Polish Annals of Medicine

Journal homepage: https://www.paom.pl

Research paper

Interactive CT/MRI 3D Fusion for cerebral system analysis and as a preoperative surgical strategy and educational tool

Michał Chlebiej[1], Andrzej Rutkowski[1], Anna Żurada[2], Jerzy Gielecki[3], Katarzyna Polak-Boroń[3]

[1] *Faculty of Mathematics and Computer Science, Nicolaus Copernicus University in Toruń, Poland*
[2] *Department of Radiology, Collegium Medicum, School of Medicine, University of Warmia and Mazury, Olsztyn, Poland*
[3] *Department of Anatomy, Collegium Medicum, School of Medicine, University of Warmia and Mazury, Olsztyn, Poland*

ARTICLE INFO

Article history
Received 2 November 2021
Accepted: 29 November 2021
Available online: 20 December 2021

Keywords
Virtual reality
Augmented reality
Magnetic resonance imaging (MRI)
Computed tomography (CT)
Cerebral vessels segmentation
3D image fusion

Doi
https://doi.org/

ABSTRACT

Introduction: The development of systems that merge existing technologies with gathered data may bring some spectacular effects that are usable both in preoperative and educational processes. Augmented reality (AR) is one of the key aspects of the new medical approach. Newly fused data sets draw from it and give users a better overall experience.

Aim: The main goal of this study was to enable the interactive presentation of patients' CT and MRI combined data with the incorporation of AR tools considering the accuracy of the data with an emphasis on vascular structures.

Material and methods: The registration method, reconstruction of the vascular system using tubular structures, and error analysis using surface distance measurements results were used in the system to provide accurate combined information about bony structures from CT volume and vascular objects and cerebral vessels from MRI.

Results and discussion: The strategies concern a series of CTI volumes that could be used to analyze bony surgical procedures. The methods are preferred, especially to the most complicated and individually modified bony structures of the skull. Removing, replacing, or modifying these bony structures or elements of the skull could be used as an analysis of operating procedures at the particular stages of the operation during neurosurgical or otolaryngological techniques.

Conclusions: Presented study regarded to an innovative system consisting of a CT and MRI datasets fusion. The distance analysis of the segmented vascular model and proposed method for stabilization of the human head combined with virtual sculpting technique. In conclusion, it was meaningful in many aspects of the scientific-technological merge.

Corresponding author: Katarzyna Polak-Boroń, Jagiellońska 52/11, 10-283 Olsztyn, Poland. Tel.: +48 889261947. E-mail address: Katarzyna.polakanat@gmail.com

1. INTRODUCTION

The fusion-based approach was successfully used in radiosurgery, neurosurgery, and hypofractionated radiotherapy.[1] It could also extend anatomical knowledge by analysing anatomical data presented in a novel way. Although, widely discussed the subject still needs to be investigated due to the great attention from medical researchers. When working with combined data, it is crucial to match datasets for the fusion process at first. The image registration area deals with this issue, and many authors discussed it in the literature.[2–4] The second fundamental part is determining what information to combine in the fusion process. There are three levels of the fusion process that can be considered: pixel/signal level, object/feature level, and symbolic/decision level.[5] This study followed the object fusion paradigm which meant extracting features (i.e., segmented masks), reconstructing tubular objects and anatomical spatial points, and delivering the combined information to the user. Perfect automatic segmentation does not exist and probably never will. In sophisticated tasks, manual intervention is often needed. Because of that, there is a demand for interactive tools to tune segmentation results. Most implementations use 2D perpendicular projections and manual selection of pixels.

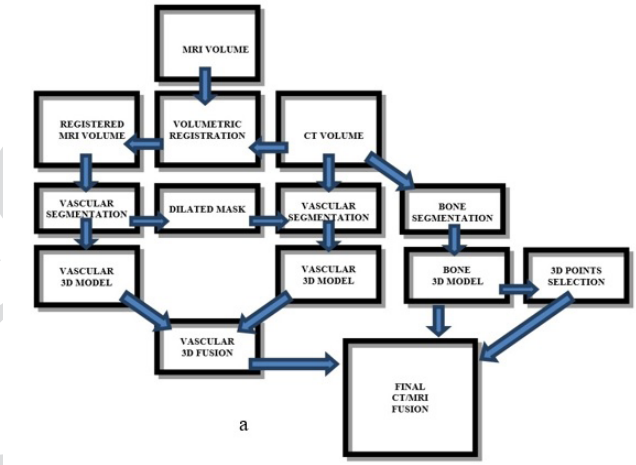
2. AIM

The primary goal of the medical image fusion method was to create a system that would be able to process combined information delivered by varied imaging modalities so that the precise visualization of the segmentation of vascular structures could be acquired. There are many segmentation approaches for brain vessel segmentation and reconstruction described in the literature, and they fall into two categories – voxel-based methods and machine learning methods. The review has not been provided by the authors. It was published by Lesage et al.[6] Several open-source libraries such as VMTK7 and Tube TK8 provide API functions for implementing vessel segmentations. The problem arises when regarding segmenting vascular structures from CT data immersed in bony structures. In such cases, general methods available in software frameworks for segmentation and reconstruction tend to fail. This study challenged these limitations.

3. MATERIALS AND METHODS

Twenty CT and MRI DICOM data examinations were reviewed. According to Bavirisetti et al.,[5] we can consider the fusion process at three levels: pixel/signal level, object/feature level, and symbolic/decision level. At intensity level, this study combined volumetric information creating a new intensity value I_3 for two input values I_1 and I_2 from registered multimodal datasets and every voxel. This kind of fusion has been widely discussed in many scientific papers and surveys.[5,9–14] The second level uses objects, labels, property descriptors derived from several sources.[15] The last method merges information on high-level. It uses local decision markers derived from objective-level fusion results for probabilistic decisions for information extraction and fusion.[16] This study followed the object fusion paradigm which meant extracting features (i.e., segmented masks), reconstructing tubular objects and anatomical spatial points, and delivering the combined information to the user. This work utilizes our reconstruction method for tubular reconstruction enabling customization of procedures proposed in Nowinski et al.[17]

The new method consisted of several stages where first was to deal with MRI and CT datasets with gadobutrol (1-molar gadolinium-based) contrast agent (Figure 1a). On both volumes, vascular structures were visible, but in CT, they could not be separated from bone structures based on



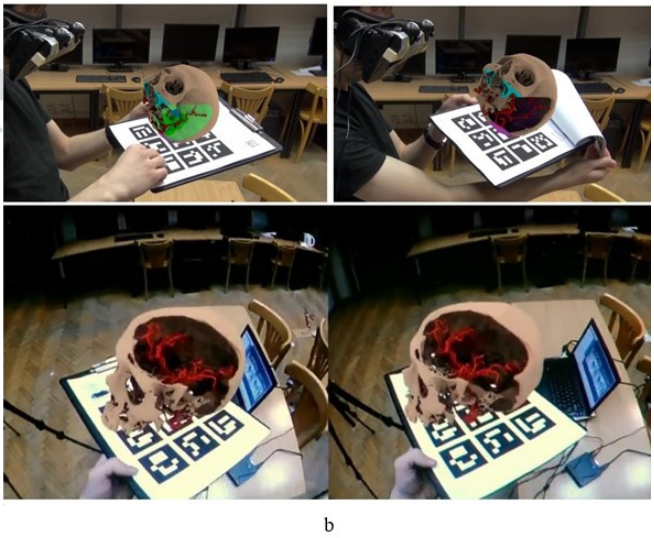


Figure 1. (a) Data flow diagram of fusion system. (b) Case studies of AR rendering system. System generated top images when connected to an external camera for presentation purposes. It represents what would be seen by the user from the point of view of that camera. A visible person is a user wearing an AR headset and operating a marker board in hands. The bottom stereo image shows a direct view of images as seen through HMD.

local image intensity values. Especially arteries immersed in bony structures like the internal carotid artery (ICA) could not be adequately segmented manually. Because of that, volumetric fusion had to be proposed which allowed for MRI cerebral system extraction and immersion with CT volumetric bone marrow. MRI modality based on T2 weighted imaging and was created with the performance of the three-dimensional turbo spin-echo (3D TSE) technique protocol. For CT modality the Aquilion PRIME Toshiba scanner with 80 detector row system to deliver advanced iterative reconstruction and Ultra Helical data was used. The first phase was performed with volume registration us-

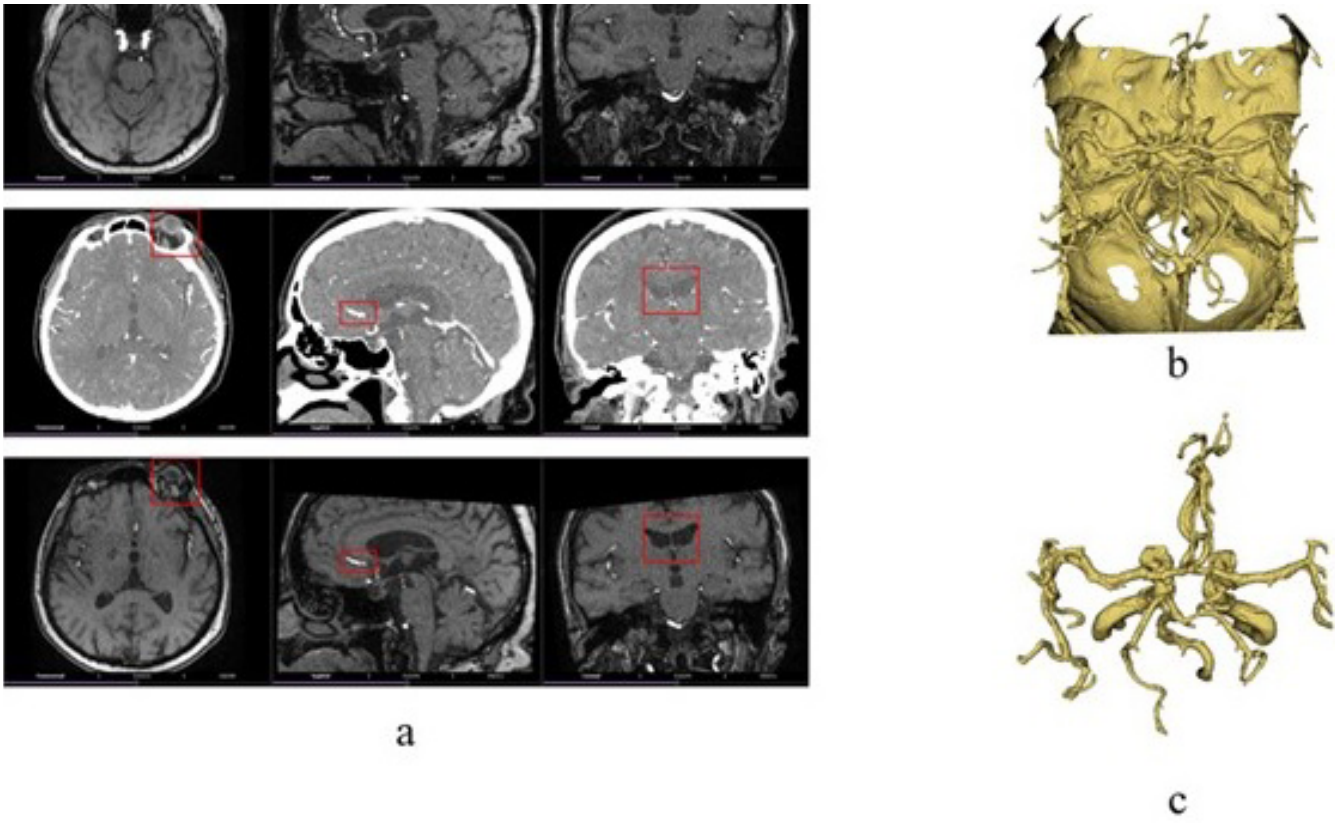


Figure 2. (a) Top (MRI), middle (CT), bottom MRI matched with CT. Red rectangles represent a visual assessment of registration quality. (b) CT segmentation, (c) CT segmentation with MRI boundaries (d)Visualization of surface distances. (e) Surface reconstructed from CT is color mapped with distance value from MRI reconstruction. The average distance value is 0.35 mm.

ing normalized mutual information as a similarity function and Powell's 18 methods as an optimization scheme. After that, a segmentation model from MRI was generated and used to limit the segmentation region for CT data – to hide bones. The next step was to reconstruct the vascular system from both CT and MRI to generate tubular segments. Such reconstructions could be compared to each other and to the source segmentation. The surface similarity measurements were calculated to ensure the correct correlation for both modalities. As the surfaces are represented by the dense set of points in our work we use Sum-of-squared-distances between two point sets as a merit function (where is a Euclidean distance between the point from one dataset and the closest point from the other dataset). After models were extracted both vascular versions could be used – from CT and extended (with ICA segments) from MRI. The last two steps involved the creation of several segmentation models (volumetric segmentation masks) of bone structures using newly developed tools for interactive 3D sculpting and passing it to the augmented reality interactive visualization tool.

3.1. 3D matching

In this step, CT was merged with MRI to deliver combined information in a matched coordinate system. Rigid-body registration (6 degrees of freedom) with normalized mutual information was used Maes et al.[2] and Garrido-Jurado et al.[19] as a similarity measure commonly used as one of the most versatile registration function for multimodal image data. As an optimization routine Powell's method was used due to its computational efficiency. It delivered the results in seconds and obtained a rigid body transformation that could be applied to the MRI dataset to transform it into the CT coordinated system. Subsequently, a fused dataset with a suitable size coordinated with two values assigned to every voxel (CT and MRI intensities) was created. After that new data could be used for visualization, processing, and measurements.

Segmentation of MRI vascular structures was straightforward. Contrast agent intensities were than any other values in 3D volume so that simple thresholding or controlled front growth segmentation could be applied. In this work, we used an iterative region growing algorithm based on Serra[20] utilizing probability maps – estimating the probability of intensity i being consistent with the volumetric sample model. Only one probability threshold value was necessary for parameterizing the segmentation procedure.

As for CT with a contrast agent, it was not trivial to segment vascular structures. Intensities regions affected by contrast agent were overlapping with the one for bony structures. The dilated mask from matched MRI was used for CT segmentation to deal with this problem. The goal of the morphological filter[21] was to expand segmentation by three voxels using a spherical structuring element. It was a sufficient operation to mask most of the bony structures enabling fast segmentation. Segmenting CT with MRI masking led to ICA arteries (Figure 2c) being segmented not correctly. It was impossible to distinguish between bony structures and ICA arteries in CT images, so they could not be considered in the reconstruction process. Final segmentation of both CT and MRI were pre-cut for the circle of the Willis region with some closest neighbors to represent corresponding vascular objects. ICA arteries were removed from CT segmentation.

At this stage, the developed method for cerebral reconstruction was applied and consisted of several phases:

(1) Initial vessel segmentation using iterative morphological growing with intensity remapping (assigning unique values to detected branches);
(2) Skeleton tracing and its optimization (removing shortest branches, reducing complexity, and smoothing);
(3) Generation and optimization of tubular structures representing separated vessels.

3.2. Matching quality analysis

In order to use fusion results, correct correlation had to be ensured. Matching procedures in volumetric space in different modalities were applied so measuring the registration quality was not trivial. The quantities of normalized mutual information or entropy measures in the registration process were minimized. Still, the value was not giving information about corresponding close tissues and their spatial position. This study recommended analysing the distance between reconstructed objects.

Figure 2 presents a color visualization of the distance between surfaces using a color lookup table. For the presented dataset maximum distance values were above 4 mm, but they emerged only in regions where there were manually cropped segmentations. For all processed datasets an average error varied in a range 0.30–0.37 mm, which was below the voxel size of the analyzed dataset (0.5 mm), and the standard deviation varied from 0.37 to 0.41 mm. A point-to-pointset distance histogram (Figure 2e) was created to show the most probable distances to occur. Significant values were marginal, and almost all surface points were in the space within an acceptable tolerance.

When assured that both reconstructed surfaces were within an acceptable distance, the bone data from CT and an extended vascular model from MRI data were used to make a final fusion for further analysis and measurements.

4. RESULTS

As the result of this study, a new fusion. This study also contributed to developing a virtual sculpting tool for modeling binary masks overlying corresponding volumetric data to create different scenarios. Figures 2d and 2e presents the results of the registration process. The user can draw polygonal regions on the screen, and everything visible in such polygonal window is removed or hidden. This study followed that idea and provided a set of tools for 3D segmentation. This study implemented several scenarios. The first one was the simple removal of all voxels that coincided with frustum volume. In the second approach, a user-selected an absolute depth (in mm), and the system removed voxels to a user-defined depth (counting from the first hit voxel – the clos-

est to camera). The third one was the relative depth mode – an object was sculpted relatively depending on the visible voxel distance from the camera. Using this model, given the ability to sculpt shapes on irregular 3D objects. The last option was beneficial in vascular structures tuning – the first hit mode – method removed a separated object closest to the camera. In practice, it meant that if a part of a structure (i.e., tubular one) was selected to be removed, it was cut out (like with a scalpel), but every object that was at the back remained untouched. Using such tools, the user could remove unnecessary information in a few interactive cuts.

5. DISCUSSION

3D visualization on the screen introduces multiple problems: scale is arbitrary and not obvious, it's sometimes very complicated to tell and normalize the object's orientation. Scientific works show that MRI scanners potentially contribute to the more common use in animals and humans experimental research.[22,23] The spatial relations of the presented parts are often not visible or distorted by projection. Interaction and manipulation become complicated for untrained operators and may raise frustration. Experimenting with stereo and head-tracked screen projection exposed how limited it was in the field of view and didn't help with the complexity of the user interface. Because of that, the study followed the head-mounted AR set with physical markers since it mitigated mentioned shortcomings of screen projections. Interaction via AR is very natural since people are experts in manipulating physical objects with their hands. The study found that for overall effect, more critical was hiding visible shortcomings of AR registration errors than the fact rendered objects were not real. The most perceptible flaws were the visible floating of virtual objects regarding the real world, sudden loss of marker tracking, misalignment with markers, and sudden object disappearance when the marker was occluded. All these problems caused a virtual object to feel flimsy, and the whole experience seemed fragile.

Because of that work focused on employing small tricks so that the simulation would seem more solid:

(1) To help with the occlusion problem, single markers were switched to whole arrays of markers, called 'boards' in the Aruco[19] library. Boards could be partially occluded, and when more than one marker on the board was visible, registration stability was vastly improved.
(2) To hide the registration errors of markers, objects were never allowed to touch markers and put a few centimeters above. The closer objects were to markers, the more visible registration errors appeared.
(3) When the user puts the marker far from the camera (mre than the arms-length), where registration errors became much more extensive, the object was moved further up from the marker and added slow cyclic animation where the object moved up and down. This animation was easy to understand and hide, otherwise perceptible floating caused registration inaccuracies.
(4) When registration was not possible from either camera, the last valid registration was used. If this situation persisted for a few seconds, visualized objects gradually faded away. This procedure helped especially with sudden movements when marker detection temporarily failed. If registration was available again after only a few frames, the user often hardly perceived registration loss. When registration was possible from one of two cameras, marker registrations were computed for the other camera using the parameters of the stereo system.
(5) When the user was looking at the marker close to sideways, the object was tilted back so that its underside was facing the user. It was a dynamic situation where most users stopped and kept the marker visible to the cameras.

These methods worked probably because the human brain predicts how things should look if given their actions, like moving an object. In this setting user quickly learned the close spatial relation of markers with virtual objects and expected it to be rigid. By moving objects away from the board or steering clear of situations when registration would go wrong, the spatial constraints of objects related to the board were softened and more room for discrepancies of user's predictions allowed with the presentation, circumventing detection of this fallacy by the brain.

5.1. Augmented reality tool

The custom-built for this study visualization passthrough AR rig consisted of Oculus Rift with a stereo camera system mounted directly on head mounted display (HMD), in front of the user's eyes. The computer used in this study was a mid-range laptop that handled capture, processing, and rendering tasks. In this setup, stereo images were captured by cameras and sent over to the computer. Visualizations were registered to the detected markers and rendered onto the source stereo image. Finally, the altered stereo image was sent to the HMD and the process was repeated. For physical interaction and embedding virtual objects in natural images, standard markers from the Aruco[19] library were used. They were chosen for ease of interaction and performance reasons. Markerless setup was also considered, but it could be precarious and use more computing power depending on the environment. Another, not trivial, advantage of using markers for interaction was that the user had a physical object to manipulate. It also drew a clear boundary of what was real and what was not. Markerless augmented reality tried to hide the fact that virtual objects were not real, but the user knew that presented results produced by the computer are not real. When the system tried to hide that fact, the contrast between what was real and what was not is lowered, which meant that minor discrepancies of the embedding became more perceptible, and visualization fell into the so-called uncanny valley trap.

5.2. Details of augmented reality fusion visualization system

The preprocessing module relied on input data sets: MRI, CT original volumes, and a set of scenarios represented as binary masks from the cutting process described earlier. All

of the collected data was combined into a fusion set. The preprocessing step was performed to create a fusion bundle consisting of colored and shaded volumes as octrees[24] prepared for real-time presentation. Fusion bundles were kept in visualization storage (local database). Each fusion bundle was associated with a predetermined ID number. In this study approach, marker boards were used. It's a construct used by the augmented reality library. Each board consisted of a set of fiducial square markers arranged in a rectangular 3 × 3 grid formation. Each marker encoded a marker ID. In this study it was always guaranteed that each marker ID was on no more than one board; every marker ID uniquely identified a marker board. After testing marker boards occurred to be more stable and resilient to partial occlusion than single markers.

The system detected a model connected to a recognized marker and rendering procedures could be applied. The main rendering loop consisted of the following steps:

(1) Image acquisition from the stereo camera;
(2) Marker detection, recognition, and grouping into marker boards;
(3) Fetching of fusion bundles corresponding to detected markers from visualization storage;
(4) Rendering the stereo-image of fusion bundles using the position and orientation computed from corresponding marker boards;
(5) Overlaying of stereo-image from the previous step over the source-image acquired earlier from the cameras;
(6) Presentation of overlay produced in the previous step to the HMD.

5.3. Human-centered interaction

The primary purpose of the presented AR system was to allow users to inspect generated fusion bundles. 3D interaction allowed for 12 spatial degrees of freedom: 6 for position and orientation of the camera system controlled by movements of the user's head and 6 for position and orientation of marker board manipulated by user's hands. Users could also choose from different scenarios by flipping the pages of multiple marker boards printed on ordinary A4 paper pages. Since object inspection was natural, users could focus on presented data and without thinking about the manipulation method itself.

Figure 1b presents a typical use setting, with two visualization examples of rendered scenarios. The system detected ID from the markers board, detected its global 3D position and spatial orientation, and applied stereo rendering routines to present final volumetric models.

The AR system was presented to over a hundred users, ranging from medical students to radiologists. Fusion bundles contained CT/MRI fusions obtained with the process described above. Most users didn't need any instructions for successful interaction. Few users asked for directions on how to rotate the view or see the objects closer and needed to be hinted that such operations could be performed naturally using hands or by the change of point of view.

6. CONCLUSIONS

(1) This study presented a novel CT and MRI fusion method of a human head based on volumetric registration.
(2) The segmentation quality was estimated using the distance analysis of the segmented vascular model from both CT and MRI datasets.
(3) The system included a proposed method for stabilization of the human head using the semi-circular plane.
(4) Virtual sculpting technique enabled visualization scenarios and augmented reality tools to explore created CT sub-models accompanied by an MRI reconstructed vascular tree.
(5) The application of AR tools met the needs of medical specialists and enabled new ways of teaching anatomy in more substantial ways.

Conflict of interest

None declared.

Funding

This research was partially supported by the Polish National Center (grant No. 2012/07/D/ST6/02479).

Ethics

The research was partially supported by the Polish National Center (grant No. 2012/07/D/ST6/02479).

References

1 Inoue HK, Nakajima A, Sato H, Noda S, Saitoh J, Suzuki Y. Image Fusion for Radiosurgery, Neurosurgery and Hypofractionated Radiotherapy. *Cureus*. 2015;7(3):e252. https://doi.org/10.7759/cureus.252.
2 Maes F, Vandermeulen D, Suetens P. Medical image registration using mutual information. *Proc IEEE*. 2003;91(10):1699–1721. https://doi.org/10.1109/JPROC.2003.817864.
3 Pol EJD, Viergever MH. Medical Image Matching – A Review with Classification. *IEEE Eng Med Biol Mag*. 1993;12(1):26–39. https://doi.org/10.1109/51.195938.
4 Sahu S, Pati UC. Intensity-based registration of medical images. *Int J Comput Vis Robot*. 2016;6(4):319–330. https://doi.org/10.1504/IJCVR.2016.079393.
5 Bavirisetti DP, Kollu V, Gang X, Dhuli R. Fusion of MRI and CT images using guided image filter and image statistics. *Int J Imaging Syst Technol*. 2017;27(3):227–237. https://doi.org/10.1002/ima.22228.
6 Lesage D, Angelini ED, Bloch I, Funka-Lea G. A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes. *Med Image Anal*. 2009;13(6):819–845. https://doi.org/10.1016/j.media.2009.07.011.
7 Piccinelli M, Veneziani A, Steinman DA, Remuzzi A, Antiga L. A framework for geometric analysis of vascular structures: Application to cerebral aneurysms. *IEEE Trans Med Imaging*. 2009;28(8):1141–1155. https://doi.org/10.1109/tmi.2009.2021652.

8 Aylward S, Pace D, Enquobahrie A, McCormick M, Mullins C, Goodlett C, Reynolds P. Tube TK, segmentation, registration, and analysis of tubular structure in images. Clifton Park 2012.

9 Hamza A Ben, He Y, Krim H, Willsky A. A multiscale approach to pixel-level image fusion. *Integr Comput Aided Eng*. 2005;12(2):135–146. https://doi.org/10.3233/ica-2005-12201.

10 Li H, Manjunath Bs. Multisensor-Image-Fusion-Using-the-Wavelet-Transform_1995_Graphical-Models-and-Image-Processing.pdf. *Graph Model Image Process*. 1995;57(3):235–245.

11 Petrovic VS, Xydeas PC. Optimising Multiresolution Pixel-level Image Fusion. *Proc SPIE*. 2001;4385:96–107. https://doi.org/10.1117/12.421097.

12 James AP, Dasarathy B V. Medical image fusion: A survey of the state of the art. *Inf Fusion*. 2014;19(1):4–19. https://doi.org/10.1016/j.inffus.2013.12.002.

13 Du J, Li W, Lu K, Xiao B. An overview of multi-modal medical image fusion. *Neurocomputing*. 2016;215:3–20. https://doi.org/10.1016/j.neucom.2015.07.160

14 Perez J, Mazo C, Trujillo M, Herrera A. Mri and ct fusion in stereotactic electroencephalography: A literature review. *Appl Sci*. 2021;11(12):5524. https://doi.org/10.3390/app11125524.

15 Sasikala M, Kumaravel N. A comparative analysis of feature based image fusion methods. *Inf Technol J*. 2007;6(8):1224–1230. https://doi.org/10.3923/itj.2007.1224.1230.

16 Tao Q, Veldhuis R. Threshold-optimized decision-level fusion and its application to biometrics. *Pattern Recognit*. 2009;42(5):823–836. doi:10.1016/j.patcog.2008.09.036.

17 Nowinski WL, Volkau I, Marchenko Y, Thirunavuukarasuu A, Ng TT, Runge VM. A 3D model of human cerebrovasculature derived from 3T magnetic resonance angiography. *Neuroinformatics*. 2009;7(1):23–36. https://doi.org/10.1007/s12021-008-9028-8.

18 Press WH, Teukolsky SA, Vettering WT, Flannery BP. NUMERICAL RECIPES The Art of Scientific Computing Third Edition. Cambridge Univ Press. 2007. https://doi.org/10.1017/CBO9781107415324.004.

19 Garrido-Jurado S, Muñoz-Salinas R, Madrid-Cuevas FJ, Marín-Jiménez MJ. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognit*. 2014;47(6):2280–2292. https://doi.org/10.1016/j.patcog.2014.01.005.

20 Serra J. Image analysis and mathematical morphology. Computer Graphics and Image Processing.1982;20:96–97.

21 Huang X, Qian Z, Huang R, Metaxas D. Deformable-Model Based Textured Object Segmentation. In: Rangarajan A, Vemuri B, Yuille AL, eds. Energy Minimization Methods in Computer Vision and Pattern Recognition. EMMCVPR 2005. Lecture Notes in Computer Science, vol 3757. Springer. 2005 https://doi.org/10.1007/11585978_9.

22 Mystkowska D, Tutas A, Jezierska-Woźniak K, Mikołajczyk A, Bobek-Billewicz B, Jurkowski M. Usefulness of clinical magnetic resonance scanners for imaging experimental changes in laboratory rodents' central nervous system. *Pol Ann Med*. 2012;19(1):43–49. https://doi.org/10.1016/j.poamed.2012.04.009.

23 Mystkowska D, Tutasa A, Jezierska-Woniar K, Mikołajczyka A, Bobek-Billewicz B, Jurkowski MK. High resolution small animals dedicated magnetic resonance scanners as a tool for laboratory rodents central nervous system imaging. *Pol Ann Med*. 2013;20(1):62–68. https://doi.org/10.1016/j.poamed.2013.07.007.

24 KAINZ W. A Review of: "The Design and Analysis of Spatial Data Structures". By H. SAMET. (Addison-Wesley Publishing Company, Inc., 1989.) *Int J Geogr Inf Syst*. 1991;5(2):253–253. https://doi.org/10.1080/02693799108927847.