IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 30 September 2022, accepted 26 October 2022, date of publication 31 October 2022, date of current version 8 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3218458

RESEARCH ARTICLE

A Deep Learning-Based Approach for Automated Coarse Registration (ACR) of Image-Guided Surgical Navigation

HAKJE YOO[1] **AND TAEYONG SIM**[2]

[1]College of Medicine, Research Institute for Medical Bigdata Science, Korea University, Seoul 02841, Republic of Korea
[2]Department of Artificial Intelligence, Sejong University, Seoul 05006, Republic of Korea

Corresponding author: Taeyong Sim (tysim@sejong.ac.kr)

This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant 2021R1I1A1A01059747; in part by the Institute of Information and Communications Technology Planning and Evaluation (IITP) funded by the Korean Government [Ministry of Science and ICT (MSIT)] through the Development of Explainable Artificial Intelligence (AI)-Based Diagnosis and Analysis Frame Work Using Energy Demand Big Data in Multiple Domains under Grant 2022-0-00106; and in part by the Technology Development Program funded by the Ministry of Small and Medium-Sized Enterprises (SMEs) and Startups (MSS), South Korea, under Grant RS-2022-00156456.

ABSTRACT Coarse registration is the first step in determining the accuracy of surgical navigation. The purpose of this study was to present an automated coarse registration (ACR) methodology to improve the convenience and accuracy. For this purpose, a deep learning model based on a convolutional neural network was used. The input variable used for learning was virtual patient point-cloud (VPPC) generated based on medical image. Output variables were values of coordinate transformation obtained in the process of sending the VPPC to the surrounding space of a medical image. The ACR model consisted of a step of extracting global features of point-clouds from medical image and patient space and a step of predicting the information of 3-dimensional coordinate transformation through global features. The coefficients of determination that evaluated the similarity between predicted and actual rotation values on the x, y, and z axes were 0.993, 0.989, and 0.990, respectively. The coefficients of determination of the predicted and actual translation values on x, y, and z were 0.993, 0.989, and 0.994, respectively. As a result of coarse registration of three phantoms using the ACR, the registration errors between the patient and the computed tomography point-cloud were 3.813 ± 0.792, 3.786 ± 0.734, and 3.653 ± 0.668 mm, which were significantly improved over the conventional method's registration error (4.671 ± 0.738, 4.865 ± 0.776, and 4.670 ± 0.455 mm). The proposed method can provide convenience in the pre-operative preparation stage by automating coarse registration. It is expected that repeatability and reproducibility can be provided by eliminating random errors that might occur by the operator.

INDEX TERMS Surgical navigation system, coarse registration, registration error, deep learning, convolution neural network, mystery curve.

I. INTRODUCTION

The main purpose of image-guided surgical navigation (IGSN) is to provide accurate information by displaying 3D position information of the tip of the surgical tool inserted into the body in real time to the operator during minimally invasive surgery [1], [2], [3]. Minimally invasive surgery is a type of surgery that causes less damage by minimizing incisions with medical devices such as IGSN [4]. The advantages of this type of surgery include less pain, shorter hospital stays, and fewer complications [5]. The IGSN is widely used in precision surgery fields such as head and neck surgery because it enables a high success rate with a reduced time for a minimally invasive surgery [6], [7], [8], [9], [10]. An important factor that determines the accuracy of the IGSN system is the "Registration" step. It means matching the space of two

The associate editor coordinating the review of this manuscript and approving it for publication was Gyorgy Eigner.

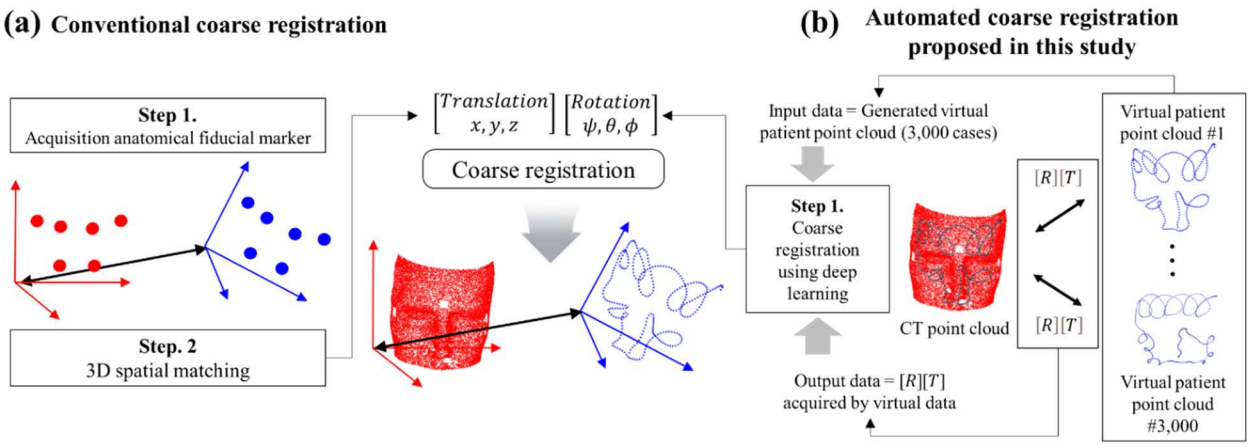

FIGURE 1. Differences between the conventional coarse registration and the automated coarse registration (ACR) proposed in this study.

point-clouds acquired for the same surgical position in the medical image and the patient space with the same coordinate system [10]. The registration is classified into surface registration and point registration [2], [3], [11], [12]. In general, point registration is widely used in neuro navigation clinical due to its relatively stable and high accuracy [3], [6], [13], [14], [15], [16]. However, it has limitations such as infection by fiducial markers implanted in the body and delay of operation time [1], [2], [3], [12], [17], [18], [19]. Accordingly, a surface registration method has been proposed [1], [10]. Related studies have been conducted to improve registration accuracy that is relatively low compared to point registration [1], [2], [9], [10], [12], [17], [19], [20], [21], [22].

Surface registration consists of coarse registration and fine registration [23]. The coarse registration aligns medical image and patient spatial information existing in different spaces into one space, which is a necessary step before performing fine registration [3], [23], [24].

The result of coarse registration directly affects the accuracy of surface registration because fine registration is the step of precisely matching the distance between two points-clouds roughly aligned by coarse registration. Conventional methods of coarse registration include a method using anatomical features of the surgical site [25] and a method using 3D coordinate information of fiducial markers placed in the corresponding position [26], [27]. However, in the manual coarse registration, the accuracy of surface registration might be lowered because of different acquisition positions of feature points that occur depending on the skill of the operator [23]. Using the method based on fiducial markers, problems such as scarring, risk of infection, and delay in operation time can occur due to implants fixed in the body. Coarse registrations that do not use anatomical features or fiducial markers have been proposed to overcome pitfalls of conventional surface registration. Serej et al. [28] have obtained anatomical features of patient's face by projecting a chessboard image to the face to improve the accuracy of surface registration. Liu et al. [23] have obtained a point-cloud for the entire face using a 3D scanner and proposed a coarse registration method based on intrinsic shape signatures. Fan et al. [2], [12] have applied a 3D scanner to overcome problems caused by fiducial markers used as reference points for coarse registration and suggested a coarse registration method based on a 3D coordinate information of point-cloud acquired on the entire head. In order to overcome limitations of the conventional coarse registration, these studies have used data acquisition equipment instead of using facial features or fiducial markers. These methods entail logistic requirements such as additional equipment and a calibration process due to that.

Therefore, the purpose of this study was to propose a deep learning-based ACR methodology to overcome limitations of the conventional coarse registration. The methodology consists of two main steps. First, input and output variables required for training of the deep learning model are generated. Input variables were virtual three-dimensional point-clouds with continuous and geometric shapes generated from computed tomography (CT) point-cloud. Output variables were relative position information of the patient point-cloud with respect to the CT point-cloud. In the second step, a deep learning model was developed to predict relative 3D rotations and translations between the generated virtual 3D point-cloud and the CT point-cloud. The purpose of this study was to propose a novel methodology based on deep learning that could automatically perform coarse registration as an initial strategy of surface registration without needing additional equipment.

The contributions of this study can be summarized as follows:

- The fiducial markers acquisition step necessary in the current coarse registration can be replaced by the "Automated coarse registration" method suggested in this study without the need for additional equipment.
- This method can overcome drawbacks like scarring, infection risk, and operation time delays that may occur due to the placement of fiducial markers in the body.

• This method eliminates the logistic issues that could arise from using additional data acquisition equipment.

• By reducing operator intervention, this automatic method can also increase reproducibility and repeatability.

In this study, we presented a methodology for improving the accuracy of coarse registration, which is the first stage, and the study was organized as follows: Section III provides an overview of the methodology proposed in this study (Section III-A), preprocessing (Section III-B), conventional coarse registration (Section III-C), proposed automated coarse registration (Section III-D), and a comparison of results obtained using the conventional and proposed coarse registration methods (Section III-E). Sections IV and V present the results and discussions for the methods discussed in Section III.

II. RELATED WORKS

A novel coarse registration methodology was proposed in this study to improve surface registration accuracy. Several previous studies attempted to improve surface registration accuracy.

To overcome the limitation that the error of the conventional surface registration method increased at the posterior parts of the head, Fan et al. [2] proposed a surface registration method based on a point-cloud of the entire head. The proposed surface registration method provided sufficient accuracy not only in the anterior but also in the posterior region of the head. Dong et al. [1] proposed selective scanning regions of the head (the sphenoid-frontal zone, parietal zone, left temporal zone, right temporal zone, and occipital zone) to improve surface registration accuracy in the surgical area of interest. The point-cloud of the regions of the head was obtained using a handheld scanner. They showed reliable accuracy in the sphenoid-frontal, parietal, and temporal areas. Yoo et al. [9] reported that the irregular intervals of point-cloud due to the general surface registration method's tracking speed affect registration accuracy. To solve the problem, irregular intervals were regenerated into regular intervals by using cubic spline interpolation. They demonstrated that using a point-cloud with regular intervals improves the accuracy of surface registration. Fan et al. [17] proposed an automatic coarse-to-fine spatial registration method to improve registration speed, ensure equivalent or higher registration accuracy, and eliminate the need for a manual coarse registration process. The proposed method showed an obvious speed advantage (17.4s and 21.4s for the elastic phantom and five patients, respectively) as well as an improvement in registration accuracy (1.17 ± 0.04 mm and range 1.64 mm to 1.83 mm for the elastic phantom and five patients, respectively). Choi et al. [20] extracted the optimal point-cloud that has the minimum distance from the point-cloud in image space. The optimal point-cloud was used to improve surface registration accuracy, and it was validated using a hemisphere mathematical model and a plastic facial phantom. The proposed registration method outperformed the optical and electromagnetic registration methods in terms of accuracy.

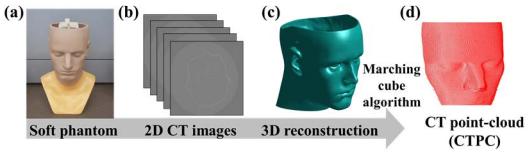

FIGURE 2. Overall process for obtaining CT point-cloud for soft head phantom based on 2D CT images. (a) A soft phantom made of silicone similar to human skin. (b) 2D medical image scanned by CT. (c) Reconstruction of 2D CT images into 3D isosurfaces. (d) CT point-cloud extracted using marching cube algorithm.

Yoo and Sim [22] proposed automated machine learning-based surface registration process to improve the registration accuracy. Using a neural network model and Bayesian optimization, they extracted a new point-cloud that matched the image space point-cloud. The proposed methodology reduced the average registration error by 57.8% when compared to conventional surface registration.

Surface registration is consisted of coarse and fine registration steps [23], and the coarse registration is an essential step before performing fine registration [3], [23], [24]. Therefore, the accuracy of coarse registration has a direct impact on the accuracy of surface registration. However, most studies have concentrated on fine-registration, the final step in surface registration.

III. MATERIALS AND METHODS

A. OVERVIEW

3-D rotation and translation information that can transform two different point-clouds located in the medical image space and the patient space into the same space is required to perform the coarse registration. Fig. 1 shows differences between the conventional coarse registration and the deep learning-based ACR methodology proposed in this study. In the conventional coarse registration, several fiducial markers as references are required to match point-clouds located in two different spaces (Fig. 1(a)). In our ACR, instead of using fiducial markers, various types of virtual patient point-clouds (VPPCs) are generated based on CT point-cloud (CTPC). The rotation/translation information between the VPPC and the CTPC is then calculated. VPPCs and 3D rotation/translation values were used as input and output variables of the deep learning model, respectively, for the development of the ACR (Fig. 1(b)).

B. ACQUISITION AND PREPROCESSING OF POINT-CLOUD IN CT IMAGE

The CTPC was used as an input variable of the deep learning model. It was also used to generate VPPCs. Fig. 2 presents the process of acquiring the CTPC. The CTPC of the head phantom was obtained by reconstructing a 2D tomographic image scanned with a CT into a 3D image and then applying a marching cube algorithm to the 3D image (Fig. 2(a)-(d)) [29], [30]. In this study, three soft head phantoms made of silicone similar to human skin hardness (shore hardness = 20A) were used [31]. 2D tomographic images of each soft head phantom were scanned with a rivo CT 385 (General Electronics Medical Systems, Milwaukee, Wisconsin, USA).

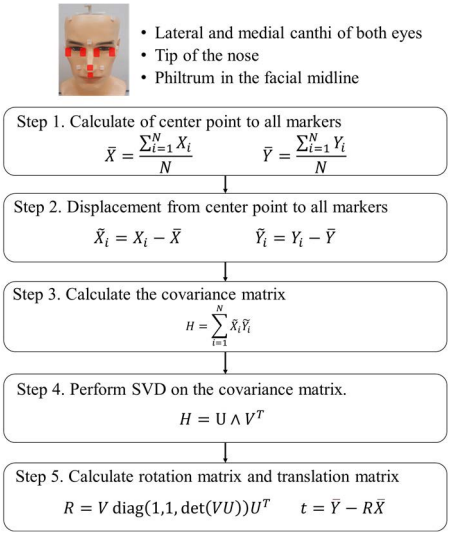

FIGURE 3. Sequential process of conventional coarse registration.

CT images had a voxel size of 0.625 mm, a resolution of 512 × 512 pixels, and a pixel spacing of 0.4883 mm. Digital Imaging and Communications in Medicine (DICOM) data of 2D medical images were loaded at pixel spacing intervals and reconstructed into 3D isosurfaces. The range used for the CTPC was from lips to the forehead except for the mouth where the surgical tube was inserted.

C. CONVENTIONAL COARSE REGISTRATION

The conventional coarse registration was performed by matching anatomical features obtained by CT scan and surgical navigation equipment. Anatomical features included lateral and medial canthus of both eyes, a tip of the nose, and the philtrum in the facial midline, which were usually the thin site of the face [9], [25].

Rotation/translation matrices were finally calculated based on six anatomical features through the following five steps in the conventional coarse registration used in this study (Fig. 3). In Step 1, X and Y mean feature coordinates of patient and medical spaces, respectively. $\bar{X}$ and $\bar{Y}$ are center points of all features ($i = 1 \sim N$) in the two spaces. $\tilde{X}_i$ and $\tilde{Y}_i$ are displacements between each feature and the center point, respectively. The covariance matrix (H) for the two features was calculated with displacements (Steps 2 and 3) and decomposed into two orthogonal matrices and one diagonal matrix using singular value decomposition (Step 4). The rotation matrix (R) was calculated using the two orthogonal matrices. The translation matrix (t) was computed with R, features of patient space, and medical features (Step 5).

D. AUTOMATED COARSE REGISTRATION PROPOSED IN THIS STUDY

1) GENERATION OF INPUT AND OUTPUT VARIABLES FOR AUTOMATED COARSE REGISTRATION

Input variables used in the ACR model were CTPC and the VPPCs distributed around that. Output variables were relative positions ($R_X, R_Y, R_Z, T_X, T_Y,$ *and* T_Z) of VPPCs with respect to CTPC. Accurate output variables corresponding to input variables are required to develop a deep learning model based on supervised learning. However, it is not possible to obtain quantitative relative coordinate transformation matrix between the point-cloud of patient space obtained in the real world and the point-cloud of the medical image space. To overcome this limitation, VPPCs were generated based on the CTPC in this study. The overall process of generating the VPPC is shown in Fig. 4.

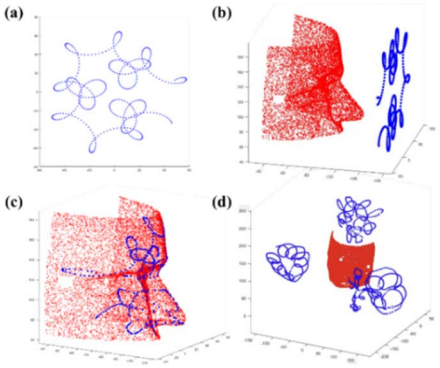

FIGURE 4. The process of generating virtual patient point clouds based on mystery curves and CT point-cloud. Red and blue indicate CTPC and VPPCs, respectively. (a) Generation of 2D mystery curve. (b) Coordinate transformation of the mystery curve in front of the CT point-cloud. (c) Generation of virtual patient points based on CT point-cloud. (d) Generation of VPPCs based on CT point-cloud.

First, a 2D trajectory in the shape of continuous with a starting point and an ending point was generated (Fig. 4(a)). Patient point-cloud used in the IGSN was acquired in a different shape each time because it was directly acquired by the operator. The mystery curve method used to design geometric shapes such as symbols, logos, trademarks, and bank-draft watermarks was applied to generate VPPCs with the same characteristics as the patient point-cloud [32]. The mystery curve is expressed as a trigonometric function of x and y based on the complex function c(t). It is converted into an exponential function to generate a geometrical curve (Equation (1)):

$$\begin{aligned} \mathrm{c}(t) &= x(t) + iy(t) \\ &= (\cos(at) + \frac{\cos(bt)}{2} + \frac{\sin(ct)}{3}) + i(\sin(at) \\ &\quad + \frac{\sin(bt)}{2} + \frac{\cos(ct)}{3}) \\ &= e^{a \cdot Time_p \cdot i} + \frac{e^{b \cdot Time_p \cdot i}}{2} + \frac{e^{c \cdot Time_p \cdot i}}{3} \end{aligned} \tag{1}$$

The mystery curve can be expressed with a period of 0 to 2π, which reflects time series characteristics of the patient point-cloud used for the IGSN. The $Time_p$ of Equation (1) represents time information of time series data. A total of 1,000 points were generated by equally dividing 0 to 2π equally into 1,000. In addition, 2D mystery curves with different trajectories were generated by applying real values with one decimal place between −15 and 15 to three

(a, b, c) frequency components of Equation (1). In the second step, the 2D mystery curve was moved to face the front of the CTPC using the orthogonal coordinates system of each point-cloud (Fig. 4(b)). In the third step, 3D virtual patient points (VP points) were generated by orthographically projecting each point of the 2D mystery curve onto the CTPC (Fig. 4(c)). Before performing the fourth step, a three-dimensional rotational movement value to be used as an output variable was obtained under the rotation/translation condition of Equation (2). Equation (2) is a mathematical process of coordinate transformation using the Euler angle. The rotation transformation is performed in the order of z, y, and x axes, followed by translation transformation [33].

$$\begin{aligned}
&\begin{bmatrix} VPPC\ Z \\ VPPC\ Y \\ VPPC\ X \end{bmatrix} \\
&= R_Z(\psi)R_Y(\theta)R_X(\phi)\begin{bmatrix} VP\ points\ Z \\ VP\ points\ Y \\ VP\ points\ X \end{bmatrix} \\
&\quad + \begin{bmatrix} T_Z \\ T_Y \\ T_X \end{bmatrix} \quad \begin{matrix} -\dfrac{\mathbf{90}\cdot\pi}{\mathbf{180}} < \boldsymbol{\psi}, \boldsymbol{\theta}, \boldsymbol{\phi} < \dfrac{\mathbf{90}\cdot\pi}{\mathbf{180}}\ (\text{rad}) \\ -\mathbf{150} < T_Z, \boldsymbol{T}_Y, \boldsymbol{Y}_X < 150\ (\text{mm}) \end{matrix} \\
&= \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \\
&\quad \times \begin{bmatrix} VP\ points\ Z \\ VP\ points\ Y \\ VP\ points\ X \end{bmatrix} + \begin{bmatrix} T_Z \\ T_Y \\ T_X \end{bmatrix} \qquad (2)
\end{aligned}$$

The coordinate transformation matrix calculated by Equation (2) was applied to VP points in order to obtain VPPC to be used as an input variable of the ACR (Fig. 4(d)). In addition, the distance from the origin of CPPC to the origin of VPPC was calculated to analyze the distance distribution of VPPCs based on the CTPC. To quantitatively analyze the distribution of VPPCs centered on the CTPC, distance analysis was performed by classifying the range of Equation (2) into inner space (0-50 mm), middle space (50-100 mm), and outer space (100-150 mm).

2) DEEP LEARNING MODEL FOR AUTOMATED COARSE REGISTRATION

A deep learning model for automatically registration of two point-clouds located in the CT space and the patient space was proposed in this study without using fiducial markers or additional equipment. Fig. 5 shows the deep learning architecture of the ACR model.

The ACR model is a structure that can predict 3D rotation/translation between two point-clouds by combining features of each point-cloud extracted from the CTPC and the VPPC. The architecture has four steps. The CTPC and VPPC used as input variable consisted of 5,000 points and 1,000 points, respectively. The initial number of channels was set to be three because 3D coordinates (x, y, and z) were used in the model.

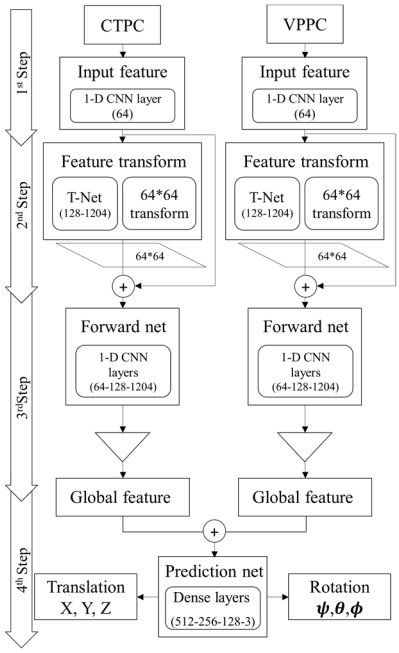

FIGURE 5. Deep learning model architecture for automated coarse registration (Step 1 to 4). In the first step, 1-dimensional CNN is used to extract features from CT point clouds (CTPC) and virtual patient point clouds (VPPC). In the second step, a transformation matrix that takes into account the rigid motion invariant of the point-cloud is obtained using T-Net and multiplied with the input feature. The network based on 1-dimensional CNN extracts global features in the third step, and a network composed of dense layers predicts translation and rotation values in the fourth step.

In the first step consisting of a 1D convolution layer, 64-channels, and a batch normalization layer, features were extracted from input variables (the CTPC and the VPPC). In the second step, the T-Net structure was applied to satisfy the characteristic of rigid motion invariant of the point-cloud, which should not affect the result even if various transformations were performed on the input variable [34]. The T-Net consisted of two 1D CNN layers (128, 1024), a max pooling layer, two dense layers, and a batch normalization layer applied between each layer. The size of the transformation matrix (64×64) obtained through the T-Net was difficult to optimize. Thus, the weight was normalized to 64 channels before transferring to the third structure. In addition, features extracted in the first step were combined after the step of feature transform to maintain initial features of input variables. In the third step, features of point-cloud were extracted to satisfy the permutation invariant property of input variables. Global features were acquired with a symmetric function [34]. Features of point-cloud were extracted through three 1D-CNNs with the number of channels of 64, 128, and 1024. Global features of 1,024 for each point-cloud were generated through max pooling. In the last step, each global feature for the CTPC and VPPC was concatenated into one feature. The feature was used as an input variable for predicting 3D rotations and translations. The last structure consisted of four dense layers reduced to 512, 256, 128, 3 to predict 3D

rotations and translations. A five-fold cross validation was performed for training and internal validation. VPPCs were not duplicated between training and validation sets. Hyper-parameters (leaky ReLU activation function, a learning rate of 0.001, and ADAM optimizer) were tuned with extensive empirical experiments.

3) EVALUATING THE PERFORMANCE OF THE ACR MODEL

The performance of the ACR was evaluated based on root mean square error (RMSE) between predicted values ($Predicted\ XYZ_i$) and actual values (XYZ_i) used as output variables. The RMSE was calculated with Equation (3). XYZ_i and $Predicted\ XYZ_i$ were the ith actual value and the ith predicted value, respectively, in Equation (1).

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (XYZ_i - Predicted\ XYZ_i)^2}{n}}, \text{[mm] or [rad]} \tag{3}$$

The coefficient of determination (R^2) was calculated to analyze the linear relationship between predicted and actual values (Equation (4)), as shown at the bottom of the next page. R^2 could quantitatively express whether there was a correlation between the two groups. It was calculated as the ratio of the sum of variances to total variances. $\overline{XYZ}$ and $\overline{Predicted\ XYZ}$ were averages of actual and predicted values, respectively.

In addition, the performance of the model was evaluated by calculating the distance from the origin of CPPC to the origin of actual/predicted VPPCs.

Additionally, differences between values predicted by the ACR model and actual values were statistically analyzed for rotations (R_X, R_Y *and* R_Z) and translations ($T_X T_Y$, *and* T_Z).

E. COMPARING DIFFERENCES OF REGISTRATION ERROR BETWEEN CONVENTIONAL COARSE REGISTRATION AND ACR

The registration error between the two methods was analyzed to compare performances of the conventional method and the proposed method. The registration error was calculated with Equation (5), which quantified average Euclidean distances between corresponding points in the patient space and medical image space. The $CT\ point_i$ was the ith point of medical images point-cloud and $Patient\ point_i$ was the ith point of patient point-cloud coordinated by coarse registration [10]. CT and patient points are made up of positional information on the x, y, and z axes, and the Euclidean distance between the points is calculated using Equation 5. The registration error is expressed as one parameter of the Euclidean distance.

$$\text{Registration error} = \frac{\sum_{i=1}^{n} \left\| CT\ point_i - Patient\ point_i \right\|}{n} \tag{5}$$

The registration error of ACR was also calculated with Equation (5). However, since ACR was performed automatically, there was no point-cloud of the medical image space that could calculate the error after performing the registration.

TABLE 1. Registration errors calculated with conventional coarse registration for each phantom and trial.

	Phantom #1	Phantom #2	Phantom #3
Trial #1	5.481 mm	4.162 mm	4.068 mm
Trial #2	5.124 mm	4.823 mm	4.902 mm
Trial #3	4.881 mm	5.989 mm	4.372 mm
Trial #4	3.629 mm	5.212 mm	4.781 mm
Trial #5	4.238 mm	4.139 mm	5.226 mm
Mean ± S.D.	4.671±0.738 mm	4.865±0.776 mm	4.670±0.455 mm

To obtain data that could exactly match the shape of the VPPC in the medical image space, the VP point extracted from the CTPC was set as the $CT\ point_i$. The difference in registration accuracies calculated for the ACR and conventional method was statistically analyzed with an independent t-test. The statistical significance level was set at 0.05 or 0.01. All statistical analyses were performed using SPSS 15.0 software (SPSS Inc., Chicago, Illinois, USA).

IV. RESULTS

A. MEDICAL IMAGE POINT-CLOUDS OF SOFT HEAD PHANTOMS USED FOR ACR DEVELOPMENT AND CONVENTIONAL COARSE REGISTRATION

A supervised learning-based ACR model was developed using medical images of soft head phantoms. Fig. 6 shows point-clouds obtained from medical images of three soft head phantoms. The CTPC of each phantom was extracted from lower lips to the forehead used in the IGSN. The number of points in each CTPC was 37,709, 24,254, and 23,940, respectively. Points obtained by using marching cube were distributed at regular intervals.

B. REGISTRATION ERROR CALCULATED BY CONVENTIONAL COARSE REGISTRATION FOR SOFT HEAD PHANTOMS

The conventional coarse registration was performed with five replicates for each soft head phantom. Six fiducial markers used as reference points for coarse registration were obtained by a skilled professional. They were attached to lateral and medial canthi of both eyes, the tip of the nose, and the philtrum in the facial midline. Registration errors calculated with the conventional coarse registration method were represented in Table 1. The average registration error was 4.735 ± 0.629 mm (range, 3.629 mm to 5.989 mm).

C. GENERATE CTPC-BASED INPUT VARIABLES AND OUTPUT VARIABLES FOR DEVELOPING THE ACR MODEL

Fig. 7(a) to (c) presents VPPC samples used as input variables to develop the ACR model. A total of 1,000 VPPCs were obtained from each of the three phantoms. They were generated in different shapes by the frequency component of a mystery curve formula [32]. All VPPCs have the same number of points. Thus, a widely distributed point-cloud has a long distance between each point, whereas a narrowly distributed point-cloud has a short distance between them. The average distance between points of all VPPCs

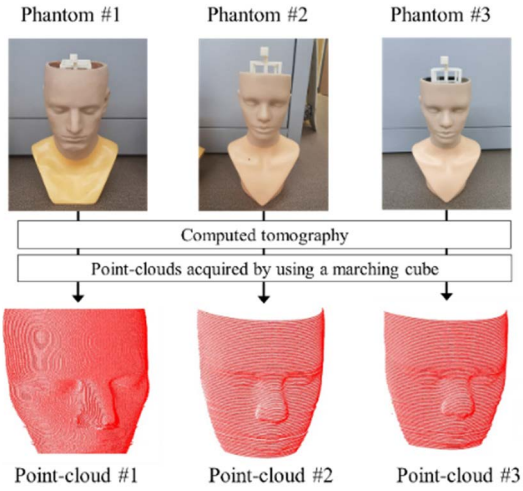

FIGURE 6. Soft head phantom produced for ACR model development and CTPC acquisition results obtained using a marching cube algorithm from CT images of phantoms.

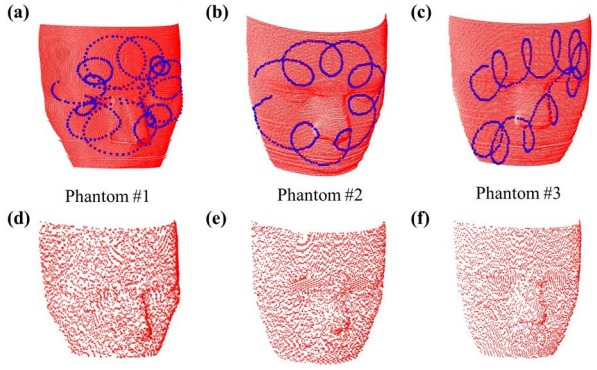

FIGURE 7. Various shapes of VPPC ((a), (b), and (c)) generated by mystery curves formula and the reduced CTPC ((d), (e), and (f)) obtained at regular intervals for use as input variables. Red and blue indicate CTPC and VPPCs, respectively.

was 6.803 ± 3.097 mm (range, 6.232 mm to 7.395 mm). The CTPC has high-resolution data obtained from medical imaging devices consisting of tens of thousands of points. For this reason, the amount of computation is increased in the training model. Therefore, in this study, the CTPC was reduced to 5,000 points by selecting at equal intervals for each phantom (Fig. 7(a), (b), and (c)). These points were then used as input variables for the ACR model.

As shown in Fig. 1, the output variables were 3D rotations and translations used to transform VP points into around the CTPC. Results of coordinate transformation of VP points into around the CTPC with rotation/translation values are

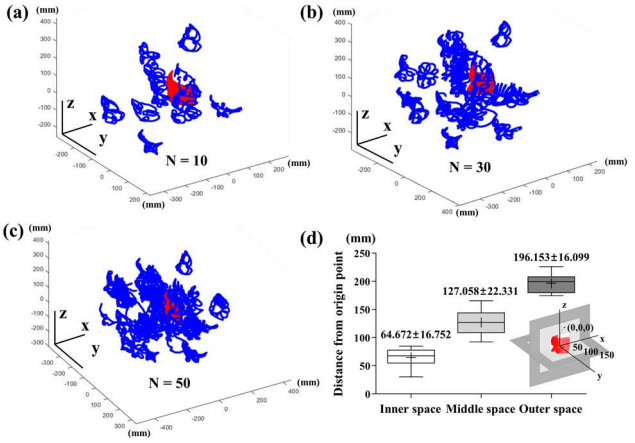

FIGURE 8. 3D space distribution of VPPCs around the CTPC according to the number of VPPCS. Visualization of distribution of (a) 10, (b) 30, and (c) 50 VPPCs around CTPC. (d) Average distance between origin and all VPPCs in three spaces.

shown in Fig. 8. Fig. 8(a) presents results of generating ten VPPCs around the CTPC. Fig. 8(b) and (c) describe results of generating 30 and 50 VPPCs, respectively. In this study, 1,000 3D translation and rotation values were obtained for each phantom. Thus, a total of 3,000 translation and rotation values were obtained. These 3D rotation and translations were used as output variables for the ACR model. Fig. 8(d) shows analysis results by dividing VPPC into three spaces based on the origin to quantitatively analyze VPPCs distribution. The total space was uniformly classified into inner space (0-50 mm), middle space (50-100mm), and outer space (100-150 mm). The average and standard deviation of the distance between the origin and VPPCs were calculated for each space. Distances between origins for VPPCs distributed in each space of inner, middle and outer spaces were 64.672 ± 16.752 mm, 127.058 ± 22.331 mm, and 196.153 ± 16.099 mm. respectively.

D. PERFORMANCE OF THE ACR MODEL DEVELOPED USING GENERATED VPPCs AND 3D ROTATION/TRANSLATION VALUES

RMSE (Equation (3)) and R^2 (Equation (4)) were calculated for actual and predicted values of 3D rotations and translations to evaluate the performance of the ACR model proposed in this study. Table 2 presents results of RMSE for 3D rotation/translation obtained by the proposed model.

Average RMSEs of the rotation obtained through 5-fold cross validation were 0.016 rad, 0.021 rad, and 0.043 rad for R_X, R_Y and R_Z, respectively. Average RMSEs of the translation through 5-fold cross validation were 3.029 mm, 2.846 mm, and 2.883 mm for T_X, T_Y and T_Z, respectively.

$$R^2 = \left[\frac{\sum_{i=1}^{n} (XYZ_i - \overline{XYZ})(PredictedXYZ_i - \overline{Predicted\ XYZ})}{\sqrt{\sum_{i=1}^{n} (XYZ_i - \overline{XYZ})^2} \sqrt{\sum_{i=1}^{n} (PredictedXYZ_i - \overline{Predicted\ XYZ}^2} } \right]^2 \quad (4)$$

TABLE 2. Results of five-fold cross validation of RMSE between prediction and actual values of 3D rotation and translation values to evaluate the performance of the ACR model.

Five-fold Cross-validation	Rotation RMSE (rad)			Translation RMSE (mm)		
	R_X	R_Y	R_Z	T_X	T_Y	T_Z
Fold 1	0.019	0.023	0.038	2.202	2.911	3.195
Fold 2	0.016	0.024	0.034	3.747	2.836	2.128
Fold 3	0.013	0.016	0.045	2.497	3.479	3.692
Fold 4	0.017	0.017	0.047	3.875	2.413	2.107
Fold 5	0.014	0.025	0.048	2.823	2.589	3.294
Mean ± S.D.	0.016 ±0.003	0.021 ±0.004	0.043 ±0.006	3.029 ±0.748	2.846 ±0.406	2.883 ±0.723

S.D., Standard deviation.; RMSE, root mean square error; ACR, automated coarse registration.

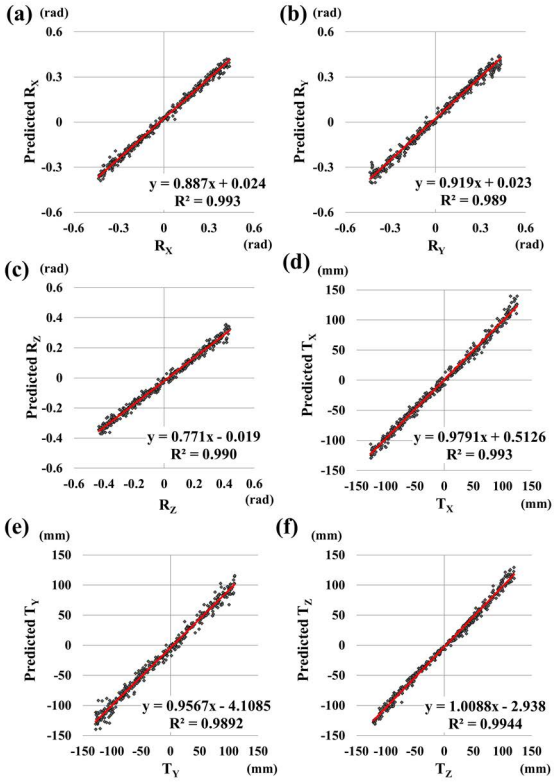

FIGURE 9. Coefficient of determination (R2) between actual and the predicted values of rotation ((a), (b), and (c)) and translation ((d), (e), and (f)) in x, y, and z axes for ACR model.

Fig. 9 shows results of correlation between actual and the predicted values obtained by the ACR model. In Fig. 9, the horizontal line represents actual values, the vertical line represents predicted values obtained through the proposed model, and R^2 means the degree of similarity between actual and predicted values. R^2 values of rotation for x, y, and z axes were 0.993, 0.989, and 0.990, respectively. R^2 values of translation for x, y, and z axes were 0.993, 0.989 and 0.994, respectively. Strong correlations were observed for all three axes. The average R^2 was more than 0.98.

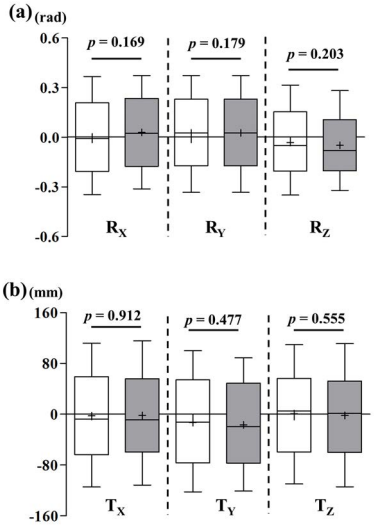

FIGURE 10. Average actual and predicted values of rotation and translation in three axes to evaluate the ACR model (White box: actual value; Gray box: predicted value; Plus point: median; whisker 5–95 percentiles).

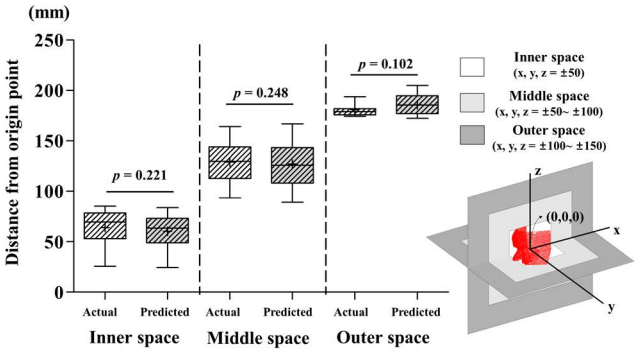

FIGURE 11. Difference between actual VPPC and predicted VPPC compared by calculating the distance from the origin of CTPC to the origin of VPPC. Horizontal line, the space classified according to the distance; Vertical line, the distance from the origin of CPPC to the origin of VPPC. Plus, median; whisker, 5–95 percentiles.

Fig. 10 shows results of comparison between actual and predicted values of the rotation/translation in the three axes. There was no statistically significant difference between the two methods.

Fig. 11 presents calculated distance from the origin of CPPC to origin of VPPCs. As a result of analyzing distances from the origin, there was no statistically significant difference between the actual distance and the predicted distance in any space.

E. REGISTRATION ERRORS BETWEEN CORRESPONDING POINTS IN PATIENT SPACE AND MEDICAL SPACE BY THE ACR AND THE CONVENTIONAL METHOD

To evaluate the performance of ACR model proposed in this study, registration errors calculated by the conventional coarse registration and our model were compared (Fig. 12). Average registration errors of the conventional

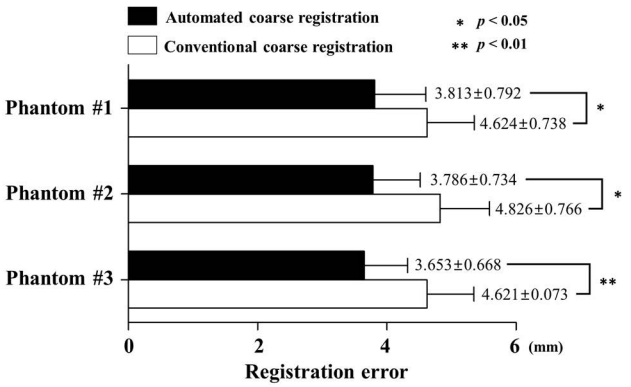

FIGURE 12. Comparison of e registration errors calculated by the conventional coarse registration and the ACR model for three soft head phantoms.

coarse registration were 4.624 mm, 4.826 mm, and 4.621 mm in the three phantoms, respectively. Average registration errors calculated by the ACR model were 3.813 mm, 3.786 mm, and 3.653 mm in the three phantoms, respectively. As a result of comparing registration errors between the two registration methods, registration errors of ACR in all phantoms were lower than those calculated by conventional coarse registration (*, $p < 0.05$; **, $p < 0.01$).

V. DISCUSSION

An ACR based on the deep learning model was proposed to improve the accuracy and convenience of conventional coarse registration. The ACR model is a supervised deep learning model that requires input and output variables. VPPCs generated based on CTPC were relocated through coordinate transformation around the CTPC and used as input variables. The output variables were 3D rotation/translation values calculated in the process of relocating the VPPCs around the CTPC. ACR was developed based on a deep learning model with input and output variables. It was used to match two different point-clouds in medical-image and patient spaces. For this, three soft phantoms were self-made. CTPCs with an average of 28,634 ± 7,861 points were obtained from each soft phantom. Eggers et al. [11] have reported that more than 100,000 points are generally obtained in a head phantom acquired by CT with a resolution of 512 × 512. The average number of points obtained from each phantom was 108,764 ± 17,159 in the present study. However, since the range of CTPC required for the ACR model was from the lower lip to the forehead, the rest of the points were removed in the preprocessing step. CTPCs were used for the conventional coarse registration. They were also used as the references for generating VPPCs and input variables of our ACR model.

The performance of the ACR model proposed in this study was compared with that of the conventional coarse registration. The error of the conventional coarse registration was 4.735 ± 0.629 mm for three phantoms, consistent with results of previous studies that presented registration errors of 4∼5 mm [2], [21], [35]. Additionally, the conventional coarse registration was necessary several anatomical positions as fiducial markers [9], [25]. However, a variation in registration error may occur because it is difficult to obtain accurate positions of fiducial markers even for a skilled professional. The VPPCs used as input variables of the ACR were generated in all different shapes based on the mystery curve formula of the exponential function [32]. The number of points constituting the VPPC was 1,000. The average distance between each point of the VPPC was 6.803 ± 1.597 mm. Considering that the frequency of a general IGSN was 40∼60Hz [36], it is expected that it will take about 20 seconds to track 1,000 points. Qyedare and Park [37] have reported that one of the widely applied rules-of-thumb for deep learning is that the dataset of input variables requires 50 to 1,000 times more data compared to the class to be classified. Therefore, 3,000 input variables were considered for three phantoms to predict six coordinate transformation values in the present study. The 3D rotation/translation values between VPPCs and CTPC calculated by Equation (2) were used as output variables. The maximum diagonal distances of the inner space (50 mm), middle space (100 mm), and outer space (150 mm) classified based on the CTPC were 86.602 mm, 173.205 mm, and 259.807 mm from the origin, respectively. As shown in Fig. 8 (d), distributions of VPPC for the three spaces were 64.672 ± 16.752 mm, 127.058 ± 22.331 mm, and 196.153 ± 16.099 mm away from the origin, respectively. These results meant that not only the VPPCs, but also the 3D rotation/translation values had various ranges.

The performance of the ACR model was analyzed using RMSE, R^2, and 3D rotation/translation values. In the RMSE analysis, the average rotation and translation errors calculated by the conventional coarse registration and ACR model were 0.026 rad and 2.919 mm, respectively. As a result of analyzing the rotation and translation values, there were no statistical differences between the actual and the predicted values in three axes (Fig. 10). In addition, strong correlations were observed for results of the coefficient of determination to analyze the linear relationship between actual and predicted values. However, R_z was relatively lower than R_x and R_y. In the point-cloud viewed from the front of the phantom, the distribution of points was widely distributed along x and z axes. On the other hand, the point-cloud viewed from above (z axes) the phantom was relatively narrow compared to the distribution viewed from the front. The reason for the low accuracy of R_z than other axes was considered to be due to insufficient coordination information of the z-axis in the dimension reduction process, which resulted in insufficient extracted features. The registration error calculated by the ACR model was 3.751 ± 0.085mm on average, which was significantly lower than the error of the conventional coarse registration (4.735 ± 0.628 mm) ($p < 0.05$ or 0.01). Hui et al. recently analyzed the error of a coarse registration algorithm based on 3 to 8 fiducial markers [38]. As a result, an error of 5.292 to 6.329 mm was presented when three fiducial points were used, and an error of 2.446 to 4.794 mm when eight

fiducial points were used. These results were comparable to or worse than our result (average error 3.751 mm). The most significant difference between ACR and conventional coarse registration, including the study of Hui et al., is that it proposed an automated method. The improvement rate of the ACR model was 20.762 ± 2.085% on average. Improvement rates for the three phantoms were 18.366%, 22.168%, and 21.753%, respectively. The ACR model, however, has technological limitation. The ACR model's input variables, VPPC, were built in a wide range of the entire front of the face to acquire facial features. For this reason, if the operator acquires the patient point-cloud narrowly, there is a risk that the registration error may increase. The ACR model does not require a physical step of the conventional method for obtaining the manual anatomical position in the preoperation. It can improve the repeatability and reproducibility because it is possible to remove human error that occurs in the step of acquiring anatomical positions manually by the operator. According to Fan et al. [17], the improvement of initial poses by coarse registration can directly improve the accuracy of fine registration. In their study, the fine registration error was reduced from 1.93 mm to 1.70 mm by improving the coarse registration. Therefore, the ACR method proposed in this study is expected to improve the accuracy of IGSN by reducing the fine registration error.

VI. CONCLUSION

The ACR model based on deep learning developed in this study is a novel method of registration that can improve the accuracy and convenience of conventional coarse registration. Our ACR model produced a soft phantom similar to human skin. Its accuracy in an environment similar to a clinical setting was evaluated. However, this study had a limitation in that the ACR was not applied to clinical trials. Additionally, in this study, the error of coarse registration was improved among two registration steps constituting surface registration. Therefore, in the future, it is necessary to verify the effectiveness by performing the second step of surface registration, fine registration. In addition, clinical trials are required to analyze the registration error in the lesion, which is the ultimate goal of surface registration. The ACR model proposed in this study not only offers convenience and improves accuracy of coarse registration, but also ensures repeatability and reproducibility of coarse registration by eliminating fiducial markers extraction step that the operator has to manually perform. These advantages indicate that the ACR model is applicable as a major technology for IGSN that requires stable surgery by maintaining consistent accuracy. Furthermore, the ACR model can improve the registration accuracy of IGSN by combining it with conventional techniques of surface registration.

REFERENCES

[1] Y. Dong, C. Zhang, D. Ji, M. Wang, and Z. Song, "Regional-surface-based registration for image-guided neurosurgery: Effects of scan modes on registration accuracy," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 14, no. 8, pp. 1303–1315, Aug. 2019.

[2] Y. Fan, D. Jiang, M. Wang, and Z. Song, "A new markerless patient-to-image registration method using a portable 3D scanner," *Med. Phys.*, vol. 41, no. 10, Sep. 2014, Art. no. 101910.

[3] M. N. Wang and Z. J. Song, "Properties of the target registration error for surface matching in neuronavigation," *Comput. Aided Surgery*, vol. 16, no. 4, pp. 161–169, Jul. 2011.

[4] H. G. Stassen, J. Dankelman, K. A. Grimbergen, and D. W. Meijer, "Man-machine aspects of minimally invasive surgery," *Annu. Rev. Control*, vol. 25, pp. 111–122, Jan. 2001.

[5] B. Jaffray, "Minimally invasive surgery," *Arch. Disease Childhood*, vol. 90, no. 5, pp. 537–542, 2005.

[6] A. I. Omara, M. Wang, Y. Fan, and Z. Song, "Anatomical landmarks for point-matching registration in image-guided neurosurgery," *Int. J. Med. Robot. Comput. Assist. Surgery*, vol. 10, no. 1, pp. 55–64, Mar. 2014.

[7] R. R. Shamir, M. Freiman, L. Joskowicz, S. Spektor, and Y. Shoshan, "Surface-based facial scan registration in neuronavigation procedures: A clinical study," *J. Neurosurgery*, vol. 111, no. 6, pp. 1201–1206, Dec. 2009.

[8] R. R. Shamir, L. Joskowicz, and Y. Shoshan, "Fiducial optimization for minimal target registration error in image-guided neurosurgery," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 725–737, Mar. 2012.

[9] H. Yoo, A. Choi, H. Kim, and J. H. Mun, "A novel surface registration for image-guided neurosurgery: Effects of intervals of points in patient space on registration accuracy," *J. Med. Imag. Health Informat.*, vol. 10, no. 6, pp. 1466–1472, Jun. 2020.

[10] H. Yoo, A. Choi, and J. H. Mun, "Acquisition of point cloud in CT image space to improve accuracy of surface registration: Application to neurosurgical navigation system," *J. Mech. Sci. Technol.*, vol. 34, no. 6, pp. 2667–2677, Jun. 2020.

[11] G. Eggers, J. Muhling, and R. Marmulla, "Image-to-patient registration techniques in head surgery," *Int. J. Oral Maxillofacial Surgery*, vol. 35, no. 12, pp. 1081–1095, Dec. 2006.

[12] Y. Fan, X. Xu, and M. Wang, "A surface-based spatial registration method based on sense three-dimensional scanner," *J. Craniofacial Surgery*, vol. 28, no. 1, pp. 157–160, 2017.

[13] K. Cleary and T. M. Peters, "Image-guided interventions: Technology review and clinical applications," *Annu. Rev. Biomed. Eng.*, vol. 12, no. 1, pp. 119–142, Jul. 2010.

[14] J. M. Fitzpatrick, "The role of registration in accurate surgical guidance," *Proc. Inst. Mech. Eng. H, J. Eng. Med.*, vol. 224, no. 5, pp. 607–622, May 2010.

[15] W. Manning and S. Zhijian, "Distribution templates of the fiducial points in image-guided neurosurgery," *Operative Neurosurgery*, vol. 66, pp. 143–151, Mar. 2010.

[16] C. R. Maurer, J. M. Fitzpatrick, M. Y. Wang, R. L. Galloway, R. J. Maciunas, and G. S. Allen, "Registration of head volume images using implantable fiducial markers," *IEEE Trans. Med. Imag.*, vol. 16, no. 4, pp. 447–462, Aug. 1997.

[17] Y. Fan, X. Yao, and X. Xu, "A robust automated surface-matching registration method for neuronavigation," *Med. Phys.*, vol. 47, no. 7, pp. 2755–2767, Jul. 2020.

[18] R. Marmulla, T. Luth, J. Muhling, and S. Hassfeld, "Automated laser registration in image-guided surgery: Evaluation of the correlation between laser scan resolution and navigation accuracy," *Int. J. Oral Maxillofacial Surgery*, vol. 33, no. 7, pp. 642–648, Oct. 2004.

[19] A. Raabe, R. Krishnan, R. Wolff, E. Hermann, M. Zimmermann, and V. Seifert, "Laser surface scanning for patient registration in intracranial image-guided surgery," *Neurosurgery*, vol. 50, no. 4, pp. 797–803, Apr. 2002.

[20] A. Choi, S. Chae, T.-H. Kim, H. Jung, S.-S. Lee, K.-Y. Lee, and J.-H. Mun, "A novel patient-to-image surface registration technique for ENT- and neuro-navigation systems: Proper point set in patient space," *Appl. Sci.*, vol. 11, no. 12, p. 5464, Jun. 2021.

[21] J.-D. Lee, C.-H. Huang, S.-T. Wang, C.-W. Lin, and S.-T. Lee, "Fast-MICP for frameless image-guided surgery," *Med. Phys.*, vol. 37, no. 9, pp. 4551–4559, Aug. 2010.

[22] H. Yoo and T. Sim, "Automated machine learning (AutoML)-based surface registration methodology for image-guided surgical navigation system," *Med. Phys.*, vol. 49, no. 7, pp. 4845–4860, Jul. 2022.

[23] Y. Liu, Z. Song, and M. Wang, "A new robust markerless method for automatic image-to-patient registration in image-guided neurosurgery system," *Comput. Assist. Surgery*, vol. 22, pp. 319–325, Oct. 2017.

[24] J.-D. Lee, T.-Y. Lan, C.-H. Huang, C.-T. Wu, and S.-T. Lee, "A coarse-to-fine surface registration algorithm for frameless brain surgery," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, pp. 836–839.
[25] P. A. Woerdeman, P. W. A. Willems, H. J. Noordmans, C. A. F. Tulleken, and J. W. B. van der Sprenkel, "Application accuracy in frameless image-guided neurosurgery: A comparison study of three patient-to-image registration methods," *J. Neurosurgery*, vol. 106, no. 6, pp. 1012–1016, Jun. 2007.
[26] H. Chen, Z.-C. Li, C.-S. Liao, Q.-S. Zhu, and J. Gu, "Development and evaluation of ultrasound-based surgical navigation system for percutaneous renal interventions," in *Proc. Int. Symp. Bioelectron. Bioinf.*, Nov. 2011, pp. 84–87.
[27] F. Meng, F. Zhai, B. Zeng, H. Ding, and G. Wang, "An automatic marker-less registration method for neurosurgical robotics based on an optical camera," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 13, no. 2, pp. 253–265, Feb. 2018.
[28] N. D. Serej, A. Ahmadian, S. Mohagheghi, and S. M. Sadrehosseini, "A projected landmark method for reduction of registration error in image-guided surgery systems," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 10, no. 5, pp. 541–554, May 2015.
[29] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," *ACM SIGGRAPH Comput. Graph.*, vol. 21, no. 4, pp. 163–169, 1987.
[30] X. Chen, L. Xu, H. Wang, F. Wang, Q. Wang, and R. Kikinis, "Development of a surgical navigation system based on 3D slicer for intraoperative implant placement surgery," *Med. Eng. Phys.*, vol. 41, pp. 81–89, Mar. 2017.
[31] V. Gupta and A. Chanda, "Biomechanics of skin grafts: Effect of pattern size, spacing and orientation," *Eng. Res. Exp.*, vol. 4, Jan. 2022, Art. no. 015006.
[32] S.-C. Pei and K.-W. Chang, "The mystery curve: A signal processing point of view [lecture notes]," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 158–163, Nov. 2017.
[33] R. P. Patera, "Attitude propagation for a slewing angular rate vector," *J. Guid., Control, Dyn.*, vol. 33, no. 6, pp. 1847–1855, 2010.
[34] R. Q. Charles, H. Su, M. Kaichun, and L. J. Guibas, "PointNet: Deep learning on point sets for 3D classification and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 652–660.
[35] C. Zhang, Y. Liu, Y. Zhang, and H. Li, "A hybrid feature-based patient-to-image registration method for robot-assisted long bone osteotomy," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 16, no. 9, pp. 1507–1516, Sep. 2021.
[36] A. Sorriento, M. B. Porfido, S. Mazzoleni, G. Calvosa, M. Tenucci, G. Ciuti, and P. Dario, "Optical and electromagnetic tracking systems for biomedical applications: A critical review on potentialities and limitations," *IEEE Rev. Biomed. Eng.*, vol. 13, pp. 212–232, 2020.
[37] T. Oyedare and J.-M.-J. Park, "Estimating the required training dataset size for transmitter classification using deep learning," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, Nov. 2019, pp. 1–10.
[38] C. Hui, W. Jie, L. Jianming, W. Shiwen, and H. Qing, "Patient registration of intelligent surgical navigation system-based on improved ICP algorithm," in *Proc. 7th Int. Conf. Signal Image Process. (ICSIP)*, Jul. 2022, pp. 430–435.

HAKJE YOO received the Ph.D. degree in biomechatronics engineering from Sungkyunkwan University, Suwon, South Korea, in 2021. He is currently a Research Professor at the College of Medicine, Research Institute for Medical Bigdata Science, Korea University. His research interests include artificial intelligence, surgical navigation, electrocardiography, cardiovascular disease, clinical support systems, signal processing, big data, and research for clinical support using deep learning and machine learning based on various data in the medical domain.

TAEYONG SIM received the Ph.D. degree in biomechatronics engineering from Sungkyunkwan University, Suwon, South Korea, in 2011. He is currently an Assistant Professor with the Department of Artificial Intelligence, Sejong University, Republic of Korea. His research interests include the application of optimization, machine learning/deep learning, and XAI techniques to biomechatronics and energy forecasting fields.