

# A Vision Based System for Guided and Collaborative Reconstruction of Fragmented Documents

Oliver Krumpek and Diana Leo

Fraunhofer IPK, dept. Machine Vision, Pascalstr. 8-9, 10587 Berlin, Germany

**Abstract.** This paper presents the development and evaluation of a collaborative system for real-time reconstruction of fragmented paper documents in the context of cultural heritage preservation. The developed system includes a collaborative robot, or cobot, that can fully manage the positioning of paper fragments using a specially designed vacuum-based suction attachment. This attachment enables gentle and precise positioning, ensuring the preservation of fragile materials. With this device, we are able to achieve a positioning repeatability of 0.57mm for fragments of 8cm<sup>2</sup>. The system offers users the flexibility to choose between manual positioning, with visual guidance, or fully automated positioning performed by the cobot. To further improve the reconstruction process, AI methods for image interpretation, specifically for segmentation and positioning tasks, were applied and evaluated for their applicability to template-based reconstruction of damaged paper fragments. Our investigation provides critical insights into the performance of different local feature matching methods under different document types, taking into account rotation, scale robustness, and the degree of damage to the fragments. With a focus on the reconstruction of damaged and optically altered archival material, SE2-LoFTR, a detector-free local feature matching method, was chosen as the preferred method for the system due to its robust performance in our experiments.

**Keywords:** Automatic Reconstruction · Assistance System · Fragmented Paper · Feature Matching · Human Machine Interaction · Collaborative Systems

## 1 Motivation

The Fraunhofer Institute for Production Systems and Design Technology (IPK) is part of the Research Alliance for Cultural Heritage (FaLKe), which comprises the Fraunhofer-Gesellschaft, the Leibniz Association and the Prussian Cultural Heritage Foundation [1]. We claim that the use of automation technology, including collaborative robots (cobots), has the potential to contribute to the feasibility of reconstruction projects. This is particularly true for projects with high repeatability of the handling tasks. Using advanced sensors and vision systems, cobots can carefully navigate and manipulate delicate materials, taking over the task of positioning fragments of different types and sizes. The integration of computer vision and artificial intelligence for automatic image analysis further enhances the capabilities for accurate reconstruction. The use of automation technology streamlines and accelerates the reconstruction process, resulting in improved efficiency.

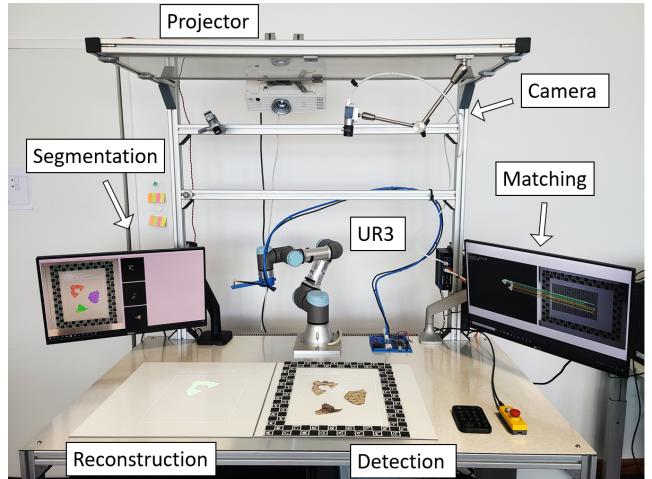


Fig. 1: Annotated image of the overall hardware setup, illustrating the key components. The system includes an industrial camera that digitises the fragments, and a projector that displays the user interface, providing real-time feedback during the reconstruction process. The user can either manually place the fragment in the found target position or delegate this task to the centrally positioned cobot, which handles the precise placement upon request.

Assistive systems that combine human expertise with machine capabilities allow experts to focus on more complex aspects of the reconstruction process such as process control, correction and validation. This human-machine synergy improves the overall reconstruction process and contributes to the preservation of cultural artefacts. Furthermore, the application of collaborative robotics and automation technology extends beyond the reconstruction of paper documents. The principles and methodologies developed in this context can be adapted to the restoration of works of art, the reconstruction of archaeological artefacts and the preservation of historical manuscripts.

The system introduced in this study was initially inspired by a case study [1] exploring the potential of automation. In its current form, the system employs a template-based puzzle-solving method specifically designed for the reconstruction of documents, such as those typically found in archives and libraries. However, this method is designed as an interchangeable module, allowing for future replacement with more advanced, fully automatic algorithms that do not require reference tem-

plates. The primary objective of this research is to develop a robust reconstruction approach tailored to the specific use case where a reference is available.

In conclusion, the integration of collaborative robots and automation technology in the reconstruction of cultural artefacts offers opportunities for increased efficiency and accuracy. By leveraging advanced technologies and combining human expertise with machine capabilities, assistive systems contribute to the preservation and appreciation of our shared cultural heritage.

## 2 Contributions

- **Contribution 1 :** We have developed a partially automated system for guided reconstruction of damaged or corrupted cultural artefacts in the form of flat fragments, such as paper documents or mosaic pieces. The proposed system combines a collaborative robot, data-driven vision systems and a projector-based interface to facilitate human-machine collaboration in the reconstruction task (see Fig. 2). Our functional and reproducible system design enables widespread adoption and use.
- **Contribution 2 :** We have developed a material-protective, vacuum-based suction attachment equipped with sensors to safely handle paper fragments with minimal physical impact. The device can be either used handheld by the restorers or mounted at the Tool Center Point (TCP) of collaborative robots. We are making this attachment publicly available for widespread use, promoting collaboration and innovation in the preservation of document fragments.
- **Contribution 3 :** Our third contribution is the evaluation of local feature matching methods for template-based document reconstruction in assistance systems. Specifically, we assessed three local feature matching methods (SIFT, SuperPoint+SuperGlue, and SE2-LoFTR) for their effectiveness in reconstructing damaged paper fragments. We evaluated their suitability based on robustness to rotation, scaling and damage intensity. Our results highlight both the strengths and limitations of these methods, providing valuable insights for selecting appropriate techniques in similar system setups.

## 3 Related Work

The concept of human-robot collaboration (HRC) has gained significant attention, particularly in the industrial sector. HRC focuses on creating cooperative partnerships between humans and robots, where they work together to accomplish tasks that leverage their respective strengths.[24] Depending on the specific task, this

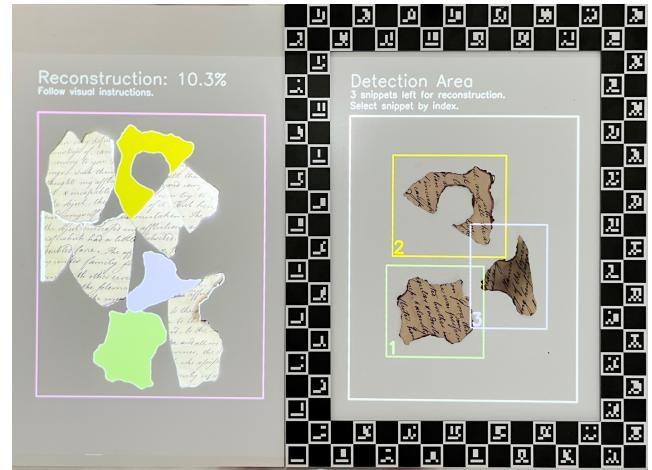


Fig. 2: Photo of the projected user interface for a guided reconstruction process. Users place snippets in the so-called detection area where paper fragments are digitised using an industrial camera and processed using AI methods for background subtraction and feature matching. The detected fragments are framed in highlighted squares (right). The virtually assembled document is visualized in the so-called reconstruction area. By selecting the identified fragments by their index, the system highlights their orientation and position, providing visual guidance for the manual reconstruction (left).

collaborative approach allows for increased productivity, improved safety, and enhanced efficiency in various industries and use cases. This is particularly true in areas such as reverse logistics and the circular economy, where the automated task of assembly and disassembly is becoming increasingly important. [7]. In addition to or in combination with HRC-systems, data driven worker assistance systems (WAS) can help to decrease the workload within increasingly complex tasks by providing the right amount of information in the right form at the right time.[2].

In the field of cultural heritage preservation, robotic-assisted systems have long been used for digitisation and virtual reconstruction tasks[17] [13]. Furthermore, the potential usage of robotic systems for physical reconstruction has already been shown by the RePAIR project [21]. This initiative specifically targets the restoration of ancient artworks in Pompeii, leveraging shape, 3D information, and decorative elements to establish connections among fragmented pieces, and using soft robotics for delicate handling and manipulation tasks. Corresponding research also exists for the general use of robots for the repair of fractured objects based on methods of reinforcement learning [26]. Similar to our system setup, the authors of[18] also used template-based matching and a robot manipulator while dealing with solid puzzle pieces. Notably, new and promising developments in the field of flexible, shape-shifting grippers and manipulators based on principles such as electroadhesion, vacuum or pneu-

matic [29], enable robotic systems to be used more extensively for tasks involving fragile and valuable objects.

More broadly, the scientific community has extensively explored automated solutions for fragment reassembly, focusing largely on virtual reconstruction without the use of robotic systems. This exploration encompasses not only theoretical exercises in computer vision and pattern recognition [4,3,12] but also practical challenges in reconstructing real-world 3D and 2D-items, like wall paintings [25,19,20] or shredded documents used in forensic investigations [14,9]. The complexity of those reassembly-tasks hinges on whether the original image is available. These are divided into template-based approaches where relationships between fragments are known [6], simplifying alignment and identification processes, and others where no original image exists, often addressed as Jigsaw-Puzzles. In this field, it has recently been shown that graph and diffusion models can be used to solve reconstruction tasks. In the context of the RePair project[21], the authors of [23] have just introduced DiffAssemble, which uses a unified model approach for 2D and 3D tasks, showing promising results on public datasets.

Li et al. [15] have demonstrated the effectiveness of local feature matching for reconstructing banknotes from numerous fragments, employing techniques that align and group fragments through agglomerative clustering.

## 4 Methods and Materials

In this section, we present the methods employed in our investigation to enable guided and automatic reconstruction of fragmented paper documents.

### 4.1 Vision and Feature Matching Methods

In order to identify the target position of the paper fragment in the reference template, our system uses proven and well-described algorithms for the extraction and matching of local features.

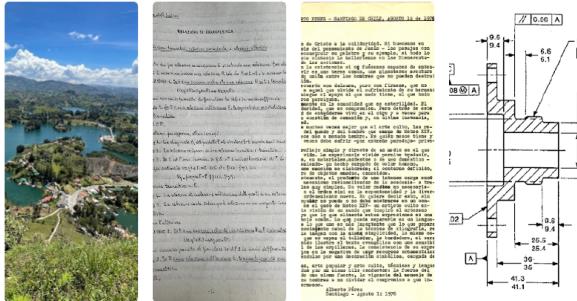


Fig. 3: Examples of document types used to evaluate the reconstruction system. From left to right: landscape image, handwritten document, typewritten document, blueprint.

We implemented and evaluated three local feature matching methods: SIFT + Nearest Neighbor Matching (NN) (BF-Matcher or FLANN Matcher), SuperPoint+SuperGlue and SE2-LoFTR. The selection of these methods has been informed by research into traditional and state-of-the-art local feature matching methods. Given the requirements of our system, we focused on algorithms with strong rotation- and scale-invariant feature extraction capabilities. SIFT [16] was chosen for its robustness in challenging scenarios involving illumination, scale, and rotation. SuperPoint [10] was implemented for its advanced keypoint detection and description capabilities, especially when combined with SuperGlue [22], known for its precision and reliability in feature correspondence. SE2-LoFTR [5], a detector-free method, was chosen for its inherent rotation invariance and streamlined workflow. Prior to the actual virtual positioning, we use the Neural Network based framework REMBG [11] for automatic background removal and robust fragment segmentation.

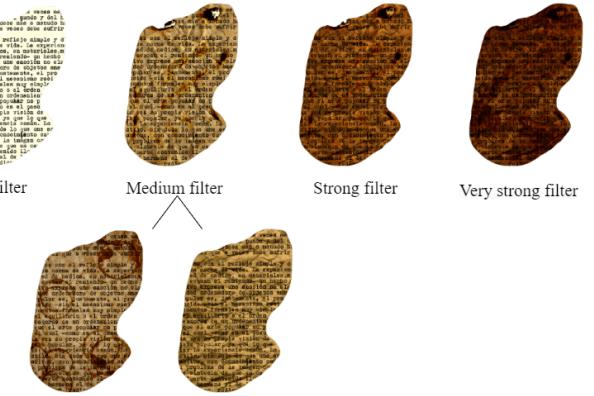


Fig. 4: Fragments with simulated damage created using filters that mimic varying levels of burns, stains, and crumpling. The medium damage level includes three variants, each simulating either burns, stains, or crumpling. These filters were applied to evaluate the system’s performance in reconstructing documents under different damage conditions.

In order to select the most appropriate and robust method for our specific application, we conducted a series of experiments to test the performance of the selected methods on fragments with increasing levels of damage. These experiments focused on the system’s potential use for reconstructing damaged and contaminated fragments, such as those affected by natural or man-made disasters. To facilitate this, we created a dataset of digitally cut fragments with varying degrees of damage by adding filters that mimic burns, stains, or crumple (see Fig. 4).

Four distinct types of documents were selected as reference images, as shown in Fig. 3: handwritten, typewritten, blueprint and outdoor landscape. The choice of these document types was strategic: typewritten and handwritten documents are characterized by their highly repetitive patterns. In contrast, the blueprint documents exhibit sparser features, primarily composed of distinct geometric shapes and lines, presenting a different set of challenges for feature matching methods due to their less textured nature. The outdoor landscape, representing the standard imagery of natural scenes, is classified as the easy document type because it offers a broader variety of distinctive features like varying textures and natural elements. These documents were selected because the system is intended to operate in the field of cultural heritage preservation, specifically targeting archives and libraries.

## 4.2 Setup and Hardware

For automatic handling of fragments with the cobot, a gripping mechanism was necessary. Therefore, a vacuum gripper was developed to gently pick up the paper fragments without causing additional damage. This device was designed to handle fragments of varying sizes (see Fig. 6). Initially, we experimented with a sponge-based approach, but this was eventually replaced by a 3D-printed structure featuring a balanced distribution of air channels, determined through experimentation (see Fig. 6). The gripper is equipped with a proximity sensor (VL53L0X Time-of-Flight) that allows the airflow to be automatically activated just before contact with the fragment, ensuring effectiveness for both pick-up and deposit processes. The required vacuum is generated and controlled using a Festo VN-14-H vacuum generator and a 3V210-08 24VDC valve (see Fig. 5).

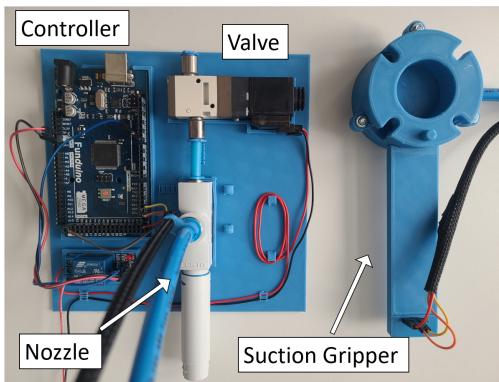
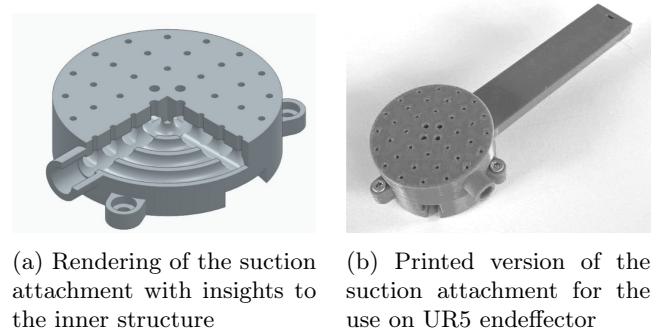


Fig. 5: Overview of mechanical and electrical components of the suction gripper. The control unit and the valve are rigidly attached to a printed carrier board and connected to the suction gripper using flexible tubing and isolated wires.



(a) Rendering of the suction attachment with insights to the inner structure

(b) Printed version of the suction attachment for the use on UR5 endeffector

Fig. 6: Design of the suction gripper

## 4.3 System Calibration and Position Measurement

In order to determine the geometric relationships of the components used in the system (see Fig. 7), several steps were first taken to calibrate the system, using essentially straightforward methods.

The relevant geometrical relations for the coordinate transformation performed during the evaluation are given by  $T_R^C$ , the constant transformation between camera and robot and  $T_P^C$ , the constant transformation between camera and projector. Both transformations can be carried out in a one-off calibration prior to the live evaluation. To calculate  $T_R^C$ , a system of the form  $AX = YB$ , also addressed as a hand-eye-calibration [8] was set up and solved.

To enable the transformation between camera and projector  $T_P^C$  we first assumed all fragments to be planar as the problem could then be reduced to the estimation of a homography between two planes and thus determined by the following expression:  $(\lambda x', \lambda y', \lambda)^T = H(x, y, 1)^T$ . Here,  $H$  describes the 2D to 2D projective transformation between the 2D plane in the 3D world and the 2D image plane. To solve for  $H$ , we find at least four corresponding points between the camera image and the projected image. This is achieved by projecting a calibration pattern onto the detection surface and detecting it in the camera image.

For manual assisted reconstruction, it was first necessary to express a description of significant features such as edges, centroids and surface information in terms of pixel values. These values could be processed for the visual interface based on the projector-to-camera calibration described above. However, in order to implement automatic positioning, it was necessary to derive positions in Euclidean space in addition to the pixel values, on the basis of which the cobot's path planning could be performed. To keep the hardware setup simple and reproducible, we avoided using an additional camera or a 3D sensor. Instead, we used a ray tracing approach to determine the 3D position of the fragments'

centre of gravity. More precisely, given the intrinsic camera parameters[27], the desired 3D coordinate is given by intersecting the ray from the camera's optical centre through the pixel position  $(u, v)$  with the plane defined by at least three 3D points (we use three corner points) of the ChArUco frame placed around the detection surface, all expressed in the Euclidean camera coordinate system (see Fig. 1).

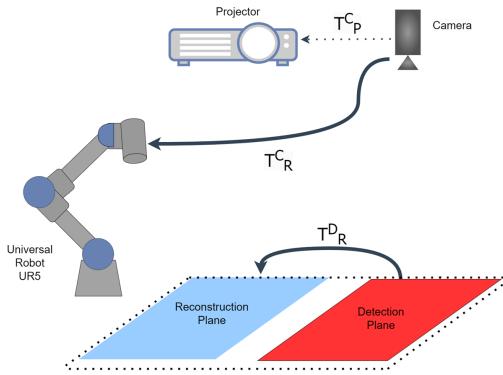


Fig. 7: System diagram illustrating the key geometric transformations, including the camera-to-robot transformation  $T_R^C$ , the camera-to-projector homography  $T_P^C$  and the transformation between the initial fragment position in the detection-area and its target position in the reconstruction-area  $T_R^D$ .

## 5 Results

Our comprehensive evaluation across two robustness experiments—Rotation and Scale—showcases the varied performance of the three selected local feature matching methods under multiple damage conditions (see Fig. 4). The evaluated local feature matching methods are: traditional SIFT+NN, SuperPoint+NN, SuperPoint+SuperGlue, and detector-free SE2-LoFTR. Each damaged fragment from the created dataset goes through the pipeline illustrated in Fig. 8. The results are categorized into three performance levels based on the Intersection Over Union (IOU) and Mean Absolute Error (MAE): "Good (MAE = [0,8]px)", "Moderate (MAE = [9,22]px)", and "Failed (MAE>22px)". The tables 1, 2 and 3 present the reconstruction failure rates SIFT, SE2-LoFTR and SuperPoint+SuperGlue respectively, across the different document types, experiments, and filter levels (medium, strong, and very strong).

### 5.1 Rotation Robustness Experiment

Each snippet, already subjected to four increasing damage levels (no damage, medium damage, hard damage, very hard damage) (see Fig. 4), was rotated through ten angles ranging from 0 to 324 degrees in increments

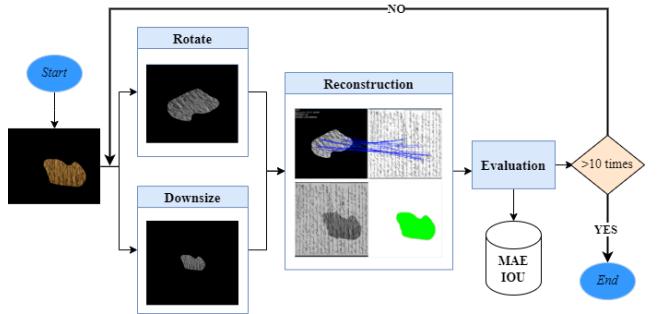


Fig. 8: Schematic representation of the experimental pipeline detailing the steps involved in the evaluation of the reconstruction of rotated (Experiment 1) or scaled (Experiment 2) document fragments. Ten rotations and ten levels of down-scaling are applied to the fragment to simulate environmental conditions, e.g. camera offsets.

of 36 degrees. SE2-LoFTR and SIFT demonstrated the most robust performance in rotation invariance across various document types and levels of damage, showcasing their inherent rotation invariance. Notably, traditional SIFT often outperformed SE2-LoFTR, particularly with less damaged documents, showcasing its enduring effectiveness despite advancements in deep learning methods like SE2-LoFTR. SuperPoint+NN and SuperPoint+SuperGlue showed more limited robustness, excelling under certain conditions but struggling in others. The handwritten document was the most challenging to reconstruct, followed by the technical drawing. The landscape document was the easiest, with only a 2% failure rate. The handwritten document posed significant challenges for the SuperPoint detector, regardless of the matcher used. Technical drawings resulted particularly difficult for SuperPoint+NN, and the most difficult to reconstruct for SE2-LoFTR.

Table 1: Failure rate by document type for SIFT.

Document	Medium		Strong		Very Strong	
	Rotation	Scale	Rotation	Scale	Rotation	Scale
Handwritten	0%	40%	0%	60%	100%	100%
Typewritten	0%	30%	0%	60%	0%	70%
Blueprint	0%	50%	0%	70%	0%	70%
Landscape	0%	20%	0%	30%	0%	70%

### 5.2 Scale Robustness Experiment

Fragments, under four level of paper damage (see Fig. 4), were downsized in ten increments, from their original size down to just 10% of their original size. The landscape document emerged as the easiest to reconstruct across all levels of downsizing and filtering. This consistency

Table 2: Failure rate by document type for SE2-LoFTR.

Document	Medium		Strong		Very Strong	
	Rotation	Scale	Rotation	Scale	Rotation	Scale
Handwritten	0%	30%	20%	50%	40%	60%
Typewritten	0%	50%	10%	80%	10%	80%
Blueprint	40%	50%	50%	60%	50%	60%
Landscape	0%	30%	0%	30%	0%	40%

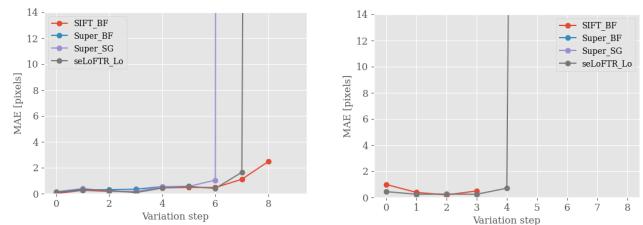
Table 3: Failure rate by document type for SuperPoint+SuperGlue.

Document	Medium		Strong		Very Strong	
	Rotation	Scale	Rotation	Scale	Rotation	Scale
Handwritten	50%	90%	100%	100%	100%	100%
Typewritten	0%	60%	20%	80%	10%	80%
Blueprint	0%	30%	20%	60%	0%	50%
Landscape	0%	20%	0%	30%	0%	30%

underscores its relative simplicity compared to other document types. In contrast, documents with highly repetitive patterns presented the most significant challenges. Among these, the handwritten document proved to be the most difficult to reconstruct, followed by the typewritten documents in terms of difficulty. SE2-LoFTR effectively handled images with repetitive patterns across all filter levels, demonstrating its robustness. SIFT, while second-best for handwritten, performed slightly better for typewritten documents. SuperPoint+NN was effective only in the no-filter scenario; it failed to perform adequately in all the other damage levels, except with the landscape document. As the level of filter damage increased, SIFT’s scale invariance diminished, whereas SE2-LoFTR showed better adaptability in handling the more challenging hard and very hard damage categories (see Fig. 9).

### 5.3 Accuracy of the Automatic Positioning Approach

In this experiment, we evaluate the repeatability of our automatic positioning approach using the UR3 cobot. To this end, three square fragments of differing sizes ( $4 \text{ cm}^2$ ,  $16 \text{ cm}^2$ , and  $64 \text{ cm}^2$ ) of  $90\text{g}/\text{m}^2$  paper with a robustly detectable feature were utilized. The fragments were each moved 100 times by the cobot from a defined constant start position to a defined target position. The height and speed of the movement were kept constant for all trials. To measure the accuracy of the positioning, the position of the feature was determined after each placement. The positioning errors are analysed by computing the mean coordinates of the detected features and determining the deviations from these mean values for each detected feature. The Root Mean Square Error (RMSE) is then calculated to quantify the overall accuracy. For the smallest fragment ( $4 \text{ cm}^2$ ), the RMSE value is the



(a) SIFT effectively reconstructs the snippet up to level 9 of downsize (90% reduction), followed by SE2-LoFTR, which works up to level 8, and SP+SG, effective only until level 6.

(b) SE2-LoFTR manages to reconstruct the hardly damaged snippet up to level 4 of downsize, followed by SIFT, which stops at level 3; the other methods fail in every level.

Fig. 9: Scale Robustness Experiment: Handwritten document, MAE graphs for unfiltered images (left) and strong filtered images (see Fig.4) (right).

highest ( $1.5254 \text{ mm}$ ); for the  $16 \text{ cm}^2$  fragment the RMSE is  $0.5735 \text{ mm}$  and for the  $64 \text{ cm}^2$  the RMSE is the lowest ( $0.4771 \text{ mm}$ ) (see Fig. 10).

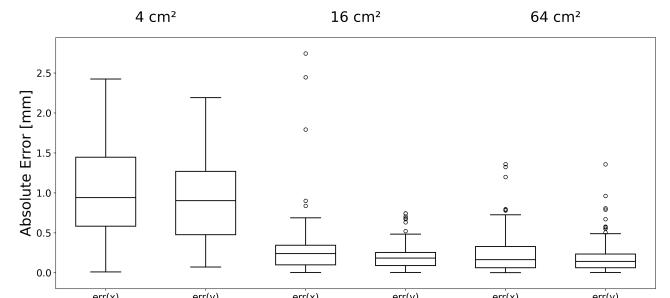


Fig. 10: Box plot illustrating the results of the experiment evaluating the accuracy of the UR3 cobot’s automatic positioning of document fragments of varying sizes ( $4 \text{ cm}^2$ ,  $16 \text{ cm}^2$ , and  $64 \text{ cm}^2$ ). The plot presents the distribution of positioning errors for each fragment size divided into their X- and Y-components.

## 6 Conclusion

The primary objective of the development was to test and analyse the possibility of automating physical reconstruction tasks, especially of fragmented documents, using a cobot and machine vision methods.

As a result of this investigation, a vision system has been implemented that uses machine learning methods to automatically isolate paper fragments and determine the target position of the fragments in a dedicated target area. Experimentally, SE2-LoFTR was chosen as the reconstruction method as it provides robust matching results in the presence of high levels of damage (see

2). Existing limitations in rotational invariance can currently be circumvented by users when positioning in the workspace. The assumption that document types with font content are a comparatively difficult reconstruction task has also been confirmed. The developments also include a projector-based visual interface that can be used to visualise the target position for manual reconstruction. Based on a measurement and coordinate transformation, it is possible to use a cobot for the reconstruction task. This prototype already allows the automation of time-consuming tasks and would be capable of reconstructing large numbers of fragments, leaving the quality check to the experienced human eye.

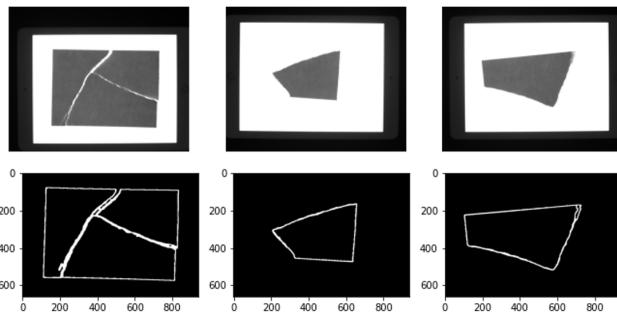


Fig. 11: Segmentation of fragments for the detection of inner and outer contours of the breaking edges. The contours were further preprocessed to fit the requirements of the cobot's path planning.

## 7 Outlook and future work

It has been shown that some of the steps in the reconstruction work can be robustly automated by using cobots and machine vision techniques. Looking to the future, however, it would be possible to enable some further steps, such as the actual bonding or welding of the fragments, on the way to a fully automated system. In this context, some preliminary tests have already been carried out, in which a laser pointer was guided along the breaking points of the already positioned fragments with the help of the cobot used (see Fig. 12). The robot's path planning was based on automatic extraction of the inner and outer contours (see Fig. 11) using an image processing pipeline and transformed into the cobot's system using the calibration described. A classical interpolation method [28] was used to derive the actual path. To validate the accuracy, the actual area of the fragment was compared to the area created by the laser path. The area-error was in average 1%. There was also an average deviation at significant points of 2.7 mm. A value that leaves room for improvement.

**Acknowledgments** We would like to thank the student project RoboGlue from the Institute of Automat-

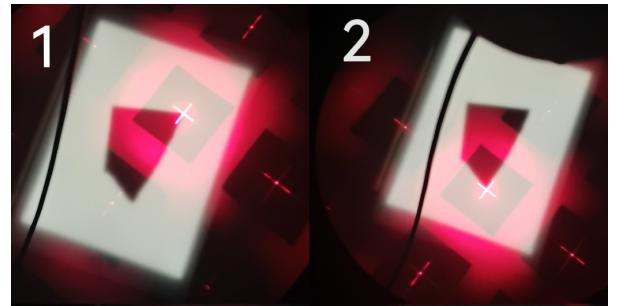


Fig. 12: Photo taken during the experiments showing the laser pointer, mounted on the Tool Center Point (TCP) of the UR3 cobot, as it follows the contours of a fragment. The cross marks the current position of the laser on the edge of the fragment.

tion Technology group of TU Berlin for the design of the suction system.

## References

- Official website of the forschungsallianz kulturerbe. <https://forschungsallianz-kulturerbe.de/>, accessed 20-09-2024
- Bornewasser, M., Bläsing, D., Hinrichsen, S.: Informatorische assistenzsysteme in der manuellen montage: Ein nützliches werkzeug zur reduktion mentaler beanspruchung? Zeitschrift für Arbeitswissenschaft **72**, 264–275 (12 2018). <https://doi.org/10.1007/s41449-018-0123-x>
- Bunke, H., Böhler, U.: Applications of approximate string matching to 2d shape recognition. Pattern Recognition **26**(12), 1797–1812 (1993). [https://doi.org/10.1016/0031-3203\(93\)90177-X](https://doi.org/10.1016/0031-3203(93)90177-X), <https://www.sciencedirect.com/science/article/pii/003132039390177X>
- Burdea, G., Wolfson, H.: Solving jigsaw puzzles by a robot. IEEE Transactions on Robotics and Automation **5**(6), 752–764 (1989). <https://doi.org/10.1109/70.88097>
- Bökman, G., Kahl, F.: A case for using rotation invariant features in state of the art feature matchers (2022)
- Cai, Y., Thalmann, D. (eds.): Proceedings of the Symposium on VR Culture and Heritage - Volume 2. ACM, New York, NY, USA (12032016). <https://doi.org/10.1145/3014027>
- Daneshmand, M., Noroozi, F., Corneanu, C., Mafakheri, F., Fiorini, P.: Industry 4.0 and prospects of circular economy: a survey of robotic assembly and disassembly. The International Journal of Advanced Manufacturing Technology **124**(9), 2973–3000 (2023). <https://doi.org/10.1007/s00170-021-08389-1>
- Daniilidis, K.: Hand-eye calibration using dual quaternions. The International Journal of Robotics Research **18**(3), 286–298 (1999)
- De Smet, P.: Semi-automatic forensic reconstruction of ripped-up documents. In: 2009 10th International Conference on Document Analysis and Recognition. pp. 703–707 (2009). <https://doi.org/10.1109/ICDAR.2009.7>

10. DeTone, D., Malisiewicz, T., Rabinovich, A.: Superpoint: Self-supervised interest point detection and description. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 337–33712. IEEE (18/06/2018 - 22/06/2018). <https://doi.org/10.1109/CVPRW.2018.00060>
11. Gatis, D.: Rembg. <https://github.com/danielgatis/rembg> (2021)
12. Goldberg, D., Malon, C., Bern, M.: A global approach to automatic solution of jigsaw puzzles. In: Proceedings of the eighteenth annual symposium on Computational geometry. pp. 82–87 (2002)
13. Hardy, H., van Walsum, M., Livermore, L., Walton, S.: Research and development in robotics with potential to automate handling of biological collections. *Research Ideas and Outcomes* **6**, 213–41 (2020)
14. Justino, E., Oliveira, L.S., Freitas, C.: Reconstructing shredded documents through feature matching. *Forensic Science International* **160**(2), 140–147 (2006). <https://doi.org/https://doi.org/10.1016/j.forsciint.2005.09.001>, <https://www.sciencedirect.com/science/article/pii/S0379073805004913>
15. Li, H., Zheng, Y., Zhang, S., Cheng, J.: Solving a special type of jigsaw puzzles: Banknote reconstruction from a large number of fragments. *IEEE Transactions on Multimedia* **16**(2), 571–578 (2014). <https://doi.org/10.1109/TMM.2013.2291968>
16. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* **60**(2), 91–110 (nov 2004). <https://doi.org/10.1023/B:VISI.0000029664.99615.94>, <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
17. Luxman, R., Castro, Y.E., Chatoux, H., Nurit, M., Siatou, A., Le Goïc, G., Brambilla, L., Degrigny, C., Marzani, F., Mansouri, A.: Lightbot: A multi-light position robotic acquisition system for adaptive capturing of cultural heritage surfaces. *Journal of Imaging* **8**(5), 134 (2022). <https://doi.org/10.3390/jimaging8050134>, <https://www.mdpi.com/2313-433X/8/5/134>
18. Ma, C.H., Lu, C.L., Shih, H.C.: Vision-based jigsaw puzzle solving with a robotic arm. *Sensors (Basel, Switzerland)* **23**(15) (2023). <https://doi.org/10.3390/s23156913>
19. McBride, J.C., Kimia, B.B.: Archaeological fragment reconstruction using curve-matching. In: 2003 Conference on Computer Vision and Pattern Recognition Workshop. p. 3. IEEE (16/06/2003 - 22/06/2003). <https://doi.org/10.1109/CVPRW.2003.10008>
20. Papaodysseus, C., Exarhos, M., Panagopoulos, M., Rousopoulos, P., Triantafillou, C., Panagopoulos, T.: Image and pattern analysis of 1650 bc wall paintings and reconstruction. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* **38**(4), 958–965 (2008)
21. Rsipvision, I.: Computer vision news 2021 october (2021), <https://www.rsipvision.com/ComputerVisionNews-2021October/22/> accessed: (30.05.2023)
22. Sarlin, P.E., DeTone, D., Malisiewicz, T., Rabinovich, A.: Superglue: Learning feature matching with graph neural networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 4938–4947 (2020)
23. Scarpellini, G., Fiorini, S., Giuliani, F., Morerio, P., Bue, A.D.: Diffassemble: A unified graph-diffusion model for 2d and 3d reassembly (2024), <https://arxiv.org/abs/2402.19302>
24. Schmidbauer, C., Hader, B., Schlund, S.: Evaluation of a digital worker assistance system to enable adaptive task sharing between humans and cobots in manufacturing. *Procedia CIRP* **104**, 38–43 (2021). <https://doi.org/10.1016/j.procir.2021.11.007>
25. Sizikova, E., Funkhouser, T.A.: Wall painting reconstruction using a genetic algorithm. *Journal on Computing and Cultural Heritage (JOCCH)* **11**, 1–17 (2016), <https://api.semanticscholar.org/CorpusID:894723>
26. Song, X., Lamb, N., Banerjee, S., Banerjee, N.K.: Reinforcement-learning based robotic assembly of fractured objects using visual and tactile information. In: 2023 9th International Conference on Automation, Robotics and Applications (ICARA). pp. 170–174. IEEE (10022023 - 12022023). <https://doi.org/10.1109/ICARA56516.2023.10125938>
27. Tsai, R.: A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal on Robotics and Automation* **3**(4), 323–344 (1987). <https://doi.org/10.1109/JRA.1987.1087109>
28. Wu, S.T., Marquez, M.: A non-self-intersection douglas-peucker algorithm. In: 16th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI 2003). pp. 60–66 (2003). <https://doi.org/10.1109/SIBGRA.2003.1240992>
29. Zaidi, S., Maselli, M., Laschi, C., Cianchetti, M.: Actuation technologies for soft robot grippers and manipulators: A review. *Current Robotics Reports* **2**(3), 355–369 (2021). <https://doi.org/10.1007/s43154-021-00054-5>