Research Article

Multiview Shooting Geometry for Multiscopic Rendering with Controlled Distortion

J. Prévotau,^{1,2} S. Chalençon-Piotin,¹ D. Debons,² L. Lucas,^{1,2} and Y. Remion^{1,2}

¹ CReSTIC-SIC EA3804, Université de Reims Champagne-Ardenne, IUT de Reims-Châlons-Charleville, rue des Crayères, BP 1035, 51687 Cedex 2, France

² TéléRelief, Pépinière d'Entreprises Henri Farman, 2 Allée Albert Caquot, 51100, France

Correspondence should be addressed to J. Prévotau, jessica.prevotau@etudiant.univ-reims.fr

Received 30 April 2009; Revised 4 October 2009; Accepted 8 December 2009

Academic Editor: Georgios Triantafyllidis

Copyright © 2010 J. Prévotau et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A fundamental element of stereoscopic and/or autostereoscopic image production is the geometrical analysis of the shooting and viewing conditions in order to obtain a qualitative 3D perception experience. This paper firstly compares the perceived depth with the shot scene depth from the viewing and shooting geometries for a couple of shooting and rendering devices. This yields a depth distortion model whose parameters are expressed from the geometrical characteristics of shooting and rendering devices. Secondly, these expressions are inverted in order to design convenient shooting layouts yielding chosen distortions on specific rendering devices. Thirdly, this design scheme provides three shooting technologies (3D computer graphics software, photo rail, and camera box system) producing qualitative 3D content for various kinds of scenes (real or virtual, still or animated), complying with any prechosen distortion when rendered on any specific multiscopic technology or device formerly specified.

1. Introduction

The three-dimensional display is on its way to becoming the next evolution in the image industry. These 3D displays hold tremendous potential for many applications in entertainment, information presentation, reconnaissance, telepresence, medicine, visualization, remote manipulation, and art. Both research and business in multiscopic display are increasing. We dispose now of numerous multiscopic rendering systems with or without glasses. Different technologies support all these systems; stereoscopy with colorimetric or temporal mixing such as anaglyph [1, 2], occultation and polarization [3], for example for projections with glasses as in some movie theaters; autostereoscopy [4, 5] such as parallax barrier and lenticular lens; for example, for 3D advertising billboards, autostereoscopic displays or lenticular printing.

As shown in Figure 1, the different rendering modes and kinds of scenes to be shot are well-known, but all these systems need content; up to now, there is no 3D shooting system specifically designed to acquire a qualitative 3D content. Some works [6–9] consist in calculating a right eye

image from a left eye image and so obtain a 3D content from a 2D-plus-depth. The main disadvantage of these methods lies in the lack of information in occluded areas which is impossible to overcome in a generic way. Yet, to comply with our demand of qualitative 3D content we focus on multi-shooting technologies. Other works [10, 11] define the projective relations between the images shot by multi-cameras in order to calibrate the different cameras and then to reconstruct the 3D shot scene from these multiple views. There is no link with any viewing device since the target is a reconstruction module. In our case, flat multiscopic viewing requires a simplified shooting layout also called “rectified geometry”.

Moreover, some works have been done to improve the control of the viewer's 3D experience in stereoscopy and computer graphics fields [12, 13]. They usually compare shooting and viewing geometries in order to choose a shooting layout fitting a given depth range in virtual space to the “comfortable” depth range of the display. We believe that choices that can be made in the shooting design may be richer than a simple mapping of depth and could differ for each observation position in the multi-view case. This

FIGURE 1: Shooting to rendering process.

requires a detailed model and a precise analysis of possible distortions for multiscopic shooting/viewing couple. Indeed, such a model will provide the characteristics of shooting which will ensure the chosen distortions on the chosen viewing device. If some authors have described the transformation between the shot and the real scene [12] in the stereoscopic case, none of them has been interested in producing an analytic multi-observer and reversible model allowing to pilot the shooting for all kind of possible distortions. Thus, we suggest a solution to produce 3D content according to the chosen rendering mode and the desired depth effect.

So we will explain how to model and quantify the depth distortion from given rendering and shooting geometries and also from a chosen rendering device and a desired depth effect, how to design the appropriate shooting layout.

This article introduces an complete analysis of the geometrical quality of 3D content based on distortion analysis by linking shooting and viewing geometries. Starting from the prior and related knowledge (i.e., viewing and shooting geometries), we will show remaining problems and model the possibilities of depth distortions between the scene perceived by a viewer and the initially shot scene. Next, we will present a shooting layout design scheme ensuring a desired depth effect (controlled depth distortion or perfect depth effect) upon a pre-determined rendering device. Finally, we will introduce three shooting technologies (which are patent pending) complying to this scheme and,

thus, achieving qualitative 3D content on formerly given rendering device: 3D computer graphics software, photo rail and camera box systems. We will present these prototypes and some of their results.

2. Prior and Related Knowledge

2.1. Viewing. 3D image rendering, with or without glasses, is known to require “stereoscopic” or “autostereoscopic” devices. All these devices make a spatial, colorimetric and/or temporal mixing over a single region of interest (ROI area physically filled by the displayed image on the rendering device) of $n \times m$ so-called “initial images” of one scene shot from several distinct viewpoints. These systems allow to optically and/or temporally separate the images reaching each eye of one or more viewers. In case of stereoscopic systems, both images are emitted in a single optical beam independently of the viewer’s position in this beam [1, 2, 14]. However, autostereoscopic systems separate the images in several distinct optical beams, organized for example, in horizontal “range” of n images ($n \geq 2$ and $m = 1$) [4, 5]. We can also imagine optical beams organized in both horizontal and vertical ranges. Then we dispose of a matrix disposition of $n \times m$ optical beams ($n \geq 2$ and $m \geq 2$), each one transporting a different image. Thus, all known devices broadcast alternately and/or simultaneously $n \times m$ images ($n \geq 2$ and $m \geq 1$) within one or several optical beams in such a way that both eyes of a correctly-placed viewer

FIGURE 2: The viewing geometry.

get different consistent images (i.e., initial images and not combinations of them). Thereby the viewer's brain rebuilds his depth perception by stereoscopy [15]. Even if the human visual system has a tolerance as for epipolar alignment, ideal positions within this tolerance correspond in particular to eyes line parallel to the display's rows. Despite this human tolerance, we calculate our images in such a way that they have a perfect epipolar alignment for well-placed eyes line.

So let's analyse the geometry of these devices (the "viewing geometry" (Figure 2) which will constrain the compatible shooting layout.

A 3D rendering device mixes $n \times m$ images sharing out a ROI of dimension W (width) and H (height). Each image (image's index $i = (i_1, i_2) \in \{1, n\} \times \{1, m\}$) is supposed to be "correctly" visible (without much mixing with others) at least from the chosen preferential position E_i . These positions are aligned upon m lines parallel to the rows of the ROI located at distance d_{i_2} , $i_2 \in \{1, \dots, m\}$, from the device ROI. The preferential positions E_i are placed on those lines according to their second index i_2 in order to guarantee that a viewer whose binocular gap is b_{i_2} (often identical to human medium binocular gap 65 mm, but possibly different according to the expected public: children, etc.), with eyes line parallel to the device rows, would have his right eye in E_i and his left eye in E_{l_i} . The right eye in E_i would catch image number i , while the left eye in E_{l_i} would catch image number l_i knowing that $l_i = i - (q_{i_2}, 0)$ with q_{i_2} being the way gap between the indexes of images composing the visible

FIGURE 3: Implied shooting geometry.

consistent stereoscopic couples with binocular gap b_{i_2} at distance d_{i_2} . Hence, associated left and right eye preferential positions E_{l_i} and E_i verify $E_i = E_{l_i} + b_{i_2}x$ and $e_i = e_{l_i} + b_{i_2}$.

We also define the lines positions vertically (because viewers of various sizes use the device) by p_{i_2} which represents the "overhang", that is, vertical gap of eyes positioning compared with the ROI center CS. If we do not know

p_{i_2} , we use a medium overhang corresponding to a viewer of medium size, which has to be chosen at design stage. Assuming u_i and u_{li} are stereoscopic homologous for images i and l_i , their perception by the right and left eye of a viewer from E_i and E_{li} leads this spectator's brain to perceive a 3D point u . The viewing geometry analysis is expressed thanks to a global reference frame $r = (CS, x, y, z4x \times y)$, chosen at the ROI center CS , with x parallel to its rows and turned towards the right of the spectators, and y parallel to its columns and turned towards the bottom.

2.2. Shooting. In order to “feed” such devices with 3D content, we need sets of $n \times m$ images from a single scene acquired from several distinct and judicious viewpoints and with specific projective geometry as the rendering upon flat multiscopic devices involves coplanar mixing of these images. This major issue is well known in multiscopy.

The image viewing is achieved according to distorted pyramids whose common base corresponds to the device ROI and the tops are the viewer's eyes or E_i positions. Given that vision axes are not necessarily orthogonal to the observed images area (ROI), the viewing of these images induces trapezoidal distortion if we don't take into account this slanted viewing during the shooting. This has an immediate consequence in order to achieve depth perception. If the trapezoidal distortions are not similar for the two images seen by a spectator, the stereoscopic matching by the brain is more delicate, or even impossible. This reduces or cancels the depth perception. This constraint, well-known in stereoscopy is called the “epipolar constraint”.

Solutions (also called toe-in camera model) of convergent systems have been proposed [16, 17], but such convergent devices manifest the constraint presented above. So, unless a systematic trapezoidal correction of images is performed beforehand (which might not be desirable as it loads down the processing line and produces a qualitative deterioration of the images) such devices do not afford to produce a qualitative 3D content. As demonstrated by [18, 19], we must use devices with shooting pyramids sharing a common rectangular base (off-axis camera model) and with tops arranged on a line parallel to the rows of this common base in the scene. For example Dodgson et al. use this shooting layout for their time-multiplexed autostereoscopic camera system [20].

Thus, aiming axes are necessarily convergent at the center of the common base and the tops of the shooting pyramids must lie on m lines parallel to the rows of the common base. Figure 3(b) shows a perspective representation of such a shooting geometry. This figure defines the layout of the capture areas (CA_i), and the centers (C_i) and specifies a set of parameters describing the whole shooting geometry completely. Figures 3(c) and 3(d) show top and full-face representations of this geometry, respectively.

The shooting geometry analysis is expressed using a shooting global reference frame $R = (CP, X, Y, Z \equiv X \times Y)$ chosen centered at the desired convergence point CP (which is also the center of the common base CB of the scene) and oriented in such a way that the two first vectors of the reference frame are parallel to the main directions of the

common base CB of the scene and so, parallel to the main directions of the capture areas. The physical size of CB is Wb and Hb . Furthermore, the first axis is supposed to be parallel to the rows of the capture areas and the second axis is supposed to be parallel to the columns of these areas.

The $n \times m$ pyramids, representative of a shooting layout, according to the principles previously explained to resolve the known issue, are specified by

- an optical axis of Z direction,
- optical centers C_i (i.e., principal points) aligned on one or more (m) line(s) parallel to the rows of the common base (so on X direction),
- rectangular capture areas CA_i .

These capture areas must be orthogonal to Z , so parallel one to another and parallel to CB and to centers lines (which are defined by their distances from CB , D_{i_2} along Z , P_{i_2} along Y and c_i along X). These capture areas are also placed at distances f_i along Z , β_i along Y and α_i along X from their respective optical center C_i . Their physical size is given by w_i and h_i . They are centered on points I_i in such a way that lines $I_i C_i$ defining the axes of sight are convergent at CP . The centers C_i and C_{li} must be on same “centers line” and with a spacing of B_i ($C_i = C_{li} + B_i X$ and $c_i = c_{li} + B_i$).

Such a shooting layout is necessary to obtain a depth effect on multiscopic device. Nevertheless, it does not ensure that the perceived scene will not be distorted relative to the shot scene. Non distortion implies that viewing pyramids are perfect counterpart of shooting pyramids (i.e., have exactly same opening and main axis deviation angles in both horizontal and vertical directions). In case of pyramids dissimilarity, the 3D image corresponds to a complex distortion of the initially acquired scene. This can be desirable in some applications to carry out some special effects, as it can be undesirable in others. This requires, that shooting and viewing must be designed in a consistent way whether we desire a depth distortion or not. Let's now model those distortion effects implied by a couple of shooting and viewing geometries.

3. Distortion Analysis and Model

In this section, we consider that we use perfect sensors and lenses, without any distortion. This assumption implies some issues which will be presented for each derived technology.

Thanks to the two previous sections we can link the coordinates (X, Y, Z) , in the reference frame R , of the point U of the scene, shot by the sensors previously defined, with the coordinates (x_i, y_i, z_i) in the reference frame r , of its counterparts u seen by an observer of the viewing device placed in an preferential position.

Assuming that the scene point U is visible on image number i , its projection U_i verifies

$$C_i U_i = \frac{-f_i}{Z + D_{i_2}} C_i U \quad i \in \{1, n\} \times \{1, m\}. \quad (1)$$

Knowing that I_i , center of CA_i , verifies

$$C_i I_i = \frac{-f_i}{D_{i_2}} C_i C P \quad i \in \{1, n\} \times \{1, m\}, \quad (2)$$

the relative position of the scene point U 's projections in the various images are expressed as

$$I_i U_i = \frac{f_i}{Z + D_{i_2}} \begin{bmatrix} -X - Z \frac{c_i}{D_{i_2}} \\ -Y + Z \frac{p_{i_2}}{D_{i_2}} \\ 0 \end{bmatrix}_R \quad i \in \{1, n\} \times \{1, m\}. \quad (3)$$

As the images are captured behind the optical centers, the projection reverses up/down and left/right axes, and the implicit axes of the images are opposites of those of the global shooting reference frame R . Moreover, the images are then scaled on the whole ROI of the rendering device. This relates U_i projections of U to their "rendered positions" u_i on the ROI:

$$C S u_{i|_r} = - \begin{bmatrix} \frac{W}{w_i} \\ \frac{H}{h_i} \\ 1 \end{bmatrix} I_i u_{i|_R} \quad \forall i. \quad (4)$$

Remarking that $f_i W b = D_{i_2} w_i$ and $f_i H b = D_{i_2} h_i$, u_i is expressed in the reference frame r as

$$u_{i|_r} = \frac{D_{i_2}}{Z + D_{i_2}} \begin{bmatrix} \left(X + Z \frac{c_i}{D_{i_2}} \right) \frac{W}{W b} \\ \left(Y - Z \frac{p_{i_2}}{D_{i_2}} \right) \frac{H}{H b} \\ 0 \end{bmatrix} \quad \forall i. \quad (5)$$

By this time, and assuming U was visible on both images l_i and i , we notice that u_{l_i} and u_i lie on the same row of the ROI. This fulfills the epipolar constraint and thus permits stereoscopic reconstruction of $u = [x_i, y_i, z_i]_r^t$ from E_{l_i} and E_i according to

$$\begin{aligned} u_{l_i} u_i &= \frac{z_i}{z_i + d_{i_2}} b_{i_2} x_i, \quad \text{which yields } z_i \text{ and} \\ E_i u &= \frac{z_i + d_{i_2}}{d_{i_2}} E_i u_i, \quad \text{which then gives } x_i, y_i. \end{aligned} \quad (6)$$

Thus, after some calculus, the relation between the 3D coordinates of the scene points and those of their

images perceived by a viewer may be characterized under homogeneous coordinates by

$$a_i \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} = \begin{bmatrix} k_{i_2} \begin{vmatrix} \mu_i & \gamma_i \\ \rho \mu_i & \delta_i \\ 1 & 1 \\ 0 & 0 \end{vmatrix} & 0 \\ 0 & \frac{k_{i_2}(\epsilon_i - 1)}{d_{i_2}} & \epsilon_i \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}. \quad (7)$$

The above equation can be seen as the analytic distortion model for observer position i which matches the stereoscopic transformation matrix given in [12]. As such this model clearly exhibits the whole set of distortions to be expected in any multiscope 3D experience, whatever the number of views implied or the very nature of these images (real or virtual). It shows too that these distortions are somehow independent one from another and may vary for each observer position i . The following detailed analysis of this model and its further inversion will offer a novel multiscope shooting layout design scheme acting from freely chosen distortion effects and for any specified multiscope rendering device.

The above model exhibits some new parameters quantifying independent distortion effects. Those parameters may be analytically expressed from geometrical parameters of both shooting and rendering multiscope devices. Their relations to geometrical parameters and impact on distortion effects are now presented

– k_{i_2} control(s) the global enlarging factor(s),

$$k_{i_2} = \frac{d_{i_2}}{D_{i_2}}. \quad (8)$$

– ϵ_i control(s) the potential nonlinear distortion which transforms a cube into a pyramid trunk according to the global reducing rate $a_i = \epsilon_i + k_{i_2}(\epsilon_i - 1)(Z/d_{i_2})$ possibly varying along Z ,

$$\epsilon_i = \frac{b_{i_2}}{B_i} \frac{W_b}{W}. \quad (9)$$

– μ_i control(s) width over depth relative enlarging rate(s), or horizontal/depth anamorphose factor,

$$\mu_i = \frac{b_{i_2}}{k_{i_2} B_i}. \quad (10)$$

– ρ control(s) height over width relative enlarging rate(s), or vertical/horizontal anamorphose factor,

$$\rho = \frac{W_b}{H_b} \frac{H}{W}. \quad (11)$$

– γ_i control(s) the horizontal "shear" rate(s) of the perceived depth effect,

$$\gamma_i = \frac{c_i b_{i_2} - e_i B_i}{d_{i_2} B_i}. \quad (12)$$

- δ_i control(s) the vertical “shear” rate(s) of the perceived depth effect by an observer whose overhang complies to what is expected,

$$\delta_i = \frac{p_{i2}B_i - P_{i2}b_{i2}\rho}{d_{i2}B_i}. \quad (13)$$

Thus we have defined the depth distortion possibilities using the previously established shooting and viewing geometries. Moreover, this model makes possible to quantify those distortions for any couple of shooting and viewing settings by simple calculus based upon their geometric parameters.

4. Shooting Design Scheme for Chosen Distortion

One can use any multiscopic shooting device with any multiscopic viewing device while giving an effect of depth to any well-placed viewer (3D movie theater for example) but Section 3 shows that distortions will not be similar for each couple of technologies. In this section, we will design the shooting geometry needed to obtain, on a given viewing device, a desired distortion: whether perfect depth or chosen distortion effect of shot scene.

Knowing how distortions, shooting and viewing parameters are related, it becomes possible to derive shooting layout from former distortion and viewing choices.

We will describe two shooting layout design schemes complying to this use of the distortion model:

- a generic scheme allowing for a precise control of each distortion parameter,
- a more dedicated one of huge interest as it is focused on “non distortion” or “perfect depth”, allowing at most control of global enlarging factor(s) k_{i2} as any other distortion parameter is set to its “non distortion value”.

4.1. Controlled Depth Distortion. To define the shooting layout using this scheme, we control global enlargement (by k_{i2}) and 4 potential depth distortions:

- (1) when $\epsilon_i \neq 1$, a global nonlinearity which results in a deformation of the returned volume onto a “pyramid trunk” (as a_i varies along Z axis),
- (2) when $\gamma_i \neq 0$, a sliding or “horizontal shear” of the returned volume according to the depth,
- (3) when $\delta_i \neq 0$ and/or when the real overhang of the observer differs from the optimal overhang, a sliding or “vertical shear” of the returned volume according to the depth and
- (4) when $\mu_i \neq 1$ and/or $\rho \neq 1$, an anamorphose producing uneven distentions of the 3 axis (X versus Z for μ_i and Y versus X for ρ).

The depth controlled-distortion is obtained by adjusting the enlarging factor(s) k_{i2} and the distortion parameters ϵ_i

TABLE 1: Proposed solutions according to the main features of scenes to be shot.

	Photo		Video
Scene	Still	Animated	Animated
Real	3D-CAM	3D-CAM	3D-CAM
	Photo Rail		
Virtual	3D computer graphics software		

(and so $a_i = \epsilon_i * k_{i2}(\epsilon_i - 1)/d_{i2}$), μ_i , ρ , γ_i and δ_i . This last condition on δ_i is more delicate because it depends on the height of the viewer which inevitably affects the effective overhang towards the device. So the chosen vertical sliding δ_i can be reached only for observer whose overhang is defined in the viewing settings for this observation position.

Thus, given the viewing settings and the desired distortion parameters, the shooting parameters can be calculated as follows:

$$\begin{aligned} P_{i2} &= \frac{p_{i2} - \delta_i d_{i2}}{k_{i2}\rho\mu_i}, & D_{i2} &= \frac{d_{i2}}{k_{i2}}, \\ W_b &= \frac{W\epsilon_i}{k_{i2}\mu_i}, & H_b &= \frac{H\epsilon_i}{k_{i2}\rho\mu_i}, \\ c_i &= \frac{e_i + \gamma_i d_{i2}}{k_{i2}\mu_i} \end{aligned}$$

f_i imposed or chosen, individually $\forall i \in \{1 \dots n\} \times$

$\{1, \dots, m\}$, by lot $\forall i_2 \in \{1, \dots, n\}$ or on the whole : (14)

$$\begin{aligned} w_i &= \frac{W_b f_i}{D_{i2}} = \frac{W f_i \epsilon_i}{\mu_i d_{i2}}, \\ h_i &= \frac{H_b f_i}{D_{i2}} = \frac{H f_i \epsilon_i}{\mu_i \rho d_{i2}}, \\ \alpha_i &= \frac{c_i f_i}{D_{i2}} = \frac{f_i (e_i + \gamma_i d_{i2})}{\mu_i d_{i2}}, \\ \beta_i &= \frac{P_{i2} f_i}{D_{i2}} = \frac{f_i (p_{i2} - \delta_i d_{i2})}{\mu_i \rho d_{i2}}. \end{aligned}$$

This depth controlled-distortion scheme allows to obtain the parameters of a shooting layout producing desired 3D content for any rendering device and any depth distortions combination.

4.2. Perfect Depth Effect. A particular case of the depth controlled-distortion is a perfect depth effect (depth perception without any distortion compared with the depth of the shot scene). To produce a perfect depth effect (whatever the enlarging factor(s) k_{i2}), we should configure the shooting in order to avoid the 4 potential distortions. This is obtained by making sure that the distortion parameters verify $\epsilon_i = 1$, $\mu_i = 1$, $\rho = 1$, $\gamma_i = 0$ and $\delta_i = 0$. The last condition $\delta_i = 0$ is more delicate, as it can be assured only for an observer complying to the defined overhang.

FIGURE 4: Images synthesis of engine photo with courtesy of PI3C, partner of PTC.

In case of shooting for perfect depth effect, the shooting parameters can be calculated as below:

$$P_{i_2} = \frac{p_{i_2}}{k_{i_2}}, \quad D_{i_2} = \frac{d_{i_2}}{k_{i_2}},$$

$$W_b = \frac{W}{k_{i_2}}, \quad H_b = \frac{H}{k_{i_2}},$$

$$c_i = \frac{e_i}{k_{i_2}}$$

f_i imposed or chosen, individually $\forall i \in \{1 \dots n\} \times$

$\{1 \dots m\}$, by lot $\forall i_2 \in \{1 \dots n\}$ or on the whole. (15)

$$w_i = \frac{W_b f_i}{D_{i_2}} = \frac{W f_i}{d_{i_2}},$$

$$h_i = \frac{H_b f_i}{D_{i_2}} = \frac{H f_i}{d_{i_2}},$$

$$\alpha_i = \frac{c_i f_i}{D_{i_2}} = \frac{e_i f_i}{d_{i_2}},$$

$$\beta_i = \frac{P_{i_2} f_i}{D_{i_2}} = \frac{p_{i_2} f_i}{d_{i_2}}.$$

This particular case is very interesting for its realism, for example: in order to convince financiers or deciders, it may be important to give the real volumetric perception of a building, or a mechanical piece, in a computer aided design (CAD) application, or medical visualization software, in a surgical simulation application.

5. Derived Shooting Technologies

5.1. 3D Computer Graphics Software. Thanks to these design schemes, we have created 3 different technologies to shoot 3D scenes: multi-viewpoint computer graphics software, photo rail and camera box system. These products have been developed under the brand “3DT 5 Solutions” and patents are pending for each of them. As shown in Table 1, we have developed 3 solutions to obtain qualitative photo or video content for any relevant kind of scene still or animated, real or virtual). We use anaglyph to illustrate our results even if their viewing on paper or 2D screen media is not optimum because the images have been calculated to be rendered on specific devices.

The common goal of our 3D computer graphics pieces of software is to virtually shoot different scenes, as photo of still or animated scene as well as computer animations of animated scene. Thanks to the previous shooting design scheme, we are able to place the virtual sensors around a usual monocular camera according to the chosen viewing device in order to obtain the desired depth effect. In this case, virtual cameras are perfect and there is no issue with distortions due to sensors or lenses.

Thus we have, by now, developed plugins and software to visualize and handle in real-time files from CAD software such as AutoCAD, Archicad, Pro/Engineer, etc. as well as medical data, such as MRI. We are going to apply this technology to other virtual scenes.

In those software pieces, we choose the rendering device parameters and the desired depth effect, and the software computes and uses its virtual shooting layout. It is possible to record different rendering devices and depth effect

FIGURE 5: Images synthesis of image of an aneurysm.

distortions; then to switch easily between these devices and these distortions.

Those pieces of software currently handle scenes up to 7 million polygons at interactive rate.

Figure 4 shows an example of images of Parametric Technology Corporation (PTC) piece (a engine) shot as the software was tuned to achieve a perfect depth effect on an autostereoscopic parallax display 57'' (optimal viewing distance 4.2 m) (Note 1).

Figure 5 shows medical data (image of an aneurysm). The software was tuned for autostereoscopic parallax display 57'' (optimal viewing distance 4.2 m) and a perfect depth effect is obtained, as it is usual with medical data in order to allow the most definite and efficient interpretation possible by the surgeon (Note 1).

5.2. Photo Rail. The goal of this photo rail is to shoot a 3D photo of still scene. By using the photo rail (Figure 6) with its controlling software it is possible to control both the usual operations of a professional digital camera and its movement along a linear axis parallel to its sensor rows.

This allows us to carry out any shooting configuration whatever the chosen depth distortion settings and viewing device if we crop the needed capture area CA_i in each digital photo in order to comply to the needed shooting geometry. With this Photo Rail, there is a possibility of distortion due to the digital camera, but distortions will be consistent for all images and of negligible magnitude, as it is professional equipment. We have not tried to correct those possible distortions but such a work could be done easily.

FIGURE 6: Photo Rail

For example, Figure 7 illustrates the shooting of a room in "Palais du TAU" [21] in Reims with a perfect depth effect for autostereoscopic parallax display 57'' (optimal viewing distance 4.2 m). We made a 3D shooting of a big hall with a significant depth (Note 1). To realize this sequence, we positioned the convergence point at 4.2 m from the photo rail. Moving the digital camera along the rail, taking pictures and storing them took us 39.67 s.

5.3. Camera Box System. The goal of this camera box system is to shoot different scenes, as photo of still or animated scene as well as video of animated scene. Thanks to the previous shooting design method, we know how to realize a camera box system containing several couples of lenses and image sensors in order to produce simultaneously the multiple images required by an autostereoscopic display with a desired depth effect. As these couples are multiple, their induced

FIGURE 7: "Palais du Tau" result of Photo Rail with courtesy of Monum.

distortions can be different. We have introduced a couple-by-couple process of calibration/correction based upon the model of Zhang [22].

We have already produced two prototypes of camera box system delivering multi-video stream in real-time (25). Their layout parameters have been defined for no distortion of specific scenes (see below) and set at manufacturing:

- The 3D-CAM1 (Figure 8(a)) which allows to shoot a life size scene (ratio $k_i = 1$) of the bust of a person to be viewed on an autostereoscopic parallax display 57" (optimal viewing distance 4.2 m). We have chosen to obtain a perfect depth effect. According to numerous viewers both novice and expert, the 3D perception is really qualitative. Unfortunately, such an experiment is impossible to produce on 2D media (Note 1). Figure 9 shows some results. Figure 9(a) corresponds to a set of 8 simultaneously shot images, Figure 9(b) is the result after their autostereoscopic mixing and Figure 9(c) is anaglyph of images indexed 3 and 4.
- The 3D-CAM2 (Figure 8(b)) enables to shoot small size objects (in the order of 10–20 cm) and to display them on an autostereoscopic lenticular display 24" (optimal viewing distance 2,8) with an enlargement factor set to $k_i = 1,85$. We have chosen to obtain a perfect depth effect. Here again the 3D perception is really qualitative according to our numerous visitors. You will find some illustrations in Figure 10. Figure 10(a) corresponds to a set of 8 simultaneously shot images and Figure 10(b) gives the result in anaglyph of images indexed 3 and 4 (Note 1).

(a) 3D-CAM1

(b) 3D-CAM2

FIGURE 8: Camera box systems.

6. Conclusion

This work firstly models geometrical distortions between the shot scene and its multiscopically viewed avatar. These distortions are related to geometrical parameters of both the shooting and rendering devices or systems. This model enables quantitative objective assessments on the geometrical reliability of any multiscopic shooting and rendering couple.

The formulas expressing distortion parameters from geometrical characteristics of the shooting and rendering devices have afterwards been inverted in order to express the desirable shooting layout yielding a chosen distortion scheme upon a chosen rendering device. This design scheme *a priori* insures that the 3D experience will meet the chosen requirements for each expected observer position. Such a scheme may prove highly valuable for applications needing reliable accurate 3D perception or specific distortion effects.

This design scheme has been applied to 3 different kinds of products covering any needs of multi-viewpoint scene shooting (real/virtual, still/animated, photo/video).

(a) One temporal frame: a set of 8 images

(b) Autostereoscopic mix

(c) Anaglyph of images 3 and 4 (Note 1)

FIGURE 9: The results of 3D-CAM1.

(a) One temporal frame: a set of 8 images

(b) Anaglyph of images 3 and 4 (Note 1)

FIGURE 10: The results of 3D-CAM2.

This work proposes several perspectives. We take an interest in the combination of real and virtual 3D scene, that is 3D augmented reality. This allows to combine virtual and real elements or scenes. In the CamRelief project, Niquin et al. [23] work on this subject and present their first results concerning accurate multi-view depth reconstruction with occlusions handling.

We are developing a configurable camera with geometric parameters possibly flexible in order to fit to a chosen rendering device and a desired depth effect. Thus, we could test different depth distortions for a same scene. Moreover, we could produce qualitative 3D content for several rendering devices from a single camera box.

We will need to do some experiments on a fussy subject as we have to validate that the perception is geometrically conform to our expectations. This will require a significant panel of viewers but also to define and set up the perception test which will permit to precisely quantify the distances between some characteristic points of the perceived scene.

Note 1. The 3D content has been produced for autostereoscopic display. Obviously, it can only be experimented with the chosen device and in no way upon 2D media such as paper or conventional display. Nevertheless, anaglyphs help the reader to notice the depth effect on such 2D media.

Acknowledgments

We would you like to thank the ANRT, the French National Agency of Research and Technology for its financial support. The work reported in this paper was also supported as part of the CamRelief project by the French National Agency of Research. This project is a collaboration between the University of Reims Champagne-Ardenne and TéléRelief. We would you like to thank Michel Frichet, Florence Debons, and the staff for their contribution to the project.

References

- [1] W. Sanders and D. F. McAllister, "Producing anaglyphs from synthetic images," in *Stereoscopic Displays and Virtual Reality Systems X*, vol. 5006 of *Proceedings of SPIE*, pp. 348–358, 2003.
- [2] E. Dubois, "A projection method to generate anaglyph stereo images," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '01)*, vol. 3, pp. 1661–1664, IEEE Computer Society Press, 2001.
- [3] R. Blach, M. Bues, J. Hochstrate, J. Springer, and B. Fröhlich, "Experiences with multi-viewer stereo displays based on lc-shutters and polarization," in *Proceedings of IEEE VR Workshop: Emerging Display Technologies*, 2005.
- [4] K. Perlin, S. Paxia, and 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.
- [5] N. A. Dodgson, "Analysis of the viewing zone of multi-view autostereoscopic displays," in *Stereoscopic Displays and Virtual Reality Systems IX*, vol. 4660 of *Proceedings of SPIE*, pp. 254–265, 2002.
- [6] K. Müller, A. Smolic, K. Dix, P. Merkle, P. Kauff, and T. Wiegand, "View synthesis for advanced 3D video systems," *EURASIP Journal on Image and Video Processing*, vol. 2008, Article ID 438148, 11 pages, 2008.
- [7] U. Güdükbay and T. Yilmaz, "Stereoscopic view-dependent visualization of terrain height fields," *IEEE Transactions on Visualization and Computer Graphics*, vol. 8, no. 4, pp. 330–345, 2002.
- [8] T. Yilmaz and U. Güdükbay, "Stereoscopic urban visualization based on graphics processor unit," *Optical Engineering*, vol. 47, no. 9, Article ID 097005, 2008.
- [9] S. Fu, H. Bao, and Q. Peng, "An accelerated rendering algorithm for stereoscopic display," *Computers and Graphics*, vol. 20, no. 2, pp. 223–229, 1996.
- [10] O. Faugeras, Q.-T. Luong, and T. Papadopoulou, *The Geometry of Multiple Images: The Laws That Govern the Formation of Images of a Scene and Some of Their Applications*, MIT Press, Cambridge, Mass, USA, 2001.
- [11] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, New York, NY, USA, 2000.
- [12] G. Jones, D. Lee, N. Holliman, and D. Ezra, "Controlling perceived depth in stereoscopic images," in *Stereoscopic Displays and Virtual Reality Systems VIII*, vol. 4297 of *Proceedings of SPIE*, pp. 42–53, June 2001.
- [13] R. T. Held and M. S. Banks, "Misperceptions in stereoscopic displays: a vision science perspective," in *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization (APGV '08)*, pp. 23–32, ACM, New York, NY, USA, 2008.
- [14] E. Peinsipp-Byma, N. Rehfeld, and R. Eck, "Evaluation of stereoscopic 3D displays for image analysis tasks," in *Stereoscopic Displays and Applications XX*, vol. 7237 of *Proceedings of SPIE*, 2009, 72370L.
- [15] A. J. Hill, "A mathematical and experimental foundation for stereoscopic photography," *SMPTE Journal*, vol. 61, pp. 461–486, 1953.
- [16] J.-Y. Son, Y. N. Gruts, K.-D. Kwack, K.-H. Cha, and S.-K. Kim, "Stereoscopic image distortion in radial camera and projector configurations," *Journal of the Optical Society of America A*, vol. 24, no. 3, pp. 643–650, 2007.
- [17] H. Yamanoue, "The relation between size distortion and shooting conditions for stereoscopic images," *SMPTE Journal*, vol. 106, no. 4, pp. 225–232, 1997.
- [18] H. Yamanoue, "The differences between toed-in camera configurations and parallel camera configurations in shooting stereoscopic images," in *Proceedings of IEEE International Conference on Multimedia and Expo (ICME '06)*, pp. 1701–1704, 2006.
- [19] A. Woods, T. Docherty, and R. Koch, "Image distortions in stereoscopic video systems," in *Stereoscopic Displays and Applications IV*, vol. 1915 of *Proceedings of SPIE*, pp. 36–48, 1993.
- [20] N. A. Dodgson, J. R. Moore, and S. R. Lang, "Time-multiplexed autostereoscopic camera system," in *Stereoscopic Displays and Virtual Reality Systems IV*, vol. 3012 of *Proceedings of SPIE*, 1997.
- [21] <http://palais-tau.monuments-nationaux.fr>.
- [22] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [23] C. Niquin, S. Prévost, and Y. Remion, "Accurate multi-view depth reconstruction with occlusions handling," in *Proceedings of the 3DTV-Conference: The True Vision—Capture, Transmission and Display of 3D Video (3DTV-CON '09)*, Postdam, Germany, May 2009.