Raycasting-Based Feature Matching for Image Registration on Complex Geometries

Adrian Kaune, Lennart Hinz, Eduard Reithmeier

Institute of Measurement and Automatic Control, An der Universität 1, 30823 Garbsen, Germany

mailto:adrian.kaune@imr.uni-hannover.de

Convolutional neural networks (CNNs) are becoming increasingly important for detecting defects in 2D image data. However, the output of such models lacks metric and spatial context. Therefore, the image data can be registered to a three-dimensional domain through a virtual reference model. A geometry-based feature matching approach for stereoscopic image pairs is proposed to improve robustness and precision of image registration.

1 Introduction

For many industrial inspection applications, a range of defect classes (such as cracks, chipping, deformations, etc.) need to be detected and quantified. Convolutional networks (CNNs) provide reliable results for image-based defect classification and segmentation at different scales. However, the defect information only refers to pixel-based quantities. One possibility for extracting metric information from the image-based segmented defects is to combine calibrated cameras and a virtual 3D reference model to a holistic model. This approach requires perfect alignment of the virtual reference model and does not account for any geometric deviations (e.g. due to wear). Therefore, misalignment similar to the one shown in the Fig. 1 can occur, leading to blurring in regions with overlapping and redundant image information. This study proposes a method to improve alignment by using the reference model to facilitate the identification of point correspondences in stereoscopic image pairs.

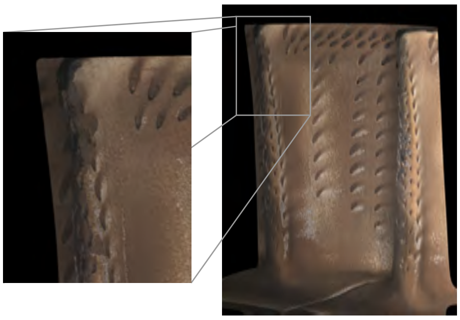

Fig. 1 *Image masked using the virtual reference model. The detail view highlights a gap between the object contour and the mask, indicating misalignment. A difference in radius is visible in the top-left corner, resulting from a geometric deviation.*

2 System Calibration

Images are acquired using calibrated cameras, with intrinsic parameters and lens distortion estimated using Zhang's calibration algorithm [1]. Extrinsic camera parameters are determined through stereo calibration. An initial estimation of the relative reference model position is performed using optical tracking markers. Fig. 2 shows a raycasting scene, where a single camera image is projected on the virtual reference model. This results in a colored point cloud.

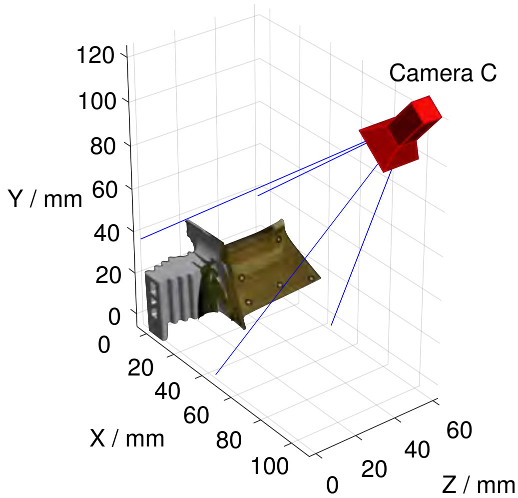

Fig. 2 *The camera rays are projected using the extrinsic and intrinsic camera parameters. The color values of the image pixels are allocated to the corresponding ray-mesh intersections.*

3 Image Registration

3.1 Stereo Camera Sparse Reconstruction

Keypoints are detected in each undistorted image of the stereoscopic pair using the SURF (Speeded-Up Robust Features) algorithm [2]. Matched keypoints are reprojected into the world coordinate system, resulting in 3D feature points. These 3D keypoints are subsequently registered to the virtual reference model using the Iterative Closest Point (ICP) algorithm [3]. The resulting transformation is applied to the reference model.

Each image yields around 30,000 keypoints. Conducting an unconstrained search to find the corresponding point in the stereo image is error-prone.

DGaO Proceedings 2025 – *http://www.dgao-proceedings.de* – ISSN: 1614-8436 – urn:nbn:de:0287-2025-A019
submitted: 30.07.2025 – published: 30.09.2025

Heuristics, including bidirectional matching and a ratio test, are commonly employed to improve robustness. Additionally, this work applies geometric constraints derived from system calibration and the reference geometry to narrow down the search space.

3.2 Deriving Constraints from Geometry

Since the camera transformation is known, keypoints can be filtered based on their scale and orientation of Haar wavelet responses. By applying the epipolar equation $x'^T F x = 0$, matches found outside the epipolar line are rejected.

3.3 Raycasting-Based Matching

All keypoints from an image pair are projected on the reference geometry. The ray-mesh intersections yield the 3D position of each keypoint. For each image, a range-limited kNN-search identifies neighbouring points from the corresponding image. A match is selected after a bidirectional matching and a ratio test. This procedure is embedded in an iterative process, where the resulting point cloud is registered on the geometry.

4 Results

For a dataset compromising 30 pairs of stereo images, the mean distance of all reprojected points to the reference geometry, as well as the mean number of matched points is evaluated (refer Fig. 3). Compared to an unconstrained matching and a conventional geometric filter, the proposed raycasting-based method reduces the mean distance, while providing a high number of correspondences.

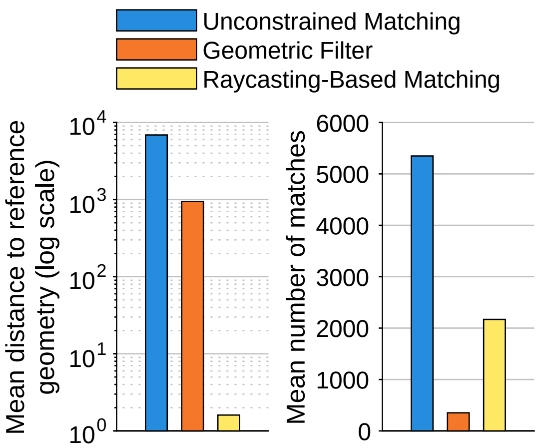

Fig. 3 *Comparison between three different methods for keypoint matching.*

An exemplary single measurement is depicted in Fig. 4.

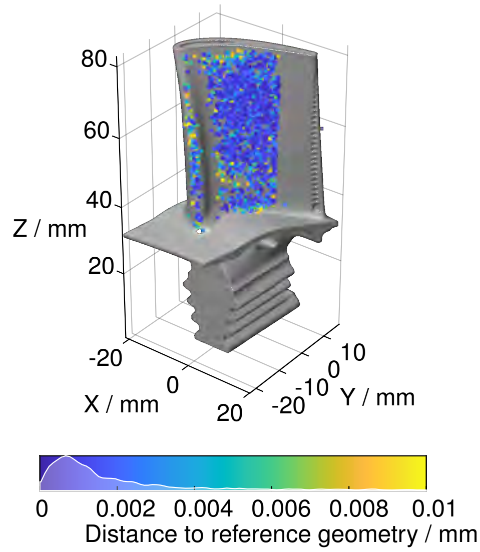

Fig. 4 *Aligned point cloud of reprojected 2D keypoints.*

Fig. 5 shows an aligned and reprojected image. Compared to the unregistered approach in Fig. 1, the edge regions exhibits a better fit. The detail view shows that the quality of fit is determined be the deviation to the reference geometry. In future works it could be investigated, whether the robust point correspondences can be used to modify the mesh or camera projection to account for geometric deviations (e.g. due to wear).

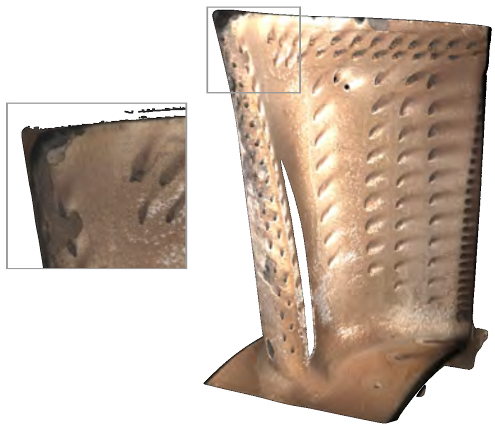

Fig. 5 *Interpolation of the holistic model from an arbitrary viewpoint.*

References

[1] Z. Zhang, “A flexible new technique for camera calibration,” IEEE Transactions on Pattern Analysis and Machine Intelligence **22**(11), 1330–1334 (2000).

[2] H. Bay, T. Tuytelaars, and L. Van Gool, “SURF: Speeded Up Robust Features,” in *ECCV 2006*, LNCS, vol. 3951, pp. 404–417 (Springer Berlin Heidelberg, 2006).

[3] Q.-Y. Zhou, J. Park, and V. Koltun, “Open3D: A Modern Library for 3D Data Processing,” arXiv:1801.09847 (2018).