Verónica García-Vázquez*, Florian Matysiak, Sonja Jäckle, Tim Eixmann, Malte Maria Sieren, Felix von Haxthausen and Floris Ernst

Catheter pose-dependent virtual angiography images for endovascular aortic repair: validation with a video graphics array (VGA) camera

<https://doi.org/10.1515/cdbme-2020-0010>

Abstract: Previous research reported catheter pose-dependent virtual angiography images for endovascular aortic repair (EVAR) (phantom studies) without any validation with video images. The goal of our study focused on conducting this validation using a video graphics array (VGA) camera. The spatial relationship between the coordinate system of the virtual camera and the VGA camera was computed with a Hand-Eye calibration so that both cameras produced similar images. A re-projection error of 3.18 pixels for the virtual camera and 2.14 pixels for the VGA camera was obtained with a designed three-dimensional (3D) printed chessboard. Similar images of the vessel (3D printed aorta) were acquired with both cameras except for the different depth. Virtual angiography images provide information from inside the vessel that may facilitate the understanding of the tip position of the endovascular tools while performing EVAR.

Keywords: electromagnetic sensor; EVAR; Hand-Eye calibration; VGA camera; virtual angiography images.

Introduction

Aortic aneurysms are a potentially lethal disease. Dilatation of the aortic diameter greater than 1.5 times its normal size leads to a weakening of the vessel wall architecture with the potential risk of rupture. Endovascular aortic repair (EVAR), a minimally invasive technique to treat aortic aneurysms, involves the deployment of a stent graft along the aneurysm with the aid of guide wires and catheters to cover the aneurysm and establish a channel for blood. The guidance of the endovascular tools from the femoral or brachial arteries to the aortic aneurysm is conducted with fluoroscopy and conventional digital subtraction angiography. These two-dimensional (2D) image modalities involve radiation exposure and may cause acute kidney injury due to the administration of contrast agent.

To reduce EVAR limitations, a recent study reported the integration of a multicore fiber with fiber Bragg gratings and two electromagnetic (EM) sensors into a stent graft system to locate the shape of the endovascular tool with respect to the patient's anatomy (in that case, a patient-specific aortic phantom) [1]. In addition, virtual angiography images were rendered based on the pose (position and orientation) of an EM-tracked catheter tip [2]. These images of the vessel lumen were visualised on a virtual 2D canvas superimposed on the real scenario with the augmented reality glasses Microsoft HoloLens (first generation). These previous studies [1, 2] are part of our ongoing research project Nav EVAR.

The process for obtaining virtual angiography images is similar to that when generating virtual endoscopies, both based on preoperative computed tomography (CT) scans. EM-tracked bronchoscopies were also reported in previous literature [3, 4]. These studies showed virtual images and also video images acquired with an endoscope. However, in [2] no validation with video images was conducted so the point of view of the virtual camera was not similar to that of a video camera.

The objective of this study was to continue our work from [2] comparing virtual angiography images with video images using an EVAR patient-specific phantom. The

*Corresponding author: Verónica García-Vázquez, Institute for Robotics and Cognitive Systems, University of Lübeck, Ratzeburger Allee 160, Lübeck, Germany, E-mail: garciavazquez@rob.uni-luebeck.de

Florian Matysiak, Division of Vascular and Endovascular Surgery, Department of Surgery, University Hospital Schleswig-Holstein, Campus Lübeck, Lübeck, Germany

Sonja Jäckle, Fraunhofer MEVIS, Lübeck, Germany

Tim Eixmann, Medizinisches Laserzentrum Lübeck GmbH, Lübeck, Germany

Malte Maria Sieren, Department for Radiology and Nuclear Medicine, University Hospital Schleswig-Holstein, Campus Lübeck, Lübeck, Germany

Felix von Haxthausen and Floris Ernst, Institute for Robotics and Cognitive Systems, University of Lübeck, Lübeck, Germany

spatial relationship between the EM sensor (coordinate system of the virtual camera) and a video graphics array (VGA) camera was computed with a Hand-Eye calibration so that the virtual camera and the VGA camera produced similar images.

Materials and methods

An Aurora Mini 6DOF Sensor (1.8 mm diameter $\times$ 9 mm length, Northern Digital Inc.) was attached to an Optikron camera (M-series VGA, USB, 0° view, LED), (Figure 1). The EM field generator chosen in this study was the Aurora Tabletop Field Generator (Northern Digital Inc.) since its thin barrier reduces the distortions produced by ferromagnetic or conductive materials located below it. Regarding the VGA camera, its specifications are as follows: housing 5.95 mm diameter $\times$ 8 mm length, matrix size 640×480 , pixel size 2.5×2.5 mm, maximum frame rate 30 frames per second, horizontal field of view approximately 60° , illumination six light-emitting diodes (LEDs). This camera was selected due to its reduced dimensions so it could move along the phantom vessel.

The virtual angiography images based on the pose of the EM sensor are generated from the vessel segmentation of a preoperative CT scan with the *NavEvar* application developed for our research project. This application was implemented using the medical image processing and visualization software MeVisLab (MeVis Medical Solutions AG). The EM data is sent from the EM tracking system to the *NavEvar* application via the open-source software PLUS Toolkit (<https://plustoolkit.github.io>) [2]. The virtual camera produces images of 512×512 pixels and the field of view (in degrees) can be adjusted. After receiving each pose of the EM sensor, the application also captures an image from the VGA camera.

The same point of view can be acquired with the virtual camera and the VGA camera after finding the transformation sT_c between the EM sensor (s , coordinate system of the virtual camera) and the VGA camera (coordinate system c), (Figure 2). sT_c was obtained with a Hand-Eye calibration, which consists in acquiring a static calibration pattern from different points of view with the VGA camera while recording the pose of the EM sensor at each acquisition (Figure 2). In this study, the calibration pattern was a black-white chessboard with 9×6 inner corners (square size 10×10 mm) printed on a paper. The

Figure 1: Electromagnetic (EM) sensor at the top of the VGA camera.

inputs of the Hand-Eye calibration are the transformations ${}^gT_s^{(i)}$ (which are the i -pose of the EM sensor with respect to the EM field generator coordinate system g) and the transformations ${}^cT_p^{(i)}$ (extrinsic matrixes of the VGA camera that transform a point in the calibration pattern p to the VGA camera coordinate system c). All transformations in Figure 2 are 4×4 matrixes, where gT_p and sT_c are fixed while ${}^gT_s^{(i)}$ and ${}^cT_p^{(i)}$ depend on the pose of the EM sensor and the VGA camera respectively.

${}^cT_p^{(i)}$, the VGA camera intrinsic matrix, its radial distortion (three coefficients) and its tangential distortion were obtained by means of a mono camera calibration with the three-dimensional (3D) coordinates of the inner corners in the calibration pattern coordinate system (p_p , theoretical values, zero z -coordinates) and their corresponding 2D coordinates (u, v)^T estimated from images of the calibration pattern. 60 images were acquired with the VGA camera at different poses, covering the whole image area, a tilt angle up to $\pm 45^\circ$ in both horizontal and vertical directions (to rightly determine the focal length) and also including fronto-parallel images (to rightly determine the lens distortion) [5]. The root-mean-square (RMS) re-projection error was computed with p_p (re-projected) and (u, v)^T to evaluate the VGA camera calibration.

Both the VGA camera calibration and the Hand-Eye calibration were implemented using the Open Source Computer Vision Library (OpenCV). This library includes five different methods of the Hand-Eye calibration [6]: Tsai's (1989), Park's (1994), Horaud's (1995), Andreff's (1999) and Daniilidis's (1999). The solution chosen was the sT_c that gave first a more stable transformation gT_p (Eq (1)) for all acquired images

$${}^gT_p = {}^gT_s^{(i)} {}^sT_c {}^cT_p^{(i)} \quad (1)$$

and second a minimum RMS error of the inner corners of the calibration pattern for all VGA camera poses. Two RMS errors were computed in this case using the errors defined in Eqs (2) and (3)

$$\text{error (mm)} = p_g - {}^gT_p p_p \quad (2)$$

$$\text{error (pixels)} = M_c {}^sT_c^{-1} {}^gT_s^{-1(i)} p_g - (u, v)^T \quad (3)$$

where the 3D ground truth p_g was obtained by means of placing the tip of an EM-tracked pointer at each inner corner of the calibration pattern, and M_c mapped from the VGA camera coordinate system to its

Figure 2: Transformations used in Hand-Eye calibration.

image plane. This last transformation takes account of the VGA camera intrinsic matrix and the distortion coefficients.

After computing the Hand-Eye calibration, the virtual angiography images are generated from the virtual camera with the transformation matrix ${}^{CT}T_c$ (Eq (4)) so that the virtual camera coordinate system matches that from the VGA camera but with respect to the CT coordinate system.

$${}^{CT}T_c = {}^{CT}T_g \quad {}^gT_s \quad {}^sT_c \quad (4)$$

where the 4×4 transformation matrix ${}^{CT}T_g$ from the EM field generator coordinate system to the CT coordinate system was computed with a marker-based registration. Each marker position in the EM coordinate system was acquired with the EM-tracked pointer while its position in the CT coordinate system was the centroid of the region identified as a marker.

After including ${}^{CT}T_g$ and sT_c in the *NavEvar* application, the virtual camera would generate similar images to those with the VGA camera. The approach to assess these settings consisted in re-projecting the 3D coordinates of several points acquired with the EM-tracked pointer (measured RMS error 0.53 mm) on the images of both cameras and, after that, comparing the 2D coordinates with their corresponding ground truth using Eq (3). The camera intrinsic matrix and the distortion coefficients of the VGA camera are known after the Hand-Eye calibration. However, these parameters should be determined for the virtual camera so that the 3D coordinates of a point in the virtual camera coordinate system can be mapped to the 2D coordinates of its image plane. This virtual camera calibration was computed with virtual images acquired from different points of view of a segmented CT scan of a 3D printed chessboard designed for this study (Figure 3, left). The dimensions of the cubes, whose top layer was painted with black ink, are $10 \times 10 \times 10$ mm. This phantom was also used in the assessment of the settings since the inner corners of the black cubes can be identified in the images from both cameras with OpenCV (2D coordinates considered as the ground truth). A CT scan of this 3D printed chessboard was acquired with a Siemens SOMATOM Definition AS+ CT scanner with the following acquisition parameters: voltage 120 kVp, exposure 168 ± 1 mAs (mean $\pm$ standard deviation), slice thickness 0.4 mm and pixel size 0.25×0.25 mm. A segmentation of a very thin layer of the top of the phantom was used to generate the virtual images of just the black squares.

After evaluating the settings, an EVAR patient-specific phantom (specifically, a rigid 3D printed aorta with iliac arteries) (Figure 3, right) was acquired with the virtual and VGA cameras. The virtual angiography images based on the pose of the EM sensor were generated with the lumen segmentation of a CT scan of the rigid aorta. This 3D image was acquired with the Siemens SOMATOM Definition AS+ CT

Figure 3: 3D printed chessboard (left) and rigid 3D printed aorta (right), both phantoms with markers for registration.

scanner and the following acquisition parameters: voltage 80 kVp, exposure 8 mAs, slice thickness 0.6 mm and pixel size 0.59×0.59 mm. In this study, a field of view of 80° was chosen to show more details from inside the vessel.

Results

The VGA camera calibration (chessboard on paper) provided an RMS re-projection error of 0.36 pixels. Horaud's Hand-Eye calibration method was selected since it yielded the most stable transformation gT_p (RMS difference to the mean gT_p : x -axis 0.55° , y -axis 0.58° , z -axis 0.54° and position 0.67 mm) and minimum RMS errors (1.50 mm and 2.80 pixels). Park's provided similar results to the previous method while Tsai's and Daniilidis's results were completely wrong. On the other hand, after registering the 3D printed chessboard with its CT scan (RMS fiducial registration error 0.49 mm), the virtual camera was calibrated for a field of view of 80° (virtual camera), giving an RMS re-projection error of 0.37 pixels.

Regarding the assessment of the settings to generate similar images with both cameras, the RMS re-projection error was 3.18 pixels for the virtual camera and 2.14 pixels for the VGA camera (Figure 4). For both cameras, the data were computed with the 2D coordinates of the 54 inner corners and 18 poses of the EM sensor/VGA camera while acquiring the 3D printed chessboard.

Figure 5 shows images from inside the vessel (rigid 3D printed aorta) acquired with both cameras. In this case, the RMS fiducial registration error was 0.42 mm.

Discussion

A 3D printed chessboard was designed in this study to assess the results of the Hand-Eye calibration, which estimated the transformation between the coordinate

Figure 4: Pointer tip positions re-projected on the images of virtual camera (left, image cropped) and VGA camera (right). 3D coordinates re-projected (red) and ground truth (blue).

Figure 5: Virtual angiography images (left) and images from the VGA camera (right). Vessel lumen (top) and vessel bifurcation (bottom).

systems of the virtual camera (specifically, the EM sensor) and the VGA camera. In some cases, the inner corners of this 3D chessboard could not be identified in the VGA images due to illumination problems (specifically, shadows). Decreasing the cube height may overcome this problem. The re-projection errors were similar to those reported in [3], slightly larger in the case of the virtual camera probably due to the camera calibration process (sources of error: 3D printing process, top layer segmentation and registration). Both re-projections errors also include the errors of the Hand-Eye calibration (larger source of error from our opinion), the EM sensor and the pointer tip.

Similar images of the vessel (rigid 3D printed aorta) were acquired with both cameras except for the different perceived depth. Small differences can be identified, probably due to the lumen segmentation. A limitation of these catheter pose-dependent virtual angiography images is that the endovascular tool often touches the vessel wall so images at that pose will be useless. A combination with the point of view along the vessel centreline is also recommended. Virtual angiography images are generated from a preoperative CT scan that may not resemble the current anatomy during EVAR due to different patient positioning, tissue deformation, respiratory motion and vessel pulsation. Despite this mismatch, these images provide information from inside the vessel, like intravascular ultrasound images (but the latter has no depth information), that may facilitate the understanding of the

position of endovascular tools while performing an EVAR procedure or for EVAR training. Intraoperative imaging may be used to update preoperative data with the current anatomy.

Research funding: This study was supported by the German Federal Ministry of Education and Research (BMBF, Nav EVAR project, number 13GW0228) and the Ministry of Economic Affairs, Employment, Transport and Technology of Schleswig-Holstein.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: Authors state no conflict of interest.

Informed consent: Not applicable.

Ethical approval: Not applicable.

References

1. Jäckle S, García-Vázquez V, Eixmann T, Matysiak F, von Haxthausen F, Sieren MM, et al. Three-dimensional guidance including shape sensing of a stentgraft system for endovascular aneurysm repair. *Int J Comput Assist Radiol Surg* 2020;15:1033–42.
2. von Haxthausen F, Jäckle S, Strehlow J, Ernst F, García-Vázquez V. Catheter pose-dependent virtual angiography images visualized on augmented reality glasses. *Curr Dir Biomed Eng* 2019;5:289–92.
3. Liu SX, Gutiérrez LF, Stanton D. Quantitative evaluation for accumulative calibration error and video-CT registration errors in

- electromagnetic-tracked endoscopy. *Int J Comput Assist Radiol Surg* 2011;6:407–19.
4. Samavati M, Ahmadian A, Abtahi H, Golnabi A, Arjmandi Asl R. A hybrid method for real-time bronchoscope tracking using contour registration and synchronous EMT data. *Iran J Radiol* 2019;16:e66994. <https://doi.org/10.5812/iranjradiol.66994>.
 5. Jakob W. Calibration best practices. Available from: <https://calib.io/blogs/knowledge-base/calibration-best-practices> [Accessed 12 May 2020].
 6. OpenCV team. camera calibration and 3D reconstruction. Available from: https://docs.opencv.org/4.1.1/d9/d0c/group__calib3d.html [Accessed 12 May 2020].