

Article

A Projection-Based Augmented Reality System for Medical Applications

Jong-Chih Chien ¹, Jiann-Der Lee ^{2,3,4,*}, Chai-Wei Chang ² and Chieh-Tsai Wu ^{3,5,6} 

¹ School of Informatics, Kainan University, Taoyuan 33857, Taiwan

² Department of Electrical Engineering, Chang Gung University, Taoyuan 33302, Taiwan

³ Department of Neurosurgery, Chang Gung Memorial Hospital, Linkou, Taoyuan 33305, Taiwan

⁴ Department of Electrical Engineering, Ming Chi University of Technology, New Taipei City 24330, Taiwan

⁵ College of Medicine, Chang Gung University, Taoyuan 33302, Taiwan

⁶ Medical Augmented Reality Research Center, Chang Gung Memorial Hospital, Linkou, Taoyuan 33305, Taiwan

* Correspondence: jdlee@mail.cgu.edu.tw; Tel.: +886-3-2118-800 (ext. 5316)

Featured Application: Projection-based Augmented Reality System for Medical Applications.

Abstract: The aim of this paper was to present the development of an Augmented Reality (AR) system which uses a 2D video projector to project a 3D model of blood vessels, built by combining Computed Tomography (CT) slices of a human brain, onto a model of a human head. The difficulty in building this system is that the human head contains, not flat surfaces, but non-regular curved surfaces. Using a 2D projector to project a 3D model onto non-regular curved 3D surfaces would result in serious distortions of the projection if the image was not uncorrected first. This paper proposed a method of correcting the projection, not only based on the curvatures of the surfaces, but also on the viewing position of the observer. Experimental results of this system showed that an average positional deviation error of 2.065 mm could be achieved under various test conditions.

Keywords: augmented reality; 2D projection; CT



Citation: Chien, J.-C.; Lee, J.-D.;

Chang, C.-W.; Wu, C.-T. A

Projection-Based Augmented Reality
System for Medical Applications.

Appl. Sci. **2022**, *12*, 12027. <https://doi.org/10.3390/app122312027>

Academic Editor: João M.

F. Rodrigues

Received: 8 October 2022

Accepted: 23 November 2022

Published: 24 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Prior to performing medical surgery, which requires both skills and experience, the surgeons usually perform re-examination of the patient's data to ensure the success of the surgery. There are many methods for visualizing patient data, and one such method involves the use of augmented reality (AR) [1]. The current state of AR uses dedicated devices, such as a head-mounted display (HMD) [2,3], which places additional burden on the doctor, because the device needs to be worn constantly. In order to alleviate the doctor of the weight of the device, a projection-based single-viewer AR system is proposed in this paper. Such a system would interactively project a 3D model of cranial blood vessels onto a physical 3D model of a human head, using a 2D video projector. The difficulties of such an AR system includes a movable observer, who may accidentally block the projection or the phantom (occlusion), projecting a 2D image onto 3D surfaces with non-regular curves, and the demand for the accuracy required for medical applications, which can be around only 2.5 mm when using a HMD (3D model to phantom registration error) [4] as an interface. The problem of the movable observer may be resolved by solving for the observer's viewing perspective, and then updating the display accordingly. Once the perspective of the observer is found, the other problems can then be corrected by adjusting the projection to the surface curvatures based on the perspective of the observer. In this investigation, the prototype of such a system was built and then tested. Such an augmented reality system can aid medical professionals by being useful in many applications; including

pre-surgery simulations, the training of medical students, visualizations of lesions within the human body, etc.

Our proposed method to calculate for the observer's perspective is through tracking the observer's eye(s). There are two categories of methods for tracking eye movements: contact-based and contactless. The methods most used in contact-based eye-tracking are the Search Coil method [5] and the Electro-oculography method [6]. The Search Coil method uses a specially designed soft contact lens, where an induction coil is added between the two soft lenses, and a fixed magnetic field is added around the eyeball. So, whenever the observer's eyeball rotates, it drives the rotation of the lens, and then the induction coil induces an electromotive force, due to the change of the magnetic flux. The magnitude of the induced electromotive force is used to calculate the deflection angle of the eyeball. However, the disadvantage of this method is that this measurement method is easily affected by the condition of the observer's eyeball, such as fluid secretion, etc., so it is not suitable for long-term application. Furthermore, the soft lens has a double-layer structure, which eventually adversely affects the observer's vision. The method of electro-oculography attaches electrodes to the skin around the observer's eyes to measure the voltage difference between the retina and the cornea. The difference in voltage between the electrodes is then used to calculate the movement pattern of the eyeball. However, the accuracy of this method can be affected by the secretion of keratin from the skin, resulting in unstable electrical signals. This implies that this method is also not suitable for public use.

In contact-less methods, the most common methods are the Purkinje image tracking method (also called Dual-Purkinje-Image, DPI) [7], the Pupil Tracking method [8], and the Infra-Red Oculography (IROG) method [9]. The DPI method uses the property that different layers of tissues of the eyeball have different refractive indices, and their reflected images are different. This method can achieve high accuracy, but the cost of the necessary equipment is prohibitively high and, thus, cannot be widely used. The Pupil Tracking method irradiates the eye with infrared and near-infrared light, and, because the pupil has a low reflectivity for infrared rays, and the iris has a high reflectivity for infrared rays, an image is generated between the pupil and the iris. The brightness increases as the iris becomes larger, while the brightness becomes smaller between the iris and the white of the eye, so it becomes easier to grasp the contour of the pupil. The direction of eyesight can then be determined by detecting the position of the pupil. Similarly, the IROG method also irradiates a row of infrared light on the iris, and, because the sclera almost completely reflects the infrared light relative to the iris and pupil, when the eye rotates, the position of the pupil and the rotation of the eye can be computed using the position and intensity of the light reflected by the sclera. However, infrared light can cause damage to the eye, and methods based on it are susceptible to errors caused by the influence of external light sources.

For the position of the observer, face detection is an effective method. There are various features for detection used in different studies, including the Haar cascades [10–12], which has been demonstrated not to be affected by scales difference or minimal occlusion by hair, but is only suitable for frontal detection, and susceptible to insufficient lighting and partial face occlusion [13]. The Histogram of Oriented Gradients (HOG) was probed [14] for this application, and though it can be used to detect faces at slightly inclined angles, in addition to frontal, it still fails at extreme angles. A Single-Shot-Multibox-detector (SSD) model, combined with deep learning, were examined in [15–17], and found to successfully detect the faces of different scale changes in the feature maps generated by the deep learning networks, but, due to the limitation of fewer intermittent definite anchor scales, the accuracy in detecting smaller faces may be too low. In determining the location of the observer, a 3D spatial layout detection device would also be very helpful, in addition to facial detection. Though electromagnetic devices have been used for spatial positioning [18], they are not really suitable for medical environments. Recent studies in medical fields have used optical spatial positioning devices, such as NDI's Polaris Vicra optical tracker [19], or Intel RealSenseTM [20] with dual-camera setup, which uses infrared cameras to track

position of the patient or the phantom, and uses the CCD cameras to track moving objects in 3D.

Image projection-based augmented reality for the application of neurosurgery was proposed fairly early. Tabrizi et al. [21], in 2005, proposed a projection-based augmented reality system, which would project the image of the region of interest in the cranial area onto the surface of the patient's head. One of the problems of such a system is the problem of registration, which is the problem of seeking to align the position of the 3D model in the computer with the real-world coordinate of the patient's head, or a phantom representing a head. In [21], the registration problem was solved by placing five feature markers on the patient's head before surgery, and also marking the positions corresponding to these five markers in the 3D model. In this way, when projecting, the computer only has to achieve perfect alignment with respect to the positions of these five markers to ensure projection accuracy. Witkowski, et.al. [22] proposed another projection-based system consisting of four parts: target tracking, head tracking, target transforming for different viewing angles, and, finally, the projection. The heart of this proposed system is in the construction of the 3D descriptors of the surface contours of the 3D model, and in the transformation of the same model under different viewing angles. In the construction of the target object, contour-based shape descriptors of the surface of the model are generated, and, then, the distance from the center of each contour to its edges at every possible angle is calculated, from 0 to 359. In target tracking, a red dot is first placed 45 mm above the mid-points between the eyes to be used for tracking. The computer then calculates the viewing angle based on the position of the red dot and its deflection from the mid-point between the eyes.

2. Methodology

The flowchart of the proposed system is shown below in Figure 1, which is composed of four major components: the separation of blood vessels from the CT images, the alignment of the 3D point cloud to the real-world coordinates, adjusting the projection surfaces with respect to the viewer, and, finally, projecting the adjusted model onto the phantom of a head for augmented reality display. Digital subtraction angiography is used to cut out the vascular portions from each CT image. The spatial position of the optical sensor or RGB-Depth camera is used to align the point cloud virtual coordinate system representing the surfaces of the head with the coordinates of the actual head in the real-world coordinate system. The first part for adjusting the display of the 3D model according to the observer's perspective is to detect the face portion of the observer, which is then combined with the extraction of facial feature points in order to not only obtain the observer's position but also his or her perspective. In order to adjust the display of the 3D model when using a 2D projector the scanlines are first projected onto the phantom following the computed perspective of the observer, and the distortions of the scanlines across the surface of the phantom are then recorded and used to calculate the slopes of changes across the surface of the phantom. The 3D model of the blood vessels is then adjusted according to the surface distortions, and, then, the projector projects the adjusted blood vessels onto the surface of the phantom head.

The setup of the proposed system is shown in Figure 2. The processing of the data is as follows. First, the areas in the images containing blood vessels and bones are manually marked and, then, Otsu threshold and region growing methods are used to separate them from the images. In order to avoid over-segmentation, the Digital Subtraction Angiograph (DSA) [23] is used, which is an algorithm that augments blood vessel centerlines with their radii, and structures such as bones are subtracted from the image in order to constrain the 3D reconstruction of the blood vessels. Finally, small nodules and noise are removed from the 3D model. An example is shown in Figure 3.

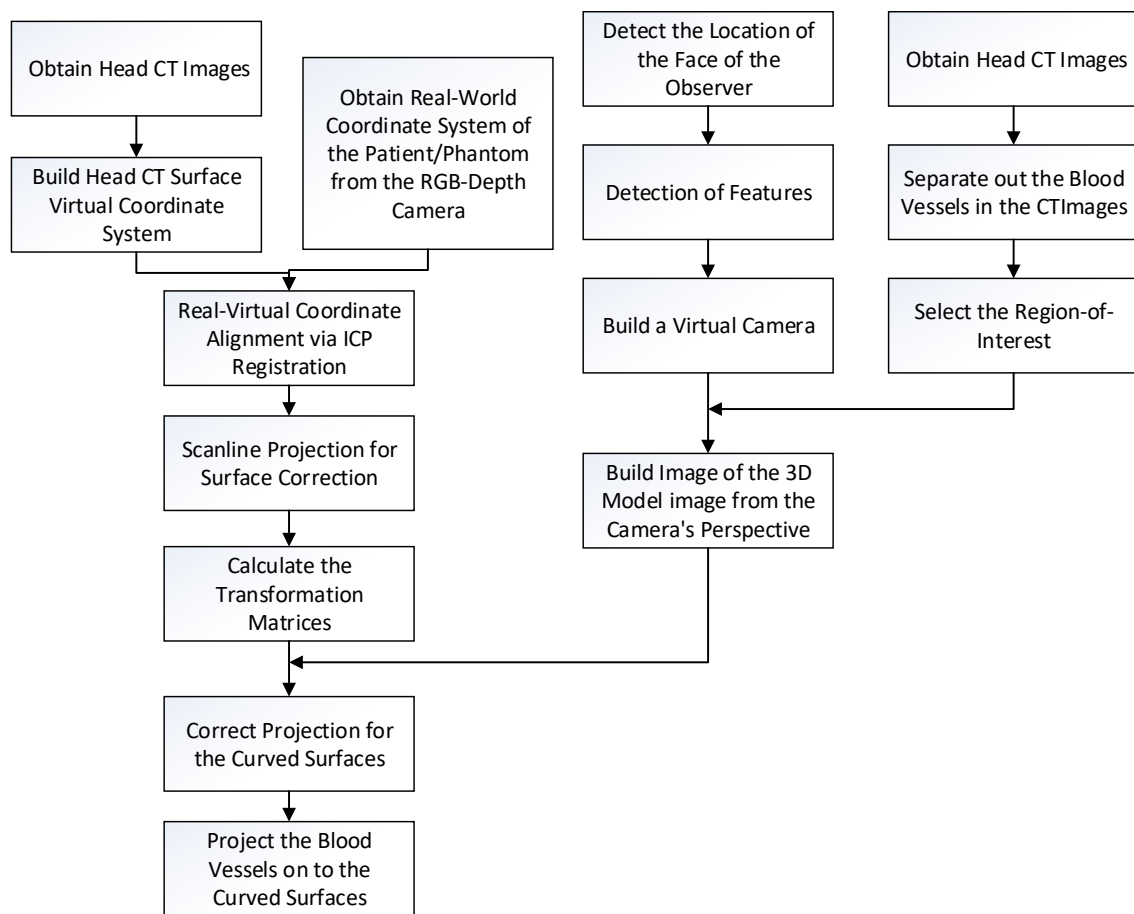


Figure 1. The flowchart of the proposed projection-based augmented reality system.

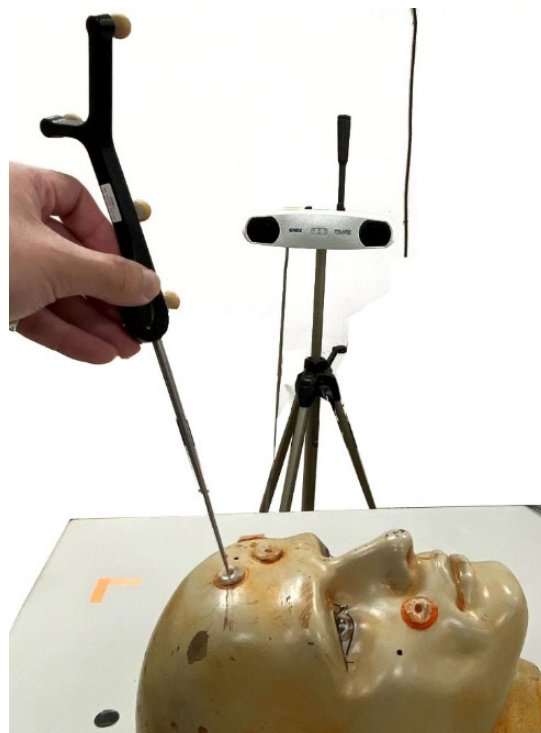


Figure 2. The hardware setup of the proposed system.

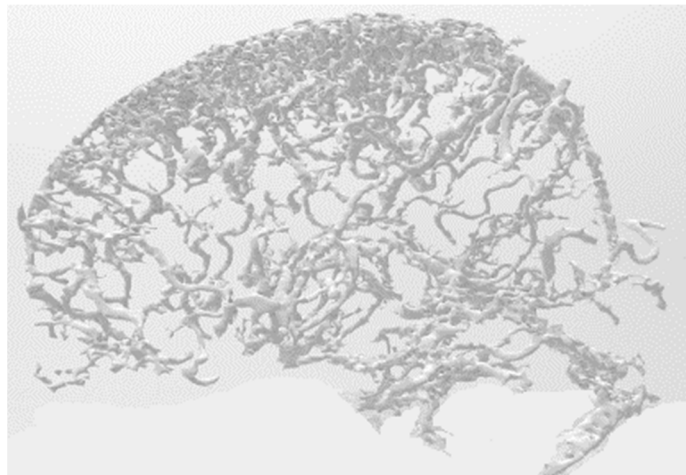


Figure 3. Example of a 3D model of cranial blood vessels.

2.1. Image Alignment

In order to align the 3D model of the blood vessels, the real-world space coordinates and the CT image space coordinate must be aligned by calculating for the transformation matrix based on the two coordinate groups. In order to achieve this goal, an Iteration Closest Point (ICP) algorithm is used. The ICP method is based on the one proposed by Besl [24], which aligns floating point data to reference point data, so that the conversion relationship between the two sets of data can be obtained and used to calculate the geometric transformation matrix. Based on the characteristics of the data, the determination of the initial point of search is important in this ICP algorithm. There is no guarantee of reaching the global optimum. So, a modification in the form of adding a perturbation mechanism should help the algorithm reach a better solution even if a random initial starting point is used. The modified ICP algorithm is as followed:

1. Let the floating point group corresponding to the patient's facial features, extracted from the camera image, be F , where $F = \{f_i(x,y,z), 1 \leq i \leq N_f\}$, and let the floating data point group corresponding to the facial features, extracted from the CT images, be R , where $R = \{r_j(x,y,z), 1 \leq j \leq N_r\}$.
2. Pick a random point from F , assume it is f , and seek its closest corresponding point in R by calculating the minimum distance between f and R , d , as:

$$d(f, R) = \frac{1}{M} \min_{j \in (1, \dots, N)} \|f_i - r_j\| \quad (1)$$

where M is the normalizing constant.

3. Calculate the median distance, Median, of all the distances. Assume the distances are re-arranged in order, then:

$$\text{Median}(d) = d_j, \text{ where } j = \begin{cases} \frac{N_f}{2}, & N_f \text{ is odd} \\ \frac{N_f+1}{2}, & N_f \text{ is even} \end{cases} \quad (2)$$

4. Assign weight to each pair, based on the distance between each pair of points.

$$W_i = \begin{cases} 1, & \text{if } d_i < \text{median} \\ \frac{\text{median}}{d_i}, & \text{otherwise} \end{cases} \quad (3)$$

5. Calculate the root mean squared error (RMS):

$$\text{RMS} = \sqrt{\frac{1}{N_f} \sum_{i=1}^{N_f} W_i \cdot d_i} \quad (4)$$

6. Calculate and record the transformation matrix, T , of current pairs. If the termination condition, based on RMS, is reached, output T as the final transformation matrix. However, if the termination condition is not reached, but the RMS value is smaller than the previous iteration, then replace the optimal transformation matrix with the current T .
7. If the termination condition is not reached, a perturbation mechanism is used; that is, applying a perturbation matrix on the current pairings. The probability of perturbation is based on the Gaussian distribution. The purpose of perturbation is allowing the search outside the current solution space, which may be small, and can allow for a locally optimum solution.
8. Restart the ICP again.

The flowchart of the modified ICP algorithm is shown in Figure 4.

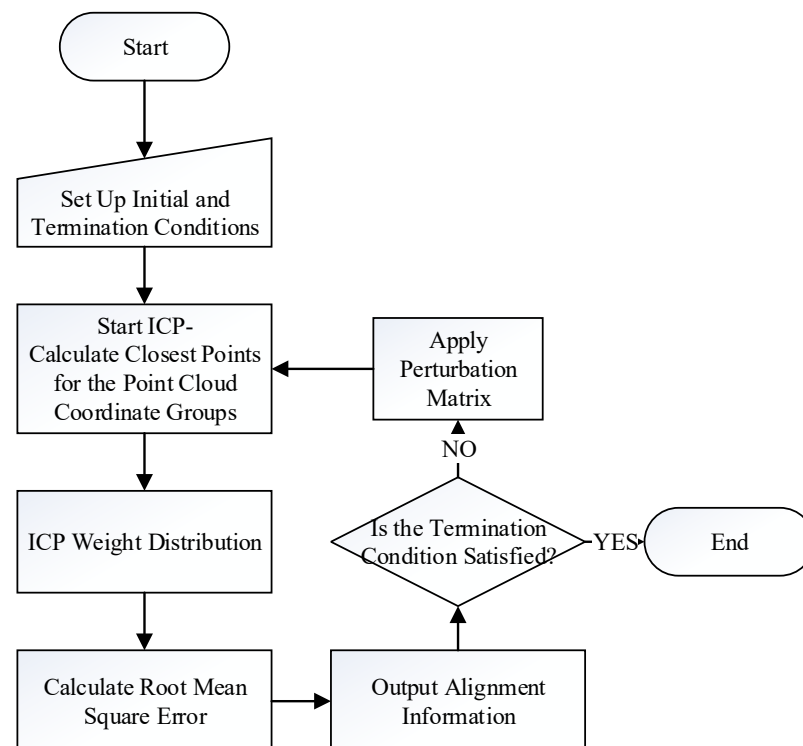


Figure 4. Flowchart of the modified ICP algorithm used in this investigation.

2.2. Capture the Position of the Observer

In this study, the RealSense™ camera was used to capture the position of the observer. The position of the observer was determined by using facial features. The Max-Margin Object Detector (MMOD) algorithm [25,26] was used to detect the observer's face in the camera images, then five-point facial key points detection was performed to locate the edges of both eyes, and the nose tip [27], as shown in Figure 5.

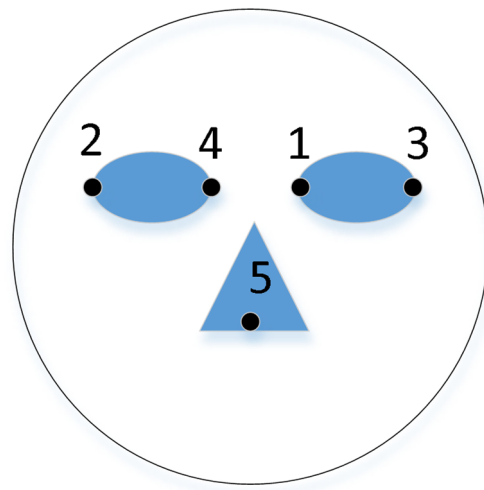


Figure 5. The five-point facial feature points used in face detection.

The Max-Margin Object Detection (MMOD) algorithm is a maximum edge object detector, based on the Convolutional Neural Network (CNN), and can help solve the problem which occurs when the observer is so far away from the camera that the features cannot be correctly identified. Once the feature points are identified, the center between the two eyes is set as the position of the observer. Once the position of the observer is determined in the real-world, a virtual camera, representing the observer, is placed in the virtual world of the 3D model and faces the position of the 3D model of blood vessels. The movements of the observer in the real-world is matched by the corresponding movements of the virtual camera in the virtual world. The position of the patient or the phantom head is not tracked, but only the location of the observer. The RealSense™ camera can rotate a total of 120 degrees, i.e., 60 degrees to each side of the patient, which is sufficient for the proposed application and so was set as the parameter used in the experiment.

2.3. Three Dimensional Model Surface Correction

The steps to surface correction are as follows:

1. Project a square matrix of scanlines onto the head of the patient or phantom.
2. Use a video camera to capture the distortions of the scanlines on the surface. An example of the scanlines projected onto a head phantom is shown in Figure 6.
3. Thin the captured scanlines in order to obtain a more accurate representation of the matrix, obtaining a matrix of regions.
4. Determine which regions are still fully closed by using the flood filling algorithm from the center of each region. This is useful for obtaining the coordinates of the intersections of the scanlines.
5. Mark out each region and obtain the coordinates of the intersections.
6. In order to reduce calculations, the user is asked to mark out regions of interest (ROI).
7. Geometric corrections are performed for each region in the ROI. An example of a blood vessel before and after adjustment is shown in Figure 7.
8. Project the result.

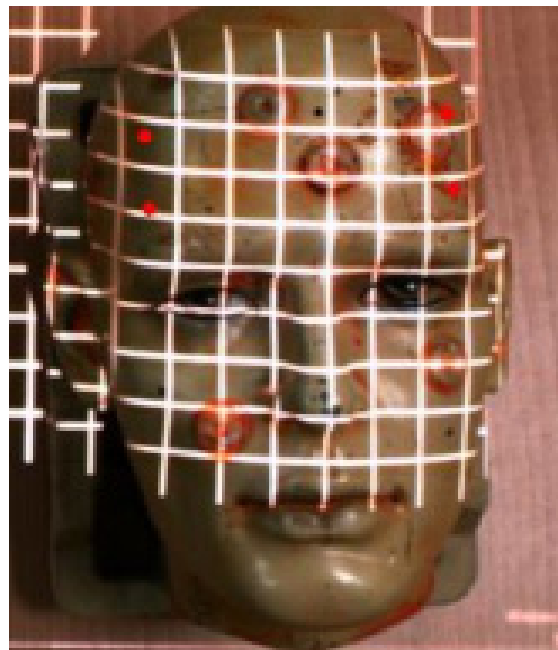


Figure 6. Project 2D scanlines onto a head phantom for calculating surface-based distortions.

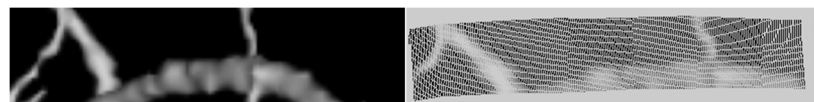


Figure 7. Image of blood vessel before (left) and after (right) geometric adjustment before projection.

A flowchart of this process is shown in Figure 8.

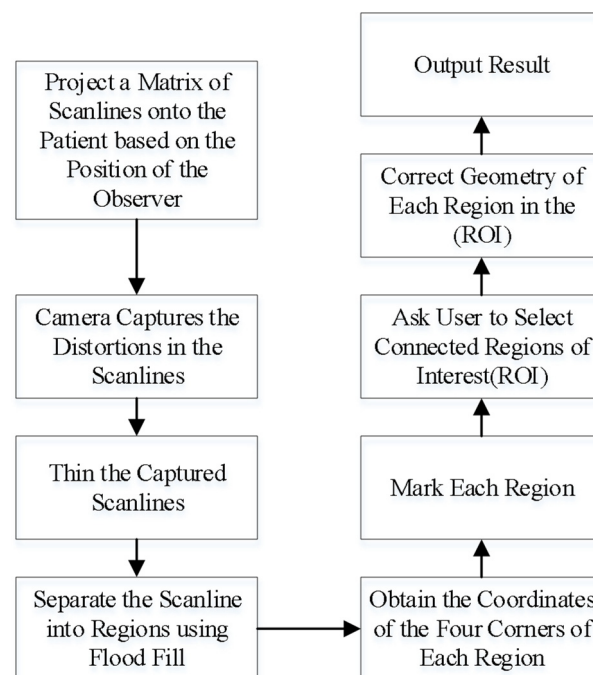


Figure 8. Flowchart of the process of projecting image onto the skin surface of a patient.

3. Experimental Results

The hardware setup for the experimental was as follows: 1. Intel™ i9 computer with 32 GB of memory and one NVIDIA® RTX2080Ti display card. 2. Two RealSense™ D435 cameras. 3. One Optoma ML750 video projector. 4. Two head phantoms. The head phantoms and the markers used in the experiments are shown in Figure 9.

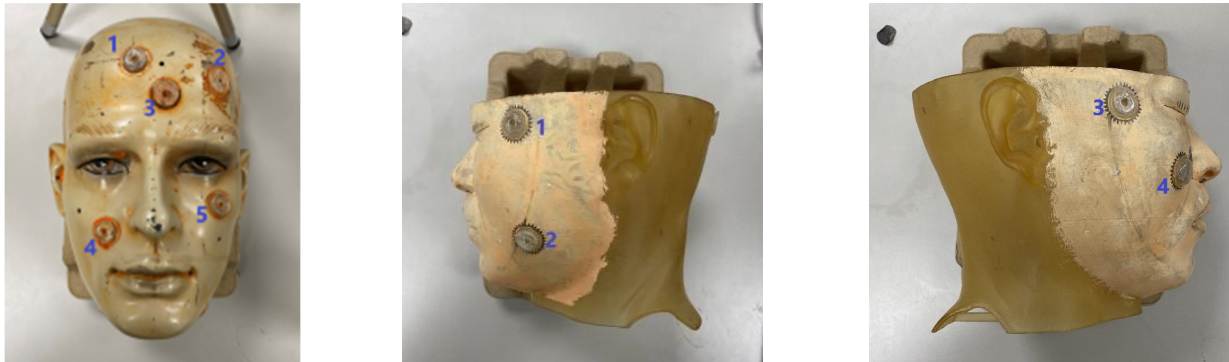


Figure 9. Head phantom 1 (left), and head phantom 2 (middle and right).

The first head phantom had five markers, P1–P5, and the second phantom had four markers (Q1–Q4). The first experiment was conducted to determine the best face detector out of the following four algorithms: the OPENCV implementations of the Haar-cascade based face detector [28], the DNN-based face detector [29], the Dlib implementation of HOG and Linear SVM face detector [30], and the MMOD face detector [31]. Evaluation was based the precision of the detection of the five feature key points mentioned above, as well as six to eight feature key points or facial landmarks which not only captured the points mentioned above, but also the surface around the mouth and the face.

The second part of the experiment was to test the proposed method for any deviation in the projection. In this experiment, both head phantoms were used. The deviations were calculated based on the distance from the locations of markers placed on the head phantoms being projected from the actual markers, and from the perspective of a stationary observer. This was followed by an experiment of the observation from a different location. Lastly, the observation moved again, and the light in the laboratory was dimmed. This part of the experiment was repeated five times, and the results shown in the tables below were the average of the five runs.

3.1. Experiments for Speed

First, we tested the speed of each implementation, by calculating the frames-per-second for display. The following figure, Figure 10, illustrates the results.

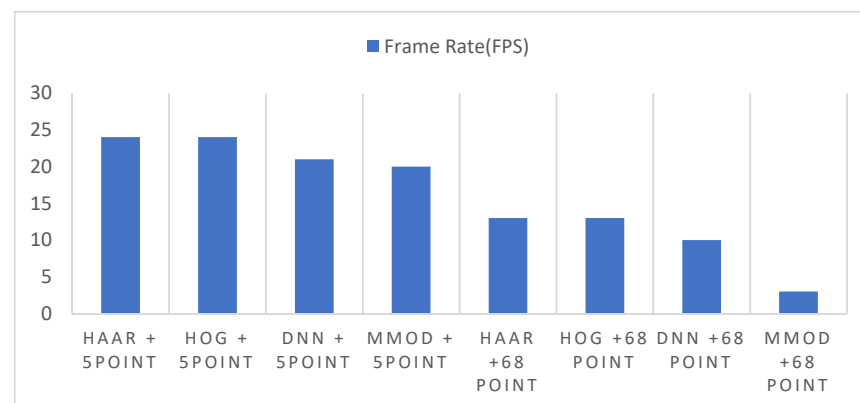


Figure 10. Speed comparison for the four methods detecting 5 and 64 feature points.

3.2. Experiments for Accuracy

Second, the accuracy of capturing the feature points was examined. For this part of the experiment, the Face Detection Data Set and Benchmark (FDDB) [32] was used, which contained 5171 faces with position labels. The results of this experiment are shown in Table 1.

Table 1. Accuracy comparison of four different methods in face detection.

Algorithm	Precision	Std. Deviation	Recall	Std. Deviation
Haar	80.23%	5.12%	44.56%	30.04%
HOG + SVM	83.78%	4.45%	48.61%	28.41%
DNN	89.33%	2.17%	70.94%	13.86%
MMOD	98.14%	1.08%	80.26%	9.79%

The second part of this set of experiments was to test the proposed method for any deviation in the projection. In this experiment, both head phantoms were used. The deviations were measured based on the distance from of the locations of markers placed on the head phantoms being projected from the actual markers, and from the perspective of an observer in front of the camera at different angles. Every experiment in this part was repeated five times, and the results shown in the tables below were the average of these five runs. Table 2 shows the average results for an observer with uncovered face.

Table 2. Deviations of the projections of markers from physical markers on head phantoms in mm.

Deg.	P1	P2	P3	P4	P5	P.Avg.	Q1	Q2	Q3	Q4	Q Avg.
0	1.51	1.58	1.51	1.54	1.5	1.528	1.61	1.43	1.48	1.58	1.525
+10	1.53	1.54	1.53	1.5	1.58	1.536	n.a.	n.a.	n.a.	n.a.	
+20	1.54	1.55	1.54	1.51	1.53	1.534	1.57	1.54	1.67	1.61	1.598
+30	1.7	1.54	1.52	1.55	2.21	1.704	n.a.	n.a.	n.a.	n.a.	
+40	2.21	1.52	1.5	*	1.52	1.688	1.68	1.71	1.66	1.7	1.688
+50	2.5	1.55	1.7	*	1.54	1.823	n.a.	n.a.	n.a.	n.a.	
+60	2.4	1.52	1.78	*	1.54	1.81	1.69	1.62	1.73	1.77	1.703
−10	1.57	1.97	1.53	1.51	1.51	1.618	n.a.	n.a.	n.a.	n.a.	
−20	1.58	1.96	1.52	1.52	1.74	1.664	1.55	1.64	1.7	1.66	1.638
−30	1.56	2.4	1.56	1.51	1.68	1.742	n.a.	n.a.	n.a.	n.a.	
−40	1.53	2.33	1.58	1.59	*	1.758	1.51	1.49	1.47	1.53	1.5
−50	1.58	2.21	1.89	1.54	*	1.805	n.a.	n.a.	n.a.	n.a.	
−60	2.2	*	2.38	1.76	*	2.113	1.73	1.68	1.74	1.76	1.728
Total Avg.						1.717					1.625

*—visually blocked, unable to measure; n.a.—no experiment was performed at this setting.

This was followed by an experiment conducted when a mask covered the observer's mouth, simulating a surgeon as observer. Again, the deviation measurements were taken at different angles, and are shown in Table 3.

Table 3. Deviations from markers on head phantoms in mm with masked observer.

Deg.	P1	P2	P3	P4	P5	P.Avg.	Q1	Q2	Q3	Q4	Q Avg.
0	1.62	1.75	1.71	1.64	1.66	1.676	1.78	1.83	1.76	1.85	1.805
+10	1.66	1.78	1.81	1.68	1.7	1.726	n.a.	n.a.	n.a.	n.a.	.
+20	1.7	1.77	1.68	1.76	1.71	1.724	1.86	1.84	1.79	1.81	1.825
+30	1.84	1.89	1.76	1.89	1.73	1.822	n.a.	n.a.	n.a.	n.a.	
+40	2.31	1.75	1.71	*	1.64	1.8525	1.91	1.88	1.93	1.94	1.915
+50	2.52	1.81	1.74	*	1.67	1.935	n.a.	n.a.	n.a.	n.a.	
+60	2.64	1.74	1.89	*	1.91	2.045	1.68	1.77	1.86	1.71	1.755
−10	1.59	1.67	1.64	1.77	1.64	1.662	n.a.	n.a.	n.a.	n.a.	
−20	1.91	1.67	1.76	1.72	1.69	1.75	1.37	1.69	1.58	1.92	1.64
−30	1.86	1.99	1.73	1.68	1.81	1.814	n.a.	n.a.	n.a.	n.a.	
−40	1.84	2.6	1.78	1.76	*	1.995	1.84	1.93	1.78	1.82	1.843
−50	1.76	2.6	1.86	1.61	*	1.958	n.a.	n.a.	n.a.	n.a.	
−60	2.1	*	2.8	1.86	*	2.253	1.76	1.71	1.77	1.83	1.768
Total Avg.						1.862					1.793

*—visually blocked, unable to measure; n.a.—no experiment was performed at this setting.

Lastly, the observer was not only masked, but the light in the laboratory was dimmed, simulating using a projector with low lumens. The results are shown in Table 4.

Table 4. Deviations from markers on head phantoms in mm with masked observer and dim light.

Deg.	P1	P2	P3	P4	P5	P.Avg.	Q1	Q2	Q3	Q4	Q Avg.
0	2.16	2.39	2.86	2.44	2.54	2.478	2.56	2.43	2.12	2.41	2.38
+10	2.34	2.76	2.45	2.67	2.78	2.6	n.a.	n.a.	n.a.	n.a.	.
+20	3.7	3.1	2.87	2.64	2.31	2.924	2.11	2.27	2.19	2.45	2.255
+30	2.51	1.97	1.88	2.34	1.92	2.124	n.a.	n.a.	n.a.	n.a.	
+40	3.41	2.67	2.46	*	2.74	2.82	2.55	2.43	2.69	2.17	2.46
+50	2.57	2.44	2.41	*	2.67	2.523	n.a.	n.a.	n.a.	n.a.	
+60	3.11	2.87	2.92	*	2.96	2.965	2.16	2.58	2.41	2.09	2.31
−10	2.21	2.34	2.4	2.41	2.39	2.35	n.a.	n.a.	n.a.	n.a.	
−20	2.37	2.6	2.78	2.61	2.21	2.514	2.76	2.88	2.43	2.22	2.573
−30	2.7	2.44	2.56	2.17	2.67	2.508	n.a.	n.a.	n.a.	n.a.	
−40	3.57	3.24	3.51	3.09	*	3.353	2.08	1.98	2.34	2.06	2.115
−50	3.87	3.8	3.66	3.91	*	3.81	n.a.	n.a.	n.a.	n.a.	
−60	3.14	*	3.02	2.86	*	3.007	2.61	2.14	1.92	2.68	2.338
Total Avg.						2.767					2.347

*—visually blocked, unable to measure; n.a.—no experiment was performed at this setting.

Examples of projection of blood vessels onto one of the head phantoms for an observer at different degrees are shown in Figure 11. In projecting the blood vessels, the colors of the background were adjusted for better viewing of the blood vessels.

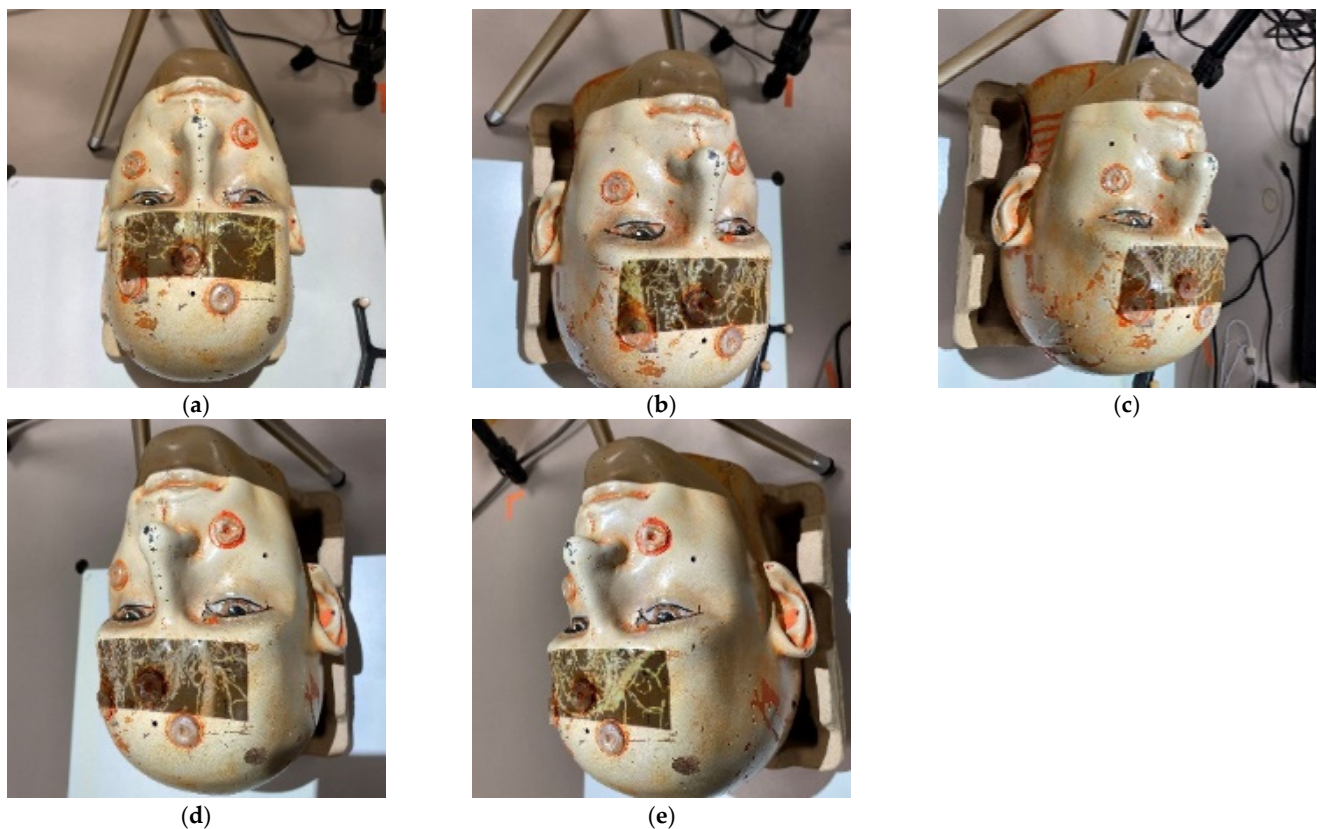


Figure 11. Example of projections of blood vessels with observer at (a) 0 degree, (b) +30 degrees, (c) +60 degrees, (d) −30 degrees, and (e) −60 degrees.

3.3. Experiments for Systems Comparison Purposes

The last experiment was designed to test the accuracy of the NDI system against our proposed system. The viewing angle was assumed to be 0, and both head phantoms were used. The first phantom was laid face up to test the five marker points, and the second phantom was laid first with the right side up, and then with the left side up. The following table, Table 5, shows the average errors of these two systems.

Table 5. Performances of NDI vs. Proposed System on the head phantoms.

System	P1	P2	P3	P4	P5	P.Avg.	Q1	Q2	Q3	Q4	Q Avg.
NDI	1.34	1.81	1.62	1.55	1.83	1.63	1.53	1.62	1.67	1.69	1.630
Proposed	1.54	1.51	1.58	1.51	1.5	1.528	1.48	1.37	1.41	1.53	1.448

4. Discussion

Figure 10 shows that performance of the MMOD system in facial features detection was the slowest of the methods tested. However, in terms of accuracy, as shown in Table 1, it was the best performer. Since the speed performance of the MMOD method was acceptable for the five feature points, it was then incorporated into the proposed system. The results of the second set of experiments showed the average positional deviation for the first head phantom, and the total average errors averaged to around 2.18 mm, while for the second head phantom the average error was around 1.95 mm, so the total average for both heads was around 2.065 mm. This value was quite acceptable for medical applications, since [4] shows that 2.5 mm for HMD is acceptable. It is noted that dimming the light resulted in slightly larger errors for both head phantoms, so it is suggested that a projector with high lumens should be used for our proposed system.

In the last experiment that was performed, which was to compare the existing NDI system with our proposed system, in terms of accuracy, on the phantoms. Even though the experiment was not exhaustive, it did show that, under regular conditions, the proposed system could perform as well as the NDI system.

5. Conclusions

In terms of setup, when compared to the NDI system [19], the proposed system is less bulky, and more portable. This study used the RGB-D camera setup to capture the features and locations of the phantom, which were then converted to cloud points for registration and alignment; a lighter process, which costs less. In order to examine the proposed system under various conditions, two different head phantoms with different marker locations were used. In addition, the observer was moved to different positions to test different viewing angles. The average values of multiple experiments were recorded, and the results showed that the proposed system could perform well, even under strict requirements. The proposed system, under tested conditions, could perform as well as the NDI system, though less bulky and costing less.

The future of this research will try to extend the location to other parts of the human body, so that the prospects of this system are not restricted to only teaching and training. The efforts put into the research of building this system instilled appreciation for the efforts of other investigators in developing methods to build a more accurate model of the human brain, such as the investigation in [33] into relationships between electroencephalogram (EEG) synchronization and emotions.

Author Contributions: Conceptualization, J.-D.L. and J.-C.C.; methodology, J.-D.L. and J.-C.C.; software, J.-D.L. and C.-W.C.; validation, J.-D.L., J.-C.C., C.-W.C. and C.-T.W.; formal analysis, J.-D.L. and J.-C.C.; investigation, J.-D.L., J.-C.C., C.-W.C. and C.-T.W.; resources, C.-W.C.; writing—original draft preparation, J.-D.L. and J.-C.C.; writing—review and editing, J.-D.L. and J.-C.C.; project administration, J.-D.L. and J.-C.C.; funding acquisition, J.-D.L. and J.-C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partly supported by the Ministry of Science and Technology (MOST), Taiwan, Republic of China, under Grants MOST110-2221-E-182-035 and MOST111-2221-E-182-018.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: CT data is not available. Data for faces can be obtained at <http://vis-www.cs.umass.edu/fddb/> (accessed on 5 August 2022).

Acknowledgments: The authors would like to thank the department of Neurosurgery at Chang Gung Memorial Hospital for their help in this investigation.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Eckert, M.; Volmerg, J.S.; Friedrich, C.M. Augmented Reality in Medicine: Systematic and Bibliographic Review. *JMIR mHealth uHealth* **2019**, *7*, e10967. [CrossRef] [PubMed]
2. Barteit, S.; Lanfermann, L.; Bärnighausen, T.; Neuhann, F.; Beiersmann, C. Augmented, Mixed, and Virtual Reality-Based Head-Mounted Devices for Medical Education: Systematic Review. *JMIR Serious Games* **2021**, *9*, e29080. [CrossRef] [PubMed]
3. Parekh, P.; Patel, S.; Patel, N.; Shah, M. Systematic review and meta-analysis of augmented reality in medicine, retail, and games. *Vis. Comput. Ind. Biomed. Art* **2020**, *3*, 1–20. [CrossRef]
4. Gibby, J.T.; Swenson, S.A.; Cvetko, S.; Rao, R.; Javan, R. Head-mounted display augmented reality to guide pedicle screw placement utilizing computed tomography. *Int. J. Comput. Assist. Radiol. Surg.* **2018**, *14*, 525–535. [CrossRef]
5. Kristin, N.H.; Margaret, R.C.; Roberts, D.C.; Charles, C. Della Santina, Low-Noise Magnetic Coil System for Recording 3-D Eye Movements. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 2–8.
6. Abdo, Y.; Yahya, E.; Ismail, H.; Saleh, M. Attention Detection using Electro-oculography Signals in E-learning Environment. In Proceedings of the 10th IEEE International Conference on Intelligent Computing and Information Systems, Cairo, Egypt, 5–7 December 2021; pp. 1–6.

7. Lu, C.; Chakravarthula, P.; Tao, Y.; Chen, S.; Fuchs, H. Improved vergence and accommodation via Purkinje Image tracking with multiple cameras for AR glasses. In Proceedings of the IEEE International Symposium on Mixed and Augmented Reality, Porto de Galinhas, Brazil, 9–13 November 2020; pp. 1–12.
8. Wan, Z.; Xiong, C.; Chen, W.; Zhang, H.; Wu, S. Pupil-Contour-Based Gaze Estimation With Real Pupil Axes for Head-Mounted Eye Tracking. *IEEE Trans. Ind. Informatics* **2021**, *18*, 3640–3650. [\[CrossRef\]](#)
9. Rigas, I.; Raffle, H.; Komogortsev, O.V. Photosensor Oculography: Survey and Parametric Analysis of Designs Using Model-Based Simulation. *IEEE Trans. Human-Machine Syst.* **2018**, *48*, 670–681. [\[CrossRef\]](#)
10. Nguyen, O.; Nguyen, K.; Pham, T.V. A comparative study on application of multi-task cascaded convolutional network for robust face recognition. In Proceedings of the 8th International Conference on Information Technology and its Application, Labuan, Malaysia, 28–29 August 2021; pp. 2–8.
11. Thai, T.; Phan, H.N.; Nguyen, D.T.; Ha, S.V. An improved single shot detector for face detection using local binary patterns. In Proceedings of the 2019 International Symposium on Communications and Information Technologies (ISCIT), Ho Chi Minh City, Vietnam, 25–27 September 2019; pp. 1–6.
12. Kadir, K.; Kamaruddin, M.K.; Nasir, H.; Sairul, I.S.; Bakti, Z.A.K. A comparative study between LBP and Haar-like features for Face Detection using OpenCV. In Proceedings of the Fourth International Conference on Engineering Technology and Technopreneurship, Kuala Lumpur, Malaysia, 27–29 August 2014; pp. 1–5.
13. Chaudhari, M.N.; Deshmukh, M.; Ramrakhiani, G.; Parvatikar, R. Face Detection Using Viola Jones Algorithm and Neural Networks. In Proceedings of the International Conference on Computing, Communication, Control and Automation (ICCUBEA), Pune, India, 16–18 August 2018; pp. 1–4.
14. Dalal, N.; Triggs, B. Histo-grams of oriented gradients for human detection. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–26 June 2005; pp. 1–8.
15. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Cheng, Y.; Berg, A. SSD: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; pp. 21–37.
16. Wang, S.; Wang, K. Real-time and accurate face detection networks based on deep learning. In Proceedings of the International Conference on Electronic Information Technology and Computer Engineering (EITCE), Xiamen, China, 18–20 October 2019; pp. 1541–1545.
17. Sun, J.; Xia, S.; Sun, Z.; Lu, S. Cross-Model Deep Feature Fusion for Face Detection. *IEEE Sensors Lett.* **2020**, *4*, 1–4. [\[CrossRef\]](#)
18. Blum, T.; Heining, S.M.; Kutter, O.; Navab, N. Advanced training methods using an augmented reality ultrasound simulator. In Proceedings of the International Symposium on Mixed and Augmented Reality, Orlando, FL, USA, 19–22 October 2009; pp. 177–178.
19. NDI. Available online: <https://www.ndigital.com/optical-measurement-technology/polaris-vicra/> (accessed on 7 September 2021).
20. Intel. Available online: <https://www.intel.com/content/www/us/en/architecture-and-technology/realsense-overview.html> (accessed on 7 September 2021).
21. Tabrizi, L.B.; Mahvash, M. Augmented reality-guided neurosurgery: Accuracy and intraoperative application of an image projection technique. *J. Neurosurg.* **2015**, *123*, 206–211. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Gierwiało, R.; Witkowski, M.; Kosieradzki, M.; Lisik, W.; Groszkowski, Ł.; Sitnik, R. Medical Augmented-Reality Visualizer for Surgical Training and Education in Medicine. *Appl. Sci.* **2019**, *9*, 2732. [\[CrossRef\]](#)
23. Frisken, S.; Haouchine, N.; Alexandra, R.D.; Golby, J. Using temporal and structural data to reconstruct 3D cerebral vasculature from a pair of 2D digital subtraction angiography sequences. *Comput. Med. Imaging Graph.* **2022**, *99*, 102076. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Besl, P.J.; McKay, N.D. A method for registration of 3D shapes, IEEE Transactions. *Pattern Anal. Mach. Intell.* **1992**, *14*, 239–256. [\[CrossRef\]](#)
25. Zhao, Z.-Q. Object Detection with Deep Learning: A Review. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 3212–3232. [\[CrossRef\]](#) [\[PubMed\]](#)
26. King, D.E. Max-Margin Object Detection. *arXiv* **2015**, arXiv:1502.00046.
27. Sagonas, C.; Antonakos, E.; Tzimiropoulos, G.; Zafeiriou, S.; Pantic, M. 300 Faces In-The-Wild Challenge: Database and results. *Image Vis. Comput.* **2016**, *47*, 3–18. [\[CrossRef\]](#)
28. Available online: https://docs.opencv.org/3.4/d2/d99/tutorial_js_face_detection.html (accessed on 15 January 2022).
29. Available online: https://docs.opencv.org/4.x/d0/dd4/tutorial_dnn_face.html (accessed on 15 January 2022).
30. Available online: <https://github.com/davisking/dlib> (accessed on 16 January 2022).
31. Edirisooriya, T.; Jayatunga, E. Comparative Study of Face Detection Methods for Robust Face Recognition Systems. In Proceedings of the 5th SLAAI International Conference on Artificial Intelligence (SLAAI-ICAI), Colombo, Sri Lanka, 6–7 December 2021; pp. 1–6.
32. Available online: <http://vis-www.cs.umass.edu/fddb/> (accessed on 1 October 2022).
33. Aydın, S. Cross-validated Adaboost Classification of Emotion Regulation Strategies Identified by Spectral Coherence in Resting-State. *Neuroinformatics* **2021**, *20*, 627–639. [\[CrossRef\]](#) [\[PubMed\]](#)