Geometrical Design Concept for Panoramic 3D Video Acquisition

O. Schreer, P. Kauff, P. Eisert, C. Weissig, J.-C. Rosenthal,

Fraunhofer Heinrich-Hertz-Institute, Berlin, Germany

ABSTRACT

The paper presents a new geometrical concept of an omni-directional and omni-stereoscopic multi-camera system. With such a system, 3D video panoramas can be captured and displayed at cylindrical 3D projection systems. The presented concept can be considered as an approximate solution of the well-known concentric mosaics that are defined as a sub-set of the plenoptic function. The paper discusses the geometrical relationship to the ideal case, while taking practical constraints for a real-working acquisition set-up into account.

Index Terms— Panoramic video, stereo, acquisition

1. INTRODUCTION

The immersive sensation by panoramic imaging lasts back to renaissance painters. The first experiments with moving panoramic images have been presented beginning of the last century. For instance, the Cinerama system, an immersive 360° projection has already been presented at the legendary Millenium World Exposition, 1900 in Paris [1]. Since then, a large variety of other systems have been developed targeting the entertainment market as well as training centers or event and exhibition technology [2]. In contrast to that, the provision of panoramic video supporting stereoscopic 3D is much more challenging and a steadily difficult task. While panoramic 3D projection is possible using today's projection techniques from 3D cinema, the acquisition of 3D video panoramas is a widely unsolved problem and is therefore investigated in this paper.

Capturing a still panoramic 2D image is possible by simply rotating a camera, warping and then stitching the images together [3]. Even digital consumer cameras have built-in this feature to create own panoramas. More difficult is the acquisition of panoramic 2D video, which requires a special camera arrangement. A common solution is to use multiple cameras, where individual single cameras look into different directions such that the resulting images can be stitched seamlessly to large panoramic views. First systems applying multiple cameras and mirrors to achieve full surround capture with high image resolution have already been used in the 60s by Ub Iwerks for Disney theme park productions [4]. Since then, many mirror-based system approaches have been proposed (e.g. [5]). Other approaches place a hyper- or parabolic mirror in front of a single camera to capture panoramic views [6] with the disadvantage of having a much lower resolution and plenty of distortions. Today, the advances and ongoing miniaturization of digital video cameras enables more compact systems and several commercial companies offer omni-directional cameras for a wide range of applications [7]. Good overview about different approaches on panoramic video acquisition is given in [8].

If now panoramic 3D acquisition comes into play, the situation is more difficult. The acquisition of static omni-stereo panoramas has already been investigated since more than 15 years. A nice overview on the major principles can be found in [9]. The basic idea is to mount cameras on a rotating bar. From literature, this concept is also known as concentric mosaics [3], a special version of the plenoptic function [10]. In this paper, the concept of concentric mosaics is extended towards video acquisition. It presents a 3D panoramic video camera that optimizes a trade-off between contradicting requirements on an adequate stereo impression, sufficiently overlapping views and correct positioning of focal points for accurate stitching. In the next section, the major principles and geometrical aspects of omni-directional and omni-stereoscopic systems for static scene are discussed. Section 3 then presents a design of an omni-stereoscopic video acquisition system that can be considered as an approximate solution to the theory of concentric mosaics. Section 4 shows first results of the proof of concept. A conclusion ends the paper.

2. THEORETICAL BACKGROUND

Since many years, the acquisition of panoramic images is a well-known and already solved problem of computer vision. As known from projective geometry, an error-free capture of panoramic 2D images requires that the focal points of the multiple camera views coincide in a common point and look in different directions. Usually, this condition is achieved by rotating a single camera at a tripod with a revolving camera head. In this ideal case, the single images can then be stitched to a panoramic 2D image without parallax errors for arbitrary scenes covering the entire depth range from zero to infinity. The extension of this 2D case towards 3D is called omni-stereoscopic imagery. The common approach is to

mount one or more cameras, looking either outwards or in tangential direction, on a rotating bar (see **Fig. 1**) [9]. For these approaches it is sufficient to use so-called slit-cameras, i.e. cameras that capture one column only. In a re-binning process, columns from the slit cameras are used to generate multiple perspectives. The column of one slit camera at each angular increment contributes to one panoramic view such that at least two slit cameras are needed to provide stereo panoramas.

For practical scenarios, a perspective camera is often used, again mounted at the end of the rotating bar looking either in tangential or normal direction. For the creation of a stereo panorama the two required columns can be taken from the image sensor where the distance between the two selected columns again defines the baseline. The two approaches are only suitable for capturing static scenes, but the distinction between the two categories "swing imaging" with radial camera orientation and "concentric mosaics" with tangential orientation helps to motivate the underlying concept of an omni-stereoscopic video system as explained in the next section.

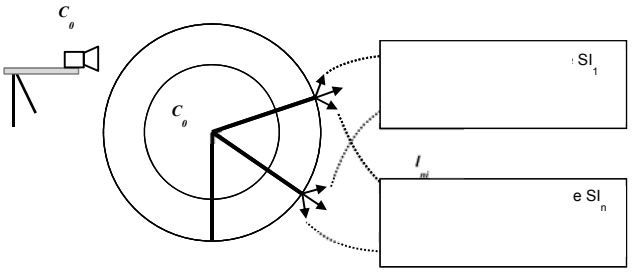

Fig. 1. Camera mounting and rebinning for swing image.

For video, the previous approaches on omni-directional or omni-stereoscopic imagery are impractical due to the need of multiple cameras for simultaneous capturing on one hand and the physical dimensions of each camera on other hand. For instance, the requirement of omni-directional imagery, that all optical centers have to coincide in one common point, is difficult to respect in video (see **Fig. 2**, left). For illustration reasons, the optical centers are drawn in front of the lens in **Fig. 2** (left), but they are considered being inside the lens. The only solution to meet this requirement perfectly is to use mirror rigs by which the virtual points of the optical centers coincide behind the mirror. Another but less perfect solution is to capture video panoramas with the star-like approach from **Fig. 2** (right). In this case, the focal points of all cameras are located on a common circle, while the optical axes are perpendicular to the arc. However, the existence of a non-zero parallax angle does not allow seamless stitching in case of close objects in the overlap area. The extension of omni-stereoscopic imagery towards video is even more complicated. This especially holds for the "swing image" approach from **Fig. 1** with a radial camera orientation. This approach supposes that the distance of adjacent cameras (see parameter S in **Fig. 3** (left) is much smaller than the stereo baseline B. However, this can only be achieved by extremely small cameras. Note that B is usually in a range of 6 cm. Hence, even if the ratio B/S can be reduced to 6, the width of the cameras must be in the range of about 1cm. Clearly, there is neither a HD camera nor a high-quality lens of this small size available. In contrast, if B is equal to or even smaller than S, the system runs in a fundamental conflict that can best be explained by a star-like arrangement as in **Fig. 3** (right).

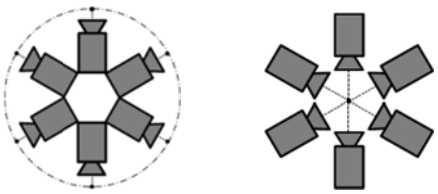

Fig. 2. left) Optimal camera arrangement; right) star-like approach.

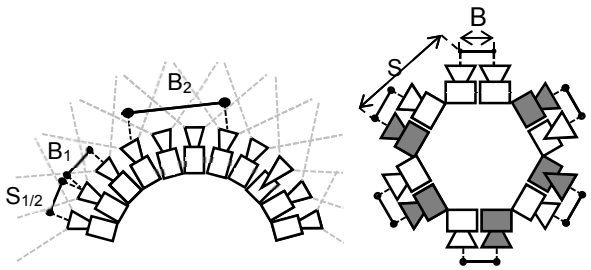

Fig. 3. left) arrangement with small micro HD cameras; right) star-like arrangement for a panoramic 3D camera setup**.**

As S is larger than B, the parallax error is larger than the stereo parallax, if the same near and far objects are also present in the overlap area. Hence, if visible parallax errors wanted to be avoided, the stereo effect is lost. The above considerations also hold for concentric mosaics with a tangential orientation. However, in contrast to the radial orientation of swing images, it can be implemented by using a mirror rig to reduce the ratio B/S to a reasonable value. The related concept will be discussed in the next section.

3. DESIGN OF AN OMNI-STEREO VIDEO CAMERA

In this section, the basic idea from concentric mosaics is further elaborated towards a practical and realistic acquisition system for omni-stereo video. The design is based on a mirror-rig that allows us to bring optical centers of all cameras close to each other (see drawing in **Fig. 4**, left). In this example, three stereo cameras are placed in front of three mirrors in a way that the virtual camera pairs cross at the radial center behind the mirrors. The geometrical arrangement of the virtual focal points behind the mirror is shown in **Fig. 4** (right). It depicts a sectional drawing at the horizontal plane that cuts the mirror pyramid at the points, where the optical axes of the real cameras intersect the mirror surfaces. In **Fig. 4**, the virtual focal

points of the same stereo pair are connected by solid bold lines (baseline B_r). Black dots relate to left cameras L_1, L_2 and L_3, whereas the grey dots refer to right cameras R_1, R_2 and R_3. As shown by the dashed lines, the field of view of each camera is framed by the opening angle α of the related mirror segment. The stereo cameras are toed-in such that they converge at the mirror surface (i.e. each camera L_i and R_i of same stereo pair looks through a window given by the mirror surface of segment *i*). In the given example, the opening angle is α=60° such that all three camera pairs cover 180° in total. According to this drawing, two major problems can be identified, the so-called stereo-gap and the stitching-gap. The reasoning for both problems and approaches to minimize them are presented in the following sections.

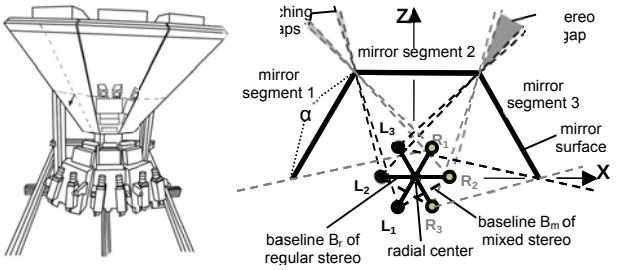

Fig. 4. Stereo arrangement using a mirror rig. (left); sectional drawing showing the virtual stereo pairs behind the mirrors (right).

3.1. The stereo-gap

The stereo-gap is shown in **Fig. 4** (right) at the mirror edge between segment 2 and 3 in dark grey. In this viewing area, the stereo information does not come from regular stereo pairs like L_2/R_2 or L_3/R_3 in adjacent segments 2 and 3, respectively, but from a mixed stereo pair L_2/R_3 from neighboring mirror segments 2 and 3. Hence, a basic constraint for seamless stereo over the entire panoramic view is that the baseline B_m of such mixed stereo pairs (see dashed bold lines in **Fig. 4** (right) must be equal to the baseline B_r of regular stereo pairs. However, a simple calculation using the Euclidian distance between the focal points of L_i and R_{i+1} shows that this constraint is not respected by the geometrical relations in **Fig. 4** (right). Hence, we obtain the following relation between B_m and B_r:

$$B_m = B_r \cdot \sqrt{(1+\cos(\alpha))/2} = B_r \cdot \cos(\alpha/2) \tag{1}$$

Eq. (1) denotes that B_m and B_r are only identical, if α is equal to zero or at least sufficiently small. This limit case refers to the theory of concentric mosaics. In practical implementations with existing high quality video cameras, however, the two baseline terms may differ considerably. The next section will show that this systematical error can be compensated by an off-center shift of the stereo pairs.

3.2 The radial off-center shift

The systematical error of the stereo gap can be compensated by shifting the virtual stereo pairs in radial direction out of the center. In practice, it can be achieved by moving the real stereo cameras towards the mirror, resulting in an off-center shift ***e*** of the related virtual stereo system as shown in **Fig. 5** (right). The corresponding arrangement of the virtual focal points behind the mirrors is depicted in **Fig. 5** (left).

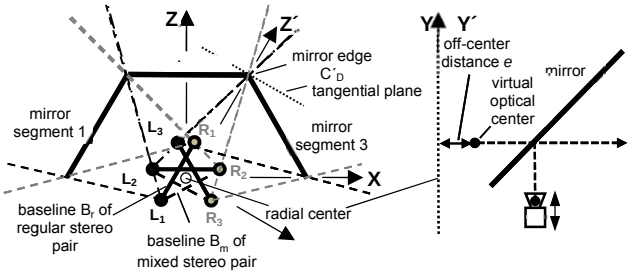

Fig. 5. Virtual stereo pairs with off-center shift (left), side view of off-center shift by moving real camera towards mirror (right).

The Euclidian distance can be calculated by considering the mixed stereo pair L_2/R_3, but now by taking into account the off-center shift ***e***. In analogy to Eq. (1) we obtain the following relation for B_m and B_r:

$$B_m = \sqrt{B_r^{\,2}\frac{(1+\cos(\alpha))}{2} + 2\cdot e^2\cdot(1-\cos(\alpha)) + 2B_r\cdot e\cdot\sin(\alpha)} \tag{2}$$

Forcing the above mentioned constraint $B_m=B_r$, results in the following expression for the off-center shift ***e***:

$$e = B_r/2\cdot\tan(\alpha/4) \tag{3}$$

Table 1 shows values of ***e***, in dependence of the opening angle ***α*** of the mirror segments. Note that for ***α***=0, the limit case of concentric mosaics results in ***e***=0.

Table 1. Off-center shifts forcing the equality of B_r and B_m

α	*0°*	*5°*	*10"*	*15°*	*30°*	*45°*	*60°*	*90°*
e	0.00 B	0.011 B	0.022 B	0.033 B	0.066 B	0.099 B	0.134 B	0.207 B

3.3 The stitching-gap

Stitching-gaps are shown in **Fig. 4** (right) at the mirror edge between segments 1 and 2 in light grey. In this camera arrangement, the optical rays from virtual focal points L_1 and L_2 (or R_1 and R_2, respectively) through the mirror edge diverge outside the mirrors and, as a consequence, stitching is not possible due to missing overlap between the two related camera images. This situation changes in case of the off-centered arrangement in **Fig. 5** (left). Here, the fact whether the optical rays diverge or converge outside the

mirror rig depends on the distance of the mirror surfaces from the radial center. This can easily be shown by using the coordinate system X′Y′Z′ in **Fig. 5** (left). Compared to the original coordinate system XYZ, it has been rotated clockwise by $\alpha/2$ around the radial symmetry axis and, thus, axis Z′ passes through the mirror edge between segments 2 and 3. Hence, the above mirror edge is located at $C'_D=(0,0,D)$ where D denotes the distance of the mirror edge from the radial center. To achieve overlap at stitching borders, D should be larger than D_{min}. A typical value is $D=2D_{min}$. The minimal distance of the mirrors relates then to the baseline B_r as follows:

$$D > D_{\min} \quad with \quad D_{\min} = B_r \cdot \cot(\alpha/2) \tag{3}$$

This relation is an important side-condition for designing practicable mirror rigs because the size of the rig also increases with D_{min}. **Table 2** shows some possible values for D_{min} depending on the segment angle α. The calculations take into account that the baseline of a stereo system of B_r=7cm is the worst-case situation for designing the mirror distance. Note that the mirror distance becomes larger than one meter for angles lower than 16°.

Table 2. Minimal mirror distance D_{min} related to segment angle α

α	3°	6°	12°	18°	24°	30°	36°	45°
D_{min}	535cm	267cm	133cm	88cm	66cm	52cm	43cm	34cm

3.4 The stitching error

Using the positions of the virtual focal points L_2 and L_3 in coordinate system X′Y′Z′, the following expressions can be defined for the inter-focal distance between L_2 and L_3:

$$\Delta X' = B_r \tan\left(\tfrac{\alpha}{4}\right) \cdot \sin\left(\tfrac{\alpha}{2}\right) = B_r(1-\cos(\tfrac{\alpha}{2})),\ \Delta Y' = 0,\ \Delta Z' = B_r \sin\left(\tfrac{\alpha}{2}\right) \tag{4}$$

It is obvious that these inter-focal distances cause parallax errors while stitching the two images from the related virtual cameras. The parallax errors can be calculated by the following expressions using the pixel metric of L_2 as reference:

$$\Delta d_u = \left| \frac{F \cdot \frac{\Delta X'}{Z'_{\min}} - u_{L_2} \cdot \frac{\Delta Z'}{Z'_{\min}}}{1 + \frac{\Delta Z'}{Z'_{\min}}} \right|, \quad \Delta d_v = \left| \frac{v_{L_2} \cdot \frac{\Delta Z'}{Z'_{\min}}}{1 + \frac{\Delta Z'}{Z'_{\min}}} \right| \tag{5}$$

The variables u_{L2} and v_{L2} denote the horizontal and vertical components of a centered coordinate system in the image plane of L_2, whereas F represents the focal length of the used cameras. u_{L2} can be calculated by the following relation:

$$u_{L_2} = F \cdot \frac{1}{2D/B_r + \tan(\alpha/4)} \tag{6}$$

The horizontal parallax error Δd_u in Eq. (5) consists of two terms, one is driven by the horizontal inter-focal distance ΔX′ and the other by the Z-difference ΔZ′ of the two virtual focal points. These two terms can compensate each other. A full compensation occurs, if $D=D_{min}$ holds. In this hard-cut situation, the points L_2, L_3 and C'_D lie at one common optical ray and, thus, the horizontal parallax error disappears. **Table 3** shows the residuals for $D=2D_{min}$ depending on the segment angle α. A usual value B_r=6 cm has been selected for the stereo baseline in this calculation. The distance of the near object has been set to Z_{min}=2m. Hence, a depth range from 2m to infinity is allowed for the stitching area. Furthermore, assuming that the horizontal field of view of the used cameras equals to the segment angle α, the focal length F is given by F=540pel/tan(α/2). This assumption refers to a mirror rig from **Fig. 4** (left) that uses HD cameras in portrait format to allow small segment angles α. Hence, u_{L3} ranges from -540 pel to +540 pel, and, accordingly, v_{L3} from -960 pel to +960 pel.

Table 3. Horizontal stitching error versus segment angle α

α	3°	6°	12°	18°	24°	30°	36°	45°
Δd_u [pel]	0.11	0.21	0.42	0.62	0.82	1.01	1.19	1.44

4. Proof of concept

To proof the concept, the system in **Fig. 5** has been investigated by computer simulations as well as by a prototype mirror rig from **Fig. 4** (left). In this context, **Fig. 6** (left) shows the set-up for the CGI experiments with the arrangement of virtual focal points according to **Fig. 5** (left). The simulations were based on six mirror segments with α=30° covering a panorama of 180° in total. The six crossing off-centered stereo systems are shown in the middle of the drawing. The corresponding six stereo views have been captured for different CGI scenes. The curved panoramic screen indicates where the six stereo views are finally re-rendered. **Fig. 7** shows the final result for an example scene. The close-up views show details of a critical region in the stitching area. A first prototype implementation of a related mirror rig is shown in **Fig. 6** (right). A segment angle of 24° has been selected for this implementation. This represents the best trade-off between size, reduction of stitching errors and good stereo quality. To achieve a precise calibration the cameras are mounted on mechanical sliders, which can be adjusted with micro-meter screws in any direction and orientation.

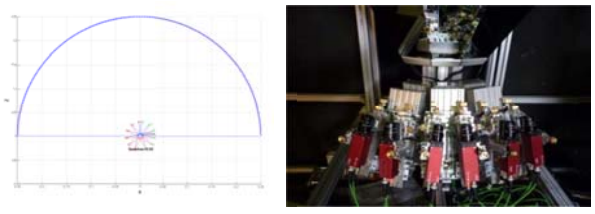

Fig. 6. Set-up for simulations on CGI basis (left), prototype of an omni-stereo mirror rig (right).

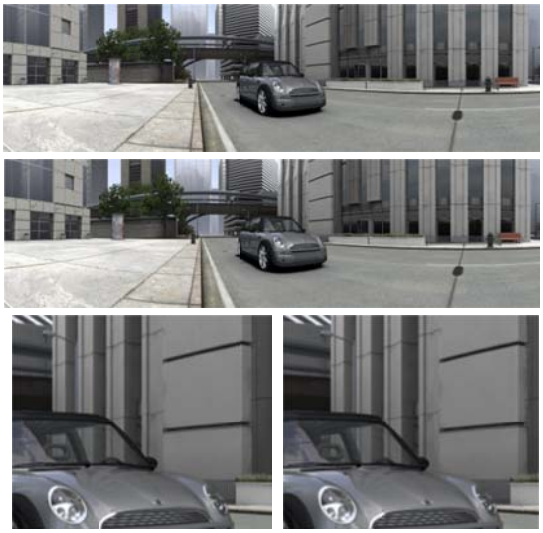

Fig. 7. Result of simulated stereo panorama with left and right view on top, close-up of critical region in stitching area (bottom)

The used HD cameras are mounted in portrait format and the spatial resolution of a resulting panorama is 7000 by 2000 pixels for 180°. **Fig. 8** shows images from regular and mixed stereo views (compare to **Fig. 5**) that have been captured during test shoots.

The sheared rectangles with the solid white lines show the effective image borders pruned by the mirror segments. The shearing is given by the fact that cameras are not positioned in the center of the mirror, but are moved horizontally by half a baseline to the left or right, respectively, and are additionally toed-in to compensate the shift. Furthermore, the left pictures in **Fig. 8** show that, in contrast to standard stereo applications, the overlap between the two views of a regular stereo system is considerably limited. The remaining stereo information has to be taken from mixed stereo pairs. By using an anaglyph overlay representation, **Fig. 9** shows how the regions from regular and mixed stereo pairs are assembled to an entire stereo panorama.

5. Conclusion

A design of an omni-stereo video acquisition system has been presented that offers an approximate solution to the theory of concentric mosaics. Several important challenges have been discussed such as the stereo-gap and the stitching-gap. By introducing a radial off center shift of stereo systems mounted in a mirror rig, it is possible to achieve a realizable solution for such a novel panoramic stereo camera, which offers a minimal stitching error while keeping a significant amount of parallax. A first prototype of such a camera has proven the correctness of the theoretical derivations.

Fig. 8. Stereo content captured by the prototype: left and right view from regular stereo (left) and from mixed stereo pair (right).

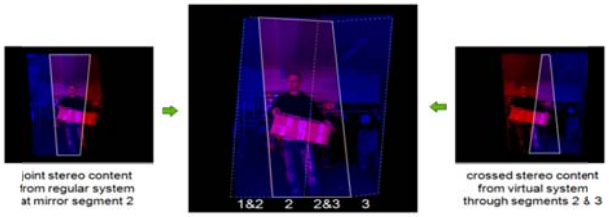

Fig. 9. Composition of 3D video panorama.

References

[1] Australian Centre for the Moving Image, “Adventures in Cybersound; Cyclorama, Cineorama, Mareorama and Myrioama”, ACMI, 2011, www.acmi.net.au/AIC/CYCLORAMA.html

[2] HPC Market Watch, “Seattle Cinerama Grand Reopening”, 2011, http://mar-kets.hpcwire.com/taborcomm .hpcwire/news/ read?GUID=15456683&ChannelID=3197.

[3] H.-Y. Shum, L.-W. He, “Rendering with Concentric Mosaics”, Proc. SIGGRAPH 99, ACM, Los Angeles, 1999.

[4] U. Iwerks, “Panoramic Motion Picture Camera Arrangement”, Canadian Patent Publication, no. CA 673633, 1963.

[5] Majumder, M. Gopi, B. Seales, H. Fuchs, “Immersive teleconferencing: A new algorithm to generate seamless panoramic video imagery”, Proc. of the 7th ACM International Conference on Multimedia, pp. 169–178, 1999.

[6] S. Baker, S. Nayar, “A theory of single-viewpoint catadioptric image formation” Int. Journal of Computer Vision, 35:175–196, 1999.

[7] Full View, “FC-1005 Camera & FC-110 Camera”, www.fullview.com/products.html

[8] K. A. Tan, H. Hua, N. Ahuja, “Multiview Panoramic Cameras Using Mirror Pyramids”, Trans. on Pattern Analysis and Machine Intelligence, Vol. 26, no7, pp.941-946, 2004.

[9] S. Peleg, M. Ben-Ezra, Y. Pritch, "Omnistereo: panoramic stereo imaging", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.23, No.3, pp.279-290, March 2001.

[10] E. H. Adelson, J. R Bergen, “The Plenoptic Function and the Elements of Early Vision”, Computational Models of Visual Processing (pp. 3-20). Cambridge, MA: MIT Press, 1991.