Title

Similarity Metric Learning for 2D to 3D Registration of Brain Vasculature.

Authors

Tang, Alice
Scalzo, Fabien

Publication Date

2016-12-01

DOI

10.1007/978-3-319-50835-1_1

Peer reviewed

HHS Public Access

Author manuscript

Adv Vis Comput. Author manuscript; available in PMC 2019 January 15.

Published in final edited form as:

Adv Vis Comput. 2016 December ; 10072: 3–12. doi:10.1007/978-3-319-50835-1_1.

Similarity Metric Learning for 2D to 3D Registration of Brain Vasculature

Alice Tang and **Fabien Scalzo**

Neurovascular Imaging Research Core Department of Neurology and Computer Science University of California, Los Angeles (UCLA)

Abstract

2D to 3D image registration techniques are useful in the treatment of neurological diseases such as stroke. Image registration can aid physicians and neurosurgeons in the visualization of the brain for treatment planning, provide 3D information during treatment, and enable serial comparisons. In the context of stroke, image registration is challenged by the occluded vessels and deformed anatomy due to the ischemic process. In this paper, we present an algorithm to register 2D digital subtraction angiography (DSA) with 3D magnetic resonance angiography (MRA) based upon local point cloud descriptors. The similarity between these local descriptors is learned using a machine learning algorithm, allowing flexibility in the matching process. In our experiments, the error rate of 2D/3D registration using our machine learning similarity metric (52.29) shows significant improvement when compared to a Euclidean metric (152.54). The proposed similarity metric is versatile and could be applied to a wide range of 2D/3D registration.

1. Introduction

During endovascular treatment for acute ischemic stroke, various images of the brain are taken at different time points before, during, and after intervention. While these images provide useful visual cues about the brain, the information obtained is often limited by the imaging technique used, thereby limiting the information available for treatment decisions. 2D-3D registration can partially address this problem by combining the advantage of 2D images with 3D ones; allowing 3D visualization of bodily structures during intervention by aligning pre-interventional 3D images (through CT, MRI, etc.) with realtime intrainterventional 2D images (e.g. fluoroscopy, DSA). Furthermore, registration also aids with comparisons of various images taken throughout the time frame, allowing further analysis of changes over time.

2D-3D coregistration has been widely used in the medical field for applications such as image-guided interventions and surgery [1] for treatments involving implants [2], spinal surgery [3], cancer therapy [4], and brain diseases. For treatment of neurovascular diseases, since 2D DSA is generally more efficiently used to visualize the vasculature during a procedure, corresponding 3D information to those 2D images can provide surgeons insights about the positioning of surgical tools in relation to the target site during treatment [5]. In the context of stroke, 2D-3D registration is an especially challenging procedure as parts of the vasculature are missing or deformed, due to occlusions in the blood vessels. In this

paper, we address these issues by presenting a 2D-3D registration algorithm that uses machine learning on local features.

Image registration can be defined as the process of finding the geometric transformation between two sets of images, which brings them to the same coordinate frame with the best possible spatial correspondence. Since manual registration can be time consuming and tedious, much work has been done on automatic ways to register medical images [6–8]. Previous works with feature-based methods use feature extraction to obtain matching geometries and correspondences [9], while intensity-based methods compare values in pixels [10]. Hybrid methods have also been employed that combine feature segmentation with pixel- wise comparisons [11]. Recent methods also performed registration in different spaces, such as the Fourier space [12].

For algorithms that rely upon image match comparison, registration is performed by optimizing a similarity measure indicating the amount of correspondence between two images. Common similarity measures that have been employed include mutual information, cross correlation, entropy, and pattern intensity, among others. Recently, machine learning has gained much interest in the medical field, with its diverse use and predictive power [13]. By learning patterns of known matches, a machine learning model can be used to predict the correspondence between two images based for example on local features, often allowing greater flexibility.

With a similarity measure established, an iterative optimization procedure can be used to find the optimal parameters which maximize or minimize a cost function based upon the similarity measure. Numerous optimization procedures have been developed and are available in the literature such as Levenberg Marquardt, Powell's method, downhill simplex, and gradient descent. Heuristic methods have also been used such as Monte Carlo random sampling, pattern search algorithm, and others.

This paper presents a registration algorithm that applies machine learning to learn the similarity of 2D and 3D local image features in a task dependent way, which can then be extended to predict the overall similarity of two images. By focusing on local image properties, the algorithm allows greater variations between the registered images. Optimization is performed with a heuristic search using the Nedler-Simplex method. We first outline the registration algorithm in section 2, and then describe the optimization process and the accuracy of this algorithm in Section 3 and 4. Then, some discussion about the algorithm and comparisons with other algorithms are presented in Section 5.

2. Methods

Image Dataset

The dataset used in this study to evaluate our framework was collected from patients admitted at a comprehensive stroke center and diagnosed with symptoms of acute ischemic stroke. The use of this dataset was approved by the local Institutional Review Board (IRB). Inclusion criteria for this study included: (1) Final diagnosis of acute ischemic stroke, (2) last known well time within six hours at admission, (3) pre-interventional Magnetic

Resonance Angiography (MRA) performed during the routine acute evaluation, (4) Digital Subtraction Angiography (DSA) of the brain performed at the end of a thrombectomy procedure. A total of 18 patients satisfied the above criteria and were included in this study. The patients had various success in revascularization. DSA images had a resolution of 1024 × 1024 pixels, while the resolution of MRA image varied across patients.

Vessel Extraction and Groundtruth

Blood vessels were detected on DSA with a Frangi filter [14]. This algorithm uses eigenvalues to compute the probability that a pixel contains a vessel by computing the similarity of a structure to an ideal tube. A simple threshold established with Otsu's method was used to create a binary mask and a cloud of 2000 points was extracted from each image. 3D MRA data was segmented semi-automatically using Osirix software to extract the major arteries of the brain. Similarly to the 2D DSA sampling, we extracted 2000 points from the segmented 3D image. The groundtruth was manually established between DSA and MRA images by a researcher in Neurology. An in-house software allowed the expert to rotate, scale, and shift the 3D MRA image volume to match the 2D DSA image. The result was a sequence of transformations that led to an ideal mapping between MRA and DSA that will later be used as groundtruth in our experiments.

Registration Model

The goal of 2D/3D registration is to find the transformation between the image coordinate system of the 3D image to a 2D image projection. The images are represented as point clouds and the transformation is assumed to be an affine transformation with constant scaling factor followed by a projection onto the xy plane. Thus, we can represent our transformation $T:\mathbb{R}^3 \rightarrow \mathbb{R}^2$ as

$$T = A(r)S\left(s_x,\ s_y\right)PR_z(\theta_z)R_y(\theta_y)R_x(\theta_x) \quad 1$$

where $R_x(\theta_x)$, $R_y(\theta_y)$ and $R_z(\theta_z)$ rotates a point respectively about x, y and z axis by θ_x, θ_y and θ_z, P projects a point from $\mathbb{R}^3$ to $\mathbb{R}^2$ by taking the first two coordinates, $S(s_x, s_y)$ shifts a point by (s_x, s_y), and A(r) scales a point by a constant factor r. This decomposition simplifies the registration process as a search for a 6-parameter vector $\mathbf{p} = [\theta_x\ \theta_y\ \theta_z\ s_x\ s_y\ r]$ that optimally aligns the images. To find these transformation parameters, the 2D-3D registration algorithm utilizes a machine learning model to learn a similarity measure between points based upon their local neighborhood.

Image Representation

The 2D/3D images are represented as point clouds in their respective spaces. An erosion filter is applied to remove outliers. In this paper, we refer to the 2D point cloud as the Target image, and the 3D point cloud as the Source image. After a transformation is applied to the 3D point cloud, we refer to the transformed 3D image as the Transformed image. Note that the Transformed image is two dimensional (as it is projected to a 2D space).

With these representative sample of the Target and Transformed image, a local 2D histogram H_i acts as a point descriptor to describe each point by its local neighborhood. The histogram can be created given parameters of histogram size, h_s, and number of bins, h_b, and is normalized by the total number of points in the histogram. Each histogram can be condensed and represented as a 1D array of size $1 \times h_b^2$. A collection of histograms for each point in an image can thus act as the image descriptor.

Machine Learning Model

We use Spectral Regression for Discriminant Analysis (SR-DA) [15] to find the parameters of a function that will output the likelihood of a match (i.e. similarity) between two points givien the point distribution in their neighborhood. During image registration, the point correspondences are two presumed point matches computed by taking the closest neighbor point in the Target image for every sampled point in the Transformed image. The average of the similarities for a representative sample of point correspondences acts as the total similarity measure for the Target and Transformed image. For each iteration of the registration process, a transformation $\hat{T}$ obtained from the parameter vector $\mathbf{p}$ will be applied to the 3D Source image to obtain a 2D Transformed image. N randomly sampled points in the Transformed image with point correspondences in the Target image will be used to construct the input feature matrix to be given to the machine learning model.

The input to the machine learning model can be represented as a matrix $X = (K_1\ K_2\ d)$. K_1 is the $N \times h_b^2$ matrix of point descriptors for the N points, $\{p_1, p_2, ..., p_N\}$, that are sampled from the Transformed image. For each point in the Transformed image, p_i, the nearest neighbor point in the Target image, q_i, will be chosen by minimizing the distance between p_i and q_i. q_i will be taken as a point correspondence, that is, the point descriptor for p_i is expected to be similar to the point descriptor for q_i. K_2 is then the $N \times h_b^2$ matrix where each row i consists of the point descriptor of the nearest corresponding point to p_i in the Target image. Lastly, d is a $N \times 1$ vector consisting of the Euclidean distance between the two point correspondences for each row, that is, $d_i = \sqrt{\left(x_{pi} - x_{qi}\right)^2 + \left(y_{pi} - y_{qi}\right)^2}$. Therefore, X is of size $N \times (2 \cdot h_b^2 + 1)$, where each row represent information of each corresponding point pairs. The similarity function is applied to each row of X, and the output can be represented as a vector $\hat{Y}$ in $\mathbb{R}^{N \times 1}$, where $\hat{Y}_i$ represents the probability of a match between p_i and q_i. The average of all values of $\hat{Y}$ gives the similarity metric between the Target and Transformed image.

To train the model, we require representative pairs of matching and nonmatching points to construct the input feature matrix X and a label vector Y with entries 1 for the matching point pairs and 0 for the nonmatching point pairs. We use data augmentation to obtain the sample of matching point pairs by generating various transformations on the Source image by adding random variations to the groundtruth parameter vector, $\mathbf{p}_{gt}$. This allows the machine learning model to be robust. The matching point pairs are also thresholded by distance in order to improve the accuracy of training by removing points present as noise or extraneous vasculature. Nonmatching point pairs were generated by applying larger variations to the parameters of $\mathbf{p}_{gt}$. These samples are represented in a training matrix, X, with corresponding label vector Y.

X is normalized column-wise and a Spectral Regression model is trained with X and Y and can then be used as a similarity function F between two candidate matches. The constructed similarity function F can be applied between point pairs in the Transformed and Target images, and the total similarity will be obtained: $\mathcal{SIM} = \overline{\hat{Y}} = F(X)$.

Registering a new set of images

The goal of 2D-3D registration is to find $\hat{T}$, an estimated transformation, based upon finding the 6 parameter vector $\mathbf{p}$ that optimizes the similarity value. Although a higher similarity value generally indicates a greater chance of matching, the negative of the similarity will be taken as our cost function in order to frame the optimization as a minimization problem: $COST = -SIM = -F(X)$. An initial parameter must be given to start the optimization, which will then proceed by the Nedler-Mead Simplex algorithm [16] as an heuristic search method to iteratively test various points in the 6-parameter space to find the optimal transformation. Optimization is stopped when the similarity or parameters are within a tolerance interval, or when 50 iterations of the algorithm have been reached.

3. Experiments

Parameter Optimization

In order to find the optimal parameters for training and applying the machine learning model, cross-validation tests were performed to compute the accuracy of the model under different parameters, using the area under the curve (AUC) as the accuracy measure. These parameters include the number of training samples, the histogram size h_s, the number of bins per histogram h_b, the regularization parameter (alpha), and the similarity smoothing parameter (t) used in the machine learning model. For every set of parameters, k pairs of training images, $P_1, P_2, \ldots, P_k$, coupled with the corresponding groundtruth transformation vector, $\mathbf{p}_{gt,1}, \mathbf{p}_{gt,2}, \ldots, \mathbf{p}_{gt,k}$, were used for training and testing. For each set of training parameters, one image, P_m, was set aside while the rest were used to train the machine learning model and compute the similarity function. The similarity function was then applied to sampled point pairs in the excluded training image, P_m, to obtain prediction $\hat{Y}_{Pm}$. This process was repeated so that all k images are excluded once. The point pair predictions $\hat{\mathrm{Y}}_{Pm}$ are then compared against known corresponding training labels Y_{Pm} and the AUC value is computed as a measure of the accuracy of the set of training parameters. This process is repeated again for each set of training parameters and corresponding AUC value computed.

Accuracy of Registration

To determine the effectiveness of the learned similarity metric, the accuracy of the registration model was tested against another simple metric which corresponds to minimizing the Euclidean norm of the error between histograms of point correspondences. For both the learned similarity (ML similarity) and the Euclidean norm similarity (Eunorm similarity), the registration process was performed to obtain an estimated parameter vector, $\mathbf{p}_{ML}$ and $\mathbf{p}_{Eunorm}$, representing the predicted transformation.

Error in this algorithm was found by transforming the 3D Source image with the ground truth parameter $\mathbf{p}_{gt}$, then finding the distance between the same points transformed with the predicted parameters. The root mean square error was found for both the ML similarity and the Eunorm similarity as a measure of accuracy.

4. Results

Cross-validation was performed to evaluate and optimize the model such that the number of training point pairs, the number histogram bins, and the regularization parameter of the regression model were optimized at each iteration of the crossvalidation. The optimal parameters found for our data set: training sample size = 2000, alpha = .001, smoothing parameter t = 6, $h_b = 6$, $h_s = 400$.

18 samples of patient pre-interventional MRA scans and post-treatment DSA scans were used for the registration processes. For each Source image, the registration process was performed with both the machine learning model similarity and the Euclidean distance similarity. Initial parameters were same for both similarity measures and varied from groundtruth parameters within 60° about the x and y axis, 150° about the z-axis, 10% shift, and 15% scaling. Root mean square error was computed between the same points in the Source image transformed using the estimated parameters, and the groundtruth parameters. Manual visualizations show that the registration process using the machine learning model similarity failed to register 2 samples of images, while using the Euclidian norm similarity failed to register 12 samples of images. The root mean square error for the Euclidean similarity metric was 152.54 ± 87.39, while the proposed ML similarity metric led to an error of 52.29 ± 64.69.

5. Discussions

The use of machine learning for image registration has many advantages and allows greater adaptation for various applications. Although we focused on registration of MRA and DSA simages in this paper, machine learning may be extended to compute the similarity function between images obtained from other modalities and acquisition parameters, provided enough training samples are available.

The disadvantage of this algorithm involves the need for training samples to exist in order to pre-train the machine learning model before the similarity function can be used for registration. The process to obtain optimal parameters for training and registration may also be challenging and time-consuming. Furthermore, the timing of registration using this algorithm varies from 1–10 minutes, which may be slow or adequate depending on the application.

Future experiments may involve testing machine learning applications for different image descriptor choices. Hybrid algorithms can be constructed by considering image features such as global histograms, gradient information, vessel shape, or transformations of the image in parameter or frequency space. It would also be advantageous to adapt the algorithm for a larger training sets in the future. Furthermore, since manual groundtruth parameters were

also prone to error, it may be beneficial to obtain multiple manual registrations from various neurology researchers.

Despite these limitations, the proposed framework could be used to register perfusion angiography [17] maps to MRA, and compare/map Computational Fluid Dynamics (CFD) parameters extracted from DSA [18] with the ones computed from 3D MRA/CTA [19,20].

6. Conclusion

Given the outcome of our local similarity machine learning method, it is shown that using machine learning for future applications of 2D-3D registration can provide adaptability beyond what has been previously obtained. Machine learning models can be trained with different datasets, such as images obtained from other modalities and methods. Eventually, a database of trained models can be collected. These models then can be easily chosen, implemented, and rapidly used for registration purposes in future clinical or research settings. Overall, using machine learning in 3D-2D registration has the potential reach for far applications into the medical field.

Acknowledgments

Prof. Scalzo was partially supported by a AHA grant 16*BGIA*27760152, a Spitzer grant, and received hardware donations from Gigabyte, Nvidia, and Intel. Alice Tang was partially supported by a UC Leads Fellowship.

References

1. Holmes D, Rettmann M, Robb R: Visualization in Image-Guided Interventions. In: Image-Guided Interventions: Technology and Applications. (2008) 45–80
2. Flood PD, Banks SA: Automated Registration of Three-Dimensional Knee Implant Models to Fluoroscopic Images using Lipschitzian Optimization. IEEE Trans Med Imaging (2016)
3. Otake Y, Wang AS, Webster Stayman J, Uneri A, Kleinszig G, Vogt S, Khanna AJ, Gokaslan ZL, Siewerdsen JH: Robust 3D-2D image registration: application to spine interventions and vertebral labeling in the presence of anatomical deformation. Phys Med Biol 58 (2013) 8535–8553 [PubMed: 24246386]
4. Fu D, Kuduvalli G: A fast, accurate, and automatic 2d3d image registration for image-guided cranial radiosurgery. Medical Physics 35 (2008)
5. Markelj P, Tomaevi D, Likar B, Pernu F: A review of 3d/2d registration methods for image-guided interventions. Med Image Anal 16 (2012) 642–661 [PubMed: 20452269]
6. Alves RS, Tavares JMRS In: Computer Image Registration Techniques Applied to Nuclear Medicine Images. (2015) 173–191
7. Tavares JMRS In: Analysis of Biomedical Images Based on Automated Methods of Image Registration. (2014) 21–30
8. Oliveira FP, Tavares JMR: Medical image registration: a review. Comput Methods Biomech Biomed Engin 17 (2014) 73–93 [PubMed: 22435355]
9. Chen X, Varley MR, Shark LK, Shentall GS, Kirby MC: An extension of iterative closest point algorithm for 3d-2d registration for pre-treatment validation in radiotherapy. In: MedVis. (2006) 3–8
10. Birkfellner W, Stock M, Figl M, Gendrin C, Hummel J, Dong S, Kettenbach J, Georg D, Bergmann H: Stochastic rank correlation: a robust merit function for 2D/3D registration of image data obtained at different energies. Med Phys 36 (2009) 3420–3428 [PubMed: 19746775]
11. Vermandel M, Betrouni N, Gauvrit JY, Pasquier D, Vasseur C, Rousseau J: Intrinsic 2d/3d registration based on a hybrid approach: use in the radiosurgical imaging process. Cellular and Molecular Biology 52 (2006) 44–53

12. Oliveira FP, Pataky TC, Tavares JMR: Registration of pedobarographic image data in the frequency domain. Comput Methods Biomech Biomed Engin 13 (2010) 731–740 [PubMed: 20526916]
13. Khandelwal P, Yavagal DR, Sacco RL: Acute Ischemic Stroke Intervention. J. Am. Coll. Cardiol. 67 (2016) 2631–2644 [PubMed: 27256835]
14. Frangi AF, Niessen WJ, Vincken KL, Viergever MA: Multiscale vessel enhancement filtering. In: MICAAI. (1998) 130–137
15. Cai D, He X, Han J: Spectral regression for efficient regularized subspace learning. In: ICCV. (2007) 1–8
16. Lagarias JC, Reeds JA, Wright MH, Wright PE: Convergence properties of the nelder-mead simplex method in low dimensions. SIAM J. Optim. 9 (1998) 112–147
17. Scalzo F, Liebeskind DS: Perfusion angiography in acute ischemic stroke. Comput Math Methods Med (2016) 1–14
18. Scalzo F, Hao Q, Walczak AM, Hu X, Hoi Y, Hoffmann KR, Liebeskind DS: Computational hemodynamics in intracranial vessels reconstructed from biplane angiograms. In: ISVC. (2010) 359–367
19. Nam HS, Scalzo F, Leng X, Ip HL, Lee HS, Fan F, Chen X, Soo Y, Miao Z, Liu L, Feldmann E, Leung T, Wong KS, Liebeskind DS: Hemodynamic Impact of Systolic Blood Pressure and Hematocrit Calculated by Computational Fluid Dynamics in Patients with Intracranial Atherosclerosis. J Neuroimaging 26 (2016) 331–338 [PubMed: 26598796]
20. Leng X, Scalzo F, Fong AK, Johnson M, Ip HL, Soo Y, Leung T, Liu L, Feldmann E, Wong KS, Liebeskind DS: Computational fluid dynamics of computed tomography angiography to detect the hemodynamic impact of intracranial atherosclerotic stenosis. Neurovascular Imaging 1 (2015) 1

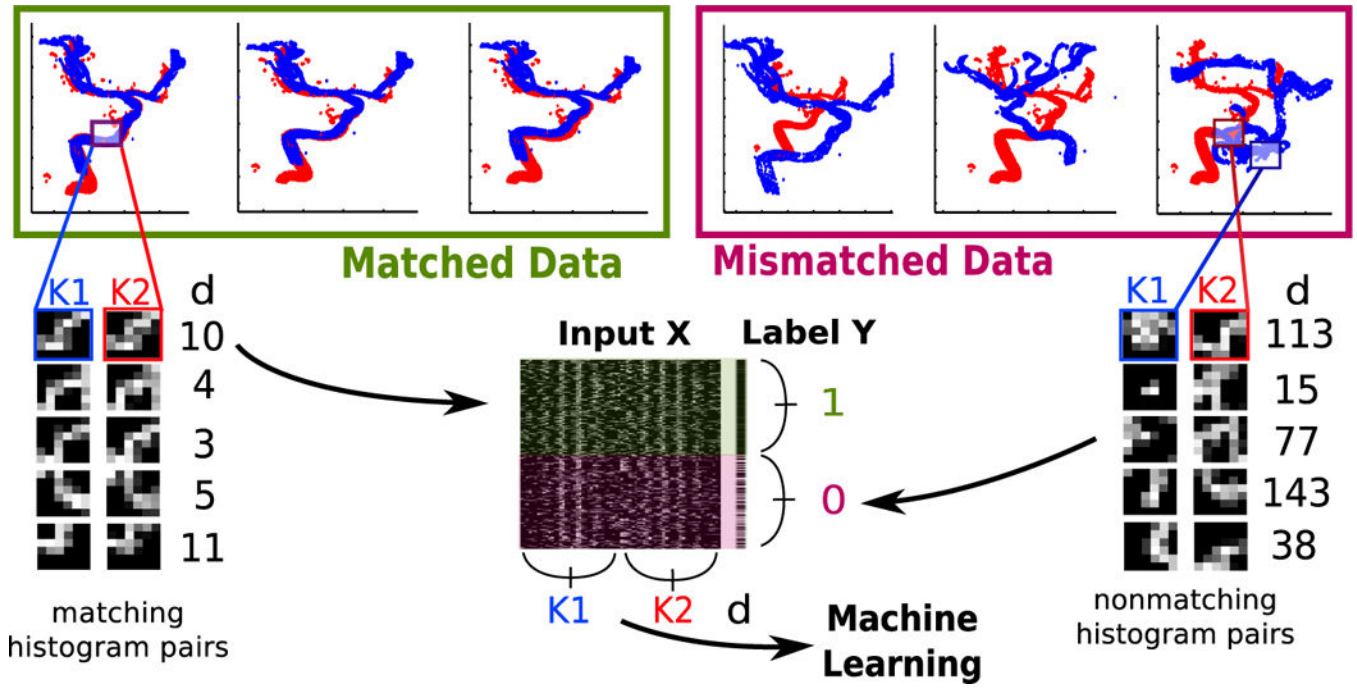

Fig. 1:

Extraction of local input features from 2D DSA and 3D MRA to train the model. Local descriptors are local point cloud histograms. Pairs of point descriptors are sampled randomly from labeled matching and nonmatching transformations.

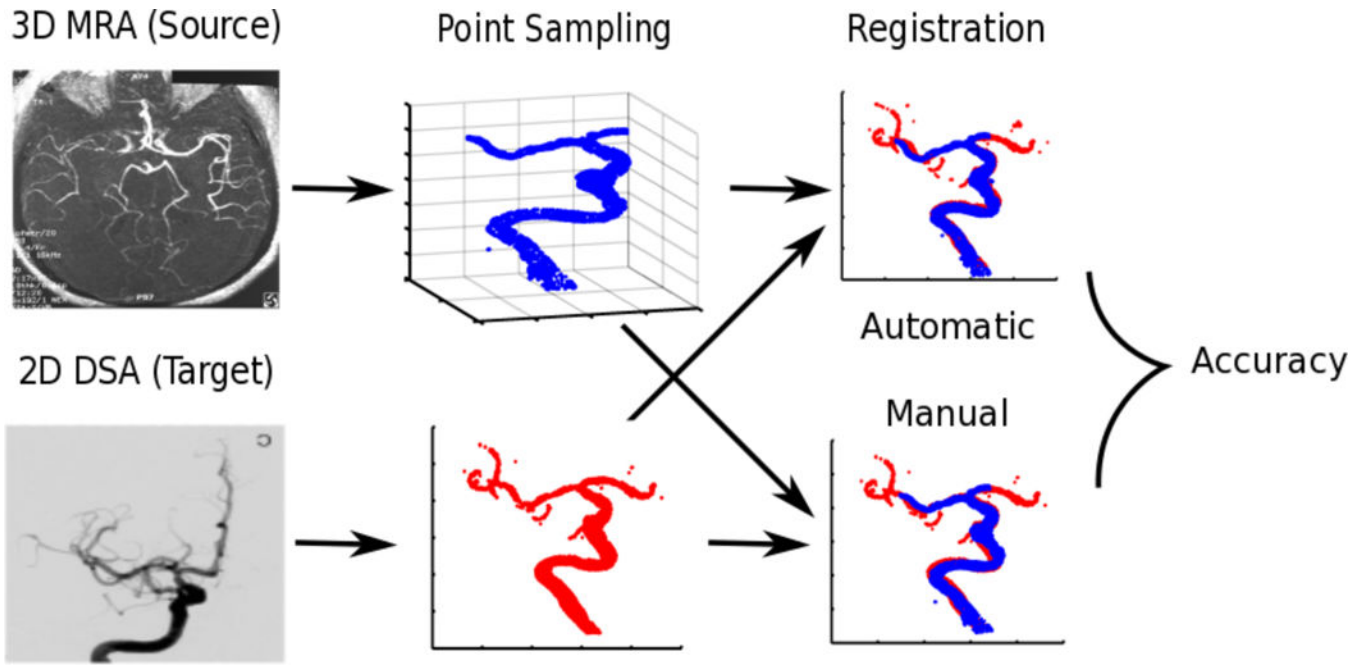

Fig. 2:
Diagram representing the registration process. The accuracy of our algorithm is determined by comparing the computed and the manual transformations.

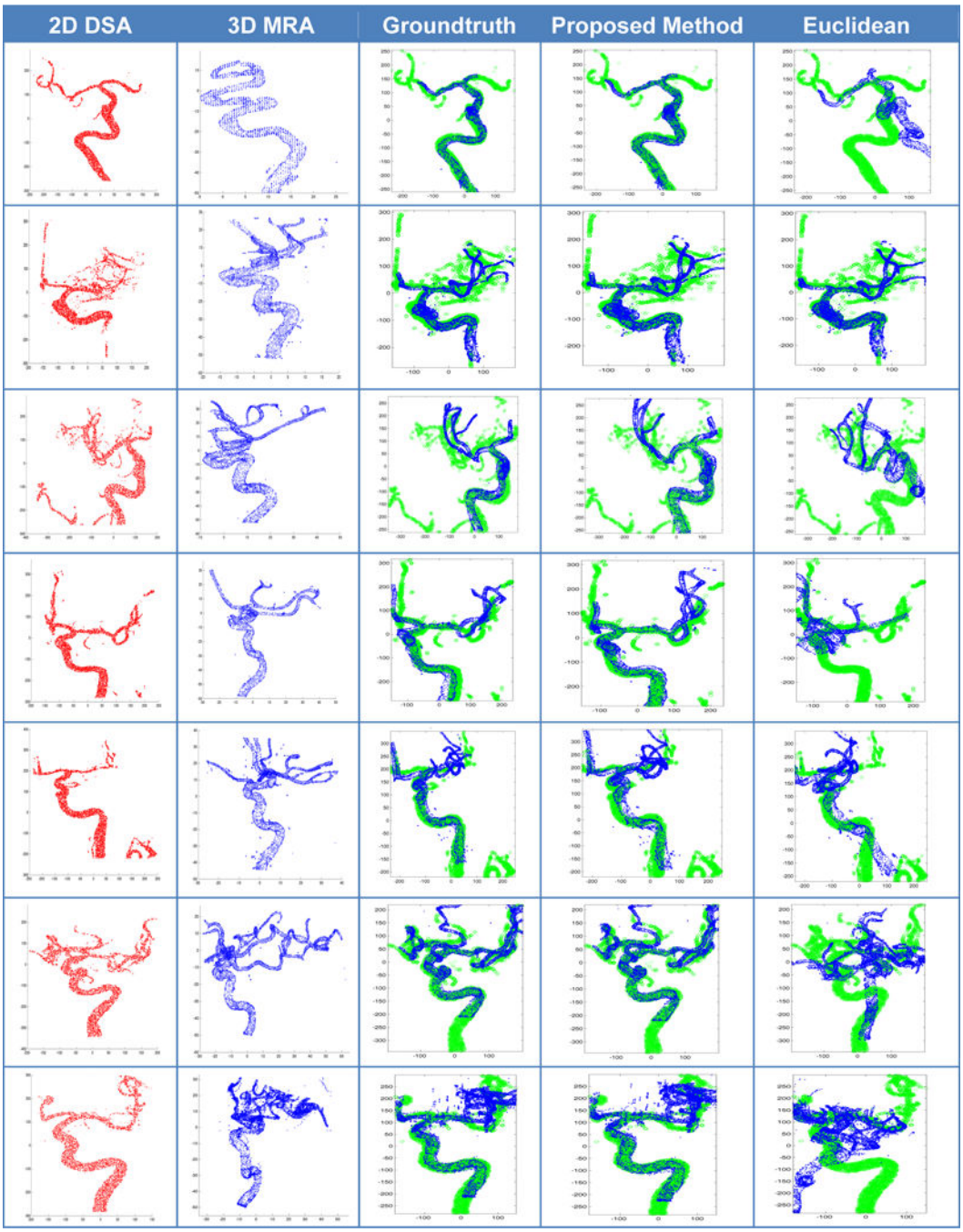

Fig. 3:
Image panel illustrating the target 2D DSA, source 3D MRA, the groundtruth, and the result of the registration using ML and Euclidean norm metrics.