Augmented Reality Assisted Brain Tumor Extraction in Mice

Adrian Schneider(✉), Peter Thalmann, Simon Pezold, Simone E. Hieber, and Philippe C. Cattin

Department of Biomedical Engineering, University of Basel, Basel, Switzerland
adrian.schneider@unibas.ch

Abstract. Computer assisted navigation is a widely adopted technique in neurosurgery and orthopedics. In general, the used tracking systems are applicable to multiple situations. However, these general-purpose devices are costly and in case of unusual laboratory applications, a dedicated solution often shows a better performance. In this paper, we propose a cost-effective 3D navigation system for the augmented reality assisted brain tumor extraction in mice, used for cancer research. Off-the-shelf camera 3D reconstruction algorithms are used to individually track a target and a surgical tool. Relative to its costs, the experiments showed an excellent navigation error of $0.48\,\text{mm} \pm 0.25\,\text{mm}$.

Keywords: Augmented reality · 3D Reconstruction · Single camera · Navigation · Micro computed tomography · Cancer

1 Introduction

According to the *World Health Organization*, cancer is a leading cause of death and its prevalence is increasing [12]. Although state-of-the-art oncology is steadily progressing, one of two patients loses the fight against cancer. Current research is tackling the illness at multiple fronts. The main effort targets the design of new drugs, proliferation, the improvement of radiotherapy methods, and the development of sophisticated surgical interventions. To improve the effectiveness of cancer therapies, a better understanding of cancer is highly important. For example, vessel parameters such as diameter and tortuosity are suspected to play a crucial role in the angiogenesis of cancer and therefore also for anti-angiogenic therapies. For the quantification of these vessel parameters, sophisticated 3D imaging techniques are necessary.

In vivo measurements are highly desired, but do not yet provide sufficient resolution. Furthermore, typical absorption-contrast micro-computed tomography (μCT) is not applicable as the contrast for soft tissues is too small for segmentation, and magnetic resonance tomography (MRT) is not yet able to visualize the smallest capillaries (diameter $\sim 5\,\mu\text{m}$) due to lack of resolution.

Simone E. Hieber and Philippe C. Cattin shared last authorship.

L.T. De Paolis and A. Mongelli (Eds.): AVR 2015, LNCS 9254, pp. 255–264, 2015.
DOI: 10.1007/978-3-319-22888-4_19

In the mouse model, one approach to overcome these issues is to use vascular corrosion casting, where the mouse is perfused with Heparin, followed by a polyurethane mixture as described in [6]. The remaining tissue is removed from the polymer specimen with a formic acid solution. Following a standard protocol, synchrotron radiation-based micro-computed tomography (SRμCT) in absorption-contrast mode [8] or, as shown recently, high-resolution laboratory CT [11] is subsequently used for imaging the specimen. However, this approach is only reliable for tumors at early stages.

A second approach is to use in-line phase-contrast SRμCT [10], a technique known for much better discrimination of soft tissues compared to standard absorption-contrast μCT, even without staining. For this technique, however, the spatial resolution of the acquired tomograms highly depends on the specimen size, such that scanning a smaller object enables achieving a higher resolution. The detectors at the synchrotron beamlines typically deliver an image with a fixed size of 2000×2000 pixel. Thus, when scanning the whole mouse brain of about 15 mm size, a spatial resolution of 7.5 µm can be reached. This is not enough to visualize the tumor's capillaries. An obvious solution is to measure only the brain part in which the tumor is located. In laboratory mice, the tumors of interest reach a diameter of approximately 2 mm. Extracting them from the brain into specimen of the size of 3 mm, the spatial resolution is increased by a factor of 5 and reaches 1.5 µm. This enables the visualization of the smallest capillaries, which feature diameters on the level of micrometers.

In this paper, we describe a novel approach to perform such a tumor extraction based on the combination of MRT and computer vision. In particular, a dedicated high resolution MRT device is used to localize the tumor within the mouse brain. Then, a single video camera is used to simultaneously track the brain and a dissection tool. To increase usability and enable a seamless integration into the surgical workflow, an intuitive augmented reality (AR) visualization technique is used. Figure 1 depicts an overview of the whole workflow.

Existing medical 3D navigation systems are highly adaptable and can be used for a multitude of navigation tasks. An accurate and common tracking technology is based on optical stereo frames operating in the infrared spectrum. Such systems achieve submillimeter accuracy in a relatively large measurement volume [2]. Therefore, they are good potential tracking solutions for our application, too. However, these devices are bulky and expensive. By contrast, we are going to present a compact navigation system that achieves comparable results for a fraction of the costs.

2 Materials and Methods

2.1 Specimen Preparation

The tumor samples are gained from mice. At the age of two months, *gliome murine cells* (GL261) are injected into the brain of a mouse. The mouse is sacrificed 12 days after the injection. At this stage, the tumor has reached a

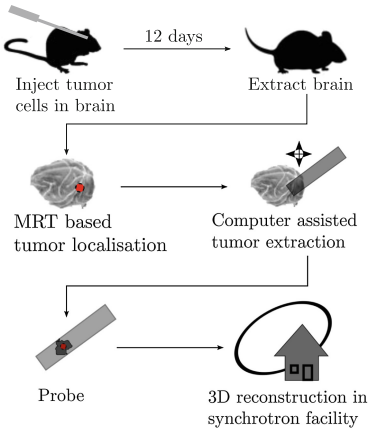

Fig. 1. Workflow from farming the tumor up to 3D reconstruction. This paper is mainly about the tumor localization and extraction. The red dot indicates the tumor (Color figure online).

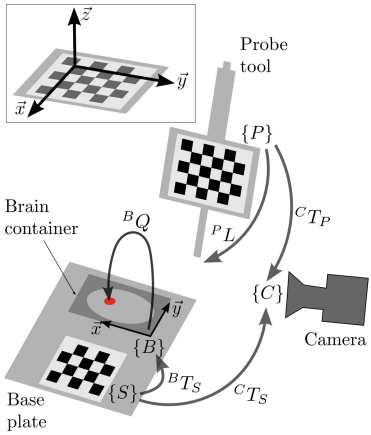

Fig. 2. Setup of the navigation system. Curved arrows denote affine transformations and the red dot indicates the tumor. In the upper left, a chessboard and its spanned coordinate system are shown (Color figure online).

diameter of approximately 2 mm. The target region – cancerous cells – is then extracted from the brain with our newly developed navigation system.

2.2 Navigation System

The setup of the proposed navigation system is shown in Fig. 2. The base plate carries a container for the whole mouse brain. To trepan the tumor from the mouse brain, the probe tool is used. A third part is a camera, which establishes an accurate spatial relationship between the base plate and the probe tool.

In the following, we are going to describe each part of the navigation pipeline in detail, covering deployed materials and algorithms. To familiarize the reader with our nomenclature, we would like to explain briefly the term *affine transformation*. An affine transformation describes a rotational and translational relationship between two coordinate systems (CS) and can be compactly represented by a 4×4 matrix in the 3D case. The notation ${}^{X}T_{Y}$ denotes an affine transformation matrix that transforms a point expressed in the $\{Y\}$ CS ${}^{Y}P$ into a position relative to the $\{X\}$ CS ${}^{X}P$. The actual transformation from ${}^{Y}P$ to ${}^{X}P$ is carried out using matrix–vector multiplication:

$$ {}^{X}P = {}^{X}T_{Y} \cdot {}^{Y}P. $$

2.3 Video Camera

The *MQ013CG-E2* digital camera from *XIMEA GmbH* (Münster, Germany) was used, which delivers images of the size of 1280×1024 pixel. As a lens, the *002915* from *Tamron* (Saitama, Japan) with a focal length of 8 mm is used. Using the notation introduced in Fig. 2, the camera image is processed in order to recover the two dynamic affine transformations ${}^{C}T_{P}$ and ${}^{C}T_{S}$, where ${}^{C}T_{P}$ is the transformation from the probe tool to the camera and ${}^{C}T_{S}$ is the transformation from the base plate to the camera.

To understand the methods of estimating a 3D transformation from a single camera image, it is necessary to have a closer look at the pinhole model, which is used to describe a projective camera. The projective transformations for ${}^{C}T_{P}$ and ${}^{C}T_{S}$ are defined as

$$ {}^{P}u \propto K \cdot {}^{C}T_{P} \cdot {}^{P}X \quad \text{and} \quad {}^{S}u \propto K \cdot {}^{C}T_{S} \cdot {}^{S}X, $$

where K are the intrinsic camera parameters, ${}^{P}u$ is the 2D pixel location of the projected 3D point ${}^{P}X$, and ${}^{S}u$ is the respective projection of ${}^{S}X$. Those pairs are called 2D–3D point correspondences and can be robustly created and recovered by using simple and detectable patterns. The $\propto$ sign indicates proportionality, because the projected pixel coordinates are usually normalized to $[u, v, 1]^{T}$.

Initially, K has to be determined. These parameters describe the projective behavior of the camera and are determined only once. A common camera calibration method is based on using multiple 2D–3D point correspondences created from planar calibration patterns [13], e. g. chessboard patterns. In addition to K, this calibration algorithm also estimates non-linear lens distortion parameters, such as *tangential distortion* and *radial distortion*.

Tracking an object with a calibrated camera is very similar to the calibration process itself. The 2D–3D point correspondences are created by extracting the 2D coordinates from the camera image and connecting them with the chosen 3D coordinates of the pattern. Based on these correspondences, an iterative method using Levenberg-Marquardt optimization is applied to estimate the extrinsic camera parameters, in particular the affine transformations ${}^{C}T_{P}$ and ${}^{C}T_{S}$.

Camera Calibration Error and Tracking Error. The camera calibration [13] and the object tracking algorithm [3] both have residual errors. Several error measures can be applied to quantify the quality of the found solution [4]. A commonly used measure is the *backprojection error* E_B, which represents the average pixel error in the image: The computed solution is used to project the i-th 3D point ${}^{Y}X_i$ of the 2D–3D point correspondences to the image plane. These projections u_i' are then compared with the recorded 2D pixel coordinates u_i taken from point correspondences. If there are N 2D–3D correspondences, then

$$ E_B = \frac{1}{N} \sum_{i=1}^{N} \left\| u_i - u_i^{'} \right\| \quad \text{with} \quad u_i^{'} \propto K \cdot {}^{C}T_{Y} \cdot {}^{Y}X_i. $$

In general, the calibration error and tracking error depend heavily on the applied equipment and the scene itself. For further detail, we refer the reader to [7], where a comprehensive error estimation for single camera tracking systems is described.

2.4 Probe Tool

The probe tool is used to trepan the tumor from the mouse brain. The user brings the tool into the right pose by following the guidance instructions of our navigation system (Sect. 2.6).

In reference to Fig. 2, the missing tool tip position ${}^{P}L$, expressed in the $\{P\}$ CS, can be found by the algorithm described in [9]. However, since the exact orientation of the tool's chessboard CS is known, a better method to determine ${}^{P}L$ is simply measuring the translation, e. g. by using a sliding calliper.

2.5 Base Plate and Brain Container

The complete brain of the mouse is surgically removed and put into the brain container. The location of the tumor Q is measured using a dedicated high resolution MRT device (*PharmaScan 47/16, Bruker BioSpin*, Ettlingen, Germany). Using the container shape, the MRT origin can be aligned with the CS $\{B\}$ easily, at the same time providing the tumor location ${}^{B}Q$ in $\{B\}$. To enhance the visibility of the container itself on the MRT image, we place it in a water bath during the scan.

The base plate has a slot to plug in the brain container. Since both geometries and the orientation of the base plate's chessboard are known, the rigid transformation ${}^{B}T_{S}$ can be measured very precisely. In a more complex situation, e. g. if the construction data of the base plate is not available, one could apply a 3D–3D registration as described in [1].

2.6 Augmented Reality Assisted Navigation

Once all affine transformations of the navigation system are known, the remaining task is to generate the navigation view. Since we use an AR user interface, it makes sense to choose the camera as the common CS. In reference to Fig. 2, the probe tool tip ${}^{P}L$ and the tumor location ${}^{B}Q$ are transformed to $\{C\}$ by

$$ {}^{C}L = {}^{C}T_{P} \cdot {}^{P}L \quad \text{and} \quad {}^{C}Q = {}^{C}T_{S} \cdot {}^{B}{T_{S}}^{-1} \cdot {}^{B}Q. $$

The tumor location ${}^{C}Q$ is visualized on the camera image using the camera calibration. Furthermore, the user gets updated with the current distance from the tool tip to the tumor, $d_T = \left\| {}^{C}L - {}^{C}Q \right\|$, by means of a dynamically scaled vertical bar that we overlay on the camera image. Figures 3 and 4 give an impression of the AR assisted navigation.

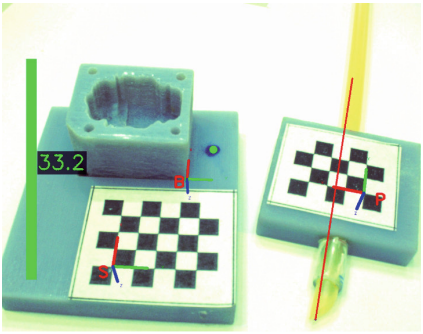

Fig. 3. AR visualization I: The base plate with the brain container is on the left; a vertical green bar displays the distance d_T, the augmented green dot indicates a simulated tumor location. The probe tool is on the right (Color figure online).

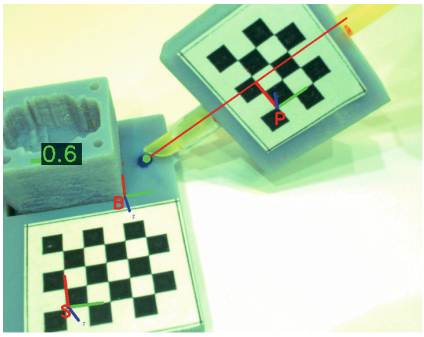

Fig. 4. AR visualization II: The chessboard coordinate system frames $\{B\}$, $\{S\}$, $\{P\}$, and the probe tool's orientation are overlaid. The distance d_T becomes smaller with the tool tip approaching the tumor location.

3 Experiments and Results

In this section, we determine the precision and accuracy of our method under several aspects. The used hardware is as presented in Sect. 2. For the following experiments, a camera calibration with a backprojection error $E_B = 0.2$ pixel was used. The calibration was performed based on 20 different chessboard poses and setting the tangential distortion to zero.

3.1 Transformation Robustness

In this experiment, our goal is to estimate the robustness of the proposed tracking method. In particular, two chessboards with a known, rigid transformation in between them are simultaneously tracked (Fig. 5). The estimated transformation between both chessboards can be quantitatively compared against the true one. This is done from several different camera poses.

The first chessboard has 4×5 fields and spans the CS $\{X\}$. The second board has 4×3 fields and spans the CS $\{Y\}$. Both have square fields, each with a length of 3 mm. The rigid transformation ${}^{X}T_{Y}$ between the two chessboard CSs was chosen so that there is no rotation, but only a translation of 30.0 mm. The transformations from each chessboard to the camera CS $\{C\}$, ${}^{C}T_{X}$ and ${}^{C}T_{Y}$, are recovered from the camera image. Therefore

$$ {}^{X}T_{Y} = {}^{C}T_{X}{}^{-1} \cdot {}^{C}T_{Y}. $$

In the following, this transformation is estimated from $N_q = 25$ different camera poses. Each transformation ${}^{X}T_{Y\,q}$ is compared with the true transformation ${}^{X}T_{Y\,t}$.

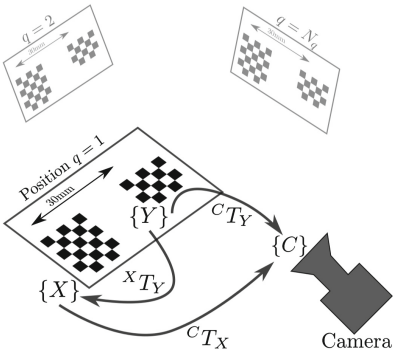

Fig. 5. Transformation robustness experiment setup.

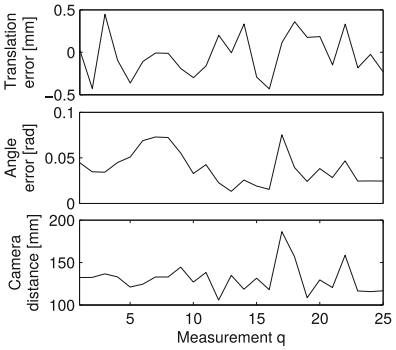

Fig. 6. Transformation robustness experiment results.

The differences can be split into a rotational part $\Delta\Theta_q$ and a translational part Δt_q. We compute $\Delta\Theta_q$ by the *inner product of unit quaternions* [5] and Δt_q by the difference of the CS offsets, as

$$\Delta\Theta_q = \arccos\left(\left|q\left(r\left({}^{X}T_{Y\,t}\right)\right)\cdot q\left(r\left({}^{X}T_{Y\,q}\right)\right)\right|\right), \quad \Delta t_q = \left\|t\left({}^{X}T_{Y\,t}\right)\right\| - \left\|t\left({}^{X}T_{Y\,q}\right)\right\|.$$

The function $q(.)$ converts a rotation matrix into a 4×1 quaternion, $r(.)$ extracts the 3×3 rotation matrix from the transformation and $t(.)$ returns the 3×1 translation component.

The distance d_q between the camera and the tracked pattern has an influence on the accuracy [7]. In our case, d_q is the average distance from both chessboards to the camera and is computed like

$$d_q = \frac{1}{2}\left(\left\|t\left({}^{C}T_{X\,q}\right)\right\| + \left\|t\left({}^{C}T_{Y\,q}\right)\right\|\right).$$

Results The results can be seen in Fig. 6. The mean translation error is 0.20 mm and has a standard deviation of 0.25 mm. The average rotation error is approximately 0.035 rad (2°). The experiments were performed within a camera distance range of 100 mm to 200 mm. However, the supposed correlation between the camera distance and the corresponding errors is not observed. The model for a theoretical error estimation of a single camera navigation system [7] does not consider an equivalent situation, thus a comparison is difficult. But it shows that the determined rotation errors and translation errors are plausible.

3.2 Navigation System Accuracy

In the above experiment, we determined the precision and robustness of tracking two chessboards. Regarding the proposed system (Fig. 2), this only corresponds

to the transformations ${}^{C}T_S$ and ${}^{C}T_P$. In the following, the accuracy of the whole navigation system is assessed.

The transformations ${}^{B}T_S$ and ${}^{P}L$ are known from the construction data or are determined by physical distance measurements. ${}^{C}T_S$ and ${}^{C}T_P$ are estimated based on the camera image. We determine the overall navigation accuracy by defining a known point in the CS $\{B\}$, denoted as ${}^{B}Q$, and by pointing the tip of the probe tool at it. This correlates with the actual task of a navigated tumor extraction. Figure 4 is a snapshot of this experiment. The green dot is the point ${}^{B}Q$. By using the complete affine transformation chain of our navigation system, the probe tool tip's location ${}^{P}L$ is then transformed to ${}^{B}L$, the tip's position expressed in the CS $\{B\}$. The Euclidean distance of the computed tip position ${}^{B}L$ from its reference location ${}^{B}Q$ can be considered as the navigation error E_N. It is given by

$$E_N = \left\| {}^{B}L - {}^{B}Q \right\| \quad \text{with} \quad {}^{B}L = {}^{B}T_S \cdot {}^{C}T_S{}^{-1} \cdot {}^{C}T_P \cdot {}^{P}L.$$

The marker for the base plate is a 4×5 chessboard that spans the CS denoted as $\{S\}$. The probe tool is tracked via a 4×3 chessboard that spans the CS denoted as $\{P\}$. Both chessboards have square fields, each with a length of 3 mm.

Results The experiment was repeated 25 times within a camera distance range of 100 mm to 200 mm. The average navigation error E_N is 0.48 mm with a standard deviation of 0.25 mm. This corresponds to the accuracy observed with commercial medical navigation systems using a pointing tool [2]. However, our measurement volume is much smaller but adequate for the problem at hand.

3.3 Brain Extraction

The navigated tumor extraction (Fig. 7) was tested on two mice and two mock objects, where the tumor was built from modeling clay surrounded by soft paraffin simulating the brain tissue. An experiment is qualitatively rated following to the amount of the tumor which was trepanned from the brain. We introduce the labels *Complete*, *Partial* and *Off Target* tumor extraction.

Results. The two experiments with the mock objects resulted in *complete* tumor extractions. Both experiments with mice led to *partial* tumor samples (Fig. 8). This difference could be explained with the observation that the real brains slightly moved and got deformed at the moment the probe tool punctured the brain surface, whereas this was not the case with the mock objects. In general, it is important to notice that our navigation system is on the edge of meeting the required accuracy for a *complete* tumor extraction. Sampling a target with a 2 mm diameter using a sampling tool that has a diameter of 3 mm requires an accuracy ≤ 0.5 mm. The experiment above showed that we reach $0.48\,\text{mm} \pm 0.25\,\text{mm}$.

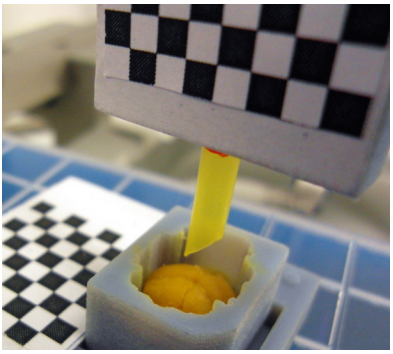

Fig. 7. Tumor extraction from a real mouse brain using the proposed navigation system. The experiment took place at the Animal Imaging Center of ETH Zurich.

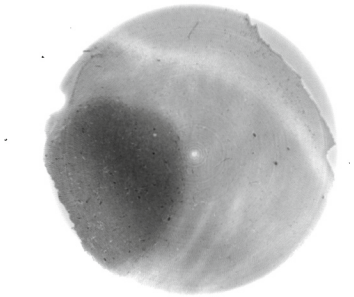

Fig. 8. Reconstructed slice with 3 mm diameter using inline phase contrast SRμCT (pixel size 1.9 μm). The dark area represents the partially extracted tumor.

4 Conclusion

In this paper we proposed an AR navigation system applicable for guided brain tumor extraction in mice. The achieved accuracy of 0.48 mm combined with its relatively low cost opens up possibilities for using the system in other fields. However, for the presented task it would be useful to enhance the accuracy. The current system is operated with one camera and the tracked patterns are planar. A better accuracy can be achieved by using two cameras, which together form a stereo frame. The drawback with the latter approach is the reduced field of view, because both cameras need to spot the patterns. Another solution could be the usage of non-planar patterns. In theory, this enhances the 3D reconstruction accuracy [7]. The disadvantage is the complexity of constructing such a pattern. Our current research is focusing on this challenge.

Acknowledgments. We would like to thank Bert Müller for his initiative and helpful discussions, Therese Bormann for her help in the container design and the 3D printing at the FHNW Muttenz, Mathias Griessen and Alex Seiler for their support in the software implementation, Sandra Burgi for preparing the murine brain, and Marco Dominietto for his useful feedback and the MRT measurements at the Animal Imaging Center of ETH Zurich.

References

1. Arun, K.S., Huang, T.S., Blostein, S.D.: Least-squares fitting of two 3-d point sets. IEEE Trans. Pattern Anal. Mach. Intell. **5**, 698–700 (1987)
2. Broers, H., Jansing, N.: How precise is navigation for minimally invasive surgery? Int. Orthop. **31**(1), 39–42 (2007)

3. Gao, X.S., Hou, X.R., Tang, J., Cheng, H.F.: Complete solution classification for the perspective-three-point problem. IEEE Trans. Pattern Anal. Mach. Intell. **25**(8), 930–943 (2003)
4. Hartley, R., Zisserman, A.: Multiple view geometry in computer vision. Cambridge University Press, New York (2010)
5. Huynh, D.Q.: Metrics for 3D rotations: comparison and analysis. J. Math. Imag. Vis. **35**(2), 155–164 (2009)
6. Krucker, T., Lang, A., Meyer, E.P.: New polyurethane-based material for vascular corrosion casting with improved physical and imaging characteristics. Microsc. Res. Tech. **69**(2), 138–147 (2006). http://dx.doi.org/10.1002/jemt.20263
7. Luhmann, T.: Precision potential of photogrammetric 6DOF pose estimation with a single camera. ISPRS J. Photogrammetry Remote Sens. **64**(3), 275–284 (2009)
8. Müller, B., Lang, S., Dominietto, M., Rudin, M., Schulz, G., Deyhle, H., Germann, M., Pfeiffer, F., David, C., Weitkamp, T.: High-resolution tomographic imaging of microvessels. Proc. SPIE - Int. Soc. Opt. Eng. **7078**, 70780B-1–70780B-10 (2008). http://dx.doi.org/10.1117/12.794157
9. Onprasert, O., Suthakorn, J.: A novel method on tool tip calibration for biomedical application. In: Proceedings of the World Congress on Computer Science and Information Engineering, pp. 650–653 (2011)
10. Paganin, D., Mayo, S.C., Gureyev, T.E., Miller, P.R., Wilkins, S.W.: Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object. J. Microsc. **206**(1), 33–40 (2002). http://dx.doi.org/10.1046/j.1365-2818.2002.01010.x
11. Thalmann, P., Hieber, S.E., Schulz, G., Deyhle, H., Khimchenko, A., Kurtcuoglu, V., Olgac, U., Marmaras, A., Kuo, W., Meyer, E.P., Beckmann, F., Herzen, J., Ehrbar, S., Müller, B.: Three-dimensional registration of synchrotron radiation-based micro-computed tomography images with advanced laboratory micro-computed tomography data from murine kidney casts. Am. J. Physiol. Heart Circ. Physiol. **9212**, 92120Y-1–92120Y-9 (2014). http://dx.doi.org/10.1117/12.2060809
12. WHO: Research for patient safety (2008). http://www.who.int/patientsafety/information_centre/documents/ps_research_brochure_en.pdf
13. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. **22**(11), 1330–1334 (2000)