

Analysis and Calibration of a Mirror Setup producing Mirror-reflected, Multi-view Videos

Gianluigi Mucciolo^{1,2}, Barbara Koneczny¹, Alessia Saggese², Nicole M. Artner¹,
Walter G. Kropatsch¹, Mario Vento²

¹ Pattern Recognition and Image Processing Group (PRIP), TU Wien, Austria

² Dept. of Information Eng., Electrical Eng. and Applied Mathematics (DIEM)
Faculty of Engineering, University of Salerno, Italy

gmucciolo@prip.tuwien.ac.at, koneczny@prip.tuwien.ac.at

<http://www.prip.tuwien.ac.at>

Abstract. *In this paper mirror-reflected, multi-view video sequences produced by a mirror setup are analysed. The mirror setup is composed of two mirror planes and a camera, which records the scene. The relation between the angle between the two mirror planes and the number of mirror-reflected views is analysed mathematically and the results are presented in this paper. Furthermore a calibration approach using a cylindrical checkerboard pattern is introduced. Using the cylindrical checkerboard pattern the relation between the central view and the mirror-reflected views can be provided and the orientation and position of the mirror planes can be specified. The knowledge about the orientation and position of the mirror planes can be used to estimate the 3D position of marker points in the central and the reflected views.*

1. Introduction

There are a lot of 3D surface-imaging technologies some of those are described in [26], where the mentioned systems are cost intensive, also are usually composed of a projector and two or more cameras. Usually motion analysis applications require 3D data with high spatial and temporal resolution. However, high speed cameras are cost intensive, a lot of space is needed, and there are synchronization requirements [19]. The two main advantages of mirror-based multiple-view vision systems can be seen in the lack of synchronization requirements and the costs of only one camera. However, these advantages have to be weighted in our case with the difficulties arising from the depth of field, since it is not

possible to focus on the real object and its reflections at the same time. The proposed ad-hoc calibration system needs to cope with the limited depth of field. The used 3D video-analysis system is similar to the system described by Frey *et al.* [6]. In order to calibrate the mirror setup (estimate the exact position and orientation of the mirrors) a manual calibration procedure has to be performed. In the current system the angle between the two mirrors has to be set to 42.9° in order to estimate the 3D position of marker points correctly. Since the angle between the two mirror planes is restricted to 42.9°, the number of reflected views is limited to two or four (depending on the position and height of the object). The purpose of this paper is to automate the calibration process, and to study the relations between the parts of the system, in order to maximize the amount of information that can be captured in a single video session.

The paper is organized as follows. Work and publications related to our approach are shown in Section 2. The hardware setup is described in detail in Section 3. The mathematical analysis of the mirror setup is given in Section 4. The used calibration device and the 3D position estimation are presented in Section 5, 6 and 7. The experimental setup is given in Section 8. The results are presented and discussed in Section 9. In 10 a summary and further work are presented.

2. Related work

According to [12, 3, 4, 8] calibration is the comparison of measurement values delivered by a device under test with those of a calibration standard of known accuracy. Such a standard could be an-

other measurement device of known accuracy, a device generating the quantity to be measured. Geometric camera calibration, estimates the parameters of a lens and image sensor of an image or video camera. These parameters can be used to correct lens distortion, measure the size of an object in world units, determine the location of the camera in the scene, or to reconstruct a 3-D scene. Well-known methods of camera calibration include algorithms introduced by Tsai [25] and Zhang [28]. With these methods, the camera parameters are estimates, by using 3-D world points and their corresponding 2-D image points. These correspondences can be obtained using multiple images of a calibration pattern, such as a checkerboard. There are also many other studies about camera calibration ([25, 28, 21, 14]) which focus on the estimation of camera parameters.

3D reconstruction is a field of research which engages a various amount of different approaches, concerning hardware setups. Research concerning facial animation and 3D object reconstruction uses multiple camera setups [1, 19] or setups using special sensors [26, 22]. Multiple camera setups always require a synchronization step, so that the recorded frames can be matched. The usage of modern 4D systems is cost intensive and the provided data might no enable the estimation of trajectories of the points of interest [26, 9]. As already mentioned the mirror setup introduces a virtual multi camera system. The application fields of this system are various. It can be used to create realistic facial animation [15] or provide a quantitative objective analysis of the progress of facial palsy patients after neuromuscular surgery [6]. There are also photogrammetric approaches which use mirrors to reconstruct the surface of not fully accessible objects or artifacts [5].

3. Hardware setup

The hardware setup consists of a mirror setup proposed by Frey *et al.* [7], a calibration grid and a commercial video camera (see Figure 1). The angle Θ between the two mirror planes is flexible, otherwise the two mirrors are fixed in the scene. The mirror-reflected views can be regarded as images taken by virtual cameras, which are in a distinct view direction comparing to the physical ones [16].

4. Analysis of the mirror setup

The aim of this section is to study the mirror system in order to find the relationship between five de-

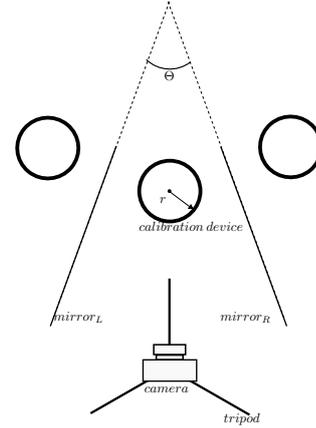


Figure 1. The mirror setup consists of two mirror planes, which intersect at an angle Θ . The calibration device is positioned inside the mirror setup. A commercial photo camera records the scene.

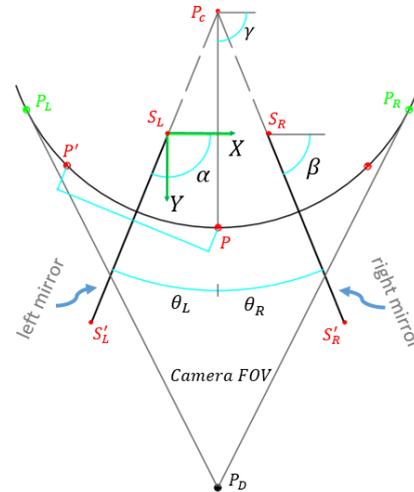


Figure 2. Representation of the mirror system as seen from above. The mirror planes are represented as bold lines.

grees of freedom of the system that are: left mirror angle, right mirror angle, object position and camera position. This is useful to know how many reflected views are visible by changing parameters. In order to study how the reflection location of the object is formed by a mirror we use the law of reflection [13]. Using the optic theory we can introduce a set of relations useful to describe the different configurations of the system. We perform a 2D analysis, as this is enough in order to achieve the objective, at this purpose the mirror system can be represented as seen from above as shown in Figure 2.

The position of an arbitrary point $P = (x_h, y_h)$, the position of left corner $S_L = (x_L, y_L)$, the right corner position $S_R = (x_R, y_R)$, left mirror angle α , right mirror angle β and the length L of the two mirrors are given, while the reference system is placed

on the left pivot so that $(x_L, y_L) = (0, 0)$. With this information the left extreme position $S'_L = (x'_L, y'_L)$ and the right extreme position $S'_R = (x'_R, y'_R)$ are obtained by projecting the length L of the mirrors on the reference system, while the angle between the mirrors is $\Theta = \alpha - \beta$.

The rotation center $P_C = (x_C, y_C)$ is obtained from the intersection between two straight lines that go through the left and right mirror.

According to the reflection theory all the reflections of the point P are placed on a circle centered in P_C with radius r given by the Euclidean distance between the rotation center and the position of P .

To obtain the position of the point $P' = (x'_h, y'_h)$ shown in Figure 2 we need the angle γ of the point P with respect to the rotation center

$$\gamma = \tan^{-1} \left(\frac{y_h - y_c}{x_h - x_c} \right) \quad (1)$$

$$\angle \Theta_R = |\alpha - \gamma| \quad (2)$$

$$\angle \Theta_L = |\beta - \gamma| \quad (3)$$

From (2) and (3) the position of P' is

$$P' = \begin{bmatrix} x'_h \\ y'_h \\ 0 \end{bmatrix} = R_z \cdot \begin{bmatrix} x_h - x_c \\ y_h - y_c \\ 0 \end{bmatrix} + \begin{bmatrix} x_c \\ y_c \\ 0 \end{bmatrix} \quad (4)$$

with R_Z rotation matrix

$$R_z = \begin{bmatrix} \cos(2 \cdot \angle \Theta_L) & -\sin(2 \cdot \angle \Theta_L) & 0 \\ \sin(2 \cdot \angle \Theta_L) & \cos(2 \cdot \angle \Theta_L) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

(4) specifies also the translation and rotation of the reference system. Other rotations are obtained by introducing the following formulas

$$r(n) = \begin{cases} 0, & n < 1 \\ r(n-2) + 2(-1)^n \Theta_L, & n \bmod 4 < 2 \\ r(n-2) + 2(-1)^n \Theta_R, & n \bmod 4 \geq 2 \\ \text{undefined}, & n > M-1 \end{cases} \quad (6)$$

$$M = \frac{360^\circ}{\Theta}, n \in \mathbb{Z} \quad (7)$$

(6) asserts that the n -th rotation can be obtained as combination of the previous rotations, the second part of the equation defines an alternation of rotations governed by the angles (2) and (3). This succession because the secondary view is obtained as a reflex of primary view and so on. (25) instead refers to the total number of reflections that are generated by a specific configuration of the mirrors, including the real object. Given also the position $P_D = (x_D, y_D)$

of the camera we find two lines, one goes from the camera center to the left tangent point P_L at circle, and another line goes from the camera center to the right tangent point P_R at circle. Using P_L, P_R, P and the circle equation we know the arc lengths l_1 from P to P_L and l_2 from P to P_R , so we can calculate the left and right boundary angle

$$\alpha_L = \sin^{-1} \left(\frac{l_1}{2r} \right) \quad (8)$$

$$\beta_R = \sin^{-1} \left(\frac{l_2}{2r} \right) \quad (9)$$

given the rotation $r(n)$ of the n -th view it is visible if

$$-\beta_R \leq r(n) \leq \alpha_L \quad (10)$$

However, when the radius of the circle where all reflections are placed exceeds the extremities of the mirror, the tangent point cannot be calculated. In this case, we first draw the lines that go from P_D to S'_L and S'_R , then we find the intersection between those lines and the circle, we chose the solutions nearest to the camera center, in this way (10) is still valid. The same happens when the circle becomes larger than the camera field of view (FOV). In this case the two lines that go from P_D must have fixed slope (see Figure 3). Given the position of an arbitrary point we can say if this point is visible or not in the mirror reflected-view, and how many times it is visible.

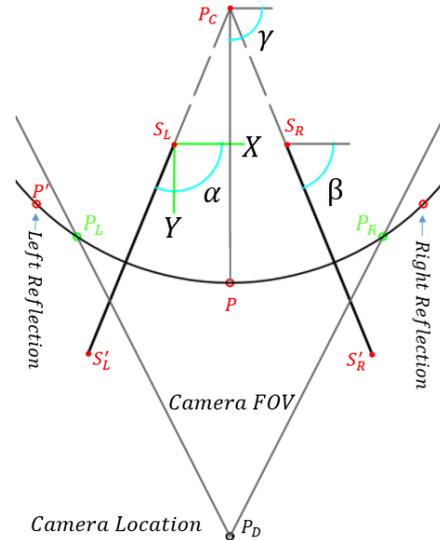


Figure 3. Left and right reflection are out of camera field of view, so you cannot see them from the camera location.

To analyze how many view are visible when parameters change, we simulate the different configurations based on the presented formulas. During

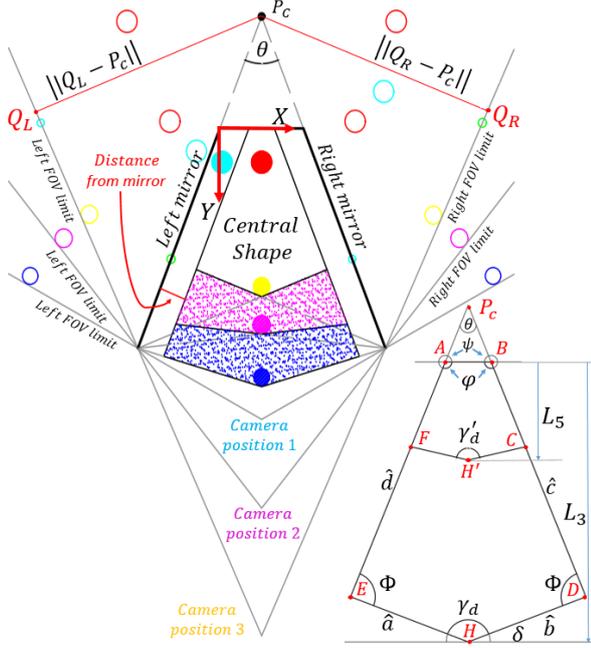


Figure 4. Mirror system. with head and camera in different locations and regions in which we can see at least three views of the face. Bottom right we can see a detail of the central shape with also the limits L_3 and L_5 until we can see three to five view.

our tests only the symmetric case is considered with $\beta = 180 - \alpha$. In this way the degrees of freedom of the system decreases from five to four: Θ is considered instead of α and β separately. The evaluation, presented in section 9.1, is used to study what happens when the object is moved. Our results denote that the number of views is constant, until the object moves inside a fixed region that is shown in Figure 4, where we see the mirrors as bold lines, the object as filled circles and the reflection as empty circles; fixed the angle Θ between the mirrors, three (purple, blue and yellow) setups are shown by changing the camera position, while at bottom right is shown a detailed central shape of chosen setup. As we can see the central shape region is defined as:

- Minimum distance between object and mirror, that is used to draw edges \hat{d} and \hat{e} in Figure 4, it is positively related to the size of the object.
- Rotations of the lines that limit the FOV of the camera around the rotation center by Θ , which are used to draw edges \hat{a} and \hat{b} in Figure 4.

Now we have to characterize this region, in order to define useful ranges. First we find the inner angles:

$$\psi = 90 - \Theta/2 \quad (11)$$

$$\phi = 90 + \Theta/2 \quad (12)$$

$$\gamma_d = \min(FOV + 2\Theta, 360 - (FOV + 2\Theta)) \quad (13)$$

The sum the inner angles of a polygon is given by $(k - 2) \cdot 180^\circ$ [18], where k is the number of the sides, in our with $k = 5$ we obtain

$$\Phi = \frac{180 \cdot (5 - 2) - 2\varphi - \gamma_d}{2} \quad (14)$$

$$\delta = 90 - \gamma_d/2 \quad (15)$$

The segments \hat{a} and \hat{b} are defined as lines that go through E and D with a certain angle δ . The interception between \hat{a} and \hat{b} gives us the position of H.

$$y = \begin{cases} \begin{bmatrix} \tan(+\delta) \begin{pmatrix} x - x'_L \end{pmatrix} \\ \tan(-\delta) \begin{pmatrix} x - x'_R \end{pmatrix} \end{bmatrix} + \begin{bmatrix} y'_L \\ y'_R \end{bmatrix}, & \gamma_d < 180 \\ \begin{bmatrix} \tan(-\delta) \begin{pmatrix} x - x'_L \end{pmatrix} \\ \tan(+\delta) \begin{pmatrix} x - x'_R \end{pmatrix} \end{bmatrix} + \begin{bmatrix} y'_L \\ y'_R \end{bmatrix}, & \text{otherwise} \end{cases} \quad (16)$$

Given the shape of the region in Figure 4, we can see at least three views as long as the object is inside the region ABDHE. To see five views of the face we need an angle Θ that is less than 45° , this can be shown by using (25), then the region until we can see five faces is obtained by finding the intersection Q_L between the left FOV limit and the line that goes from the rotation center with fixed slope given by $90 + 2\Theta$

$$y = \tan(90 + 2\Theta) (x - x_c) + y_c \quad (17)$$

the same is done with the right FOV limit by using $90 - 2\Theta$ as slope for the line from the rotation center

$$y = \tan(90 - 2\Theta) (x - x_c) + y_c \quad (18)$$

so given also Q_R the limit of the shape until we can see five faces is obtained from

$$L_5 = \min(\|Q_L - P_C\|, \|Q_R - P_C\|) + y_c \quad (19)$$

$$\gamma'_d = \frac{L_5}{L_3} (\gamma_d - 180) + 180 \quad (20)$$

in this way the region ABCH'F until five views are visible is completely defined. The slopes $\tan(90 \pm 2\Theta)$ are obtained by taking into account the rotations of secondary reflections.

5. System calibration with Color chessboard

As stated in Table 2, we know that by changing the angle between the mirrors we can see at least three up to five views of the patients head from different view

points. This means that a planar calibration devices is not suitable for our purpose since its views are not visible five times. The same applies to the calibration grid currently used, which is not properly designed to use the configurations where five views are visible. Furthermore, currently the calibration process is done in semiautomatic way, and it is not reliable with an angle between the mirrors that is different from 42.9° [6, 27]. In the following sections a new approach which automates the calibration is proposed.

5.1. Calibration device

As we said in Section 5 there are many problem with the current calibration grid, in order to solve them, and to recall the shape and size of the human head, we propose to use a novel three-dimensional calibration device which uses colors to simplify points associations (see Figure 5).

Our device is made of a cylindrical surface on which we draw a checkerboard pattern. Since the human head has a circumference of about 60 cm also our cylinder must have the same size, the height of the cylinder is 13 cm, while each patch has a dimension of 25 mm. Thanks to a smallest device we can make the most of setups. As to the color selection, the RGB space can be represented as a unit-cube rectangular Cartesian space of three perpendicular axes, where each axis represents the excitation of one color [11]. The vertices of the unit-cube are black (0,0,0), Blue (0,0,1), Magenta (1,0,1), Red (1,0,0), Green (0,1,0), Yellow (1,1,0), White (1,1,1), Cyan (0,1,1). As it is shown in Figure 5, we chose white and black as background of the checkerboard pattern, while the other colors are chosen by following the maximum distance between the vertices of the unit-cube. In this way, a Hamming distance of two between the colors that belong to a same angular sector is obtained. The order of colors on calibration device, from left to right, is shown in Table 1

Table 1. Order of colors on cylinder

	col	bg	col	col	bg	col
r	1	0	1	1	1	0
g	0	0	1	0	1	0
b	1	0	0	0	1	1
r	0	0	0	1	0	1
g	0	1	0	1	1	1
b	0	1	0	1	0	1
	bg	col	bg	bg	col	bg

while the transition from a dark background

checkerboard to a white background checkerboard defines the origin of the axes. Chosen color combination has high error resilience because we can distinguish e.g. magenta from red, and cyan from blue, also considering the background.

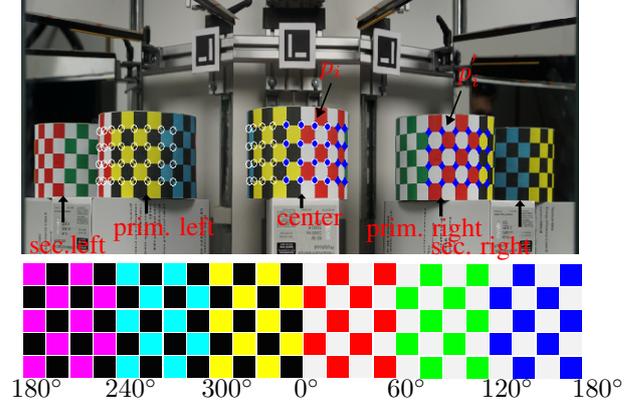


Figure 5. The calibration device is placed inside the mirror setup. The calibration device consists of alternating color coded checkerboard patterns, attached to a cylinder with a radius of 9.55 cm.

6. Segmentation and Modeling of cylinder

The whole task is summarized below. We start with the preprocessing of the image, followed by color segmentation (Section 6.2), view identification (Section 6.3), modeling, corner detection (Section 6.4), and corner association (Section 6.5).

6.1. Preprocessing and Segmentation of the image

The objective of the preprocessing is to avoid artefacts such as shadows and highlights, that could decrease the performance of the segmentation. Methods like normalized RGB, $c_1c_2c_3$ and $l_1l_2l_3$ colour model [17] are usually used. However, results obtained with those methods are not reliable to our purpose since they represent the white and black of checkerboard with different color. To this purpose we introduce our own normalization described below

$$\begin{bmatrix} r \\ g \\ b \end{bmatrix} = \begin{bmatrix} \min(1, Red(2 - Luminance)) \\ \min(1, Green(2 - Luminance)) \\ \min(1, Blue(2 - Luminance)) \end{bmatrix} \quad (21)$$

where *Red*, *Green*, *Blue* area indicates the color bands and *Luminance* is the luminance matrix taken from $L^*a^*b^*$ color representation. With this normalization, darkest regions are multiplied with higher gain, while brightest region area multiplied with lower gain. An image with adjusted light condition is obtained in this way.

6.2. Color segmentation

The purpose of the color segmentation is to isolate the cylindrical pattern within the scene. The main problem is that most of segmentation techniques are designed for monochrome images and subsequently partly extended for the segmentation of color images by using RGB color coordinates or some of their transformations [2]. Methods like color threshold, or K-means are not reliable to our purpose because small changes of image have different results, and we do not know how many clusters there are in the image because of the background. Manual selection of the color seeds is also not suitable for our purpose because the user must select color samples every time and the background belongs to a user defined class, without any predictability.

In order to remove anything that does not belong to the calibration cylinder the proposed solution uses background subtraction [23]. In our method the background estimation uses a linear color space independent by luminosity condition, at this purpose we chose the $L^*a^*b^*$ color representation. We fit the a^* and b^* histograms of the background image with normal distribution and then the 95% confidence interval is used as threshold on the image with the cylinder. Figure 6 shows hue histograms before and after background subtraction. The colors are separated by using a peak finding algorithm which uses the minimum peak distance, and the width of smoothing window as input parameters. In this way the estimation of background is done globally by using histograms instead of the distance between the colors, so small changes in pixel values due the noise are not relevant.

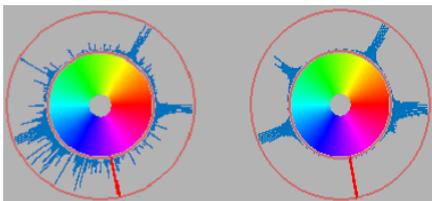


Figure 6. Hue histograms before and after background subtraction

6.3. View identification

The purpose of the identification is to identify inside the original image the left secondary view, left primary view, central view, right primary view, and right secondary view that are showed in Figure 5. The identification is done by splitting the original image in at least 3 up to 5 windows. In this operation, we assume that at least two colors for every reflec-

tion are visible. The view identification is carried out by using the colors and the relative positions of the checkerboards. Knowing that the odd reflections generated by the mirror are left-right reversed we can identify the primary and secondary reflections by using the order of the colors (Section 5.1) and by counting the occurrence of the same color on each side.

6.4. Modelling and Corner Detection

The main problem of the modeling task, and of course also for corner detection, is that images in our cases may appear blurred. This happens because we cannot focus at the same time on the real cylinder and on its images because the real object and the reflected ones are at different distances with respect to the camera position. Thus, the patches in the segmented image tend to be grown together due to blur. For correct identification of the corners an erosion step is applied [20], then the model of the cylinder is obtained by masking and interpolation operations. The contours of the blobs are used to retrieve the edges positions (see Figure 7). The interpolation returns a model used in later stages for corner detection and association. The state of the art methods assume to have a black-white planar checkerboard ([12, 3, 4, 8]). These assumptions are not verified in our cases. We use standard Harris corner detector [10], together with the model of the cylinder. The model is used to obtain for each corner between four patches its close position, then the position of the corner is refined by using Harris.

6.5. Corner association

The objective of this section is to associate each corner from a particular view with each corner in all other views. At this purpose the corners are sorted with respect to the center of the axes defined in Section 5.1 and then organized in matrices, one for each color and type of reflection. Each matrix is flipped if it refers to an odd reflection. The index of the matrices, as well as the color, encodes the association between the corners on the cylinder and its reflections. The detected corners of the checkerboard pattern are used as input parameters (p_i and p'_i) for the 3D position estimation.

7. Position estimation in 3D

Lin *et al.* [16] proposed a robust and in-expensive method to estimate the 3D position of markerpoints using multi-view video sequences. The main advantage of this method is that it takes the mirror proper-

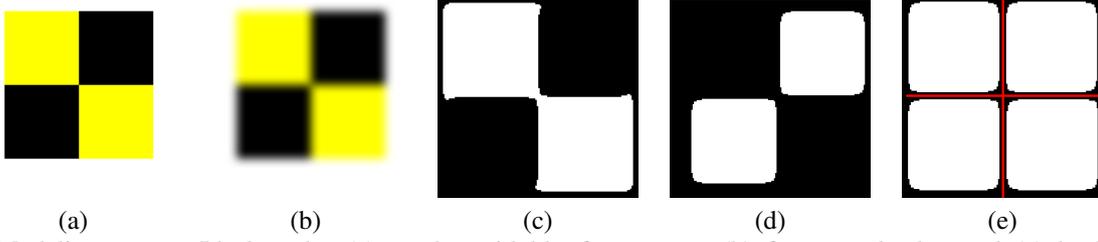


Figure 7. Modeling process. Ideal patches (a), patches with blur from camera (b), foreground color mask (c), background mask only for black and white patches (d), resulting mask after erosion (e) and contour interpolation.

ties into account. Furthermore the angle between the mirror planes can be chosen arbitrarily. The orientation and location of the mirror planes, as well the 3D position estimation of the corners of the calibration device, was estimated according to Lin *et al.* [16]. The location and orientation of each of the two mirror planes can be described by a plane equation

$$ax + by + cz = d \quad (22)$$

The unit normal of the plane is $u = (a, b, c)^t$. The 3D position of an arbitrary point $m_i = (x_{mi}, y_{mi}, z_{mi})$ and the corresponding virtual 3D position $m'_i = (x'_{mi}, y'_{mi}, z'_{mi})$ of the mirror reflection can be presented as follows

$$m_i = \frac{z_{mi}}{f} \cdot p_i \quad (23)$$

$$m'_i = \frac{z'_{mi}}{f} \cdot p'_i \quad (24)$$

$p_i = (x_{pi}, y_{pi}, f)$ and $p'_i = (x'_{pi}, y'_{pi}, f)$ represent the projections onto the mirror plane. The focal length is given by f . The vectors m_i , m'_i and u are co-planar. According to Lin *et al.* each marker and rest stationary point for rigid body calibration a matrix M can be formed,

$$Mu = 0 \quad (25)$$

$$M = \begin{bmatrix} (y_{p1} - y'_{p1})f & (x_{p1} - x'_{p1})f & (x_{p1}y'_{p1} - y_{p1}x'_{p1}) \\ \vdots & \vdots & \vdots \\ (y_{pn} - y'_{pn})f & (x_{pn} - x'_{pn})f & (x_{pn}y'_{pn} - y_{pn}x'_{pn}) \end{bmatrix} \quad (26)$$

By using the least square method the vector u with the least error can be estimated.

z_{mi} and z'_{mi} can be estimated by solving the system of linear equations given in (27).

$$\begin{bmatrix} \left(\frac{2a^2-1}{2f}\right)x_{pi} + \left(\frac{ab}{f}\right)y_{pi} + ac & \frac{x'_{pi}}{2f} \\ \left(\frac{ab}{f}\right)x_{pi} + \left(\frac{2b^2-1}{2f}\right)y_{pi} + bc & \frac{y'_{pi}}{2f} \\ \left(\frac{ac}{f}\right)x_{pi} + \left(\frac{bc}{f}\right)y_{pi} + \frac{2c^2-1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} z_{mi} \\ z'_{mi} \end{bmatrix} = d \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad (27)$$

The corners of the checkerboard in the frontal and reflected view were p_i and p'_i (see Figure 5). For the estimation of the position and location of the mirror only one frame was necessary as the position and location of the mirror planes does not change through the video sequence. The camera should be approximately oriented along the symmetry axis of the mirror pair, so that an equal amount of markerpoints and the corresponding reflections are visible in both mirrors. As long as this is true the camera position has no influence on the accuracy of the result.

According to (27) and (23) the 3D position of an arbitrary point can be estimated. The estimated 3D positions of the calibration cylinder in the frontal and reflected view can be seen in Figure 8.

8. Experimental setup

The cylinder was positioned inside the mirror setup. The scene was recorded with a commercial photo camera. The intrinsic camera parameters were obtained using the calibration toolbox by Bouquet [3]. A dataset of 50 images was created by varying the position of the camera, the angle between the mirrors and the cylinder position; in this way, a set of images with variable number of reflections from three to five is obtained. The ground truth is generated by three people in order to not bias results. For each image the segmentation, the position of the corners, and the size of the largest patch are manually defined. The same evaluation method is also used in [8]. For a segmentation, denoting with P the result image without the background, and with S the ground truth the following are defined

- $TP = |P \cap S|$ $TN = |U| - |P \cup S|$
- $FP = |P / S|$ $FN = |S / P|$

where $|U|$ is the size of the whole image. The result of comparison is obtained by using precision (PR), recall (RC) and F1-score (F1)[24].

The previously described corner detection procedure is evaluated in same way, the algorithm provide the corners which are measured by comparing the position of the corner obtained by manual labeling with the position returned by the algorithm. The pixel error is normalized to the maximum size of the patch which is also manually defined for each image in the dataset. We use the mean μ and standard deviation σ of the Euclidean distance between the labeled position and the predicted one as evaluation measurement, also we report the mean and the total number of match obtainable theoretically from our cylinder.

9. Results and Discussion

The results of the proposed methodology are presented and discussed in the following sections.

9.1. Analysis of the mirror setup

In this section we simulate different setups of the system based on the formulas presented in Section 4. The length L of the mirrors is 65 cm, $(x_R, y_R) = (0, 28.5)$ and the used camera has a FOV of 50° . The range from 15° to 120° degree is used as value of Θ because outside this range only the real face is visible. The camera is placed in the middle between the mirrors with $x_D = x_R/2$ while y_D changes according with the angle Θ and the FOV of the camera

$$y_D = L \cos\left(\frac{\Theta}{2}\right) + \frac{|x'_R - x'_L|}{2} \coth\left(\frac{FOV}{2}\right) \quad (28)$$

To reduce the number of setups we only consider the angle between the two mirrors (symmetric case) and the position y_h of the object along Y axis shown in Figure 2 while $x_h = x_R/2$ is fixed in the middle, also we only consider the frontal and side face because we do not have interest to characterize how many faces are visible from behind. $f^* = L \cos(\Theta/2)$ is the projection of L on Y axis, which value is related to the angle between the mirrors. The Y Position of the object is the product between f^* and a gain factor obtained empirically from our application. Table 2 shows that the range with highest number of reflection is with Θ between 28° to 45° ,

9.2. Segmentation

To test the performance of our solution we used the method described in Section 8. For segmentation the comparison returns an average precision of 0.9976, the recall is 0.9733 and the F1-score is 0.9853, if it is possible to characterize the background through a normal distribution, and therefore

Table 2. Total number of view that can be visible by changing the system parameters.

Angle [$^\circ$]	Head Position	N $^\circ$ views
88 $^\circ$:120 $^\circ$	\sim	1
45 $^\circ$:88 $^\circ$	$y_h < 0.8 \cdot f^*$	3
	$y_h > 0.8 \cdot f^*$	1
28 $^\circ$:45 $^\circ$	$y_h < 0.4 \cdot f^*$	5
	$y_h > 0.4 \cdot f^*$	3
	$y_h < 0.6 \cdot f^*$	3
	$y_h > 0.6 \cdot f^*$	1
15 $^\circ$:28 $^\circ$	$y_h < 0.5 \cdot f^*$	3

Table 3. Segmentation results. (PR= precision; RC= recall; F1-S= F1-Score)

N $^\circ$ view	PR	RC	F1-S
3	0.9975	0.9688	0.9829
4	0.9975	0.9763	0.9868
5	0.9977	0.9717	0.9845

Table 4. Corner detection. (μ = mean error; σ = std error; M/T M = Mean/Total number of Match)

N $^\circ$ view	μ	σ	M/T M
3	0.2628	0.2569	40 / 44
4	0.2258	0.2679	44 / 48
5	0.2499	0.2672	52 / 56

there would be no other objects with same color characteristics within the scene. Table 3 shows the performance for each group of pictures in the dataset.

About the detection of corners the μ error in pixels above whole dataset is 1.5576 and the σ of the error is 1.7121. Table 4 shows the error reported in mm for each group of images, where the size of each patch is 25 mm. The error is lowest near the center of the device, where we have a black to white transition, and increases with the distance from the center because of the projections. Another drawback is that the more is the distance from the center, the more little are the patches on the side borders, so that some corners are lost, as shown in Table 4.

9.3. Position estimation in 3D

The result of the 3D position estimation using the proposed methodology can be seen in Figure 8. The angle Θ between the two mirror planes was estimated to be 41.9° . The real angle between the mirror planes was at 42.9° . The difference between the estimated and the real angle can be explained by a systemic error, since the two mirror planes cannot be fixed at a defined angle. The calculated 3D position m_i of the central view is indicated with red and blue stars in

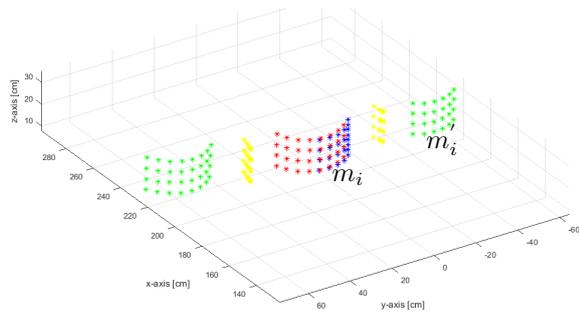


Figure 8. Using the corresponding marker points p_i and p'_i the calibration device can be represented in 3D by m_i and m'_i (best seen in color).

the Figure 8. The 3D positions of the calibration device in the reflected view m'_i are indicated with green stars. The mirror planes (yellow stars) are situated in at $0.5 \cdot \overline{m_i m'_i}$, as expected according to Section 4.

10. Conclusion

In this paper, we studied the mirror system in order to find the relationship between the left mirror angle, right mirror angle, object position, and camera position. This is useful in order to know how many faces could be visible by changing these parameters, and to maximize the amount of information that can be captured in a single video session. In previous sections, we found that our setup provides better results for angles of the two mirrors between 28° and 45° . In this range, we can see at least three up to five images of the head of the patient from different viewpoints, and we have a lot of redundant information which is useful for reconstruction of a 3D environment, and for physicians who need to measure the progress of facial palsy after a surgery. Also, our new calibration device performed well for the proposed automatic calibration approach. Evaluations show that with this colored cylinder, we can find at least 40 up to 52 matches instead of 35 obtainable from the calibration grid. Also, it permits to use all the five views of the patients faces. The matches obtained using the proposed calibration device can be used to estimate the position and orientation of the mirror planes. As mentioned in Section 7 the knowledge about the unit normal vector of the mirror planes enables the estimation of the 3D position of corresponding points in the central and the reflected view in the image. This enables a 3D surface reconstruction or estimation of 3D trajectories of interest points.

In the future, we plan to analyze a different color orders, to maximize the contrast between two adja-

cent sectors, and to build a higher cylinder to obtain more matches and less background to have less limits on it during the segmentation task.

References

- [1] T. Beeler, B. Bickel, P. Beardsley, B. Sumner, and M. Gross. High-quality single-shot capture of facial geometry. *ACM Trans. Graph.*, 29(4):40:1–40:9, 2010. 2
- [2] S. Bhattacharyya. A brief survey of color image preprocessing and segmentation techniques. *Journal of Pattern Recognition Research*, 1(1):120–129, 2011. 6
- [3] J. Bouguet. *Camera Calibration Toolbox for Matlab, Computational Vision at the California Institute of Technology*. 2012. 1, 6, 7
- [4] G. Bradski and A. Kaehler. *Learning OpenCV: Computer vision with the OpenCV library*. ” O’Reilly Media, Inc.”, 2008. 1, 6
- [5] A. Fasano, M. Callieri, P. Cignoni, and R. Scopigno. Exploiting mirrors for laser stripe 3d scanning. *Fourth International Conference on 3-D Digital Imaging and Modeling, 2003. 3DIM 2003. Proceedings.*, pages 243–250, 2003. 2
- [6] M. Frey, P. Giovanoli, H. Gerber, M. Slameczka, and E. Stüssi. Three-dimensional video analysis of facial movements: A new method to assess the quantity and quality of the smile. *Plastic and Reconstructive Surgery*, 104(7):2032–2039, 1999. 1, 2, 5
- [7] M. Frey, A. Jenny, P. Giovanoli, and E. Stüssi. Development of a new documentation system for facial movements as a basis for the international registry for neuromuscular reconstruction in the face. *Plastic and Reconstructive Surgery*, 93(7):1334–1349, 1994. 2
- [8] A. Geiger, F. Moosmann, . Car, and B. Schuster. Automatic camera and range sensor calibration using a single shot. *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 3936–3943, 2012. 1, 6, 7
- [9] B. Guenter, C. Grimm, D. Wood, H. Malvar, and F. Pighin. Making faces. *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*, pages 55–66, 1998. 2
- [10] C. Harris and M. Stephens. A combined corner and edge detector, 1988. 6
- [11] G. Hasting and A. Rubin. Colour spaces-a review of historic and modern colour models. *African Vision and Eye Health*, 71(3):133–143, 2012. 5
- [12] J. Heikkila and O. Silven. A four-step camera calibration procedure with implicit image correction. *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pages 1106–1112, 1997. 1, 6

- [13] M. Katz. *Introduction to Geometrical Optics*. WORLD SCIENTIFIC, 2013. 2
- [14] J.-S. Kim, P. Gurdjos, and I. Kweon. Geometric and algebraic constraints of projected concentric circles and their applications to camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):637–642, 2005. 2
- [15] I.-C. Lin and M. Ouhyoung. Mirror MoCap: Automatic and efficient capture of dense 3d facial motion parameters from video. *The Visual Computer*, 21(6):355–372, 2005. 2
- [16] I.-C. Lin, J.-S. Yeh, and M. Ouhyoung. Extracting 3d facial animation parameters from multiview video clips. *IEEE Computer Graphics and Applications*, 22(6):72–80, 2002. 2, 6, 7
- [17] E. Maggio and A. Cavallaro. Feature extraction. In *Video Tracking: Theory and Practice*, page 40. 2011. 5
- [18] J. Proctor. Irregular polygons. In *Revision Plus GCSE Edexcel Maths Foundation Revision Guide: OCR A*, pages 63–64. Letts Educational, 2008. 4
- [19] T. Putze. Geometric modelling and calibration of a virtual four-headed high speed cameramirror system for 3-d motion analysis applications. *Optical 3-D measurement techniques VII*, pages 167–174, 2005. 1, 2
- [20] M. Ruffi, D. Scaramuzza, and R. Siegwart. Automatic detection of checkerboards on blurred and distorted images. *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3121–3126, 2008. 6
- [21] J. Salvi, X. Armangu, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 35(7):1617 – 1635, 2002. 2
- [22] M. Shreve, M. Pamplona, T. Laguev, D. Goldgof, and S. Sarkar. High-resolution 3d surface strain magnitude using 2d camera and low-resolution depth sensor. *Pattern Recognition Letters*, 50:34 – 42, 2014. Depth Image Analysis. 2
- [23] A. Sobral and A. Vacavant. A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos. *Computer Vision and Image Understanding*, 122:4–21, 2014. 6
- [24] M. Sokolova, N. Japkowicz, and S. Szpakowicz. Beyond accuracy, f-score and ROC: A family of discriminant measures for performance evaluation. *AI 2006: Advances in Artificial Intelligence: 19th Australian Joint Conference on Artificial Intelligence, Hobart, Australia, December 4-8, 2006. Proceedings*, pages 1015–1021, 2006. 7
- [25] R. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal on Robotics and Automation*, 3(4):323–344, 1987. 2
- [26] C.-H. Tzou, N. Artner, I. Pona, A. Hold, E. Placheta, W. Kropatsch, and M. Frey. Comparison of three-dimensional surface-imaging systems. *Journal of Plastic, Reconstructive & Aesthetic Surgery*, 67:489–497, 2014. 1, 2
- [27] C.-H. Tzou, I. Pona, E. Placheta, A. Hold, M. Michaelidou, N. Artner, W. Kropatsch, H. Gerber, and M. Frey. Evolution of the 3-dimensional video system for facial motion analysis: Ten years’ experiences and recent developments. *Annals of Plastic Surgery*, 69(2):173–185, 2012. 5
- [28] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000. 2