NEURO-FUZZY STEREO CAMERA CALIBRATION ARCHITECTURES USED IN INTERACTIONS WITH MEDICAL VOLUMES

Marius DANCIU, Radu ORGHIDAN, Mihaela GORDAN, Aurel VLAICU

Technical University of Cluj-Napoca Str. C.Daicoviciu nr.15, 0264/401309, danciu.marius@com.utcluj.ro

<u>Abstract:</u> Computer assisted diagnosis/surgery systems need improved solutions for visualization and interactions, as the virtual probe tools. This work presents such an interaction tool and its applications in medical imaging. The novelty is represented by the use of neuro-fuzzy systems in 3D virtual probe positioning; two such configurations are presented and compared to the crisp positioning system. We prove that the neuro-fuzzy solutions outperform the crisp solutions in terms of computational complexity and positioning accuracy. The proposed framework integrates several functionalities: positioning of arbitrary cutting planes; 3D editing, measurements and annotations; stereo visualization. Being a flexible architecture, other functionalities could be integrated in the future.

Keywords: neuro-fuzzy, 3D interaction, 3D editing, stereo calibration

I. INTRODUCTION

Camera calibration is essential in computer vision applications that require metric measurement of the scene such as 3D reconstruction, 3D interaction, object inspection or scene mapping. The most commonly used model for camera calibration is the pinhole camera model that generates the image through a projective transformation from a 3D Euclidean space onto the image plane.

Camera calibration can be described as the process of finding the camera model and implicitly estimation of its parameters.

Explicit camera calibration implies that the calibration process provides accurate physical parameters of the camera. This solution can be seen as a white box-type model. A white box model is the most detailed type of model, and is as close as possible to a complete description of the real system.

Calibration by means of white box-type models has some disadvantages due to the complexity of mathematically approximating the optics of a camera and the associated set-up necessary for the calibration.

Implicit calibration, on the other hand, provides a set of rules that emulate the camera behavior without actually needing the camera parameters. Because of this behavior we can call this a black-box model, such a model is used where the response of a system is not broken down into its underlying mechanisms, and is represented by an empirical description or set of transfer functions and parameters that do not describe any internal physics. This type of model presents a series of advantages. Since black box models usually consist of a set of rules and simple equations, they are easy to be optimized and can run very rapidly. Also it requires minimal computation power due to its relative simplicity. Because a high accuracy but also a low computing overhead is necessary for our calibration process, a neurofuzzy model appears to be the most appropriate solution.

Neuro-fuzzy systems act as universal approximators, i.e. any non-linear function can be approximated with an arbitrary accuracy using a properly tuned structure (fuzzy sets and fuzzy rules) [1], [2].

In the recent years, the advantages of using machine learning techniques for camera calibration have been observed and applied by many researchers. For example, there have been proposed methods that use a black box model to calibrate a single camera [6], [9] or a stereo configuration of cameras [10]. Liu Wan-Yu [7] presented an approach for camera calibration based on the neural network optimized by Genetic algorithms. Dong-Min [8] performed camera calibration using a nonlinear modeling function of an artificial neural network and reports better results than in the case of the Tsai's calibration method [5].

BP neural networks were also used by Li [3] for the calibration of a structured light system and by Ge Dong-Yuan [4] for camera calibration.

A coplanar camera calibration method based on neural network is proposed by Xiabo et al [11], neural network is used here to learn the relationship between the image information and the 3D information to correct aberrance of camera.

Fuzzy neural systems were also used for moving PTZ camera accurately to the target object in the work from [12]. Their system can overcome image center shift in zooming in object, and object deviating from the image frame and outside of the center region when enlarging image. Therefore, fuzzy neural system simulates PTZ camera to move image frame accurately so as to make full use of intelligent monitoring system.

In our previous work we indicated that neuro-fuzzy systems can give better results than crisp camera calibration in terms of accuracy and noise resilience [13].

The approach presented in this paper extends the onestage neuro-fuzzy architecture used for calibrating a stereo camera system to a cascaded multiple stage neurofuzzy architecture, compares the performances of these neuro-fuzzy architectures as well as of the crisp one, and presents their application in medical imaging visualization and volume manipulation.

Traditional interaction devices, such as keyboard or mouse used as a way to interact with windows, icons, menus and pointers (WINP) are not the most appropriate tools for some tasks involving 3D perception and interaction. Ergonomic human-computer interfaces are crucial when dealing with medical applications. The traditional 2D interaction devices, that offer clinicians the possibility to navigate and edit volumetric medical data slice by slice, are no longer satisfactory in daily practice; 3D anatomical objects are mentally reconstructed - as a natural cognitive ability of humans but this is not always desirable. Nowadays, many new software applications can help radiologists in the process of 3D reconstruction of anatomical structures from DICOM images. In this case, 3D stereo visualization or augmented and virtual reality, where applicable, provides radiologists with a more natural way of evaluating and learning patient anatomy, facilitating diagnosis, therapy design, and monitoring.

Following these considerations, it is important to have 3D interaction techniques to manipulate virtual representations of human organs, and this is currently an emerging medical imaging field. Actually, according to Bowman et al. [14], 3D interaction should be defined as the type of human-computer interaction in which user's tasks are done in a 3D context – therefore simple nonobstructive interfaces are required.

Many research teams investigated and developed new interaction and visualization modalities – some of them specifically designed for medical applications. Gratzel et al. [15] presented a non-contact mouse for surgeon – computer interaction using gestures. Feied at al. [16] also developed a hands-free system for visualization of medical images during clinical procedures, to avoid radiologists' contact with the keyboard and the mouse, as potential contamination sources.

De Paolis et al. [17] propose another system that helps surgeons to model patient's organs and interact with them in a more effective way for pre-operative planning and also during the real surgical procedure. The finger movements are detected by means of an optical tracking system and are used to simulate the touch of the virtual interface. Gallo and Ciampi [18] developed a 3D interface for medical data exploration. 3D medical data can be manipulated in a semi-immersive virtual environment, using a data glove with an attached infrared led, tracked by a Wii remote sensor.

Another medical data visualization system is proposed by Cooperstock and Wang [19], which supports 3D display of the content and implements interaction techniques similar to the manipulation of physical objects. Reitinger et al [20] developed a system for planning liver surgical procedures, called LiverPlanner, with the benefits of stereoscopic visualization through HMD and 3D interaction for refinement of medical images through a system that tracks both the body position and the interaction device using an optical tracking system.

However most of these solutions require either a complicated setup, are not flexible enough to be easily used by an untrained user (in terms of a natural simple interaction) or are too expensive in terms of the camera and calibration equipment.

Our goal was to design and implement a flexible, fast and accurate alternative to the existing solutions, by integrating the neuro-fuzzy system in the 3D positioning of the virtual interaction probe. For this purpose, we developed an interaction pen-like device tracked by two web cameras, device that is integrated in a framework for 3D manipulation and 3D editing of medical data.

Structure of the remaining paper is as follows. Section II presents stereoscopic camera modeling. In Section III, we briefly describe the initial neuro-fuzzy architecture and the proposed cascaded neuro-fuzzy architecture. Section IV presents evaluation and experimental results, and the conclusions are drawn in Section V.

II. STEREOSCOPIC CAMERA MODELING

In the area of stereo imaging, accurate calibration of the camera pair is crucial.

There are many techniques to obtain 3D information from a scene, using computer vision such as: modifying the intrinsic camera parameters, i.e. depth from focus/defocus and depth from zooming, considering an additional source of light projected onto the scene, i.e. photometric stereo or structured light, considering additional surface information, i.e. shape from shading, shape from texture and shape from geometric constraints, multiple views, such as stereo and shape from motion. We focus here on the stereo methods i.e. based on multiple views. Each point of a given image determines a ray that crosses the scene and intersects in a unique 3D position with the corresponding rays obtained from other views of the scene. This is known as the triangulation principle.

The triangulation is the process of finding the 3D position of a point from two other points placed along a baseline. In case of stereovision systems, each point of a given image, determines a ray that crosses the scene and intersects in a unique 3D position with the corresponding rays obtained from other views of the scene. Obviously, both the calibration and the correspondence problem must be solved.

As a preliminary step, a simulation procedure has been setup in order to test the performance and robustness of the proposed neuro-fuzzy calibration method. In Figure 1, the two cameras point to a cloud of points.



Figure 1. Stereo configuration for 3D reconstruction.

The projection of the scene points is modeled for each camera by equation (1) as represented in Figure 2.



Figure 2. Geometrical relation between 3D points and 2D representation of points in a 2D image.

$$P_i = {}^{2D} K_C \cdot {}^C K_W \cdot {}^W P \tag{1}$$

Where:

^wP is an object point represented with respect to the world coordinate system,

 P_i is the image point corresponding to the projection of ^WP,

 ${}^{C}K_{W}$ is the transformation matrix that relates the camera coordinate system with respect to the camera coordinate system,

 ${}^{2D}K_{c}$ is the projection matrix that models the internal geometry of the camera by a set of intrinsic parameters:

$${}^{2D}K_{C} = \begin{bmatrix} \alpha_{u} & 0 & u_{0} & 0\\ 0 & \alpha_{v} & v_{0} & 0\\ 0 & 0 & 1 & 0 \end{bmatrix}$$
(2)

$$\alpha_u = -f \cdot k_u, \quad \alpha_v = -f \cdot k_v$$

So, the projection of an object point ^WP onto the camera (obtaining the point P_i) is defined in equation (3): So, we obtain,

$${}^{C_{j}}P_{i} = \begin{bmatrix} s \cdot x_{i} \\ s \cdot y_{i} \\ s \end{bmatrix} \qquad {}^{C_{j}}P_{i} \text{ is the ith image point of the camera Cj.}$$

The 3D points are mapped onto the 2D points by the camera projection, therefore obtaining a bijective relation between the two sets. This relation can be modeled by a neuro-fuzzy system [13].

III. STEREO CALIBRATION USING NEURO-FUZZY ARCHITECTURES

A. The basic uncorrelated neuro-fuzzy (ANFIS) calibration architecture

The model used for camera calibration consists in fuzzy inference systems, represented in special artificial neural network architecture, called Takagi-Sugeno fuzzy systems, [21]-[23].

Sugeno-type FIS structure is generated using subtractive clustering and requires separate sets of input and output data.

Since fuzzy systems are universal approximators, it is expected that their equivalent neural network representations will possess the same property.

The reason for using a fuzzy system in terms of neural network is to utilize the learning capability of neural network to improve performance, such as adaption of fuzzy systems.

The initial proposed method [13] finds a set of rules that define the functionality of the cameras involved in the stereo configuration.

The initial step consists in producing a reliable training set of input-output data that captures all the system features and covers the full input variable space. The data set is then subdued to a subtractive clustering procedure resulting an initial first order Takagi-Sugeno (T-S) fuzzy system. In this step the structure of the fuzzy system is set (number of fuzzy sets and fuzzy rules), and initial values of the fuzzy system parameters (nonlinear parameters for the input fuzzy sets and linear parameters for the output fuzzy sets) are computed. Usually this initial fuzzy system is not accurate enough. To improve the accuracy, the initial fuzzy system is then trained using an adaptive neuro-fuzzy training method (ANFIS) and the training data set. During the training, both types of fuzzy system parameters (nonlinear and linear ones) are adapted. ANFIS uses a combination of least-squares and back-propagation gradient descent-based learning procedure for tuning initial parameters. The resulting fuzzy model is tested from the accuracy point of view. If the model accuracy is acceptable, the modeling procedure stops and provides the desired neuro-fuzzy model. If this is not the case, the procedure can be resumed either by retraining the fuzzy system with different parameters of the training procedure (e.g. an increased number of training epochs), generating a new initial fuzzy system with a different structure, or even by using a new data set.

In our case, with 4 inputs, 2 from each camera, a

$$\begin{bmatrix} s \cdot x_i \\ s \cdot y_i \\ s \end{bmatrix} = \begin{bmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} WP_x \\ WP_y \\ WP_z \\ WP_z \\ 1 \end{bmatrix}$$
(3)

fuzzy rule can be expressed as:

If $(X_1 \text{ is } A_k)$ and $(X_2 \text{ is } B_j)$ and $(X_3 \text{ is } C_l)$ and $(X_4 \text{ is } B_m)$ then $Y_n = p_n X_1 + q_n X_2 + r_n X_3 + s_n X_4 + t_n$

The resulting structure of uncorrelated neuro-fuzzy (black box representation) model is given in Figure 4. Three fuzzy logic systems are involved, one for each dimension of the 3D space. All these fuzzy logic systems take as inputs the coordinates of the two 2D cameras: ${}^{c_1}x_i$, ${}^{c_1}y_i$ and ${}^{c_2}x_i$, ${}^{c_2}y_i$. The output of the neuro-fuzzy systems are the coordinates of the reconstructed points in 3D space: ${}^{3D}x_i$ - on X, ${}^{3D}y_i$ - on Y, and , ${}^{3D}z_i$ - on Z.



Figure 4. Uncorrelated neuro-fuzzy (ANFIS) calibration architecture

In Figure 4 and Figure 5, the block "Initial FiS" refers to the initial fuzzy inference system for ANFIS training, whereas FIS refers to the tuned FIS structure whose parameters are set according to a minimum training error criterion.

B. A cascaded architecture of neuro-fuzzy systems for exploiting data correlation

The Multi-Input Multi-Output Adaptive Neural-Fuzzy Inference Systems (MANFIS) extend the notion of a single output system, ANFIS, to produce multiple outputs.

One way to generate a MANFIS model is to place as many ANFIS models side by side to produce multiple outputs, but in this case no modifiable parameters are shared by the juxtaposed ANFIS models as it can be seen in Fig. 4. Each ANFIS model has an independent set of fuzzy rules, which makes difficult to realize possible certain correlations between outputs.

Another way of generating multiple outputs and at the same time of using correlations between apparently independent parallel ANFIS models is by using a cascaded architecture that uses the initial outputs from the first step as inputs for the second step along with the same initial training data. This strategy provides the possibility of exploiting hidden correlations in the training data and refining the operation of the complete 3D position prediction system. Several ways of combining the independent outputs of the three individual prediction systems for the x, y and z coordinates may be found. A simple and effective solution (as proven by our experiments) is to use as additional input to the refining second stage of the cascade, the norm of the 3D vector formed by the origin and the initial initial estimated coordinates x, y and z. Provided the initial estimation error is reasonably small, this way of combining the initially estimated x, y and z coordinates offers a reliable input value and also restricts the searching region to a sphere whose radius (equal to the length of each 3D vector) is roughly known with good accuracy.

Again, FIS generator blocks - generates a Sugeno-type FIS structure using subtractive clustering and requires separate sets of input and output data as input arguments followed by ANFIS blocks that tune the initials FIS.

This approach of using a cascaded architecture

effectively exploits the correlation underlying the set of coordinates of 3D points and gives better results in terms of accuracy and noise resilience, this architecture is presented in Figure. 5.



Figure. 5. A cascaded architecture of neuro-fuzzy systems for exploiting data correlation

IV. MEDICAL VOLUME VISUALIZATION AND INTERACTION USING NEURO-FUZZY TOOLS

Our proposal emerged from the clinicians' need for more natural solutions to human-computer interaction (HCI) in medical image visualization and editing. Starting from the study of Luigi Gallo et al [18], we identify some requirements that must be taken into account when new solutions are to be proposed:

• non-obstructive hardware components should be preferred;

• intuitive interaction techniques are ideal for a satisfactory degree of friendliness;

• near real time interactivity should be provided to give a reasonable interaction perception;

• input devices should be ergonomic, wire-free and easy to handle.

Finally, for medical applications, not all 3D tasks (navigation, selection, manipulation) [14] are to be performed, as there is generally only one object in the scene and therefore the most frequent task is the editing of parts of this object.

In general, a 3D interaction process implies following steps: modeling and calibration of the camera setup, followed by the detection, tracking and accurate positioning of the interaction device.

In the initial step two web-cameras are involved and a checker-board, for producing a reliable training set of input-output data that captures all the system features and covers the full input variable space (that represent our virtual bounding box).

Fig. 6 represents the entire process of image acquisition from each of the two cameras and detection of interest points, which are the corners of the checkerboard, followed by the actual calibration with cascaded neuro-fuzzy systems architecture.



Figure 6. Implemented system for camera calibration.

At this moment cameras are calibrated with cascaded neuro-fuzzy systems architecture and we obtained FIS structures that allows for reconstruction of 3D points in the defined virtual bounding box.

This represents the essential step in developing of our interaction tool.

The next step in the detection of our interaction tool was the synchronization of the video streams of the two cameras using triggers that acquire one frame at a time from the video streams of the pair of camera Figure 7.



Figure. 7. Synchronization process of video streams.

Afterwards, threshold segmentation was applied on the video frames to detect the light spot of an infrared led. The led has a wide angle of radiation so that it is detected by the two cameras simultaneously, even if the led is not pointing directly to the camera. The led is positioned on top of the interaction tool, further used as a marker in the

interaction process. To avoid any complications introduced by the complex background, we attach infrared filters to the two webcams. As the spot light is narrow, its center of gravity is an accurate approximation of the real marker position. This helps the 3D tracking because the marker may be approximated with negligible error to a single 3D point. The spatial position of the led is found by the neuro-fuzzy inference system described in the previous section, once we provide the pairs of 2D coordinates detected in the synchronized frames of the video streams. The procedure is illustrated in Fig. 8.



Figure 8. Detection and 3D tracking of the "virtual probe" used for interaction.

The integration of the interaction module with a medical image/volume visualization system was done through a configuration based on the MATVTK framework [24]. MATVTK was developed as a flexible working environment that provides fast computing, registration and segmentation of medical data. This configuration is presented in Figure 9. Basically, MATVTK allows an easy integration of the flexible VTK's data-flow approach with the powerful computing tools provided by MATLAB.



Figure. 9. Integrated system architecture for visualization and 3D interaction.

Our system implements a hybrid interaction with medical data, by integrating the standard input devices

(keyboard and mouse), for defining the volume manipulation behavior along with the virtual probe used as 3D interaction tool. We selected this approach as the clinicians are also used to work with classical input devices and only need the extra-capabilities to interact in 3D scene for further processing.

In general, the interactive 3D editing methods, involve correction tools for refining the results obtained by some automatic processing methods. There are few approaches in the literature on this topic, such as the ones presented in [25] – [27]. We integrated so far three 3D interactive editing tools to refine some medical volume segmentation, but due to its open architecture, our system can be easily extended with new methods These three tools are 3D erasing, 3D filling and 3D morphological operations:

3D erasing: assuming that the interaction tool of the virtual probe has a spherical shape of radius R, the voxels of the 3D volume contained by the sphere may be removed. The center of the sphere, denoted by the 3D point $O_c(x_c, y_c, z_c)$, is the mapping of the real 3D position of the infrared led placed at the top of the 3D interaction tool – tracked in real time, inside the virtual display scene of the 3D medical volume, figure 10.



Figure 10. 3D erasing process.

Mapping between real coordinates and virtual scene coordinates was realized by registering the virtual bounding box resulted at the calibration step and the medical volume presented in the virtual scene. The 3D position of the probe in the virtual scene always takes into account the position of the virtual camera in the virtual scene, see Fig. 11, Therefore, we can correctly interact with the volume even when it is rotated.

If V denotes the medical volume of the analyzed organ (previously segmented) and x, y, z, some voxel coordinates, the erasing algorithm is simply described by:

$$V'(x, y, z) = \begin{cases} 0, (x - x_c)^2 + (y - y_c)^2 + (z - z_c)^2 < R^2 \\ V(x, y, z), otherwhise \end{cases}$$
(5)
$$V(x, y, z) = \begin{cases} 0, outside of the segmented volume \\ 1, inside of the segmented volume \end{cases}$$

where V' is the processed volume.

3D filling: Having a binary segmentation of the volume, with V(x, y, z) = 1 inside the organ, the volume can be filled by setting to 1 the voxels inside the sphere.

3D morphological operations: are typically used in post processing. Dilation, erosion, opening and closing are the basic operations, often applied in image processing. Initially developed for 2D binary images, they may easily be extended to 3D using a cube/sphere-like structuring element, in terms of the minimum and maximum operations defined inside the cube to replace the current voxel on which the cube/sphere is centered.

Some additional functionalities of our system that could be developed with the help cascaded neuro-fuzzy architecture used for 3D positioning are:

Virtual cutting plane– There are situations when cutting planes have to be positioned in arbitrary positions different then standard orthogonal planes (transversal, sagittal, coronal). This objective is difficult to reach in 3D space, when using traditional interaction tools, because plane is described either by three points or by a normal vector to the plane and a point. While the first method is somehow cumbersome, the second one is easy to implement with our solution. The approach is as follows:

-the arbitrary cutting plane is described by the point $p(x_0,y_0,z_0)$ – selected with the interaction tool

-Let \overline{n} (a₁, a₂, a₃), be the normal vector perpendicular on the cutting plane with the origin in the center of the volume (the same point where virtual camera of the scene is focusing).

-Using this normal vector and the point selected with proposed interaction tool you can determine the best cutting plane, Figure 11.

$$\vec{n} = \vec{O}_c - \vec{T}_c$$
(6)
Where:

 O_c - camera center of the virtual scene where the medical volume is represented

 T_c – vector that indicates camera's orientation (towars our volume).

n - normal vector of the plane



Figure 11. Cutting plane representation using only a point and (the normal vector.

3D annotation – for pointing zones of interest inside the medical volume and attaching information at the desired spatial location;

3D measurements of the distances between two or more points inside of medical volume

V. EVALUATION AND EXPERIMENTAL RESULTS

The first step of the experimental stage is the simulation of a stereo configuration, which is eventually, calibrated using the proposed method. The advantage presented by the simulation is that the reconstruction accuracy can be precisely calculated since the ground truth is known. The resulting black box model is then used for finding depth using the 3D objects generated by simulation and shown in Fig. 12.

Figure. 12. The Cylinder reconstruction using corresponding images from the two cameras and estimation of standard error and standard deviation.

The 3D error was calculated for all the points of the 3D object as the distance between the calculated point and the ground truth. In order to estimate the robustness to noise, different amounts of noise were applied to the 2D image points and the 3D position was estimated.

Note that the error is very small when there is low noise. This fact has been confirmed when the method has been applied to real data.

Table I shows that the error obtained in case of a crisp calibration has the range between 0.30 cm and 18.91 cm being clearly higher than the error obtained using the neuro-fuzzy calibration which is between 0.05 cm and 2.34 cm.

The proposed cascaded neuro-fuzzy architecture provides even better results than neuro-fuzzy one especially in noisy environments, Figures 13 a and b.

 TABLE I.
 Noise Effects On The 3d Reconstruction

SNR	2D Standard Deviation for Image 1	2D Standard Deviation for Image 2	3D error (cm)		
			Crisp	Neuro- Fuzzy	Cascaded Neuro- fuzzy
40	0.0065	0.0065	0.30	0.05	0.05
35	0.0114	0.0115	0.29	0.11	0.09
30	0.0207	0.0204	1.23	0.22	0.18
25	0.0371	0.0374	2.46	0.47	0.33
20	0.0652	0.0656	5.11	0.74	0.57

15	0.1152	0.1150	8.01	1.33	1.02
10	0.2090	0.2070	18.91	2.24	1.76
2.5 uption across (fum), 83 and 96 at a function of the functi	to a construction of the c		18 18 16 16 16 16 10 10 10 10 10 10 10 10 10 10	20 25 Signal to Noise Ra	
		<i>a</i>)		<i>b</i>)	

Figure 13 a. Fuzzy calibration: standard deviation and 3D reconstruction error; b. Cascded Neuro-Fuzzy calibration: standard deviation and 3D reconstruction error

As expected, in all three cases the error is proportional with the noise as can be seen from Table1. The proposed cascaded neuro-fuzzy method is clearly more robust to noise compared to the other two. This conclusion is also confirmed by the tests realized using real images.

The hardware setup used for our preliminary experiments involves a laptop equipped with an Intel i5 processor, 4 GB DDR3, NvidiaGforce GTS 360M video graphics, one additional monitor, anaglyph glasses, two Logitech Pro 9000 webcams and the 3D interaction tool (a stick with an infrared led of 120° radiation angle attached). The system's setup during the operation is shown in Fig. 14.

Figure 14. The setup of the proposed 3D medical visual interaction system.

For the calibration part, experiments realized with real data are presented in Fig.15. and Fig. 16. Fig. 15 presents the images captured by the two cameras together with the detected points of interest which are the corners of the checker-board. We detect 132 points on the checker-board and the checker-board is placed at eleven different positions, with a relative distance of 5 cm between two consecutive positions. Therefore, we obtain a total of

1452 pairs of points with the help of automatic corner extraction engine provided by Bouguet[28].

Each of the points is imaged by the two cameras, so the set of image points is formed by ${}^{C_1}P_i$ and ${}^{C_2}P_i$,

Figure. 15 Pair of images and detected points of interest.

 $W \mathbf{D}$

The 3D points ${}^{W}P_{i}$ are calculated using the 2D points and the obtained results are shown in Figure 16. In the training step for all the tests either with synthetic or real images for re-training of FIS we use 20000 epochs.

The error is calculated by comparing the distance between the retrained and the real points.

The mean error and the standard deviation obtained measuring objects located at within a distance of 30 cm and 80 cm from the cameras is presented in Table II.

TABLE II. COMPARISONS BETWEEN PREVIOUS IMPLEMENTATION AND THE NEW PROPOSED CASCADED NEURO-FUZZY IMPLEMENTATION

	Neuro-fuzzy	Cascaded
	conngulation	configuration
Standard deviation	0.27	0.27
Mean error	0.74	0.69

As expected, the new architecture provides better results in terms of accuracy than previous implementation.

Figure 16. 3D reconstruction accuracy using cascaded neuro-fuzzy architecture.

The comparison between crisp and cascaded-neurofuzzy calibration is presented in Table III.

TABLE III. COMPARISON BETWEEN CRISP AND NEURO-FUZZY CALIBRATIONS

Features	Crisp Calibration	Cascaded NeuroFuzzy Calibration
Robustness to noise	Low	High
Accuracy outside of a bounding box	High	Low
Model simplicity	Low	High
Adaptability to new camera configurations	Low	High

Tests that involve the proposed interaction tool are presented in the next lines.

Figure 17 presents functionality of adding **arbitrary cutting planes** using the proposed interaction tool. As we stated before, we need to position only one point in the displayed volume and the best fitting plane is constructed (the one that contains the positioned point and that is perpendicular with the normal vector that gives camera's orientation).

Figure 17. Virtual cutting planes positioned with the help of proposed interaction tool is presented in the above images.

The measurement tool that uses any two points positioned in the virtual scene volume is presented in Figure 18. This method helps the measurement of any distance from within a medical volume.

Figure 18. 3D distance measurement inside a medical volume.

Another processing tool that uses the proposed interaction tool, tool that is guided by cascaded neurofuzzy systems is the one that allows you to refine regions of interests with the help of morphological operations as can be seen in Figure 19.

Figure 19. Volume refinement using the proposed interaction tool and morphological operations.

3D editing (addition/deletion) is also implemented using the proposed tool. Figure 19 presents the process of refinement that involves a 3D segmented liver that contains some misclasified regions (false positive voxels). In case of the deletion function, the voxels that lie inside the sphere are deleted.

Figure 19. Volume refinement using the proposed interaction tool and deletion/addition tool.

II. CONCLUSIONS

This paper proposes an improved version of our previous work [29], by means of a novel calibration method for the 3D positioning and measurement that uses a cascaded neuro-fuzzy architecture.

The proposed architecture enables fast and reasonably accurate 3D measurements in contrast with calibration methods. It also provides better noise resilience in contrast with our previous work. The obtained results are encouraging and prove that the method is reliable, robust and easy to implement.

The presented system opens many future work directions, for example, the fuzzy calibration can be extended to a high number of cameras leading to better accuracy, less occlusions and wider view angles.

Even though the experimental results show a better accuracy for the neuro-fuzzy calibrated system, we noticed that the fuzzy configuration performs well only in the virtual bounding box in which the system was trained. On the other hand, if the system is used in applications like 3D interactions as in our case, this limitation is not necessary a disadvantage.

We also proposed and presented a virtual probe based 3D visual interaction system for medical volumes editing, applicable especially to the refinement of segmentation results of medical data. Unlike the existing methods, we use our proposed neuro-fuzzy logic-based approach for 3D reconstruction of the interaction device – a virtual probe equipped with an infrared led. The proposed approach is integrated in a flexible framework, providing some basic editing and visualization functionalities, which will be extended in our future work.

ACKNOWLEDGMENT

This paper was supported by the project "Doctoral studies in engineering sciences for developing the knowledge based society-SIDOC" contract no. POSDRU/88/1.5/S/60078, project co-funded from European Social Fund through Sectorial Operational Program Human Resources 2007-2013.

This paper was also supported by the project "Development and support of multidisciplinary postdoctoral programmes in major technical areas of national strategy of Research - Development -Innovation" 4D-POSTDOC, contract no. POSDRU/89/1.5/S/52603, project co-funded by the European Social Fund through Sectoral Operational Programme Human Resources Development 2007-2013.

REFERENCES

[1] J.-S. R. Jang, "ANFIS: Adaptive-network-based fuzzy inference system," *IEEE Transaction on System, Man, and Cybernetics*, vol. 23, No. 3, pp. 665-685, 1993.

[2] L. Behera, K. Indrani, "Intelligent systems and control. principles and applications," *Oxford University Press*, ISBN-10: 0-19-806315-6, 2009.

[3] F. Li, L. Zhenzhong, D. Baorui, "Camera calibration for linear structured light system based on BP neural network," *Third International Workshop on Advanced Computational Intelligence (IWACI)*, 2010 pp. 465 - 468, 2010.

[4] Ge Dong-Yuan, Yao Xi-fan, "Camera calibration and precision analysis based on BP neural network," 2nd International Congress on Image and Signal Processing, pp. 1–5, 2009.

[5] R. Tsai, "A versatile camera calibration technique for highaccuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE Journal of Robotics and Automation*, vol. 3, Issue 4, pp. 323, 1987.

[6] J. Heikkila and O. Silven, "A four-step camera calibration procedure with implicit image correction," *CVPR* '97, pp. 1106–1112, 1997.

[7] L. W.Yu, X. Kai, "A camera calibration method based on neural network optimized by genetic algorithm," *IEEE International Conference on Systems, Man and Cybernetics*, pp. 2748 – 2753, 2007.

[8] D.-M. Woo, D.-C. Park, "An efficient method for camera calibration using multiLayer perceptron type neural network," *International Conference on Future Computer and Communication*, pp. 358 – 362, 2009.

[9] L. Berthouze, F. Chavand, C. Barret, "A camera neural model," *IEEE International Conference on Systems, Man, and Cybernetics*, vol. 3, pp. 2124 - 2126, 1996.

[10] D. Yongtae, "Application of neural networks for stereocamera calibration," *International Joint Conference on Neural Networks*, vol. 4, pp. 2719 - 2722, 1999.

[11]C. Xiaobo, G. Haifeng, Y. Yinghua, Q. Shukai, "A new method for coplanar camera calibration based on neural network," *Computer Application and System Modeling (ICCASM), 2010 International Conference on*, vol. 2, pp. V2-617 - V2-621, Oct. 2010.

[12] C.-L. Lai, S.-T. Tsai, "Application of using fuzzy-neural network on ptz camera," *Machine learning and cybernetics*

(ICMLC), 2010 international conference on, vol. 5, pp. 2577 – 2582, july 2010.

[13] R. Orghidan, M. Danciu, A. Vlaicu, G. Oltean, M. Gordan, C. Florea, "Fuzzy versus crisp stereo calibration: a comparative study," *7th Int'l Symposium on Image and Signal Processing and Analysis (ISPA 2011)*, Dubrovnik, Croatia, pp. 627-632, September 4-6, 2011.

[14] D. A. Bowman, E. Kruijff, J. J. Jr. LaViola, and I. Poupyrev, "An Introduction to 3D User Interface Design," *Computer Science*, Virginia Tech.*pp.* 96-108, 2001.

[15]C. Gratzel, T. Fong, S. Grange, C. Baur, "A Non-contact mouse for surgeon–computer interaction,".*Technology and Health Care Journal, IOS Press*, vol. 12, no. 3, pp. 245-257, 2004.

[16]C. Feied, M. Gillam, J. Wachs, J. Handler, H. Stern, M. SMITH, "A real-time gesture interface for hands-free control of electronic medical records," *AMIA Annual Symposium Proceedings*, Washington, DC, USA, February 2006.

[17]L. Paolis, M. Pulimeno, G. Aloisio, "Advanced visualization and interaction systems for surgical pre-operative planning," *Journal of Computing and Information Technology*, vol. 18, no 4, pp. 385-392, 2010.

[18]L. Gallo and M. Ciampi, "Wii remote-enhanced handcomputer interaction for 3D medical image analysis," *Current Trends in Information Technology (CTIT), 2009 International Conference*, pp. 85-90, 2009.

[19] J. R. Cooperstock and G. Wang, "Stereoscopic display technologies, interaction paradigms, and rendering approaches for neurosurgical visualization," In *Stereoscopic Displays and Appliations*, San Jose, CA, USA, january 2009.

[20] B. Reitinger, A. Bornik, R. Beichel, D.Schmalstieg, "Liver surgery planning using virtual reality," *IEEE Computer Graphics and Applications*, vol. 26, no. 6, pp. 36-47, nov./dec. 2006.

[21] J.-S. R. Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System", *IEEE Transaction on System, Man, and Cybernetics*, vol. 23, No. 3, pp. 665-685, 1993.

[22]C. Tran, A. Abraham, L. Jain, "Decision support systems using intelligent paradigms", *Computing Research Repository*, cs.AI/0405052, http://arxiv.org/abs/cs.AI/0405052, 2004.

[23]L. Behera, K. Indrani, "Intelligent systems and control. principles and applications," *Oxford University Press*, ISBN-10: 0-19-806315-6, 2009.

[24]E. Birngruber, R. Donner, G. Langs, "MatVTK - 3D visualization for Matlab," *MICCAI 2009 Workshop paper*, August 25, 2009.

[25] D. Lemire "Streaming maximum-minimum filter using no more than three comparisons per element," *Nordic Journal of Computing*, vol. 13, no 4, pp. 328-339, 2006.

[26]J. R. Cooperstock and G. Wang, "Stereoscopic display technologies, interaction paradigms, and rendering approaches for neurosurgical visualization," *In Stereoscopic Displays and Applications, San Jose, CA, USA*, january 2009.

[27] B. Reitinger, A. Bornik, R. Beichel, D.Schmalstieg, "Liver surgery planning using virtual reality," *IEEE Computer Graphics and Applications*, vol. 26, no. 6, pp. 36-47, Nov./Dec. 2006.

[28] Bouguet J. Y. Camera Calibration toolbox for Matlab. www.vision.caltech.edu/bouguetj.

[29] M. Danciu, M. Gordan, R. Orghidan, E. Sorantin, C. Florea, A. Vlaicu, "A Fuzzy Logic-Guided Virtual Probe for 3D Visual Interaction Applications", IEEE International Conference on e-Health and Bioengineering (EHB 2011), 24-26 November, 2011, Iasi, Romania, pp. 431-434.