World Academy of Science, Engineering and Technology
International Journal of Mechanical and Mechatronics Engineering
Vol:5, No:11, 2011

Calibration of Parallel Multi-View Cameras

M. Ali-Bey, N. Manamanni, and S. Moughamir

***Abstract*—**This paper focuses on the calibration problem of a multi-view shooting system designed for the production of 3D content for auto-stereoscopic visualization. The considered multi-view camera is characterized by coplanar and decentered image sensors regarding to the corresponding optical axis. Based on the Faugéras and Toscani's calibration approach, a calibration method is herein proposed for the case of multi-view camera with parallel and decentered image sensors. At first, the geometrical model of the shooting system is recalled and some industrial prototypes with some shooting simulations are presented. Next, the development of the proposed calibration method is detailed. Finally, some simulation results are presented before ending with some conclusions about this work.

***Keywords*—**Auto-stereoscopic display, camera calibration, multi-view cameras, visual servoing

I. Introduction

NOWADAYS, three-dimensional television (3DTV) knows a real revolution thanks to the technological headways in visualization, computer graphics and capture technologies. Depending on the technology adopted, the 3D visualization systems can be either stereoscopic or auto-stereoscopic. In stereoscopy, viewing glasses are required and different technologies are used to separate the left-eye and right-eye views using anaglyphs or color multiplexing [1], [2], occultation and polarization multiplexing [3], time sequential presentation using active shuttering glasses [4]. In auto-stereoscopy, the display devices do not need any special viewing glasses since they are direction-multiplexed devices equipped by parallax barriers or lenticular systems [4]-[6].

To supply these display devices with 3D contents, the more interesting and used methods are based on the synthesis of multiple viewpoint images from 2D-plus-depth data for stereoscopic display [7] and auto-stereoscopic display [8]. The transformation between viewing and capturing space with controlling perceived depth in stereoscopic case is described in [9]. A generalized multi-view transformation model between viewing and capturing space with controlled distortion is proposed in [10]. A time varying concept of this architecture to capture dynamic scenes is reported in [11], [12] and a study of the rendering quality assessment is reported in [13]. 3DTV technologies advances and evaluation can be found in [14], [15].

The authors would like to thank the National Agency of Research (ANR) within Cam-Relief project and the Champagne-Ardenne region within CPER CREATIS for their financial support.

The Authors are with the CReSTIC-URCA, Moulin de la Housse BP. 1039, 51687 Reims Cedex2, France. Corresponding author: Mohamed ALI-BEY, phone: +33-3-26-91-86-17; Fax: +33-3-26-91-31-06; (e-mail: mohamed.ali-bey@univ-reims.fr, noureddine.manamanni@univ-reims.fr, said.moughamir@univ-reims.fr).

In the present paper we are interested in the calibration of multi-view cameras with parallel and decentered image sensors. In fact, we are confronted to the camera calibration problem every time we need to know the intrinsic and/or extrinsic parameters of the considered shooting system. In view of the multiple possible cases of study, a lot of works in the literature are devoted to this subject. We are particularly interested in the calibration method proposed by Faugéras and Toscani in [16], [17] then reported in [18] and that we extend here to the case of a multi-view camera with parallel and decentered image sensors.

This paper is organized as follows: in Section 2 some recalls about the viewing/shooting geometrical process for auto-stereoscopic visualization in the case of parallel and decentered shooting configuration are given. Then, an appropriate perspective projection model is derived. In section 3 the calibration method development is detailed. Next, the simulation results are presented in section 4. Finally, some conclusions and prospects end the paper.

II. Shooting/Viewing Geometrical Process

The shooting/viewing geometrical process model consists in some geometric transformations from the capturing space to the rendering one. Thus, three groups of parameters can be defined: a rendering parameters group imposed by the auto-stereoscopic viewing geometry, a second group defining the geometrical structure of the 3D camera model, and a third one controlling the distortions that affect the 3D rendering. Knowing the parameters of these three groups and the relations between them, one can define a capturing configuration satisfying both parameters imposed by the visualization device and those of the wished distortions.

Thereafter, one recalls succinctly the different parts of this geometric process and the associated parameters developed in our laboratory by [10]. In the first subsection, a multi-view rendering geometry of auto-stereoscopic display device is described with the viewing parameters definition. Then, the shooting geometry of parallel and decentered configuration is described defining the capture parameters. After that, the relations between the capturing and the viewing parameters are given to define the distortion controlling parameters.

A. Multi-View Rendering Geometry

The considered display device is an auto-stereoscopic screen as depicted in Fig. 1, where H and W represent respectively the height and the width of the device.

To perceive the 3D rendering, the observers should be at a preferential positions imposed by the screen and determined by a viewing distance d, a lateral distance o_i and a vertical

distance δ^o corresponding to a vertical elevation of the observer's eyes. Let O_i and O_{ri} be respectively, the position of the left and the right eye and b is the human binocular gap. The perceived point noted m results from the viewed points m_i and m_{ri} respectively, by the left and the right eye. A viewing frame $r = (C_r, x, y, z)$ is associated to the device in its centre C_r for expressing the viewing geometry.

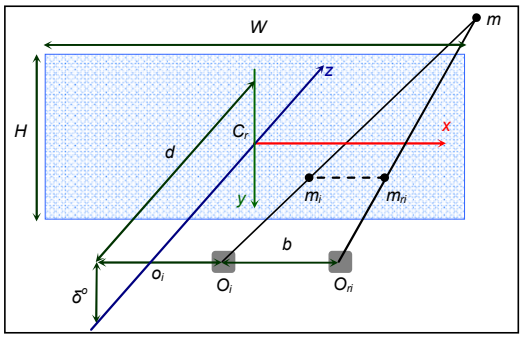

Fig. 1 Viewing geometry

B. Shooting Geometry

The geometry of a parallel multi view shooting with decentered image sensors configuration is presented in Fig.2. To help the reader to well analyze this shooting geometry, we have chosen to represent the perspective, the front and the top views. Note that the explanation given thereafter can be transposed for each of the given views.

Let $R = (C_p, X, Y, Z)$ be the frame associated to the scene plane CB whose dimensions are $W_b \times H_b$ (Fig.2-a). The shooting system is composed of n sensor/lens pairs with focal length f. It is characterized by parallel optical axes Z_{ci} distanced uniformly by an inter-optical distance B. The position of the corresponding optical centers C_i is defined in the frame R by a lateral position p_i, a vertical position $-P$ and a convergence distance $-D$ according to the Z direction. The optical centers are lined up parallel to the image sensors which are coplanar between them. Each image sensor represented by its principal point I_i with dimensions $w \times h$ is decentered regarding to its corresponding optical center C_i, by a lateral distance a_i and a vertical distance e. Note that all the sighting axes (I_iC_i) converge to a common point C_p situated in the center of the scene plane which is distant from the optical centers' line by the convergence distance D.

For more clarity a practical scheme of a five points of view shooting system is depicted in Fig.3. It shows clearly the structural parameters of the capture. Note that five point of view is the minimum needed by the existing commercial auto-stereoscopic display devices. Hence, we have chosen to illustrate a five points of view system in order to avoid the scheme to be too cumbersome.

C. Transformation Parameters

The transition from the shooting space to the viewing one is expressed by the transformation between the captured point homogenous coordinates $M(X, Y, Z, 1)_R$ and those of the perceived point $m(x, y, z, 1)$ for more details see [10]:

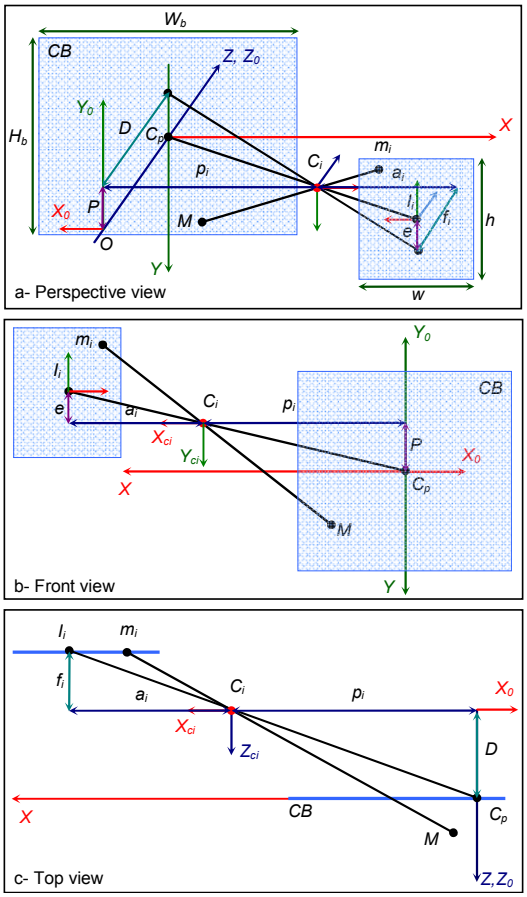

Fig. 2 Shooting geometry

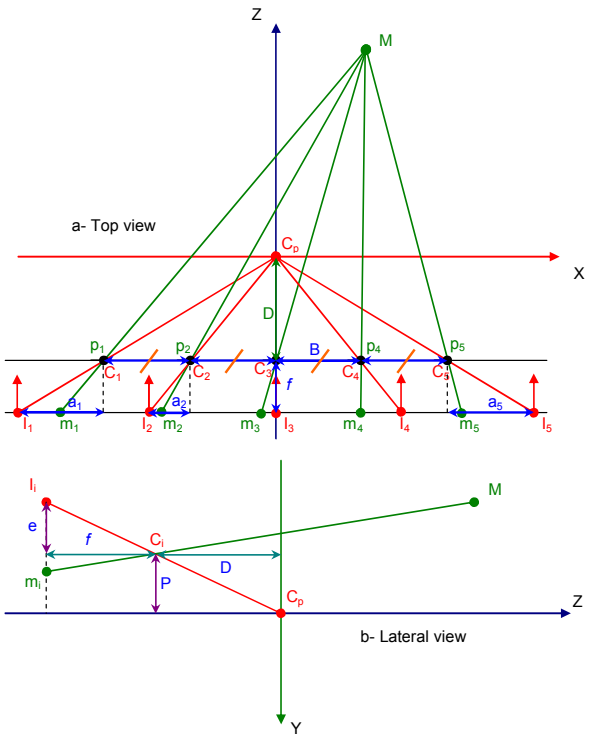

Fig. 3 Five viewpoints shooting system

$$\alpha\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} k & \mu & \gamma & 0 \\ & \rho\mu & \delta & 0 \\ & & 1 & 0 \\ 0 & 0 & k(\varepsilon-1)/d & \varepsilon \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

Where the transformation parameters quantifying independent distortion effects are defined as follows: $k = d / D$ is the global enlargement factor, $\varepsilon = (b / B)(W_b / W)$ controls the nonlinearity of the depth distortion according to the global reduction rate $\alpha = \varepsilon + k(\varepsilon - 1)Z / d$, $\mu = b / kB$ controls the relative enlargement width/depth rate, $\rho = (W_b / H_b)(H / W)$ controls the relative enlargement height/width rate. $\gamma = (p_i b - o_i B) / (d\ B)$ controls the horizontal clipping rate and $\delta = (\delta^o B - P\, b\, \rho) / (d\ B)$ controls the vertical clipping rate.

D.Specification of Multi-View Shooting Layout

Knowing the viewing, capturing and distortion parameters presented previously, one can specify a capturing layout satisfying the transformations and taking into account both the parameters imposed by the display device (Fig.1) and the parameters of the desired distortion effects *k, ε, μ, ρ, γ* and *δ*. Then, the geometrical parameters of the specified capture layout are pulled and expressed as presented in Table1.

TABLE 1
THE SHOOTING PARAMETERS

Width and height of the scene plane:	$W_b = \frac{W\varepsilon}{k\mu},\ H_b = \frac{H\varepsilon}{k\rho\mu}$
Convergence distance:	$D = \frac{d}{k}$
Width and height of the image sensors:	$w = \frac{W_b f}{D} = \frac{W f \varepsilon}{\mu d}, h = \frac{H_b f}{D} = \frac{H f \varepsilon}{\mu \rho d}$
Vertical position of the lenses line:	$P = \frac{\delta^o - \delta d}{k\rho\mu}$
Lateral position of the i^{th} optical centre:	$p_i = \frac{o_i + \gamma d}{k\mu}$
Lateral decentring of the ith image sensor:	$a_i = \frac{p_i f}{D} = \frac{f(o_i + \gamma d)}{\mu d}$
Vertical decentring of the image sensors:	$e = \frac{P f}{D} = \frac{f(\delta^o - \delta d)}{\mu\rho d}$
Focal length of each sensor-lens pair:	$f = \frac{D F}{D + F}$

The last relation in Table 1 is pulled from the well known Descartes relation: $1/f = 1/D + 1/F$, where F is the lens focal.

Note that to obtain a perfect 3D rendering without distortions, it is sufficient to choose the distortion parameters as follows: $\varepsilon = 1$, $\mu = 1$, $\rho = 1$, $\gamma = 0$ and $\delta = 0$.

Based on this analysis some industrial applications such as 3D-CAM1 and 3D-CAM2 prototypes were developed in collaboration with our industrial partner 3DTV-Solutions Society (Fig.4). These prototypes are able to capture images of eight points of view simultaneously and which can be displayed, after interlacing, on an auto-stereoscopic screen in real-time. The aim of the present work is to develop a calibration method for this type of cameras.

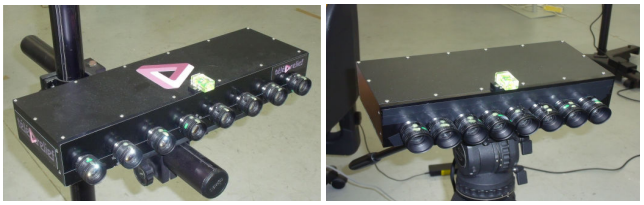

Fig. 4 3D-CAM1 and 3D-CAM2 prototypes

III. PERSPECTIVE PROJECTION MODEL

The geometric model of this multi-view shooting system is characterized by vertical and lateral decentering of image sensors regarding to their respective optical centers. In addition, the system of shooting presents a rectified geometric configuration where the optical axes of the different points of view are parallel and the horizontal axes x_{Ii} are parallel to the line connecting the optical centers C_i. The perspective projection model consists in expressing the image coordinates (*u, v*) of a material point *M* in terms of its well known cartesian coordinates (*X, Y, Z*).

The matching of the image coordinates and the cartesian coordinates of the point *M* known in the pattern frame will serve to identify the intrinsic, structural and extrinsic parameters of the geometrical model of the camera. For that, a sufficient number of correspondences 2D-3D should be known. To make easier this matching we use a standard object as pattern.

The relationship between the cartesian coordinates expressed in the pattern frame and the image coordinates expressed in the image frame is described by a transformation pattern/image that can be decomposed into a transformation pattern/camera followed by a perspective projection and a transformation camera/image.

A. Transformation Pattern/Camera (i)

This transformation allows the passage from the pattern frame F_O to the frame F_{Ci} linked to the i^{th} point of view (i^{th} optical center). In other words, it expresses the coordinates of the point *M* in the frame F_{Ci}. It is defined by three rotations and three translations:

$$^{Ci}P_M = {}^{Ci}T_O\ {}^{O}P_M \quad (2)$$

With:

$$^{Ci}T_O = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} {}^{Ci}R_O & {}^{Ci}P_O \\ 0 & 1 \end{bmatrix} \quad (3)$$

Where ${}^{Ci}R_0$ expresses the rotation matrix of the frame F_O regarding to the frame F_{Ci} and ${}^{Ci}P_O$ is the position vector of the origin *O* of the frame F_O regarding the frame F_{Ci} with:

$$\left[t_x,\ t_y,\ t_z\right]^t = \left[p_i,\ P,\ D\right]^t \tag{4}$$

B. Perspective Projection

It expresses the coordinates of the image point in the i^{th} optical center frame F_{Ci} using a perspective projection. In the absence of distortions, the point *M* is projected onto the image plane at a point *m* of homogeneous coordinates ${}^{Ci}P_m = ({}^{Ci}x_m, {}^{Ci}y_m, {}^{Ci}z_m, 1)$. Using Thales' theorem, one can write from (Fig. 3 and 4) that:

$$\frac{{}^{Ci}z_m}{{}^{Ci}z_M} = \frac{{}^{Ci}x_m}{{}^{Ci}x_M} = \frac{{}^{Ci}y_m}{{}^{Ci}y_M} \tag{5}$$

Knowing that ${}^{Ci}z_m = -f$ and using (5) it yields:

$$\begin{cases} {}^{Ci}x_m = -f\ {}^{Ci}x_M / {}^{Ci}z_M \\ {}^{Ci}y_m = -f\ {}^{Ci}y_M / {}^{Ci}z_M \\ {}^{Ci}z_m = -f \end{cases} \tag{6}$$

This can be written in matrix form using the homogeneous coordinates as follows:

$$\begin{bmatrix} s\,{}^{Ci}x_m \\ s\,{}^{Ci}y_m \\ s\,{}^{Ci}z_m \\ s \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & \frac{1}{f} & 0 \end{bmatrix} \begin{bmatrix} {}^{Ci}x_M \\ {}^{Ci}y_M \\ {}^{Ci}z_M \\ 1 \end{bmatrix} \tag{7}$$

And the cartesian coordinates of *m* are: $s\,{}^{Ci}x_m / s$, $s\,{}^{Ci}y_m / s$ and $s\,{}^{Ci}z_m / s$. Let us note this matrix as $T_{Perspective}$.

C. Transformation Camera(i)/Image

It consists of two transformations: a transformation camera(i)/sensor(i) followed by a transformation sensor/image.

Transformation Camera(i)/Sensor(i)

It allows the passage from the frame F_{Ci} to the frame F_{Si} whose origin is I_i to express the coordinates of the image point *m* in the sensor frame F_{Si} :

$${}^{Si}P_m = {}^{Si}T_{Ci}\ {}^{Ci}P_m \tag{8}$$

$${}^{Si}T_{Ci} = \begin{bmatrix} {}^{Si}R_{Ci} & {}^{Si}P_{Ci} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & a_i \\ 0 & -1 & 0 & -e \\ 0 & 0 & 1 & f \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{9}$$

Transformation Sensor/Image

It allows the passage from the sensor frame F_{Si} to the image frame $F_{\mathrm{Im}i}$ to express the image point coordinates in pixels. This transformation is defined by:

$${}^{\mathrm{Im}i}P_m = {}^{\mathrm{Im}i}T_{Si}\ {}^{Si}P_m \tag{10}$$

The element relative to the axis perpendicular to the image plane being always zero, we can ignore the third line of ${}^{\mathrm{Im}i}T_{Si}$ and write this transformation as follows [18]:

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} 1/l_u & 0 & 0 & u_{i0} \\ 0 & -1/l_v & 0 & v_{i0} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} {}^{Si}x_m \\ {}^{Si}y_m \\ {}^{Si}z_m \\ 1 \end{bmatrix} \tag{11}$$

Where (u_{i0}, v_{i0}) are the coordinates of the principal point I_i and l_u, l_v denote the pixel dimensions according to the directions *u* and *v* respectively.

D. The Global Transformation Pattern/Image

The global transformation pattern/image is obtained by the multiplication of all these transformations:

$${}^{\mathrm{Im}i}T_O = {}^{\mathrm{Im}i}T_{Si} \left[{}^{Si}T_{Ci}\ T_{\mathrm{Percpective}} \right] {}^{Ci}T_O \tag{12}$$

This transformation consists of three parts involving three categories of parameters: The first transformation involves the intrinsic parameters of the image sensors. The second transformation involves the structural parameters defining de location of the image sensors regarding to their corresponding optical centers. The third transformation involves the extrinsic parameters defining the pose of the i^{th} camera regarding to the pattern frame including the position of each optical centre in the shooting system.

By multiplying all the terms by *f* (this do not changes the result since the homogeneous coordinates are defined within about a multiplicative factor) the global transformation pattern/image is given as in relation (13) where $k_u = \frac{1}{l_u}$ and $k_v = \frac{1}{l_v}$:

Where:

$$\begin{bmatrix} su_i \\ sv_i \\ s \end{bmatrix} = \begin{bmatrix} f k_u r_{11} + (k_u a_i + u_{i0}) r_{31} & f k_u r_{12} + (k_u a_i + u_{i0}) r_{32} & f k_u r_{13} + (k_u a_i + u_{i0}) r_{33} & f k_u t_x + (k_u a_i + u_{i0}) t_z \\ -f k_v r_{21} + (k_v e + v_{i0}) r_{31} & -f k_v r_{22} + (k_v e + v_{i0}) r_{32} & -f k_v r_{23} + (k_v e + v_{i0}) r_{33} & -f k_v t_y + (k_v e + v_{i0}) t_z \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} {}^0x_M \\ {}^0y_M \\ {}^0z_M \\ 1 \end{bmatrix} \tag{13}$$

IV. The Calibration Method

The calibration problem amounts to determining the shooting system parameters i.e., the parameters of the transformation (13). For a multi-view shooting system as the one considered here, determining the intrinsic an extrinsic parameters is performed in two main steps: the first one is to determine the intrinsic and extrinsic parameters of each point of view, the second one amounts to determining the pose of each camera regarding to another.

A. Calibration of a Point of View

A procedure adopted in several works including [18] consists in the determination of the elements of a matrix M identifying the transformation (13). For that we perform a sufficient number of point matching 2D-3D. The matrix M is defined as follows:

$$\begin{bmatrix} su_i \\ sv_i \\ s \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} {}^0x_M \\ {}^0y_M \\ {}^0z_M \\ 1 \end{bmatrix} \tag{14}$$

That can be written under a condensed form:

$$M = \begin{bmatrix} m_1 & m_{14} \\ m_2 & m_{24} \\ m_3 & m_{34} \end{bmatrix} \tag{15}$$

Where m_1, m_2 and m_3 denote respectively the first three elements of the three rows of the matrix M. The transformation (13) also can be written under a condensed form as:

$$T = \begin{bmatrix} f k_u r_1 + (k_u a_i + u_{i0}) r_3 & f k_u t_x + (k_u a_i + u_{i0}) t_z \\ -f k_v r_2 + (k_v e + v_{i0}) r_3 & -f k_v t_y + (k_v e + v_{i0}) t_z \\ r_3 & t_z \end{bmatrix} \tag{16}$$

By identifying M with the global transformation ${}^{Imi}T_O$, one obtains a set of relationships that allow calculating the shooting system's parameters depending on the elements of the matrix M, it yields:

$$\begin{cases} k_u a_i + u_{i0} = m_1 \cdot m_3 \\ k_v e + v_{i0} = m_2 \cdot m_3 \\ f k_u = \|m_1 \wedge m_3\| \\ f k_v = \|m_2 \wedge m_3\| \\ r_3 = m_3 \\ r_1 = \dfrac{1}{f k_u}\left(m_1 - (k_u a_i + u_{i0}) r_3\right) = \dfrac{1}{\|m_1 \wedge m_3\|}\left(m_1 - (m_1 \cdot m_3) m_3\right) \\ r_2 = \dfrac{1}{-f k_v}\left(m_2 - (k_v e + v_{i0}) r_3\right) = \dfrac{1}{-\|m_2 \wedge m_3\|}\left(m_2 - (m_2 \cdot m_3) m_3\right) \\ t_z = m_{34} \\ t_x = \dfrac{1}{f k_u}\left(m_{14} - (k_u a_i + u_{i0}) t_z\right) = \dfrac{1}{\|m_1 \wedge m_3\|}\left(m_{14} - (m_1 \cdot m_3) m_{34}\right) \\ t_y = \dfrac{1}{-f k_v}\left(m_{24} - (k_v e + v_{i0}) t_z\right) = \dfrac{1}{-\|m_2 \wedge m_3\|}\left(m_{24} - (m_2 \cdot m_3) m_{34}\right) \end{cases} \tag{17}$$

These relationships are not sufficient to calculate the intrinsic and structural parameters of the camera. This is not the case for single cameras since the number of parameters to be identified is not high. Therefore, the parameters f, a_i and e can be determined by using relation (4) and certain relationships characterizing the multi-view shooting system:

$$f = \frac{D\,F}{D+F};\ a_i = \frac{p_i\, f}{D};\ e = \frac{P\, f}{D} \tag{18}$$

B. Determination of the Elements of the Matrix M

The transformation (14) can be written under linear equations set form by dividing the first and the second line on the third line of (14), it yields:

$$\begin{cases} u_i = \dfrac{m_{11}\,{}^0x_M + m_{12}\,{}^0y_M + m_{13}\,{}^0z_M + m_{14}}{m_{31}\,{}^0x_M + m_{32}\,{}^0y_M + m_{33}\,{}^0z_M + m_{34}} \\ v_i = \dfrac{m_{21}\,{}^0x_M + m_{22}\,{}^0y_M + m_{23}\,{}^0z_M + m_{24}}{m_{31}\,{}^0x_M + m_{32}\,{}^0y_M + m_{33}\,{}^0z_M + m_{34}} \end{cases} \tag{19}$$

Each point (x_i, y_i, z_i) projected in (u_i, v_i) provides two linear equations according to the elements of the matrix M. To determine the 12 elements of the matrix M, 6 points at least are needed. It yields:

$$\begin{cases} {}^0x_M m_{11} + {}^0y_M m_{12} + {}^0z_M m_{13} + m_{14} - u_i\,{}^0x_M m_{31} - u_i\,{}^0y_M m_{32} - u_i\,{}^0z_M m_{33} = u_i m_{34} \\ {}^0x_M m_{21} + {}^0y_M m_{22} + {}^0z_M m_{23} + m_{24} - v_i\,{}^0x_M m_{31} - v_i\,{}^0y_M m_{32} - v_i\,{}^0z_M m_{33} = v_i m_{34} \end{cases} \tag{20}$$

For n points, $2n$ equations are obtained and can be put in a matrix form:

$$A_{2n\times 11}\ x_{11} = b_{2n} \tag{21}$$

To deduce a non trivial solution for such homogeneous system, an element m_{ij} should be fixed. It is possible to choose $m_{34} = 1$ as a constraint, which amounts to divide all the matrix M elements by m_{34}. The parameters of the camera model are then determined to within about a multiplicative factor, namely $t_z = m_{34}$.

Faugéras and Toscani proposed a method for identifying the parameters of the matrix M using the constraint $\|m_3\| = 1$ [16], [17]. Indeed, one can see in (17) that $m_3 = r_3$. It is then possible to verify that $r_{31}^2 + r_{32}^2 + r_{33}^2 = 1$, hence we obtain:

$$\|m_3\|^2 = m_{31}^2 + m_{32}^2 + m_{33}^2 = 1 \tag{22}$$

The equation (21) can be written as follows:

$$B_{2n\times 9}\ x_9 + C_{2n\times 3}\ x_3 = 0 \tag{23}$$

With:

$$B_{2n\times 9}=\begin{bmatrix} & & & & \vdots & & & & \\ x_i & y_i & z_i & 1 & 0 & 0 & 0 & 0 & -u_i \\ 0 & 0 & 0 & 0 & x_i & y_i & z_i & 1 & -v_i \\ & & & & \vdots & & & & \end{bmatrix}$$

$$C_{2n\times 3}=\begin{bmatrix} & \vdots & \\ -u_i x_i & -u_i y_i & -u_i z_i \\ -v_i x_i & -v_i y_i & -v_i z_i \\ & \vdots & \end{bmatrix}$$

$$x_9=\begin{bmatrix} m_1 & m_{14} & m_2 & m_{24} & m_{34} \end{bmatrix}^t,\ x_3=\begin{bmatrix} m_3 \end{bmatrix}^t$$

This system can be solved by considering it as a least-square minimization problem. Indeed, the equation (23) can be written in the following form:

$$B_{2n\times 9}\ x_9 + C_{2n\times 3}\ x_3 = e \tag{24}$$

e represents an error vector. Then, the criterion to minimize is the following:

$$Q=\left\| B_{2n\times 9}\ x_9 + C_{2n\times 3}\ x_3 \right\|^2 \tag{32}$$

With the constraint:

$$\left\| x_3 \right\|^2 = \left\| m_3 \right\|^2 = 1 \tag{25}$$

The criterion can be written as follows:

$$Q=\left\| B\ x_9 + C\ x_3 \right\|^2 + \lambda\left(1-\left\| x_3 \right\|^2\right) \tag{26}$$

After some development we can get:

$$Q = x_9^t B^t B\, x_9 + x_3^t C^t C\, x_3 + x_9^t B^t C\, x_3 + x_3^t C^t B x_9 + \lambda\left(1 - x_3^t x_3\right) \tag{27}$$

By differentiating with respect to x_3 and x_9, and by requiring that the partial derivatives with respect to x_3 and x_9 are zero, we obtain the two following equations:

$$\begin{cases} B^t B\, x_9 + B^t C\, x_3 = 0 \\ C^t C\, x_3 + C^t B\, x_9 - \lambda\, x_3 = 0 \end{cases} \tag{28}$$

Hence, one has:

$$x_9 = -\left(B^t B\right)^{-1} B^t C\, x_3 \tag{29}$$

$$D x_3 = \lambda\, x_3 \tag{30}$$

$$D = C^t C - C^t B\left(B^t B\right)^{-1} B^t C \tag{31}$$

By substituting in (27) one obtains:

$$Q = x_3^t D x_3 = \lambda x_3^t x_3 = \lambda \tag{32}$$

One can remark that the matrix *D* is symmetric and positive 3 × 3 and its eigenvalues are real and positive. Equation (30) requires x_3 to be an eigenvector of *D* associated with eigenvalue λ. Equation (25) requires that x_3 be normalized.

To minimize the criterion we follow this procedure:

- Calculate the eigenvalues of the matrix *D* and choose the smallest value.
- Calculate the corresponding eigenvector, namely x_3.
- Normalize x_3
- Calculate x_9

The sign of the matrix *M* is determined by choosing $m_{34} = t_z > 0$.

C. Determining the pose between two cameras

Since the considered configuration is based on a rectified geometry, the transformation between two cameras *i* and *j* is expressed as follows:

$${}^{Cj}T_{Ci}=\begin{bmatrix} 1 & 0 & 0 & p_j - p_i \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{33}$$

The sought parameter here is the inter-optical distance separating two adjacent points of view *i* and *j* corresponding to the translation along x_{Ci} or x_{Cj} axis:

$$B_{ij} = p_j - p_i \tag{34}$$

V. Simulation Results

To validate the proposed calibration method we use a cube with chess board as faces as pattern in simulation. To simulate the multi-view camera we use the theoretical projection model presented above with theoretical values for the parameters to compute the image coordinates of the pattern's points. The interest is to determine the parameters of the identification model and to evaluate the precision of the used calibration method developed previously. Fig. 5 presents the cube used as pattern. The eight points of view images and the interlaced 3D image are depicted in Fig. 6. The 3D image can be visualized on an auto-stereoscopic screen.

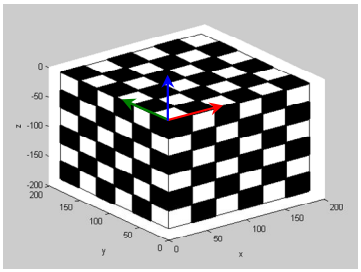

Fig. 5 The Pattern (cube with chess board as faces)

The estimated parameters for the different points of view are summarized in the Table II.

The obtained results show a slight error on the focal length f and the scale factors k_u and k_v. A maximal error of 16 pixels is observed on the coordinates of the principal points of the image sensors u_0 et v_0. There is also a small error of a few tens of microns on the lateral and vertical decentering respectively a_i and e_i. The error is almost zero on the extrinsic parameters of the camera. The results are satisfactory and the proposed method is applicable and effective for the considered shooting system.

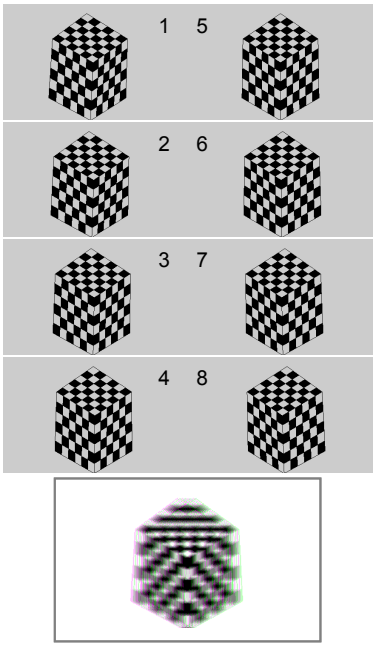

Fig. 6 The eight points of view and the 3D interlaced image

TABLE II
ESTIMATED PARAMETERS AND ESTIMATION ERRORS

	Theoritical model	Cam1	Cam2	Cam3	Cam4	Cam5	Cam6	Cam7	Cam8
f	16.0716	16.0735	= [1]	=	=	=	=	=	=
$\tilde{f}$		0.0019	=	=	=	=	=	=	=
k_u	430.4381	430.3886	=	=	=	=	=	=	=
$\tilde{k}_u$		0.0495	=	=	=	=	=	=	=
k_v	413.1571	413.1095	=	=	=	=	=	=	=
$\tilde{k}_v$		0.0476	=	=	=	=	=	=	=
u_0 (pixel)	640	623.4819	627.6114	631.7409	635.8705	640.0	644.1295	648.2591	652.3886
$\tilde{u}_0$ (pixel)		16.5181	12.3886	8.2591	4.1295	0.0	-4.1295	-8.2591	-12.3886
v_0 (pixel)	384	374.8529	=	=	=	=	=	=	=
$\tilde{v}_0$ (pixel)		9.1471	=	=	=	=	=	=	=
a_{model}	-	1.4924	1.1193	0.7462	0.3731	0	-0.3731	-0.7462	-1.1193
$a_{estimated}$		1.5309	1.1482	0.7655	0.3827	$-6.8\ 10^{-9}$	-0.3827	-0.7655	-1.1482
$\tilde{a}_i$		-0.0385	-0.0289	-0.0193	-0.0096	$6.8\ 10^{-9}$	0.0096	0.0193	0.0289
e	0.8610	0.8832	=	=	=	=	=	=	=
$\tilde{e}$		0.0222	=	=	=	=	=	=	=
$t_x(p_{i\_model}$	-	333.3571	250.0179	166.6786	83.3393	0	-83.3393	-166.6786	-250.0179
$p_{i\_estimated})$		333.3571	250.0179	166.6786	83.3393	$-1.49\ 10^{-6}$	-83.3393	-166.6786	-250.0179
$\tilde{p}_i$		0	0	0	0	$1.49\ 10^{-6}$	0	0	0
t_y (P)	192.3214	192.3214	=	=	=	=	=	=	=
$\tilde{P}$		0	=	=	=	=	=	=	=
t_z (D)	3500	3500	=	=	=	=	=	=	=
$\tilde{D}$		0	=	=	=	=	=	=	=
r_{11}	0.7071	0.7071	=	=	=	=	=	=	=
r_{12}	-0.7071	-0.7071	=	=	=	=	=	=	=
r_{13}	0	0	=	=	=	=	=	=	=
r_{21}	0.3536	0.3536	=	=	=	=	=	=	=
r_{22}	0.3536	0.3536	=	=	=	=	=	=	=
r_{23}	0.8660	0.8660	=	=	=	=	=	=	=
r_{31}	-0.6124	-0.6124	=	=	=	=	=	=	=
r_{32}	-0.6124	-0.6124	=	=	=	=	=	=	=
r_{33}	0.5	0.5000	=	=	=	=	=	=	=
$B(p_j\text{-}p_i)$	83.3393	83.3393	83.3393	83.3393	83.3393	83.3393	83.3393	83.3393	

[1] = means the same value as the left cell.

VI. Conclusion

In this paper a calibration method for a particular type of multi-view camera presenting parallel and decentered image sensors is proposed. At first, the shooting system is presented and an appropriate perspective projection model is derived. This model was used to simulate the production of 3D images for auto-stereoscopic display. Then, a calibration method for this type of cameras was developed. It constitutes an extension of the approach adopted by Faugéras and Toscani in 1986 and 1987. The obtained simulation results are satisfactory and the method is affectively applicable for the calibration of such cameras. It should be noted that other more recent methods could be extended to the case of the shooting system considered in this paper, we cite in particular the methods reported in [19], [20]. The model considered assumes a rectified geometry, however, a small rotation angle of image sensors around the axis perpendicular to their image plane is often found in prototypes developed in collaboration with our industrial partner 3DTV-Solutions. Also, it would be interesting to consider this parameter.

References

[1] W. Sanders, D. F. McAllister, "Producing anaglyphs from synthetic images," in *2003 Stereoscopic Displays and Virtual Reality Systems X, vol. 5006 of Proceedings of SPIE*, pp. 348–358.

[2] E. Dubois, "A projection method to generate anaglyph stereo images," in *Proc. 2001 IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP '01),* vol. 3, pp. 1661–1664,

[3] R. Blach, M. Bues, J. Hochstrate, J. Springer, B. Fröhlich, "Experiences with multi-viewer stereo displays based on lc-shutters and polarization," in *Proceedings of 2005 IEEE VR Workshop: Emerging Display Technologies*.

[4] L. M. J. Meesters, W. A. IJsselsteijn, P. J. H. Seuntiëns, "A Survey of Perceptual Evaluations and Requirements of Three-Dimensional TV," *IEEE Trans. on Circuits and Systems for Video Technology,* vol. 14, no. 3, pp. 381 – 391, 2004.

[5] K. Perlin, S. Paxia, J. S. Kollin, "An autostereoscopic display," in *Proceedings of the 27th ACM Annual Conference on Computer Graphics (SIGGRAPH '00)*, vol. 33, pp. 319–326, 2000.

[6] N. A. Dodgson, "Analysis of the viewing zone of multi-view autostereoscopic displays," in *2002 Stereoscopic Displays and Virtual Reality Systems IX, vol. 4660 of Proceedings of SPIE,* pp. 254–265.

[7] U. Güdükbay, T. Yilmaz, "Stereoscopic view-dependent visualization of terrain height fields," *IEEE Trans. on Visualization and Computer Graphics,* vol. 8, no. 4, pp. 330–345, 2002.

[8] K. Müller, A. Smolic, K. Dix, P. Merkle, P. Kauff, T. Wiegand, "View synthesis for advanced 3D video systems," *Hindawi Publishing Corporation, EURASIP Journal on Image and Video Processing,* 2008.

[9] J. Graham, L. Delman, H. Nicolas, E. David, "Controlling perceived depth in stereoscopic images," in *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems VIII*, 4297:42-53, June, 2001.

[10] J. Prévoteau, S. Chalençon-Piotin, D. Debons, L. Lucas, Y. Remion, "Multi-view shooting geometry for multiscopic rendering with controlled distortion," *Hindawi, Int. Journal of Digital Multimedia Broadcasting (IJDMB), Advances in 3DTV: Theory and Practice,* 2010.

[11] M. Ali-Bey, N. Manamanni, S. Moughamir, "Dynamic Adaptation of Multi View Camera Structure," in *Proc. 2010 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON'10),* Tampere, Finland.

[12] M. Ali-Bey, S. Moughamir, N. Manamanni, "Towards Structurally Adaptive Multi-View Shooting System," in *2010 18th Mediterranean Conference on Control and Automation (MED),* Marrakesh, Morocco.

[13] M. Ali-Bey, S. Moughamir, N. Manamanni, "Auto-stereoscopic rendering quality assessment depending on poisioning accuracy of image sensors in multi-view capture systems," in *6th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications VISIGRAPP-IMAGAPP,* 5-7 March, 2011, Vilamoura, Algarve, Portugal.

[14] A.A. Alatan, and al, "Scene Representation Technologies for 3DTV—A Survey," *IEEE Trans. on Circuits and Sys. Video Tech.,* Vol. 17, NO. 11, pp: 1587 – 1605, Nov. 2007.

[15] E. Stoykova, and al, "3-D Time-Varying Scene Capture Technologies—A Survey*," IEEE Trans. on Circuit. Sys. Video Tech.,* Vol. 17, NO. 11, pp:1568 – 1586, Nov. 2007

[16] O. Faugéras, G. Toscani, "The calibration problem for stereo," in *Proceedings of the 1986 International IEEE Conference on Computer Vision and Pattern Recognition,* pages 15–20, Miami Beach, FL.

[17] O. Faugéras, G. Toscani, "Camera calibration for 3D computer vision," in *1987 International Workshop on Machine Vision and Machine Intelligence,* pp. 240-247). Tokyo. Japan. February.

[18] R. Horaud, O. Manga, "Vision par ordinateur : outils fondamentaux," Editeur Hermès - Lavoisier, 2nd Edition, 1995.

[19] Z. Zhang, "A Flexible New Technique for Camera Calibration," Technical Report MSR-TR-98-71, Microsoft Research, 1998.

[20] P. F. Sturm, S. J. Maybank, "On Plane-Based Camera Calibration: A General Algorithm, Singularities, Applications," in *1999 IEEE Conference on Computer Vision and Pattern Recognition*, Fort Collins, USA, p. 432-437.