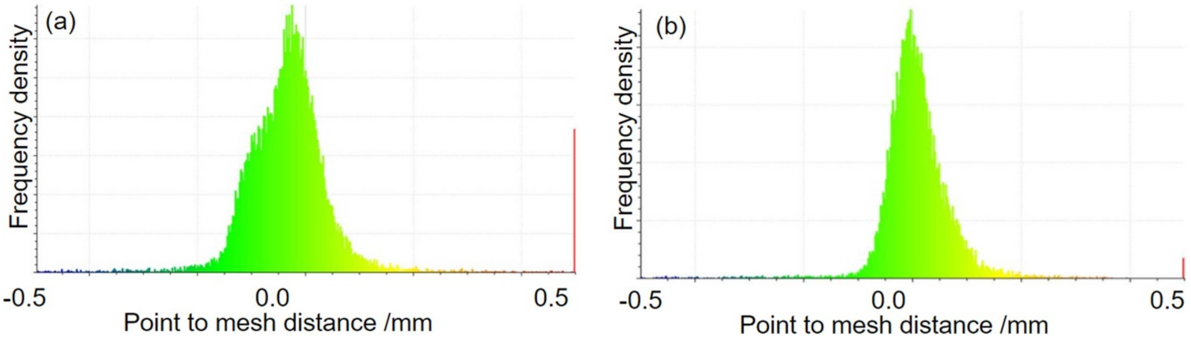

Figure 13. Comparison of the distribution of PTM distances for artefact 2 (a) without background masking (b) with background masking.

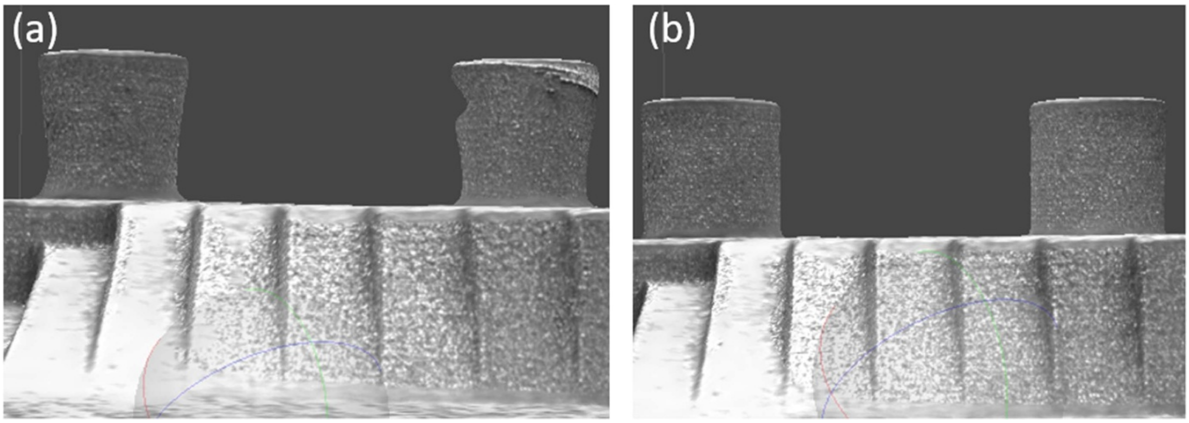

Figure 14. Comparison of the reconstruction of cylindrical features on artefact 2 (a) without background masking (b) with background masking.

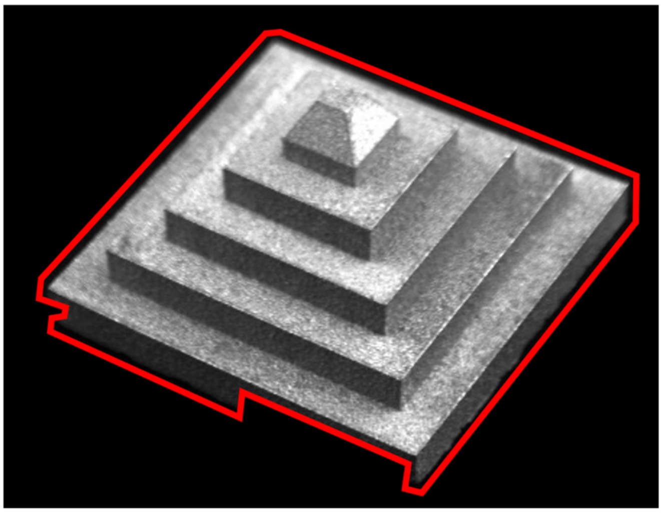

Figure 15. Example failure case of the background removal, masking contour shown in red.

The histograms in figures 10 and 12 show that the distribution of all the PTM distance data has some skew. This is likely due to a small error in the scale applied to the photogrammetric portion of the data. This scale is based on the stereo baseline, the distance between the optical centres of the stereo cameras. This is difficult to measure directly so was established through reconstructing a ball bar of calibrated size. However, the tighter spread of the data when background masking was applied, as shown by the lower standard deviations, is still strong evidence of greater agreement with CMM measurements despite this potential scaling error.

Although the background removal approach is relatively robust to a range of objects, as was shown in figure 5, it is not perfect and there are occasional viewing angles which cause the masking process to erroneously remove some of the object data. Figure 15 shows an example of this from the data used to reconstruct artefact 1.

The erroneous masking shown in figure 15 occurs when the bottom edge of the object is in shadow due to the lighting conditions present within the measurement system. The shadow effectively blurs the boundary of the object and causes the edge detection part of the pipeline to fail. However, because this only occurs from very few viewing angles, enough of the surrounding views detect and triangulate points in the area of missing data that the result on the final measurement result is minimal. These erroneous cases could be prevented entirely by changing the lighting conditions in the measurement volume to be as diffuse and uniform as possible.

It is worth noting that the exact time taken to mask each image (approximately 1.5 s per image) is largely dependent on the implementation of the presented algorithm. It is likely that an optimised and compiled version of the algorithm could operate much faster than the Python implementation used here, especially with the many file IO operations required. There is also an obvious hardware dependence; in this case all image processing and reconstruction operations were performed on the same PC with an Intel Xeon W 2123 CPU, 32G GB of RAM.

Recent research has explored the use of DL methods for both edge extraction (see review [24]), and end-to-end background removal (see review [25]). However, many of these methods as they are currently implemented are either inaccurate such as extracting only bounding boxes [26] or are developed for specific applications and as such would not generalise well across any possible measurement artefact [27]. Furthermore, as our method has been shown to be effective with only traditional methods we avoid the computational overhead required to train a DL model. Avoiding ‘black box’ style neural networks also makes our results simple to interpret and understand.

5. Conclusions

In this paper we have proposed an image processing pipeline for the removal of background pixels from images taken within a photogrammetric measurement system. This pipeline is dependent on there being no closed contours in the background portion of any images captured in the measurement data. The background masking problem can, therefore, be reduced to the extraction of the closed contour of largest area within the image. It was shown that this approach can effectively remove the background from a large range of example objects.

To test the impact of using masked images directly within photogrammetry measurements, two test artefacts were reconstructed both with and without background masking applied to the input images. It was shown in both cases that applying imaging masking reduced reconstruction times and memory usage, increased the number and density of surface points reconstructed, and dramatically reduced the number of superfluous background points reconstructed.

The impact on the measurement result was investigated by comparing them to the measurement data gathered through repeat tactile measurement using a CMM. It was found that applying background masking lowers the number of outlying points reconstructed and reduces the standard deviation in the PTM distances when the photogrammetry and CMM data are registered together. This improvement in measurement agreement with CMM is likely due to the static background degrading the triangulation quality of the points when background masking is not applied.

6. Future work

In order to get the most efficient and optimal measurement strategy, it would be advantageous to integrate the approach shown here with the view-planning optimisation approach given in Zhang *et al* [28]. However, initial testing has shown that directly applying masking to the optimised imaging positions results in poor measurement results. Therefore, a modified version of the view optimisation algorithm will be developed to take the reduced information within each image into account.

Data availability statement

The data generated and/or analysed during the current study are not publicly available for legal/ethical reasons but are available from the corresponding author on reasonable request

Acknowledgments

This work was supported by the EPSRC (Grant Nos. EP/M008983/1 and EP/L016567/1) and Taraz Metrology Ltd. The authors also thank Dr Tomáš Koutecky from Brno University of Technology for the design of artefact 2 and Nathan Roberts and Dr Mohammed Isa the University of Nottingham for conducting the CMM measurements.

ORCID iD

J Eastwood https://orcid.org/0000-0003-4485-940X

References

[1] Luhmann T, Robson S, Kyle S and Harley I 2007 *Close Range Photogrammetry* (New York: Wiley)

[2] Isa M A, Piano S and Leach R K 2020 Laser triangulation *Advances in Optical Form and Coordinate Measurement R K* Leach (Bristol: IOP Publishing)

[3] Zhang S 2016 *High-Speed 3D Imaging with Digital Fringe Projection Techniques* (Boca Raton, FL: CRC Press)

[4] Eastwood J, Zhang H, Isa M, Sims-Waterhouse D, Leach R K and Piano S 2020 Smart photogrammetry for three-dimensional shape measurement *Proc. SPIE* 11352 1135A

[5] O'Riordan A, Newe T, Dooly G and Toal D 2018 Stereo vision sensing: review of existing systems *Proc. Int.Conf. on Sensing Technology (ICST)* pp 178–84

[6] Fan R, Wang L, Bocus M J and Pitas I 2020 Computer stereo vision for autonomous driving (arXiv:2012.03194)

[7] Woloszyk K, Bielski P M, Garbatov Y and Mikulski T 2021 Photogrammetry image-based approach for imperfect structure modelling and FE analysis *Ocean. Eng.* **223** 108665

[8] Rupnik E, Daakir M and Deseilligny M P 2017 MicMac–a free, open-source solution for photogrammetry *Open Geospat. Data Softw. Stand.* **2** 1–9

[9] Sathirasethawong C, Sun, Lambert A and Tahtali M 2019 Foreground object image masking via EPI and edge detection for photogrammetry with static background *Proc. Int. Conf. on Sensing Technology (ICST)* (*Lake Tahoe, USA*) pp 345–57

[10] Moulon P, Monasse P, Perrot R and Marlet R 2016 OpenMVG: open multiple view geometry *Proc. Reproducible Research in Pattern Recognition RRPR* (*Cancún, Mexico*) pp 60–74

[11] Agisoft 2019 Metashape: standard edition, v. 1.5.5 (available at: www.agisoft.com)

[12] Taraz Metrology Ltd 2022 *Measurement Systems* (available at: https://taraz-metrology.com/measurement-systems/)

[13] Bradski G 2000 The OpenCV library DDJ 2236121

[14] Donoho D L 1995 De-noising by soft-thresholding *IEEE Trans. Inf. Theory* **41** 613–27

[15] Starck J L, Candès E J and Donoho D L 2002 The curvelet transform for image denoising *IEEE Trans. Image Process.* **11** 670–84

[16] Zada S, Tounsi Y, Kumar M, Mendoza-Santoyo F and Nassim A 2019 Contribution study of monogenic wavelets transform to reduce speckle noise in digital speckle pattern interferometry *Opt. Eng.* **58** 034109

[17] Tounsi Y, Kumar M, Nassim A and Mendoza-Santoyo F 2018 Speckle noise reduction in digital speckle pattern interferometric fringes by nonlocal means and its related adaptive kernel-based methods *Appl. Opt.* **57** 7681–90

[18] Tounsi Y, Kumar M, Nassim A, Mendoza-Santoyo F and Matoba O 2019 Speckle denoising by variant nonlocal means methods *Appl. Opt.* **58** 7110–20

[19] Hao F, Tang C, Xu M and Lei Z 2019 Batch denoising of ESPI fringe patterns based on convolutional neural network *Appl. Opt.* **58** 3338–46

[20] Banterle F, Corsini M, Cignoni P and Scopigno R 2012 A low-memory, straightforward and fast bilateral filter through subsampling in spatial domain *Comput. Graph.* **31** 19–32

[21] Zhou P, Ye W, Xia Y and Wang Q 2011 An improved canny algorithm for edge detection *J. Comput. Inf. Syst.* **7** 1516–23

[22] Rosebrock A PyImageSearch 2015 *Zero-Parameter, Automatic Canny Edge Detection with Python and OpenCV* (available at: https://pyimagesearch.com/2015/04/06/zero-parameter-automatic-canny-edge-detection-with-python-and-opencv/)

[23] Mitutoyo 2018 MICROCORD CRYSTA-APEX S series *High-performance, low-price CNC coordinate measuring machine that meets global standards* p E16004 (available at: www.mitutoyo.co.uk/application/files/3415/5888/3870/E16004.pdf)

[24] Hou S M, Jia C L, Wanga Y B and Brown M 2021 A review of the edge detection technology *Proc. STAIQC* **1** 26–37

[25] Bouwmans T, Javed S, Sultana M and Jung S K 2019 Deep neural network concepts for background subtraction: a systematic review and comparative evaluation *Int. J. Neural Netw.* **117** 8–66

[26] Xiao Y, Tian Z, Yu J, Zhang Y, Liu S, Du S and Lan X 2020 A review of object detection based on deep learning *Multimedia Tools Appl.* **79** 23729–91

[27] Heidler K, Mou L, Baumhoer C, Dietz A and Zhu X X 2021 HED-UNet: combined segmentation and edge detection for monitoring the Antarctic coastline *IEEE Trans. Geosci. Remote Sens.* **60** 1–4

[28] Zhang H, Eastwood J, Isa M, Sims-Waterhouse D, Leach R K and Piano S 2021 Optimisation of camera positions for optical coordinate measurement based on visible point analysis *Precis. Eng.* **67** 178–88