A Marker-based Method for precise 3D Registration between CT-Data and photogrammetric Datasets

Sebastian Preußel [1], Cyrill Milkau [2], Hans-Gerd Maas [1]

[1] Institute of Photogrammetry and Remote Sensing, Dresden University of Technology, Germany – (sebastian.preussel1, hans-gerd.maas)@tu-dresden.de
[2] Faculty of Spatial Information, HTW Dresden – University of Applied Sciences, Germany – cyrill.milkau@htw-dresden.de

Keywords: Computed Tomography, Photogrammetry, 3D Point Clouds, Fiducial Markers, Registration

Abstract

In order to enable photogrammetric tracking of objects from a computed tomography (CT) dataset with a multi-camera system, a transformation between the CT data space and a photogrammetric reference frame is required, typically based on control points. To achieve a robust and precise registration between CT and photogrammetric datasets, this work proposes a marker-based approach. The main goal is to use a marker model that allows straightforward segmentation and control point estimation in CT voxel space, while also supporting reliable and precise control point estimation in the photogrammetric images. As a proof-of-concept, spherical markers were investigated, since they allow centre estimation in both domains. In the CT data, marker centres were determined by intensity-based thresholding followed by sphere fitting, while in the photogrammetric data they were estimated by intensity-based thresholding, edge detection, circle fitting, and multi-image spatial intersection. Two different marker models were tested. The results show that the proposed method is feasible and yields sub-millimetre standard deviations of unit weight for both marker types. However, since a sufficient stochastic model is not yet available, the reported accuracy measures may be optimistic and should therefore be interpreted with caution. Future work will address these limitations, in particular uncertainty modelling as well as remaining lighting and contrast issues.

1. Introduction

This study is presented within the context of improving radiotherapy treatment through intelligent patient positioning and monitoring. Thereby, the approach involves high-accuracy, markerless face tracking with precision of approximately 1 mm in object space. We use self-trained Convolutional Neural Networks across two-dimensional (2D) images and three-dimensional (3D) Computed Tomography (CT) datasets, leveraging an interconnected human body model based on bony- and facial landmarks. By applying a series of chained transformations, the primary objective is to accurately estimate the absolute 3D position of a tumour. In order to determine the level of precision for this absolute tumour position, an independent validation method is necessary.

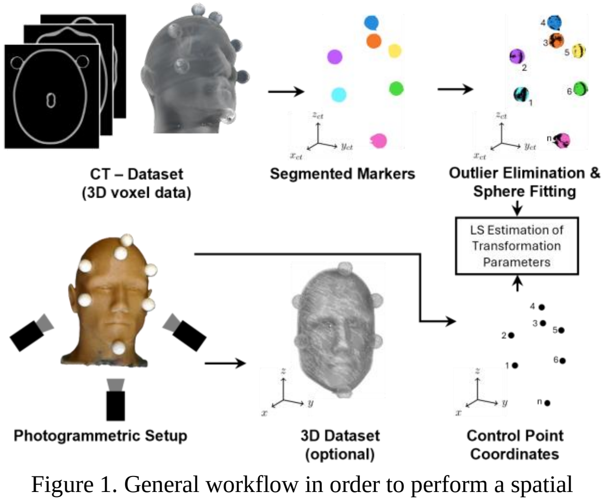

Figure 1. General workflow in order to perform a spatial transformation for dataset registration of CT and photogrammetric data.

We introduce a separate, marker-based approach grounded in conventional photogrammetry, operating entirely without machine learning or artificial intelligence-based methods, to serve as a reliable baseline for comparative testing. The core methodology relies on estimating transformation parameters from the CT space to the photogrammetric object space using physical markers, which are detectable and segmentable in both modalities. Spherical markers used as control points have the advantage that their centre can be estimated precisely both photogrammetrically, using forward intersection of the projected circle, and in voxel data, using spherical fitting. By transforming between these spaces via an overdetermined set of marker centres, we can accurately calculate the necessary stochastic parameters. To achieve the high precision required, this study addresses two primary challenges: Developing a robust image segmentation algorithm capable of handling variable lighting and contrast situations, and designing an optimal marker model. This marker model must ensure robust segmentation, guarantee visibility within CT data, and enable accurate sphere fitting in the 3D CT datasets alongside corresponding circle fitting in the 2D photogrammetric images. The general simplified workflow is displayed in Figure 1. Figure 2 shows the tested marker models.

Figure 2. The tested marker models – 3D printed (left) and ceramic (right). Minor irregularities, which vary from marker to marker, can be observed in the 3D-printed model.

The approach shares similarities with methods for CT dataset registration, which have been explored in the literature. For example, Cao et al. (2020) developed a method for automatic registration of CT images to physical markers for augmented reality applications. In the medical field, the registration of CT images is common practice, and fiducial marker-based methods are frequently employed. Typically, those are performed between images, as presented by Deegan et al. (2015). Volumetric registration methods have been explored in the field of 3D digitisation. For example, Zhan et al. (2021) use marker-less primitive fitting to align photogrammetric point clouds with CT volumes.

2. Materials and Methods

For this proof-of-concept, two distinct marker sets were developed and examined: a custom 3D-printed 3/4-sphere and a ceramic bearing ball functioning as a complete sphere. Photogrammetric measurements were acquired and evaluated utilising a joint three-camera setup (Daheng Imaging *MARS-2440-35GTC*, with a resolution of 5328 × 4608 pixels at a pixel size of 2.74 µm, combined with 12 mm optics). The setup is displayed in Figure 3. Corresponding 3D voxel data was captured using an industrial CT scanner (Procon *CT-XPRESS*), which provided the necessary datasets to test and refine robust segmentation algorithms and sphere-fitting methods.

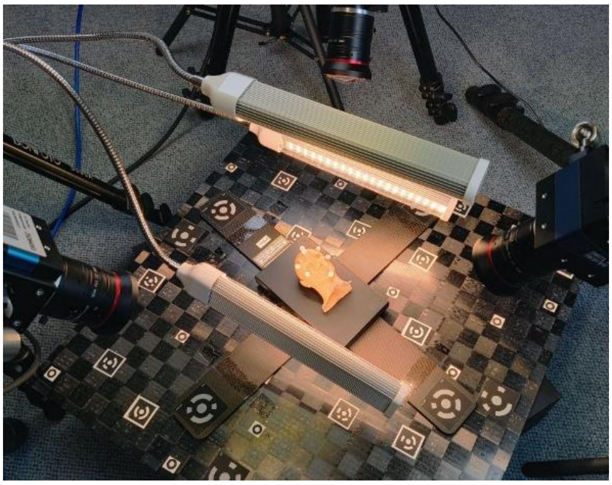

Figure 3. Setup with three obliquely positioned cameras, a coded marker board and illumination.

2.1 Preparations

Before the CT scans were performed, the 3D-printed markers were applied to the 3D-printed head model, while ceramic markers were attached to the plastic head model. These head models are hereafter referred to as $\mathbf{H_1}$ and $\mathbf{H_2}$, respectively. For performance reasons, the CT datasets were downsampled by a factor of 4 with linear interpolation, which resulted in a voxel size of 0.12 mm for $\mathbf{H_1}$ and 0.09 mm for $\mathbf{H_2}$. The photogrammetric setup and measurement preparation were realised by configuring the cameras at oblique intersection angles, approximately 120° apart in a circular arrangement with about 30 cm spacing between them. This configuration was selected to ensure strong intersection geometry and sufficient base-to-height ratios, which are critical considerations for maintaining high measurement accuracy. Camera calibration and pose estimation were performed via a simultaneous joint bundle adjustment utilising the software AICON 3D Studio v12.5 (AICON 3D Systems GmbH, 2018). This was achieved by observing a reference object with known approximate coordinates, utilising a planar marker board with 19 coded markers and a reference cross to properly establish spatial scale. In preparation for the measurements, the local contrast of the plastic head model $\mathbf{H_2}$ was artificially enhanced by applying black permanent marker directly around the white markers to ensure reliable image segmentation. Figure 4 shows sections of the resulting measurement images. Illumination of the markers is key for the segmentation, hence we used lighting from several lamps around the object, to ensure consistency.

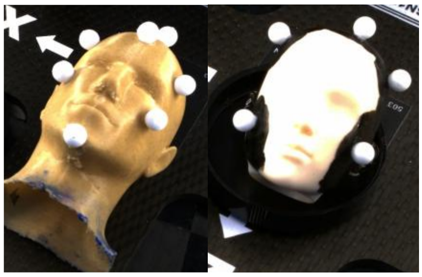

Figure 4. 3D-printed markers applied to the 3D-printed head model $\mathbf{H_1}$ (left) and ceramic markers applied to the plastic head model $\mathbf{H_2}$ (right).

2.2 2D Marker Detection and Measurement

The basic idea of the image measurement algorithm is derived from gradient-based ellipse fitting methods via the star-operator (Luhmann, 1986). The methodology consists of several sub-steps, which are identical for both types of markers. The fundamental methodology for the image processing was implemented semi-automatically in MATLAB. While the perspective projection of a spherical target onto a 2D image plane results in an ellipse (Liebold and Maas, 2026), for the current stage we are using a circle fit, considering the small marker size and central image positions. The algorithm consists of several distinct sub-steps that remain identical for both the 3D-printed and ceramic spherical markers.

Initially, a Region of Interest (ROI) is defined around the respective head model within each measurement image to isolate the markers and effectively minimise false detections. In the resulting binary image (Luhmann, 2023), clusters get segmented via a basic connected components clustering algorithm. Due to lighting hotspots in the surrounding area, incorrectly assigned clusters may occur occasionally (also see section Limitations and Discussion). These get eliminated by filtering by pixel count, suitable values can be assumed based on test images. Around each valid cluster, a search window gets defined via its corresponding bounding box with an additional 10-pixel padding, as displayed in Figure 5. This way we further restrict the subsequent circle detection to relevant areas.

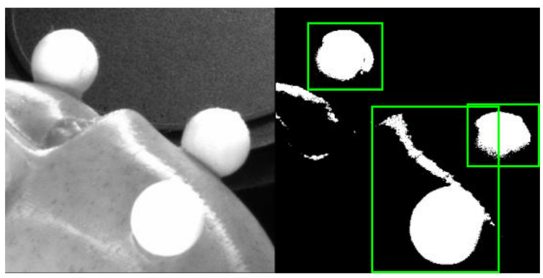

Figure 5. Measurement image detail (left) and the corresponding binary image with detected marker clusters (right).

Following this, a Hough transformation-based circle detection algorithm – which occasionally requires specific input parameter adjustments based on the image – is employed to provide a reliable initial guess for the marker boundaries (Atherton and Kirbyson, 1999). A buffer is then established around this detected Hough circle to serve as a constrained search window for the next phase, extending 5 px inward from the circle boundary and 20 px outward in the radial direction. This discrepancy proved to be necessary, since due to lighting conditions (shading) the detected Hough circles tend to be rather too small. From the estimated centre of the circle, search rays are cast outward in 128 discrete azimuthal directions. Simultaneously, a Gaussian blur (σ = 1.0 px) is applied to the image, thereafter a gradient magnitude image gets calculated based on the result via the Sobel-operator. Along each individual ray, the gradient magnitude at subpixel locations gets sampled using 2D spline interpolation. From these sampled points, the ones with the maximal gradient value get saved. Their positions are refined by fitting a 3-point quadratic to the two neighbouring points (as described by Yu and Xie, 2014). When all the cluster points are acquired in this manner, a Random Sample Consensus (RANSAC) circle fitting procedure is performed to detect in- and outliers of the assumed circles. The resulting point set of inliers is used for the final centre estimation by least squares circle fitting. The approach provides robustness, as it allows to estimate centres of partly obscured markers, presuming that the overall lighting and contrast conditions are appropriate. Figure 6 illustrates the individual steps.

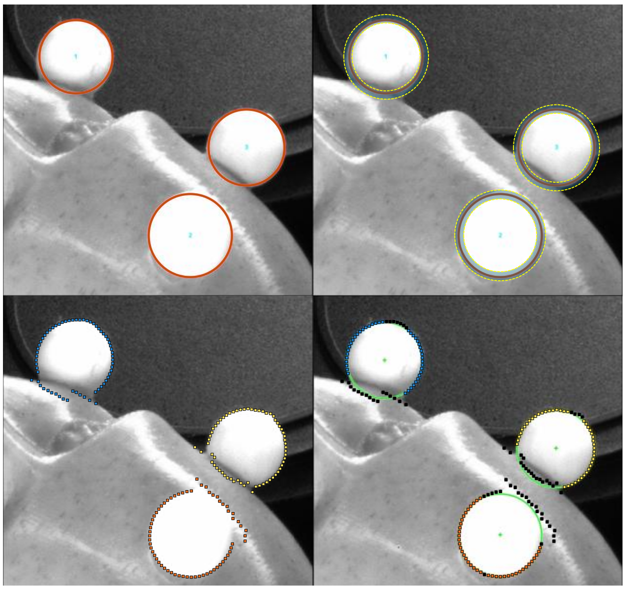

Figure 6. Top left to bottom right: Hough circle detection, buffer definition (not to scale), gradient based edge detection and least-squares circle fit for final centre estimation with outliers filtered via RANSAC displayed in black colour.

By acquiring these 2D centre positions across multiple views, we can accurately estimate the 3D spatial coordinates of the markers via an overdetermined intersection from the three cameras. Furthermore, it allows us to yield stochastic information regarding measurement accuracy and reliability by providing the resulting variance-covariance matrix $\boldsymbol{C}_{xx}$ and the a posteriori standard deviation of unit weight σ_0.

2.3 Voxel Dataset Processing

Since single-material markers exhibit a homogeneous X-ray attenuation, they appear in the reconstructed CT volume as regions with nearly uniform voxel intensities. In volumetric CT data, such homogeneity makes intensity-based simple thresholding a suitable first step. For the 3D-printed markers, this procedure is in principle straightforward and has been tested successfully. However, it could not be applied to the corresponding head model $\mathbf{H_1}$ used in this study, because the model and the markers were manufactured from the same material and therefore did not provide sufficient contrast in the CT data. Instead, the markers in this dataset had to be segmented manually. For the ceramic markers on $\mathbf{H_2}$, thresholding was successful, yet outliers occurred due to scattering effects (Figure 7). Additionally, we faced the problem, that solid spheres appear as filled spheres in the 3D voxels, thus making the fitting problem more challenging (compared to, e. g. surface scans). Consequently, in order to transform this into a least-squares fitting problem, it was necessary to sample surface points. Both of these issues were addressed simultaneously during the centre estimation step. In preparation, the voxel dataset was converted into a point cloud by sampling the centre of each voxel and assigning it an associated intensity value. This allowed us to handle the issues via point cloud processing.

Figure 7. Ceramic marker point cloud with outliers from the CT-voxel-dataset visualised in 3D.

After segmentation, the marker centres get estimated using the same procedure for both ceramic and 3D-printed markers. First, the median position of each segmented cluster is computed and used as an initial estimate of the respective marker centre. Starting from this point, rays get cast in uniformly distributed 3D directions, analogous to a Fibonacci sphere, in order to sample the marker boundary well spread out from multiple sphere segments. In the present workflow, 1000 rays were used. Each ray was modelled as a cylindrical search region, and for all segmented points located inside this cylinder, the point with the greatest distance along the ray from the median centre was selected, provided that its distance lays within the interval defined by the known marker radius of 2.5 mm ± 0.25 mm. In this way, a set of candidate surface points was obtained over the marker edge. The method is illustrated schematically in a two-dimensional slice together with a visualisation of the ray distribution in Figure 8.

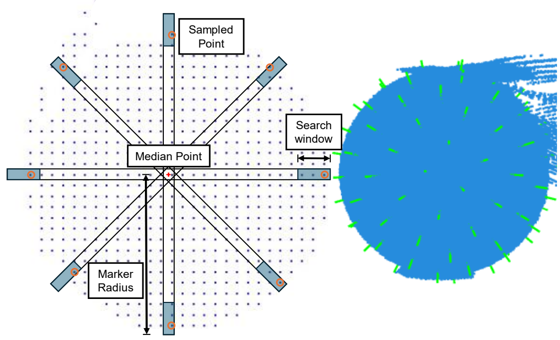

Figure 8. Sampling of hull points with the cylinder search method on a sphere slice with outliers (left, not to scale) and visualisation in the 3D marker point cloud (right).

Sampling outer hull points this way allowed us to continue with a RANSAC approach to eliminate outliers in both the sampled and regular points, while finding the outermost points of the marker simultaneously: With a pre-defined radius buffer, find the sphere configuration with the maximum of inliers. We used these best-configuration inliers for a least squares sphere fit, which again allows for a qualitative assessment of the stochastic parameters. The result for a cluster processed this way is shown in Figure 9.

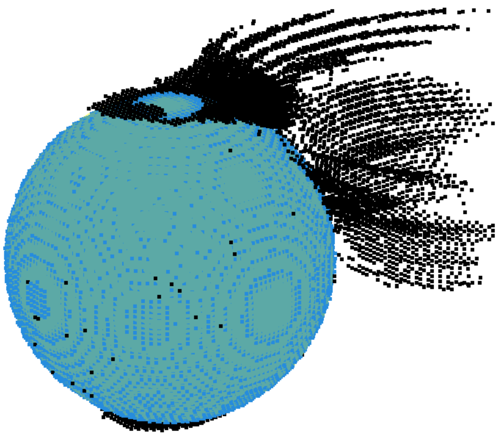

Figure 9. Marker with detected outliers (black), hull inliers (blue) and fitted sphere (green).

3. Results

3.1 Camera Orientation

The results of the interior and exterior camera orientation parameters (IOR and EOR) estimation obtained from the joint bundle adjustment based on a 120-image dataset with a coded marker board combined with a reference cross are summarized in Table 1. Overall, we reached a standard deviation of unit weight σ_0 of 0.47 pixels. This is considered sufficient, given the enforcement of global consistency across all observations from three cameras. The estimated parameters achieved a sub-millimetre / sub-milliradians level of accuracy, indicating a precise calibration of both IOR and EOR. As these results were considered sufficiently accurate for the purposes of the subsequent analysis, the estimated camera poses, including both EOR and IOR, were treated as fixed for the subsequent photogrammetric intersection of the marker centres.

Camera №	s_x [mm]	s_y [mm]	s_z [mm]	s_ω [mrad]	s_φ [mrad]	s_κ [mrad]
1	0.075	0.090	0.114	0.318	0.286	0.156
2	0.060	0.053	0.095	0.274	0.181	0.165
3	0.106	0.050	0.119	0.230	0.401	0.255

Table 1. Achieved empirical standard deviations s_i of the EOR parameters via the bundle adjustment.

3.2 Image and voxel space measurements

Regarding the circle fitting in image space, the ceramic markers demonstrated superior results compared to the 3D-printed markers, as displayed in Table 2. The depicted standard deviations are bundled as mean values of all measured markers per camera image. The disparity between the marker types most likely arose from the superior, near-perfect spherical geometry of the ceramic bearing markers, whereas the 3D-printed markers exhibited minor surface bumps and physical deformations. Nevertheless, subpixel accuracy in the image measurements could not be achieved for either marker type. A similar pattern could be observed regarding the sphere centre estimation via sphere fitting in the voxel space, as displayed in Table 3. While the estimation of the sphere centres reached subvoxel accuracy here for both types, the ceramic workers performed significantly better due to their more accurate geometry.

	Camera №	$\overline{s_x}$ [px]	$\overline{s_y}$ [px]	$\overline{s_r}$ [px]	$\overline{\sigma_0}$ [px]
$\mathbf{H_1}$	1	0.104	0.108	0.077	2.170
	2	0.103	0.106	0.074	2.258
	3	0.105	0.108	0.077	2.170
$\mathbf{H_2}$	1	0.096	0.099	0.072	2.106
	2	0.054	0.053	0.039	1.427
	3	0.034	0.036	0.025	0.950

Table 2. Means of the empirical standard deviations s_i and standard deviation of unit weight σ_0 of the circle fit after image segmentation.

	$\overline{s_x}$ [µm]	$\overline{s_y}$ [µm]	$\overline{s_z}$ [µm]	$\overline{s_r}$ [mm]	$\overline{\sigma_0}$ [mm]
$\mathbf{H_1}$	4.714	5.214	4.742	0.386	0.076
$\mathbf{H_2}$	2.380	2.320	2.440	0.168	0.054

Table 3. Means of the empirical standard deviations s_i and standard deviation of unit weight σ_0 after sphere fitting in voxel space.

3.3 Marker intersection

For the 3D-printed markers attached to $\mathbf{H_1}$, the triple-camera intersection yielded an image measurement standard deviation of unit weight σ_0 of 4.2 µm in the mean of all markers, which corresponds to approximately 1.5 pixels in image space in regards of measurement accuracy. Considering the known disadvantages of these markers, as discussed later in the limitations section, this level of precision is regarded as satisfactory. For the ceramic makers on $\mathbf{H_2}$, we receive a σ_0 of 5.0 µm in the mean, which is equivalent to about 1.8 pixels. The resulting standard deviations of the estimated 3D marker coordinates are summarized in Table 4.

	Marker №	s_x [mm]	s_y [mm]	s_z [mm]	σ_0 [µm]
$\mathbf{H_1}$	1	0.061	0.061	0.077	4.64
	2	0.068	0.067	0.084	5.24
	3	0.027	0.028	0.036	2.13
	4	0.070	0.069	0.088	5.20
	5	0.044	0.045	0.058	3.47
	6	0.062	0.062	0.079	4.65
	7	0.053	0.054	0.068	4.15
	mean	0.055	0.055	0.070	4.21
$\mathbf{H_2}$	1	0.080	0.078	0.102	6.02
	2	0.080	0.079	0.106	6.02
	3	0.067	0.065	0.086	4.86
	4	0.056	0.057	0.077	4.23
	5	0.054	0.053	0.072	3.88
	mean	0.067	0.066	0.089	5.00

Table 4. Empirical standard deviations s_i and standard deviation of unit weight σ_0 of the intersected markers.

3.4 Spatial transformation

When comparing the inner network distances within the two datasets for each head model, we noticed a small difference in scale. This may have been caused by earlier processing steps, such as downsampling the CT data or rounding errors during the conversion of the CT dataset into a point cloud. Another possible reason is a slight error of scale introduced by the CT scanner itself due to calibration issues. Taking that into consideration, we applied an overdetermined 7-parameter Helmert transformation. Initial parameters were obtained utilising the Kabsch algorithm (Kabsch, 1976). It should be noted that $\mathbf{H_1}$ includes seven markers, and therefore offers higher redundancy than $\mathbf{H_2}$, which only provides five markers. After the adjustment, we obtained a σ_0 of 0.04 mm for $\mathbf{H_1}$ and 0.06 mm for $\mathbf{H_2}$, which indicates high consistency between the two datasets. The results are displayed in Table 5 and show a sufficient overall fit of the data.

	s_{tx} [mm]	s_{ty} [mm]	s_{tz} [mm]	$s_{r\omega}$ [rad]	$s_{r\varphi}$ [rad]	$s_{r\kappa}$ [rad]	s_{scale}
$\mathbf{H_1}$	0.040	0.036	0.048	0.018	0.001	0.018	0.001
$\mathbf{H_2}$	0.062	0.068	0.065	0.004	0.002	0.004	0.002

Table 5. Empirical standard deviations s_i of the seven transformation parameters.

4. Limitations and Discussion

A primary limitation of the current proof-of-concept is the absence of a comprehensive stochastic model. In the current stage, all weight matrices are set to the identity matrix. This simplification implies that the achieved accuracies, particularly regarding the spatial camera intersections and the 7-parameter transformation results, must be considered overly optimistic. Our upcoming work will focus on resolving this issue. We intend to implement advanced weighting concepts based on the stochastic outputs derived from the bundle adjustment, circle fit accuracies, sphere fit accuracies in voxel space, and the triple intersection process.

A further challenge encountered during marker segmentation in image space, where suboptimal illumination and contrast negatively impacted marker detection and subsequent sphere fitting. To mitigate this in future iterations, utilising colour-coded markers and performing segmentation within the corresponding image RGB channels may offer a robust solution and yield improved results. Furthermore, reflective surfaces on the head models lead to bright hotspots, which made the edge detection prone to errors (Figure 10).

Additionally, for the current implementation, we refrained from applying correction terms described by Liebold and Maas (2026), as our experimental setup relied on small markers located in central image positions. Nevertheless, integrating these terms in the future could potentially refine the marker centre positions, yielding slightly more accurate image coordinates for the intersection phase. Finally, contrary to our initial expectations, the intersection results for the 3D-printed markers outperformed those of the ceramic markers slightly. Although the ceramic bearings possess a significantly more near-to-perfect spherical geometry, they exhibited problematic reflectivity. This high reflectivity generated generally poorer contrast conditions, which ultimately rendered the image segmentation and circle fitting processes notably less reliable than those for the 3D-printed markers.

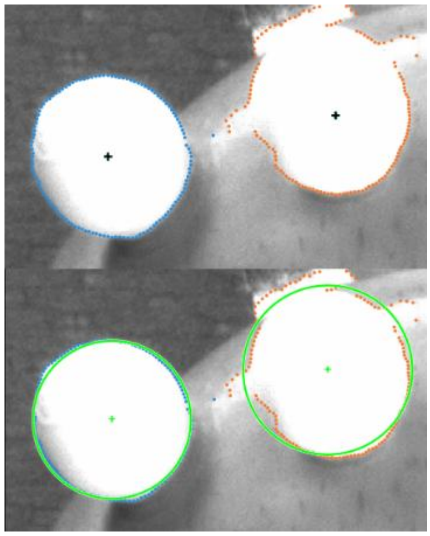

Figure 10. Reflective hotspots and bad lighting conditions complicate the edge detection and circle fitting. Especially for the right marker, a lot of edge outliers occur (top). This leads to worse results for the circle fit (bottom).

5. Conclusion and Prospect

In this work, we were able to realise a combined method based on two marker models that enables the registration of CT voxel data with photogrammetric measurements. Overall, the effectiveness of the proposed workflow could be demonstrated. Nevertheless, the present results must be interpreted with caution, particularly with respect to accuracy measures, since the uncertainties introduced by the chained transformations and measurement steps are not yet modelled explicitly. For future work, a stochastic model that accounts for these effects, for example within a Gauss-Helmert framework, will therefore be employed. At present, the image measurement process is semi-automatic, but there is considerable potential for further automation, extending even to full marker tracking. In this context, the implementation of a real-time tracking approach via a *C++* implementation is a current objective. Although both marker models proved functional in the given experimental environment, each of them exhibited certain disadvantages. Future work will therefore focus on replacing them with an improved marker design. In particular, a retro-reflective marker in combination with ring flash illumination appears to be a promising solution for image measurement. Furthermore, different edge detection algorithms will be evaluated regarding robustness against varying lighting conditions and contrast. Finally, the precision and reliability of the proposed marker-based method will be compared with marker-free approaches based on facial and bony landmarks.

Acknowledgement

This paper was developed within the framework of the project SEMECO-B6 "*Intelligent Patient Positioning in Precision Radiotherapy*"[1] (funding number 100699409). The project was funded by the European Regional Development Fund (ERDF/EFRE) and co-financed by the Free State of Saxony.

Co-funded by the European Union

This project is co-financed from tax revenues on the basis of the budget adopted by the Saxon State Parliament.

[1] https://semeco.info/projekte/intelligent-patient-positioning-in-precision-radiotherapy; last visited 30/03/2026

References

AICON 3D Systems GmbH, 2018. AICON 3D Studio.

Atherton, T.J., Kerbyson, D.J., 1999. Size invariant circle detection. Image and Vision Computing 17 (11), 795–803. https://doi.org/10.1016/S0262-8856(98)00160-7.

Cao, A., Dhanaliwala, A., Shi, J., Gade, T., Park, B., 2019. Image-based marker tracking and registration for intraoperative 3D image-guided interventions using augmented reality. https://doi.org/10.48550/arXiv.1908.03237.

Deegan, T., Owen, R., Holt, T., Fielding, A., Biggs, J., Parfitt, M., Coates, A., Roberts, L., 2015. Assessment of cone beam CT registration for prostate radiation therapy: fiducial marker and soft tissue methods. Journal of Medical Imaging and Radiation Oncology 59 (1), 91–98. https://doi.org/10.1111/1754-9485.12197.

Kabsch, W., 1976. A solution for the best rotation to relate two sets of vectors. Acta Crystallographica Section A 32 (5), 922–923. https://doi.org/10.1107/S0567739476001873.

Liebold, F., Maas, H.-G., 2026. Spherical target eccentricity correction in photogrammetric applications. ISPRS Journal of Photogrammetry and Remote Sensing 231, 761–777. https://doi.org/10.1016/j.isprsjprs.2025.11.022.

Luhmann, T., 2023. Nahbereichsphotogrammetrie. Grundlagen – Methoden – Beispiele, 5th edition. Wichmann, H Wichmann Verlag, Berlin.

Luhmann, T., 1986. Ein Verfahren zur rotationsinvarianten Punktbestimmung. Photogrammetrie - Fernerkundung - Geoinformation Bildmessung und Luftbildwesen 54 (4), 147–154.

Yu, K.L., Xie, Z.Y., 2014. Polynomial Interpolation Sub-Pixel Edge Detection Method Based on Improved Canny Operator. Applied Mechanics and Materials 563, 203–207. https://doi.org/10.4028/www.scientific.net/AMM.563.203.

Zhan, K., Fritsch, D., Wagner, J.F., 2021. Photogrammetry and Computed Tomography Point Cloud Registration using Virtual Control Points, 265–270. https://doi.org/10.5194/isprs-archives-XLIII-B2-2021-265-2021.