Augmented-reality visualization for improved patient positioning workflow during MR-HIFU therapy

Citation for published version (APA):
Manni, F., Ferrer, C. J., Vincent, C. E. C., Lai, M., Bartels, L. W., Bos, C., van der Sommen, F., & de With, P. H. N. (2021). Augmented-reality visualization for improved patient positioning workflow during MR-HIFU therapy. In C. A. Linte, & J. H. Siewerdsen (Eds.), *Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling* Article 115982C (Proceedings of SPIE; Vol. 11598). SPIE. https://doi.org/10.1117/12.2580838

DOI:
10.1117/12.2580838

Document status and date:
Published: 15/02/2021

Document Version:
Accepted manuscript including changes made at the peer-review stage

PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Augmented-reality visualization for improved patient positioning workflow during MR-HIFU therapy

Manni, Francesca, Ferrer, Cyril, Vincent, Celine E., Lai, Marco, Bartels, Lambertus W., et al.

Francesca Manni, Cyril J. Ferrer, Celine E. C. Vincent, Marco Lai, Lambertus W. Bartels, Clemens Bos, Fons van der Sommen, Peter H. N. de With, "Augmented-reality visualization for improved patient positioning workflow during MR-HIFU therapy," Proc. SPIE 11598, Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling, 115982C (15 February 2021); doi: 10.1117/12.2580838

SPIE.

Event: SPIE Medical Imaging, 2021, Online Only

Augmented-reality visualization for improved patient positioning workflow during MR-HIFU therapy

Francesca Manni[a], Cyril J. Ferrer[b], Celine E.C. Vincent[a], Marco Lai[a,c], Lambertus W. Bartels[b], Clemens Bos[b], Fons van der Sommen[a], and Peter H.N. de With[a]

[a]Eindhoven University of Technology, Eindhoven, The Netherlands
[b]Department of Radiology, Image Sciences Institute, University Medical Center Utrecht, Utrecht, The Netherlands
[c]Philips Research, Eindhoven, The Netherlands

ABSTRACT

MR-guided high-intensity focused ultrasound (MR-HIFU) is a non-invasive therapeutic technology which has demonstrated clinical potential for tissue ablation. The application of this therapeutic approach facilitated to be a promising option to achieve faster pain palliation in patients with bone metastasis. However, its clinical adoption is still hampered by a lack of workflow integration. Currently, to ensure sufficient positioning, MR images have to be repeatedly acquired in between patient re-positioning tasks, leading to a time-consuming preparation phase of at least 30 minutes, involving extra costs and time to the available treatment time. Augmented reality (AR) is a promising technology that enables the fusion of medical images, such as MRI, with the view of an external camera. AR represents a valid tool for a faster localization and visualization of the lesion during positioning. The aim of this work is the implementation of a novel AR setup for accelerating the patient positioning during MR-HIFU treatments by enabling adequate target positioning outside the MRI scanner. A marker-based approach was investigated for fusing the MR data with video data for providing an augmented view. Initial experiments on four volunteers show that MR images were overlaid on the camera views with an average re-projection error of 3.13 mm, which matches the clinical requirements for this specific application. It can be concluded that the implemented system is suitable for MR-HIFU procedures and supports its clinical adoption by improving the patient positioning, thereby offering potential for faster treatment time.

Keywords: high-intensity focused ultrasound, MR-HIFU, augmented reality, patient tracking, image fusion.

1. INTRODUCTION

High-intensity focused ultrasound (HIFU) allows ablation of pathological tissues by non-invasively delivering a high level of acoustic energy *in situ*.[1] The technology relies on high-intensity ultrasound waves, focused on a small region, depositing high levels of energy, and results in a local and precise ablation without damaging the surrounding tissue. Recent advances in medical imaging have resulted in a renewed clinical interest on this technology.[2,3] Clinical development has been enabled by integration of magnetic resonance imaging giving guidance to high-intensity focused ultrasound imaging (MR-HIFU), which has shown promising results for pain palliation.[2–4] The MR-HIFU technology is currently being used for palliative treatment of bone metastases, as a second-line treatment after unsuccessful radiation therapy and is occasionally used for tumor control.[5] MR-HIFU provides real-time temperature mapping, enabling the visualization of the thermal dose within the targeted tissue and giving a clear indication of the tissue damage.[5,6] However, improvements and a better workflow integration remain an open issue. Bone metastasis treatment with HIFU consists of several treatment phases.[2] From the available screening CT and MR images, a treatment strategy is envisioned by the radiologist, which aims at finding the optimal acoustic window. However, during the HIFU treatment day, this positioning strategy has to be translated from pre-interventional images to the patient on the HIFU table. This step is of utmost importance, since it will enable to sonicate safely to the target area. Because the radiologist is currently lacking sufficient

Further author information: (Send correspondence to F.M.)
F.M.: E-mail: F.manni@tue.nl

Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling,
edited by Cristian A. Linte, Jeffrey H. Siewerdsen, Proc. of SPIE Vol. 11598, 115982C
 · doi: 10.1117/12.2580838

guidance during the positioning, the patient has to be re-positioned multiple times with repeated acquisitions of MR images in between the re-positioning tasks. After having confirmed the optimal patient positioning, the treatment planning is finalized and the sonication starts. The extra burden of multi-cycle MR image acquisitions and re-positioning tasks lead to a long preparation phase, which is critical for the time budget available for patient treatment. In our positioning approach, 3D real-time image overlay can enable validating patient positioning and facilitate adequate ultrasound beam path imaging without acquiring extra MR images on the treatment day. In this research, we aim to accelerate the patient positioning for MR-HIFU procedures, by implementing a real-time augmented reality (AR) setup that displays the anatomical structures, which are segmented from the MR image on the patient view in relationship with the HIFU beam path. In this context, AR systems were broadly applied to aid surgical guidance by using either marker- or markerless-based approaches. They are used in numerous medical applications to provide patient tracking and an augmented visualization through image fusion.[7,8] Wang *et al.*[9] performed MR image fusion on CT scans by automatically detecting and tracking markers located on the patient skin, achieving a patient-to-image registration. Similarly, in Lai *et al.*[10,11] CT images were overlaid on top of the endoscopic view, by using an optical tracking system for endoscope tracking to conduct experiments on a phantom model, reaching a good accuracy in tracking and rendering the augmented view. Thus, the fusion of MR information on optical cameras by directly tracking markers on a patient, can be used in this specific application to assess the patient position and correct for it. Therefore, this approach can offer an additional advantage in the current clinical practice.

We present a method for MR image fusion on a stereo camera view. This will enable to display the locations of the target lesion, thereby reaching adequate and faster patient positioning. In order to fuse the MR data on top of the stereo camera view, we employ a marker-based image fusion approach, by using fiducial markers attached on the target area. The framework is evaluated and tested on four volunteers enrolled on a human subject study. The results show an accuracy that matches the application requirements, proving the feasibility of fusing MR information on optical cameras in real time and allowing a more direct and faster assessment of the patient position. We can conclude that integrating AR during MR-HIFU can potentially represent a valuable aid for improving the clinical workflow and patient safety.

2. METHODS

A ZED stereo camera[12] (Stereolabs Inc., San Francisco, CA, USA) was employed to perform patient tracking with a marker-based approach. The ZED camera embeds two high-resolution cameras (2560 × 720 pixels) rigidly fixed inside the stereo camera frame. Intrinsic and extrinsic camera parameters are estimated using the ZED proprietary software. The stereo camera enables video recording with a frame rate equal to 30 frames per second (fps), and provides a depth map with a depth range of up to 25 meters. The fiducial markers used in this study are visualized in Figure 1. The markers were printed in black and circular shaped on top of a white background, giving an encapsulated positive MRI contrast agent at the center which enabled dual detection with both optical cameras and MRI. Figure 1a) and Figure 1b) depict the marker and its positioning on the leg of a subject. In the experimental acquisition, seven markers were positioned non-symmetrically on the right leg of each subject, such that at least three markers are always visible from the camera view. Then, the target area was found by detecting the HIFU table position and subsequently obtaining the transducer location. To do so, the ArUco markers were used and fixated to the surface of the HIFU table as shown in Figure 2a) and Figure 2b). After having performed the marker detection using a Canny edge detector, the 3D model of the visible markers was reconstructed in space and the transformation between the camera and the markers could be found. The markers, which were visible from both the MR system and the camera system, were used for the image fusion. The leg of the subject was manually segmented from the MR and the markers, positioned on the surface of the leg, were matched with the markers detected from the video camera. The external views were then combined with the internal 3D view, acquired via the MR system, to construct a 3D augmented-reality view of the combined external and internal views. This process is now specified in a more formal way. The transformation from the MR image to the markers in the optical space is denoted as T_M^{MR} and the transformation from the marker to the camera as T_C^M. Then, the transformation from the MR image to the camera which is represented as T_C^{MR}, can be found by computing:

$$T_C^{MR} = T_M^{MR} \times T_C^M . \tag{1}$$

This transformation between the MR image and the camera system allows to register the MR data with the camera view and to render anatomical structures. The virtual view of the internal structure will be subsequently used to guide the radiologist in better positioning the patient. A schematic diagram of the marker-based workflow is depicted in Figure 3. At the left, the experimental phase is illustrated, resulting in the MRI and image data acquisition. The acquired images are then combined to create the AR view, as shown at the right side in Figure 3. Furthermore, following a similar approach, an acoustic window can be defined at the transducer location and can be overlaid on the camera view, together with the MR image and used for patient positioning. We specify the transformation from the acoustic window to the ArUco markers as T^{W}_{Ar} and the transformation from the ArUco markers to the camera by T^{Ar}_{C}. Then the transformation from the transducer to the camera T^{W}_{C} can be found as follows:

$$T^{W}_{C} = T^{W}_{Ar} \times T^{Ar}_{C}. \tag{2}$$

This transformation allows to overlay the acoustic window on top of the video images.

Figure 1. a) Marker position on the subject leg to be visible on the MR and video acquisitions. b) Magnified view on the marker employed to achieve visual tracking.

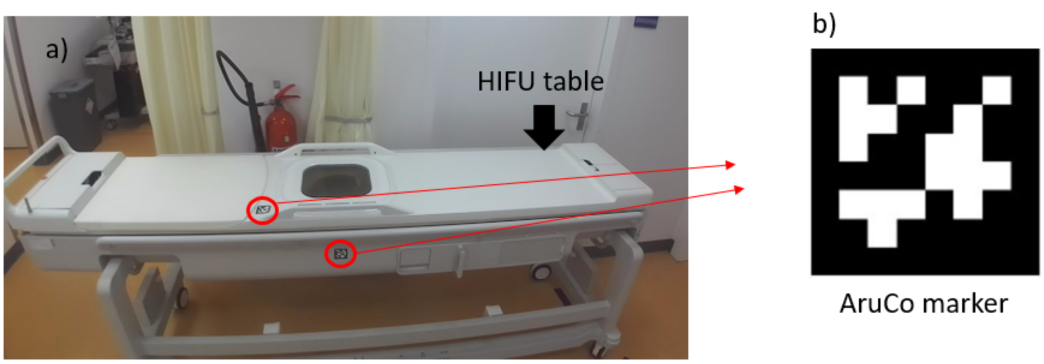

Figure 2. a) HIFU table with the AruCo markers attached to determine the table position. b) Magnified view of an AruCo marker.

2.1 Experimental dataset

The experiments were conducted on volunteers positioned on the HIFU table. Data from four volunteers were gathered at the University Medical Center Utrecht, the Netherlands. Images from the right leg of each volunteer

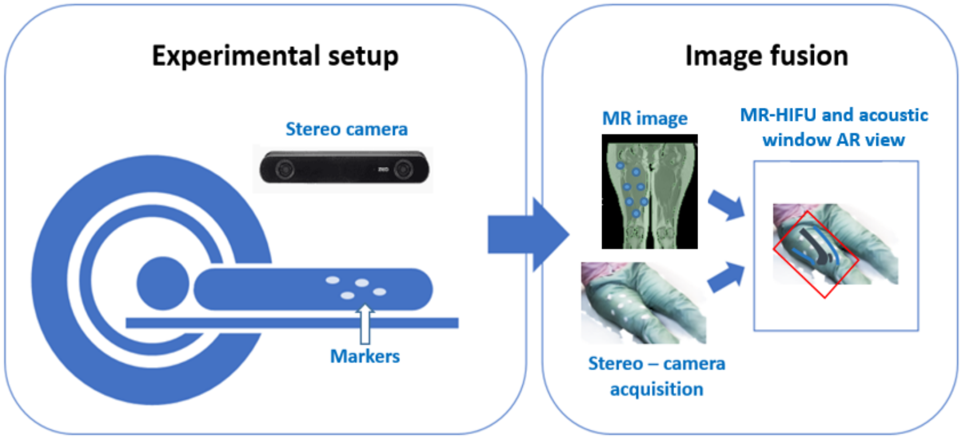

Figure 3. Marker-based approach. Left: experimental setup with the optical cameras mounted at the top of the MR-HIFU table. Right: augmented reality output of the MR video data and the acoustic window fusion.

were recorded and considered as target area. During the acquisitions, the volunteer was asked to place seven MRI markers on the leg. Then, the volunteers laid down on the HIFU table and positioned their legs over the transducer. Each volunteer was asked to somewhat move up and down and to rotate, to simulate movements made during positioning of the patient. Finally, an MRI scan of the lower legs was taken. Figure 4 shows an example of the stereo camera acquisition with a volunteer laying on the HIFU table and the markers used for tracking attached onto the right leg.

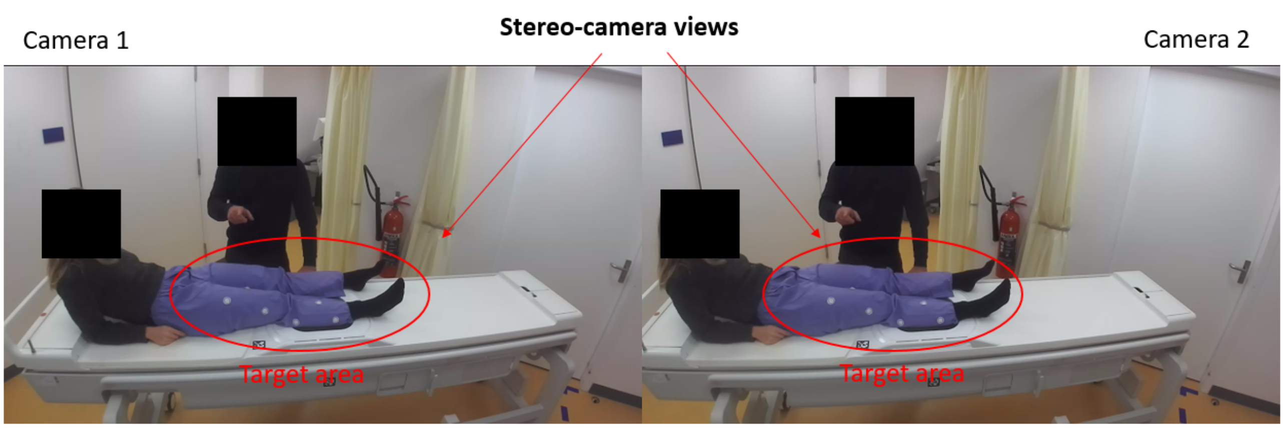

Figure 4. Example of the stereo camera acquisition with the volunteer laying on the HIFU table and the markers attached on the right leg.

3. EXPERIMENTAL RESULTS

3.1 Workflow for achieving the augmentation

The system was tested on the right leg of each subject as a target area. Thus, to simulate a realistic MR-HIFU procedure, the right leg is tracked. Evidently, the same setup can be applied to the left leg. The leg of the patient and the markers positioned on its surface are manually segmented from the MR images. Afterwards, the AruCo markers are attached to the table in close proximity to the transducer area, in order to define the acoustic window location in the camera view. The markers attached to the leg can be seen and eventually tracked by the camera, in order to always define the subject position in space and can serve as ground truth. By matching the markers segmented from the MR images with the 3D markers position in space, the transformation from the MR to the markers is defined, following Equation (1), allowing the MR to be overlaid on the camera view. In the same way, the acoustic window is overlaid on the camera following Equation (2). The MRI augmented view, as well as the acoustic window, are displayed on the camera view and can be used to assess the position of the target area inside the leg with respect to the acoustic window, and eventually for adjusting the leg position prior to starting the HIFU tumor treatment. The MR and the video images are overlaid and displayed as shown in Figure 5a) and Figure 5b) (corresponding to the first and second volunteer), where the segmented model (in yellow) is rendered on top of the video image. The volunteer study was performed with the approval of the institutional review board of the University Medical Center of Utrecht (NL53099.041.15), and written informed consents were obtained from the volunteers.

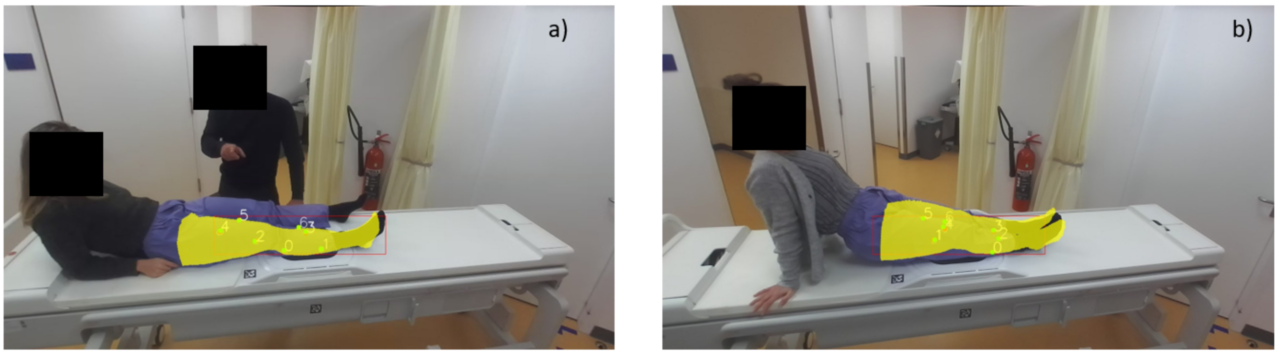

Figure 5. Two examples of overlaid views in a) and b). This involves matching of the 3D markers on a 2D image view with the 3D image fusion on a 2D camera view for the first and second volunteer.

3.2 Statistical analysis of the re-projection error

Preliminary experiments were performed by using the marker-tracking approach from the optical cameras. The markers were segmented from the MR image and overlaid on the stereo views. The accuracy of the overlay is assessed for each subject by calculating the (backward) re-projection error (expressed in mm). For the conversion from pixels to mm, we use the ratio between the AruCo markers in the camera view (expressed in pixels) and their real dimensions (in mm). On average, a re-projection error of 3.13 ± 1.43 mm was found, when tracking the markers and using them to perform the MR image overlay with the camera views. Figure 6 illustrates the boxplots of the errors in the image overlay for each subject. The number of data points includes the error computed from each marker through the video recording. It can be noticed that the error appears to be nearly constant and behaves in stable way among the volunteers. A median error within a range of 3–4 mm was found for the examined subjects. The error can be correlated to the amount of markers visible in the camera view, and to the distance between camera and markers, as well as to the tilting angle of the markers with respect to the image plane of the camera. The upper whisker in Figure 6 increases and reaches its maximum at the third subject of 2.5 mm, while it decreases for the last subject, whereas the lower whisker does not move from the range 1.0–1.5 mm. Therefore, it can be concluded that the system yields a quite high accuracy in MR image overlay on the stereo camera view. The one-way analysis of variance (ANOVA) was used to assess whether there are differences between the means of the studied subjects. It was found that there were no statistically significant differences between the re-projection means of the subjects on the basis of the one-way ANOVA analysis (F-score = 1.47, $p = 0.22$).

4. DISCUSSION

We have presented a first prototype system that combines an optical tracking system based on stereo cameras, with preliminary MR images, to improve patient positioning during MR-HIFU tumor ablation. This approach holds potential to provide a faster and more efficient positioning of the patient, therefore simplifying the current

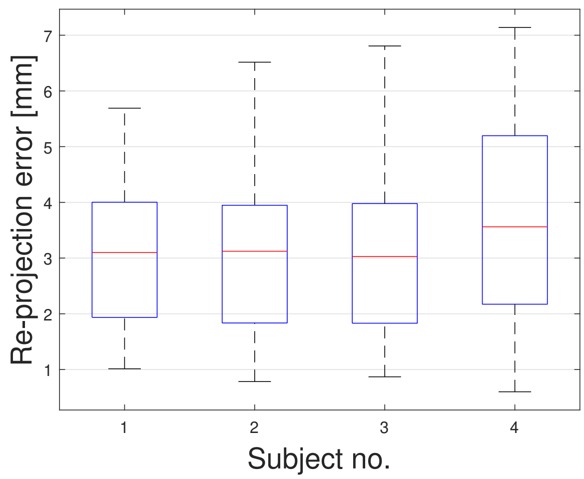

Figure 6. Boxplots of the re-projection errors in the image overlay for the four subjects.

workflow for time efficiency. In the current clinical setting, once the patient is positioned on the table, the operator can operate the HIFU transducer position and move it maximally 1 cm in order to fine-tune the target position to the exact area of the lesion. Consequently, a registration-accuracy error in the order of magnitude of 1 cm would meet the clinical requirements. With our approach, we obtained a re-projection error smaller than 4 mm on average for all the subjects, with an image fusion on the camera view that matches the clinical requirements for this specific application. The system was tested by recording data from the right leg of the volunteers. No statistically significant differences were found on the re-projection errors ($p>0.05$), leading to the conclusion that the error is consistent for the examined subjects. Future studies will involve the investigation on different anatomical locations (i.e., the entire pelvic area, shoulders and ribs).

Prior to proceeding with clinical experiments on patients, the system needs further improvements. A real-time implementation is necessary, to facilitate its use for the patient re-positioning. Also, an automatic or semi-automatic method for leg and lesion segmentation from MR images could be implemented, e.g. following a deep learning approach.[13] Furthermore, the ZED stereo camera can be further exploited, developing a markerless approach based on the depth images provides by the camera. By combining the video and depth images, a point cloud of the scene can be found and overlaid with the 3D MR model. In particular, existing algorithms such as iterative closest point (ICP), allows to find the transformation between two point clouds, that minimizes their distance.[14,15] After having matched the two point clouds, the optimal transformation can be found using the video data to optimally overlay the 3D MR model. All the adopted improvements from the future will be benchmarked with the existing system.

5. CONCLUSIONS

In this study, a first prototype system that improves patient positioning during MR-HIFU tumor ablation was developed. The first preliminary experiments have been successfully conducted and illustrate the feasibility of marker tracking with two optical cameras, in order to facilitate fast and accurate patient positioning in MR-guided HIFU ablation procedures. The image views were then registered with MR images for obtaining the overlay and visualizing anatomical structures with a re-projection error of 3.13 mm on average. We aim to extend the experiments and prove that AR can improve the clinical workflow, thereby preventing errors in patient positioning. The presented method will strengthen the clinical application for MR-HIFU systems, opening up new scenarios for minimally-invasive tumor therapies.

ACKNOWLEDGMENTS

This study was funded by the Applied Data Science Grant, Utrecht University, Utrecht, The Netherlands.

REFERENCES

[1] Zhou, Y.-F., "High intensity focused ultrasound in clinical tumor ablation," *World journal of clinical oncology* **2**(1), 8 (2011).

[2] Scipione, R., Anzidei, M., Bazzocchi, A., Gagliardo, C., Catalano, C., and Napoli, A., "Hifu for bone metastases and other musculoskeletal applications," in [*Seminars in interventional radiology*], **35**(04), 261–267, Thieme Medical Publishers (2018).

[3] Maloney, E. and Hwang, J. H., "Emerging hifu applications in cancer therapy," *International Journal of Hyperthermia* **31**(3), 302–309 (2015).

[4] Ferrer, C. J., Bos, C., de Senneville, B. D., Borman, P., Stemkens, B., Tijssen, R., Moonen, C., and Bartels, L., "A planning strategy for combined motion-assisted/gated mr guided focused ultrasound treatment of the pancreas," *International Journal of Hyperthermia* **36**(1), 701–710 (2019).

[5] Mihcin, S. and Melzer, A., "Principles of focused ultrasound," *Minimally Invasive Therapy & Allied Technologies* **27**(1), 41–50 (2018).

[6] Senneville, B., Mougenot, C., Quesson, B., Dragonu, I., Grenier, N., and Moonen, C., "MR thermometry for monitoring tumor ablation," *European radiology* **17**, 2401–10 (10 2007).

[7] Wu, J.-R., Wang, m.-l., Liu, K.-C., Hu, M.-H., and Lee, P.-Y., "Real-time advanced spinal surgery via visible patient model and augmented reality system," *Computer methods and programs in biomedicine* **113** (01 2014).

[8] Nakao, M., Oda, Y., Taura, K., and Minato, K., "Direct volume manipulation for visualizing intraoperative liver resection process," *Computer Methods and Programs in Biomedicine* (01 2014).

[9] Wang, M. and Song, Z., "Automatic localization of the center of fiducial markers in 3D CT/MRI images for image-guided neurosurgery," *Pattern Recognition Letters* **30**, 414–420 (03 2009).

[10] Lai, M., Skyrman, S., Shan, C., Babic, D., Homan, R., Edström, E., Persson, O., Burström, G., Elmi-Terander, A., Hendriks, B. H., et al., "Fusion of augmented reality imaging with the endoscopic view for endonasal skull base surgery; a novel application for surgical navigation based on intraoperative cone beam computed tomography and optical tracking," *Plos one* **15**(1), e0227312 (2020).

[11] Lai, M., Shan, C., Babic, D., Homan, R., Terander, A. E., Edstrom, E., Persson, O., Burstrom, G., and de With, P. H. N., "Image fusion on the endoscopic view for endo-nasal skull-base surgery," in [*Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling*], Fei, B. and Linte, C. A., eds., **10951**, 350–357, International Society for Optics and Photonics, SPIE (2019).

[12] "ZED camera." https://www.stereolabs.com/zed/.

[13] Deniz, C. M., Xiang, S., Hallyburton, R. S., Welbeck, A., Babb, J. S., Honig, S., Cho, K., and Chang, G., "Segmentation of the proximal femur from mr images using deep convolutional neural networks," *Scientific reports* **8**(1), 1–14 (2018).

[14] Chen, Y., "Medioni," *Object modeling by registration of multiple range images. Im age andV isionC o mput ing* **10**(3), 145–1 (1991).

[15] Zhang, Z., "Iterative point matching for registration of free-form curves and surfaces," *International journal of computer vision* **13**(2), 119–152 (1994).