

## A prototype for the creation and interactive visualization of 3D human face models

Radu Orghidan<sup>1,a</sup>, Mihaela Gordan<sup>1,b</sup>, Marius Danciu<sup>1</sup>, and Aurel Vlaicu<sup>1</sup>

Technical University of Cluj-Napoca, Romania

{<sup>a</sup>Radu.Orghidan, <sup>b</sup>Mihaela.Gordan}@com.utcluj.ro

**Keywords:** 3D Reconstruction, Immersive interaction, Self-calibration, Structured Light, Vanishing points.

**Abstract.** This paper presents a complete solution for the creation and manipulation of the 3D representation of a human face model using off the shelf components. First, the 3D model is obtained by means of a structured light system that is calibrated using only the vanishing points extracted from a simple planar surface. Then, an immersive interaction technique is used to manipulate the 3D model. The interaction tool is located in 3D space using a fuzzy technique with the advantages of a low memory usage, real-time operation and low positioning errors as compared to classical solutions. Experimental results including accuracy evaluation of the reconstruction and of the interaction tool are presented. The resulting system can find applications in expression based human-computer interaction, virtual assisted cosmetic surgery, as well as in the teleconferencing, being in line with the current trends of intelligent user interfaces.

### Introduction

The creation, visualization and interaction with 3D models are crucial steps in many applications. The digitized 3D world must be readily accessible to human users, therefore, ergonomic human-computer interfaces are necessary. Traditional interaction devices such as the keyboard or the mouse are not the most appropriate tools considering that, in daily practice, 3D objects are approached in a much more flexible way. We describe here a complete solution for the creation and manipulation of the 3D representation of a human face using off the shelf components. First, the 3D model is created using active shape acquisition techniques. Then, the face model is manipulated using an immersive technique based on a hand-held device that produces a realistic human-computer interaction.

**Active shape acquisition techniques for 3D model creation.** Computer vision applications usually require several points of view for depth calculation in a real scene. Commonly, shape acquisition techniques can be classified in two groups: passive and active techniques.

The passive techniques obtain depth information without any physical interaction with the measured scene. Among these techniques, one of the most common is the stereoscopy. A stereoscopic system is composed by at least two cameras and provides the necessary view-points for triangulation. For using such a system, two conditions must be accomplished: (a) an accurate correspondences between the images must be determined and (b) the calibration of the cameras composing the system must be solved.

The techniques using structured light [1], briefly noted as SL, are designed for scene depth acquisition. A normal SL system is composed by a stereoscopic pair in which a projector acts like a reverse camera [2,3]. The the projector introduces landmarks on the 3D objects alleviating the correspondence problem. Thus, the SL method is considered as an active acquisition technique. In order to use an active technique, the calibration of the projector must be performed because its model is required for the triangulation.

Despite the model equivalence, the classical methods for camera calibration cannot be applied for the projector because it doesn't provide an image of the scene. Moreover, in the case when the projector is not orthogonal to the screen, a particular distortion, called *the keystone effect*, appears. Traditionally, the calibration of a SL system is preformed by first calibrating the camera. Also, objects with known

structure and location are used. The camera is used for the estimation of the 3D position of the features projected by the light projector. A major problem of this methodology is that the approximations of the camera model is propagated to the projector's model. A better turnaround is to calibrate the camera and the projector independently. Martynov [4] used an uncalibrated camera for performing projector calibration. The camera was used to find out the equivalent image of the projector from the camera-projector transformation. Using this transformation (which is actually a homography) the projector can be modeled as a reversed camera. Then, it can be calibrated like a classic camera. However, the method described by Martynov et al. requires a series of iterations for determining the best fitted points for the projected image. Besides, it is assumed that the points are uniquely found on the camera image, i.e. in the image that was taken with the camera, which is not a trivial task because usually the projected pattern is distorted.

A pinhole model can be calibrated without a calibration object if the special type of points that appear in perspective projections, the *vanishing points* (VPs), are used. In the Manhattan world assumption, the VPs can be calculated [5] and used. The VPs that are linked to the orthogonal directions in a scene [6,7] can be used to determine the focal length and the rotation of the camera with respect to the world coordinate system. We can conclude that the VPs embed crucial data for modelling a camera using the common pinhole model. There are methods [8,9,10,5] that can be used for accurately detecting the VPs.

**Human-computer interfaces for interactive visualization.** In real applications, the end users want to have a natural, immersive, way for analysing the 3D objects. Bowman et al. [11], defined the 3D interaction as a specific human-computer interaction in which user tasks are done in a 3D context. For this purpose, the interfaces must be designed to be simple and non-obstructive. Gratzel et al. [12] built a non-intrusive mouse dedicated to surgeons using only gestures. Feied et al. [13] proposed a hands-free configuration for analysing medical images avoiding the contact with the potentially aseptic keyboard or mouse. Gallo and Ciampi [14] proposed a 3D interface for exploring medical data by means of a data glove tracked by a Wii remote sensor. Later on, De Paolis et al. [15] described a system designed for helping the surgeons to effectively model and interact with human organs for surgery planning during the real procedure. Cooperstock et al. [16] proposed a 3D displaying system and implements interaction in the same natural way as the manipulation of physical objects. A method for surgical procedures planning, named LiverPlanner, was proposed by Reitingner et al. [17] and tracks simultaneously the body position and the interaction device using an optical tracking system.

This paper targets the class of medical applications dealing with the manipulation of 3D face models and offers the premises for the development of an extensive set of medical volume editing tools. Our goal was to design and implement a flexible, fast and accurate alternative to the existing solutions, by integrating a fuzzy logic system in the 3D localization of the virtual interaction probe. For this purpose, we developed an interaction pen-shaped device that is tracked by two web cameras. The proposed setup is used for the manipulation of the 3D data obtained by the non-intrusive 3D reconstruction techniques using a SL system.

The mathematical background for the SL calibration is presented in the next section. The proposed fuzzy model for 3D interaction is described in detail in the following part of the paper. Some experimental results are presented in order to validate the calibration algorithm and to show some of the interaction capabilities of the immersive manipulation tool. The paper ends with the conclusions that are detailed in the last section.

## Calibration of a structured light configuration using vanishing points

**Pinhole calibration using vanishing points.** As previously stated, the camera and the projector can be both modeled using the pinhole model. the camera matrix  $K$  can be expressed as shown in Eq. 1, considering the following intrinsic parameters: the focal distance  $f$ , the image center  $(u_0, v_0)$ , the skew coefficient  $\gamma$  and the aspect ratios,  $\alpha_u$  and  $\alpha_v$ .

$$\mathbf{K} = \begin{bmatrix} \alpha_u f & \gamma & u_0 \\ 0 & \alpha_v f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

The camera matrix can be written in a simpler form assuming that  $\gamma = 0$  and  $\alpha_u = \alpha_v = 1$ . Therefore, the remaining intrinsic parameters are the image center and the focal distance. The set of extrinsic parameters is formed by three rotations and three translations, a pair for each orthogonal axis. The pinhole model is said to be calibrated if the nine parameters are known.

Guillou et al. [7] described a method for the calibration of a pinhole device from two VPs calculated using the orthogonal edges of the scene structure. We use a similar approach here.

Consider that the world and the camera coordinate systems are at  $O_w$  and  $O_c$ , respectively. The camera projection center is set at  $O_c$  and the center of the image, denoted by  $O_i$ , is the orthographic projection of  $O_c$  on the image plane. The two axes  $x_w$  and  $y_w$ , respectively, determine the VPs:  $V_1$  and  $V_2$ , as illustrated by Fig. 1. The projections of the VPs in the image plane are  $V_1 = (v_{1i}, v_{1j})$  and  $V_2 = (v_{2i}, v_{2j})$ . The projection of  $O_i$  on the line  $(V_1 V_2)$  is  $V_i$ .

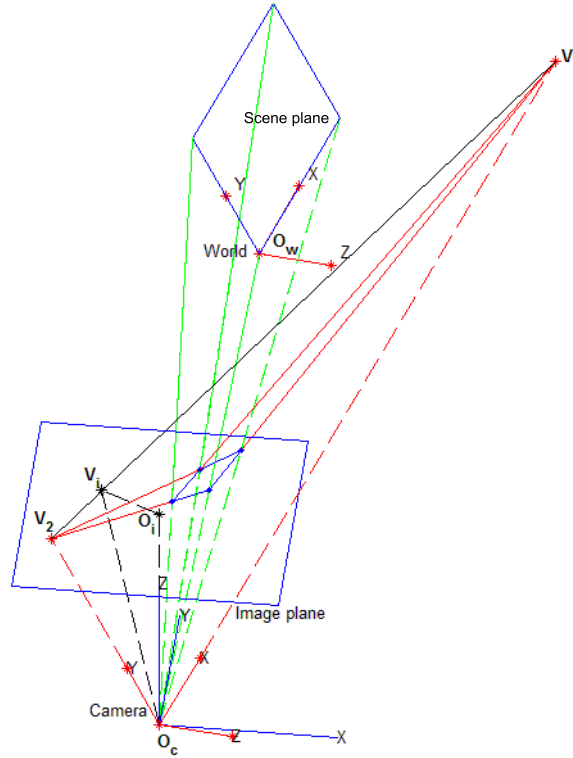


Fig. 1: The vanishing points can be used to calculate the focal distance, the location and the orientation of a camera with respect to a 3D plane.

Without loss of generality, we can set the principal point at the image center and the aspect ratio equal to one, i.e.  $\alpha_u = \alpha_v = f$ . The two VPs can be used to calculate the intrinsic and the extrinsic parameters of the pinhole model. The position of the principal point  $(u_0, v_0)$  can be obtained because it is located at the intersection of the optical axis with the image plane. Hartley [18] shown that its position is crucial for the calibration.

Considering that  $O_c$  and  $O_i$  are on the optical axis, as shown in Fig. 1, the focal distance  $f$  can be calculated as follows:

$$f = \|O_c O_i\| = \sqrt{\|O_c V_i\|^2 - \|O_i V_i\|^2}. \quad (2)$$

Consequently,  $O_i V_i$  can be estimated as:

$$\|O_c V_i\| = \sqrt{\|V_1 V_i\| \cdot \|V_i V_2\|}. \quad (3)$$

The matrix  ${}^C R_W$  encodes the world to camera rotation. A new coordinate system, centered at  $O_c$  and having the same orientation as the world system, can be build considering the vectors  $\mathbf{X}'_c = \overrightarrow{O_c V_1}$ ,  $\mathbf{Y}'_c = \overrightarrow{O_c V_2}$  and  $\mathbf{Z}'_c = \mathbf{X}'_c \times \mathbf{Y}'_c$ .

The vectors  $\mathbf{X}'_c$ ,  $\mathbf{Y}'_c$ , and  $\mathbf{Z}'_c$  are:

$$\begin{aligned} \mathbf{X}'_c &= \frac{\overrightarrow{O_c V_1}}{\|\overrightarrow{O_c V_1}\|} = \left[ \frac{v_{1i}}{\|\overrightarrow{O_c V_1}\|} \quad \frac{v_{1j}}{\|\overrightarrow{O_c V_1}\|} \quad \frac{f}{\|\overrightarrow{O_c V_1}\|} \right]^T \\ \mathbf{Y}'_c &= \frac{\overrightarrow{O_c V_2}}{\|\overrightarrow{O_c V_2}\|} = \left[ \frac{v_{2i}}{\|\overrightarrow{O_c V_2}\|} \quad \frac{v_{2j}}{\|\overrightarrow{O_c V_2}\|} \quad \frac{f}{\|\overrightarrow{O_c V_2}\|} \right]^T \\ \mathbf{Z}'_c &= \mathbf{X}'_c \times \mathbf{Y}'_c \end{aligned} \quad (4)$$

Thus, the rotation matrix  ${}^C R_W$  obtained is:

$${}^C R_W = \begin{bmatrix} \frac{v_{1i}}{\sqrt{v_{1i}^2 + v_{1j}^2 + f^2}} & \frac{v_{2i}}{\sqrt{v_{2i}^2 + v_{2j}^2 + f^2}} & z'_{cx} \\ \frac{v_{1j}}{\sqrt{v_{1i}^2 + v_{1j}^2 + f^2}} & \frac{v_{2j}}{\sqrt{v_{2i}^2 + v_{2j}^2 + f^2}} & z'_{cy} \\ \frac{f}{\sqrt{v_{1i}^2 + v_{1j}^2 + f^2}} & \frac{f}{\sqrt{v_{2i}^2 + v_{2j}^2 + f^2}} & z'_{cz} \end{bmatrix}. \quad (5)$$

The translation between the two reference systems is the last part of the extrinsic set of parameters to be calculated. This can be done using a segment with a known length using the rotation vectors.

The calibration of a SL system is presented in Fig. 2. The SL configuration is shown in Fig. 2(a). First, the distortion (visible in Fig. 2(b)) of the projector is removed. The distortion removal is obtained by transforming the initial image of the projector, shown in Fig. 2(c), such that the VPs of the projected pattern coincide with the VPs of the screen, see Fig. 2(d). Then, the VPs of the new projector image, shown in Fig. 2(e), are used for the calibration of the projector.

**The vanishing points of a projector image** The camera is used to observe the projected pattern. Therefore, the pattern is altered by two distortion superposed: the keystone effect and the camera perspective transformation. Despite this double distortion, a transformation from the screen to the projector can be determined [19]. An orthogonal projection to the screen has the same VPs as the screen and is bounded by the points  $P'_i$ . Let us denote by  $V_x$  and  $V_y$  the VPs of the camera image along the  $X$  and  $Y$  world axes, respectively. The keystone effect is removed by enforcing the co-linearity between the points  $(P'_1, P'_2, V_x)$  and  $(P'_4, P'_3, V_x)$  along the  $X$  direction and between the points  $(P'_1, P'_4, V_y)$  and  $(P'_2, P'_3, V_y)$  along the  $Y$  direction.

Considering  $P'_1 = P'_1$ , the other three points can be obtained from the intersections of the following lines:  $P'_2 = (P'_1, V_x) \cap (P_2, V_y)$ ,  $P'_4 = (P_4, V_x) \cap (P_1, V_y)$  and  $P'_3 = (P'_4, V_x) \cap (P'_2, V_y)$ .

The transformation that can alleviate the two distortions can be determined as a planar homography between the two sets of points,  $P_i = [x_i, y_i]^T$  and  $P'_i = [x'_i, y'_i]^T$ :

$$\begin{bmatrix} x'_i \\ y'_i \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix}. \quad (6)$$

The projector image can be compensated by applying the homography from (6). Finding the correspondence between the camera image and the projector image is the last step in order to calculate vanishing points  $V_x$  and  $V_y$  from the edges of the screen. Let us consider that the camera and the projector reference systems, shown in Fig. 2(a), are placed at  $\{C\}$  and  $\{P\}$ , respectively, and the two devices point towards a planar surface aligned with the  $XY$  plane of the world coordinate system

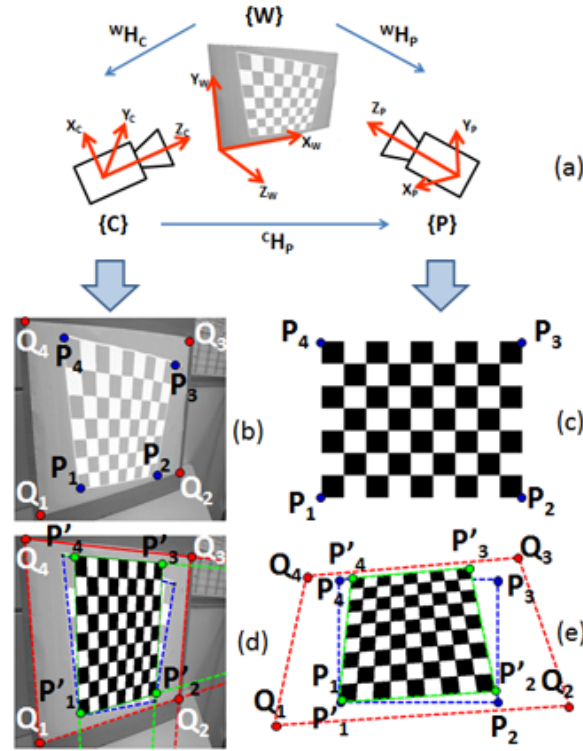


Fig. 2: Diagram of a SL system calibration.

$\{W\}$ . Then, the homography  ${}^W H_C$  between the screen plane and the camera image can be calculated using at least four points. Similarly, the homography  ${}^C H_P$  between the images of the projector and the camera is determined. Thus, the homography between the screen and the projector can be calculated:

$${}^W H_P = {}^W H_C \cdot {}^C H_P. \quad (7)$$

Using the VPs of the screen, the points  $P'_i$  are determined on the camera image, see Fig. 2(d). From the homography  ${}^P H_C = \text{inv}({}^C H_P)$ , the points  $P'_i$  can be also be positioned on the projector image. Fig. 2(e) shows the image of the projector containing the estimated position of the points  $Q_i$  and  $P_i$ . The transformation of the pattern can be modeled as the homography (6) and will rectify any image from the given projector-screen setup. The transformed image of the projector, shown in Fig. 2(e), contains two VPs that are linked to the projector-camera-screen configuration.

**Structured light used for 3D reconstruction** We use here SL to introduce landmarks in the scene for solving the correspondence problem in the case of non-textured objects. A coded pattern, formed by means of a spatial neighborhood coding technique using a De Bruijn sequence, is projected. A particular De Bruijn pattern [20] is used. This pattern is using 4 hue levels grouped in 64 vertical slits separated by black stripes. The correspondences are, therefore, identified in a single shot. Such a sequence of order  $m$  in conjunction with  $n$  different symbols forms a single dimensional structure of length  $n \cdot m$  containing unique instances of substrings of length  $m$ . The background is eliminated by setting a threshold for the lower intensity values.

The image is decoded by performing an horizontal scan. The intensity peak for each stripe is calculated using a second derivative filter applied to the signal on a given scan line. A filter is used for enhancing the differences in intensity and the detection is obtained with sub-pixel accuracy. The segmentation takes into account the regions where the second derivative is less than one, which are set to 0, while the others are set to 1. The color of a stripe can be precisely classified among the 4 Hue levels of the projected pattern. The decoding process is based on identifying the hue value of each stripe in a given scan-line. The projected and the perceived stripes are matched using the colors

of neighbor stripes, with a window of size equal to 3. Many correspondences can be detected by iteratively applying the process for all the rows of the image. In the ideal situation, all the stripes of the scan-line are correctly identified and the complete decoding is performed. In real situations, however, due to occlusions or incorrectly labeled stripes, classification errors happen which eventually decrease the global accuracy of the reconstruction. The correspondence between the projected stripes and the detected ones is obtained through a dynamic programming algorithm [21], using the RGB components of the stripes.

### An ANFIS trained stereo configuration for 3D tracking

A 3D interaction using two cameras implies the calibration of the stereoscopic configuration followed by the detection and tracking of the target device. In our case, a black box type of model formed by a set of rules, i.e. an Adaptive Neuro Fuzzy Inference System (ANFIS), is used for modelling the stereo configuration. The resulting model is easy to optimize and runs very fast, with minimal computing power. A fuzzy system benefits of the fact that it deals with uncertainties, preserving, thus, the ambiguities and postponing the crisp calculations until it is unavoidable. Therefore, the error doesn't propagate and the obtained accuracy improves [22].

The camera calibration is modeled here by a neural network based on Takagi–Sugeno fuzzy inference system. The method estimates a set of rules defining the functionality of the cameras involved in the stereoscopic configuration. In our case, a fuzzy rule can embed the four inputs, as shown in Fig. 3.

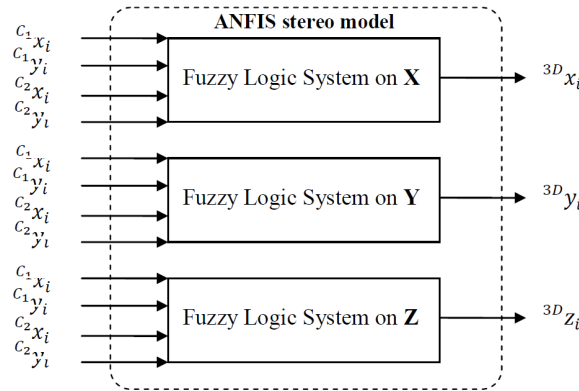


Fig. 3: Diagram of the black box model using ANFIS.

Three fuzzy logic systems are defined, one for each dimension of the Euclidian space. Each fuzzy logic system receives at the input the coordinates of the two 2D points:  $(C^1x_i, C^1y_i)$  and  $(C^2x_i, C^2y_i)$  and provides at the output the coordinates of the 3D points in space:  ${}^{3D}x_i$  - on X,  ${}^{3D}y_i$  - on Y, and,  ${}^{3D}z_i$  - on Z.

### Experimental results

**Range estimation using the calibrated SL system** The proposed calibration method can be evaluated by performing 3D reconstructions of objects placed in front of the SL system. As explained, the pattern is decoded using a De Bruijn sequence implemented by a multi-slit pattern of 4 colors arranged in 64 stripes.

Another important feature of a tracking system, directly related with the calibration, is the accuracy - which can be calculated using a planar surface. In our experiments, the SL system is located at around 1.5m from a plane that is then moved and placed at several known distances. The system calibration is performed for each position. The Delaunay triangles with the area higher than a given threshold are used to automatically remove the outliers from the 3D reconstruction of the surface. A plane fitting is

performed using the cloud of points. The reconstruction accuracy is measured as the average distance error from each point to the plane. The average error from the points to the plane is estimated as a percentage of the distance from the SL system to the plane in the scene and results to be near to 3%.

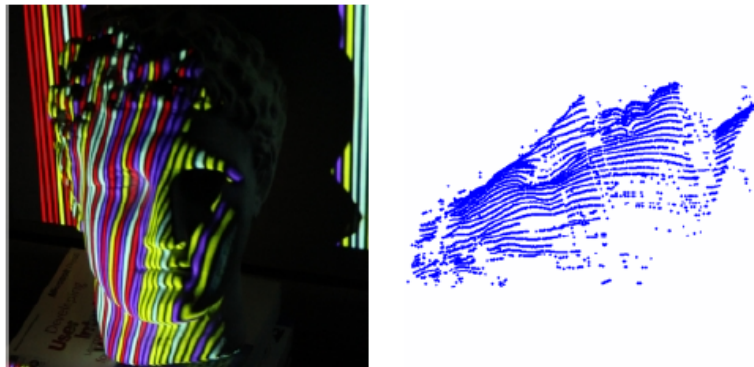


Fig. 4: Reconstruction of a statue head by means of the proposed SL configuration.

### ANFIS training for 3D interaction

The experimental setup is using two Logitech Pro 9000 web cams, placed along a baseline, that are calibrated using the previously described ANFIS method. The calibration is performed using the images of a set of 3D coplanar points from planes located at different distances from the stereo configuration. The calibrated system is used to find depth.

The initial step consists in producing a reliable training set of input-output data that covers the full input variable space. A subtractive clustering procedure was initially applied to the data set, giving a first order Takagi-Sugeno fuzzy system. If the simple fuzzy system obtained is not accurate enough, its accuracy can be enhanced by an adaptive neuro-fuzzy training method (ANFIS). The accuracy of the fuzzy model is tested iteratively and if acceptable, the training procedure is stopped. Otherwise the procedure is resumed either by re-training the fuzzy system with other training parameters (e.g. with more training epochs), or by determining a new data set.

A classical calibration procedure was compared [22] with the proposed fuzzy calibration for a stereo configuration of cameras. The ground truth was provided by a set of points with known positions. For distances of about 30 cm, the error obtained from the crisp calibration is around 0.3 cm, whereas the fuzzy calibration error is less than 0.1 cm, being obviously much lower.

The interaction tool is simultaneously detected done in the video streams of the two cameras, as shown in Fig. 5. A threshold is used on the video frames to detect the light spot of an infrared led positioned on top of the interaction tool, used in the interaction process as a *3D mouse*. The cameras are provided with infrared filters in order to avoid any interferences introduced by a potentially complex background. The center of gravity of the infrared spot proved to be an accurate approximation of its position. Therefore, the marker's position is considered to be, with negligible error, a single 3D point, found by the stereoscopic configuration system using the pairs of 2D coordinates simultaneously detected in the two video streams.

All the components working together are presented in Fig. 6. First, the SL system, formed by the projector and camera 1, is calibrated using the calibration screen. The SL system is used for obtaining the 3D model. Then, the two cameras are calibrated and are used to track the interaction device in order to manipulate the 3D model.

### Conclusions

This paper presents a complete system for the creation and the manipulation of 3D objects.

The properties of the vanishing points are used to calibrate the SL system. The camera and the projector are calibrated independently, thus, the error propagation present in other calibration methods of SL configurations is eliminated. Moreover, the calibration pattern is reduced to a planar surface



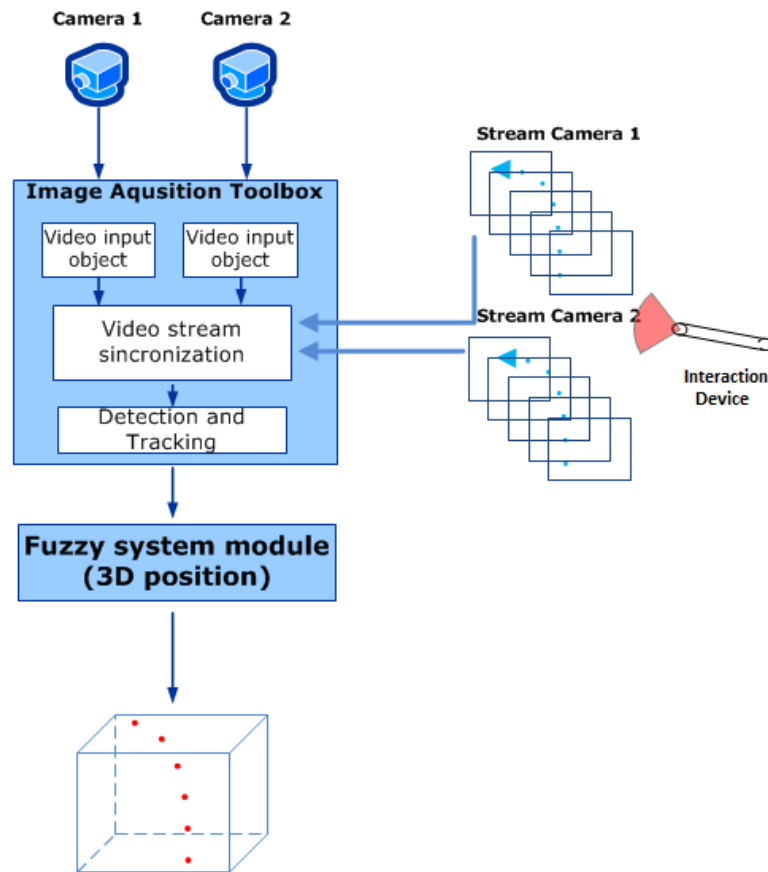


Fig. 5: Marker tracking for 3D interaction.

leading to a simplified calibration setup since there is no need of specially tailored calibration objects. Thus, the process is easy to handle even by inexperienced users. Despite the calibration simplicity, the method provides a good resolution. The mean error obtained experimentally was around 3% of the distance from the SL to the target screen with a low standard deviation. The measured accuracy proves that the method is suitable for computer vision applications that need good depth estimation and a fast self-calibration. The method works for single shot reconstructions which means that it can be used for dynamic scenes. Moreover, the calibrated model can be applied in a flexible manner for a variety of applications as it is not tied to a specific SL algorithm.

The 3D visual interaction system uses a novel fuzzy logic-based approach. A virtual probe equipped with an infrared LED is sufficient as a reliable interaction device. The keyboard and mouse are used for the selection of some processing options and parameters, but the immersive interaction with the volume is done only in visual manner. The most significant advantages of the proposed interaction device are the increased 3D positioning accuracy offered by the fuzzy logic stereo calibration, its flexibility, low hardware complexity and noise resilience.

This work sets the bases for a whole range of 3D computer vision applications. Regarding the 3D reconstruction complexity, it can be enhanced by registering several 3D one-shot reconstructions in order to obtain richer 3D objects. On the other hand, the interaction tool can be linked to 3D volume editing application. Since the proposed system uses off-the-shelf components and is easy to setup even by inexperienced users it can provide the ground for a wide range of entertainment and professional applications.



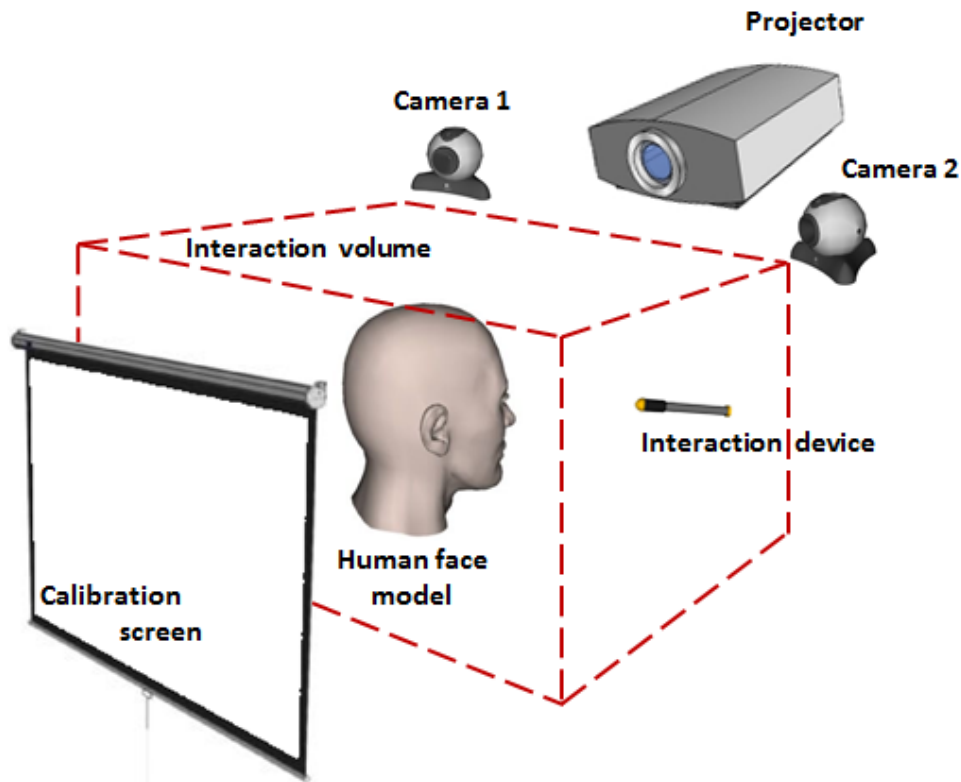


Fig. 6: Schematic representation of the configuration used for 3D object reconstruction and manipulation.

### Acknowledgements

This paper was 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. Salvi, S. Fernandez, T. Pribanic, and X. Llado. A state of the art in structured light patterns for surface profilometry. *Pattern Recognition*, 43(8):2666--2680, 2010.
- [2] X. Zhang and L. Zhu. Projector calibration from the camera image point of view. *Optical Engineering*, 48:117--208, 2009.
- [3] M. Kimura, M. Mochimaru, and T. Kanade. Projector calibration using arbitrary planes and calibrated camera. *CVPR*, 2007.
- [4] Ivan Martynov, Joni-Kristian Kämäräinen, and Lasse Lensu. Projector calibration by "inverse camera calibration". In *Scandinavian Conf. on Image Analysis (SCIA)*, 2011.
- [5] J.P. Tardif. Non-iterative approach for fast and accurate vanishing point detection. *IEEE 12th International Conference on Computer Vision*, pages 1250--1257, 2009.
- [6] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision -- 2nd Edition*. Cambridge University Press, 2004.

- [7] E. Guillou, D. Meneveaux, E. Maisel, and K. Bouatouch. Using vanishing points for camera calibration and coarse 3D reconstruction from a single image. *The Visual Computer*, 16:396--410, 2000.
- [8] Carsten Rother. A new approach for vanishing point detection in architectural environments. *Journal Image and Vision Computing (IVC; Special Issue on BMVC 2000)*, vol. 20, no. 9-10:647-656, January 2002
- [9] J. Kogecka and W. Zhang. Efficient computation of vanishing points. *IEEE International Conference on Robotics and Automation*, 1:223--228, 2002
- [10] Michael Bosse, Richard Rikoski, John Leonard, and Seth Teller. Vanishing points and three-dimensional lines from omni-directional video. *The Visual Computer*, 19:417--430, 2003.
- [11] D.A. Bowman, E. Kruijff, J. J. Jr. LaViola, and I. Poupyrev. An introduction to 3d user interface design. *Computer Science, Virginia Tech.*, pages 96--108, 2001
- [12] C. Gratzel, T. Fong, S. Grange, and C. Baur. A non-contact mouse for surgeon computer interaction. *Technology and Health Care Journal, IOS Press.*, 12, 2004
- [13] C. Feied, M. Gillam, J. Wachs, J. Handler, H. Stern, and M. Smith. A real-time gesture interface for hands-free control of electronic medical records. *AMIA Annual Symposium Proceedings, Washington, DC, USA*, 2006.
- [14] L. Gallo and M. Ciampi. Wii remote-enhanced hand-computer interaction for 3d medical image analysis. *Current Trends in Information Technology (CTIT), International Conference*, pages 85-90, 2009.
- [15] L. Paolis, M. Pulimeno, and 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, no 4:385--392, 2010.
- [16] 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, 2009.
- [17] B. Reitinger, A. Bornik, R. Beichel, and D. Schmalstieg. Liver surgery planning using virtual reality. *IEEE Computer Graphics and Applications*, 26, no. 6:36--47, 2006.
- [18] R. I. Hartley and R. Kaucic. Sensitivity of calibration to principal point position. In *Proc. Seventh European Conference on Computer Vision*, pages 433--446. Springer, 2000.
- [19] Rahul Sukthankar, Robert G. Stockton, and Matthew D. Mullin. Smarter presentations: Exploiting homography in camera-projector systems. *Proceedings of International Conference on Computer Vision*, 1:247--253, 2001.
- [20] J. Pages and J. Salvi. A new optimised de bruijn coding strategy for structured light patterns. *17th International Conference on Pattern Recognition*, 4:284--287, 2004
- [21] L. Zhang, B. Curless, and S. M. Seitz. Rapid shape acquisition using color structured light and multi-pass dynamic programming. *Proceedings of the 1st International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT)*, pages 24--36, 2001.
- [22] R. Orghidan, M. Danciu, A. Vlaicu, G. Oltean, M. Gordan, and C. Florea. Fuzzy versus crisp stereo calibration: a comparative study. *Image and Signal Processing and Analysis (ISPA)*, 2011 7th International Symposium on, pages 627--632, 2011