Creation of High-Resolution Video Panoramas of Sport Events

Christoph Fehn[†], Christian Weissig[†], Ingo Feldmann[†], Markus Müller[†], Peter Eisert[†], Peter Kauff[†], and Hans Bloß[‡]

[†]Fraunhofer Institute for Telecommunications
Einsteinufer 37, 10587 Berlin, Germany
Tel.: +49 30 31002-{611,571,290,612,614,615}, Fax: +49 30 392 72 00
{Christoph.Fehn,Christian.Weissig,Ingo.Feldmann,Marcus.Mueller,Peter.Eisert,Peter.Kauff}@hhi.fhg.de

[‡]Fraunhofer Institute for Integrated Circuits
Am Wolfsmantel 33, 91058 Erlangen, Germany
Tel.: +49 9131 776-520, Fax: +49 9131 776-588
bls@iis.fhg.de

Abstract

This paper describes an approach for creating high-resolution video panoramas of large-scale sport events. In an exemplary football scenario, we used two ARRIFLEX D-20 "film-style" digital cameras to capture both sides of the playing field as well as large parts of the stands. From the recorded left and right view images, we created a joint panoramic view with a resolution of 5016x1400 pel (5k) using sophisticated image processing algorithms. These highly immersive football panoramas were screened with our modular, high-resolution multi-projection system during the FIFA World Cup 2006 in a 600 seat CinemaxX movie theater in Berlin providing the viewers with the feeling of actually watching the game from a good seat on the stands.

1. Introduction

Attending a live event, be it a musical concert, a theatrical performance, or a football match, is undeniably one of the most exciting ways to appreciate any entertainment [17]. In order to bring the powerful experience of 'being there' also to those users who are not able to participate, *immersive media* must be able to reproduce the psychological and physical cues that an observer is using to feel 'present' in the real world. Such cues include, amoung others, high-quality, high-resolution imagery, a wide field-of-view (FOV), multi-channel surround sound, etc. [18].

An important first step towards the realization of immersive audio-visual applications is the ongoing transition from the traditional, analog world of the movie theater to the future digital cinema [24]. The rapidly increasing availability of high-end digital projectors will soon enable feature film reproduction in 2K and 4K as well as the display of live events in HD quality. Some applications, however, will require even larger aspect ratios and higher resolutions than currently supported by the DCI's (Digital Cinema Initiative) Digital Cinema System Specification [4].

In this paper we present such an application. From two high-resolution videos captured with two ARRIFLEX D-20 "film-style" digital cameras, we created a 5016x1400 pel (5k) panorama of a football match that provides the viewers with the feeling of actually watching the game from a good seat on the stands. To show such large-scale panoramic imagery, we developed a modular, high-resolution multi-projection system that can be flexibly configured in terms of number of projectors as well as screen layout [31].

This paper is organized as follows. At first, Section 2 provides an overview over related work. This is followed by Section 3, where we describe the dual-camera setup that was used to acquire our input imagery. Then, in Section 4, we explain the various image processing steps that are required for the creation of the football panorama. Thereafter, our modular, high-resolution multi-projection system is introduced in Section 5. Finally, in Section 6, we conclude our work and present some ideas for future directions of research.

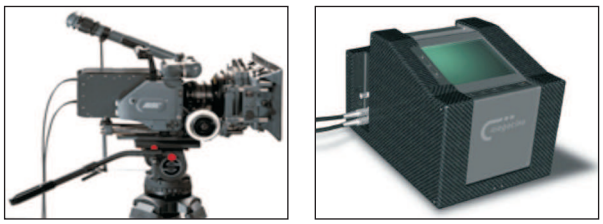
2. Related Work

Panoramic video is based on mosaicing approaches originaly developed for use with still imagery [22, 29]. While such static panoramas have become quite popular since the introduction of Apple's Quicktime VR format in 1995 [2], only a few works target the creation of dynamic high-resolution video panoramas. These are often based on multiple cameras with special optics that produce a wide (or even omnidirectional) FOV.

For example, Pintaric *et al.* [21] developed a system with five SD cameras mounted on a parabolic mirror to produce 360° panoramic videos with a total resolution of 3520x480 pel. The japanese company Mega Vision [20] commercialized a custom-build lens system with a maximum 90° horizontal opening angle that can be operated with three off-the-shelf HD cameras resulting in an aspect ratio (AR) of up to 48:9 (3xHD).

3. Video Acquisition

In order to be able to acquire moving imagery with sufficient spatial resolution and a wide field-of-view (FOV), we captured the football match with two ARRIFLEX D-20 "film-style" digital cameras (see Fig. 1 (a)) [1]. At the heart of the D-20 is a single 6 million pixel CMOS (Complementary Metal-Oxide Semiconductor) sensor with a Bayer mask. As the CMOS sensor chip has the same size as as a Super 35 mm film aperture, the D-20 can be used with the same lenses as conventional 35 mm film cameras. In our scenario, we used 20 mm Ultra Prime lenses with a horizontal opening angle of 62.8°.

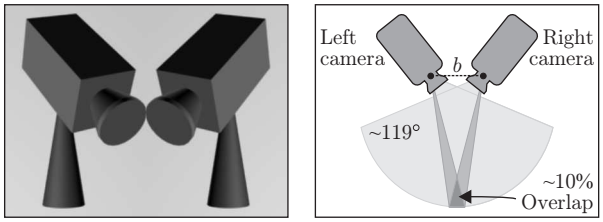


(a) (b)

Figure 1. The equipment used for video acquisition: (a) ARRIFLEX D-20 "film-style" digital camera from ARRI [1]; (b) Megacine portable recording and storage device.

While the D-20 can be used in Video Mode to output a progressive 1920x1080 pel 10bpp (bits/pel) HD signal through single link HD-SDI (the debayering is then done in real-time in the camera [26]), we ran our cameras in Film Mode to output the raw "Bayer data" at a bit depth of 12bpp via dual link HD-SDI. (The interleaved color samples were later debayered to a resolution of 2880x2160 pel using a higher quality offline post processing algorithm.) The high-resolution footage from each of the two D-20 cameras was saved to a so-called *Megacine*, a portable recording and storage device developed by Fraunhofer IIS (see Fig. 1 (b)). It combines 16 notebook (3.5 inch) hard-disks in a customized RAID setup resulting in a total storage capacity of 1 Terabyte – or, equivalently, a recording time of up to 50 minutes of raw "Bayer data" video.

Figure 2 visualizes in more detail our dual-camera setup. The two ARRIFLEX D-20s were aligned over cross such that each camera captured one side of the football field. The overlap between the two views was about 10% (i.e., around 288 pel) leading to a total FOV of approximately 119°. Although we tried to bring the cameras as close together as possible, our setup was not perfect because the centers of projection (nodal points) of the left and right view cameras were not located at the exact same point in 3D space (see also Section 4.2) [27]. The baseline b that resulted from the form factor of the lenses was in the area of about 11-12 cm.



(a) (b)

Figure 2. Dual-camera setup: (a) 3D visualization; (b) Schematic view showing the baseline b of the system as well as the area of overlap.

For a better illustration, two captured left and right view images (after debayering and color grading (see Section 4.1)) are shown in Fig. 3.

4. Creation of Panoramas

To successfully construct a high-quality panorama from two or more individual images, two different types of adjustments are typically required. The *geometric adjustments* transform the input imagery into a joint panoramic image plane and they correct for small parallax errors that can result from non-perfect camera setups. The *photometric adjustments* correct for mismatches in color in order to generate smooth transitions between the registered views.

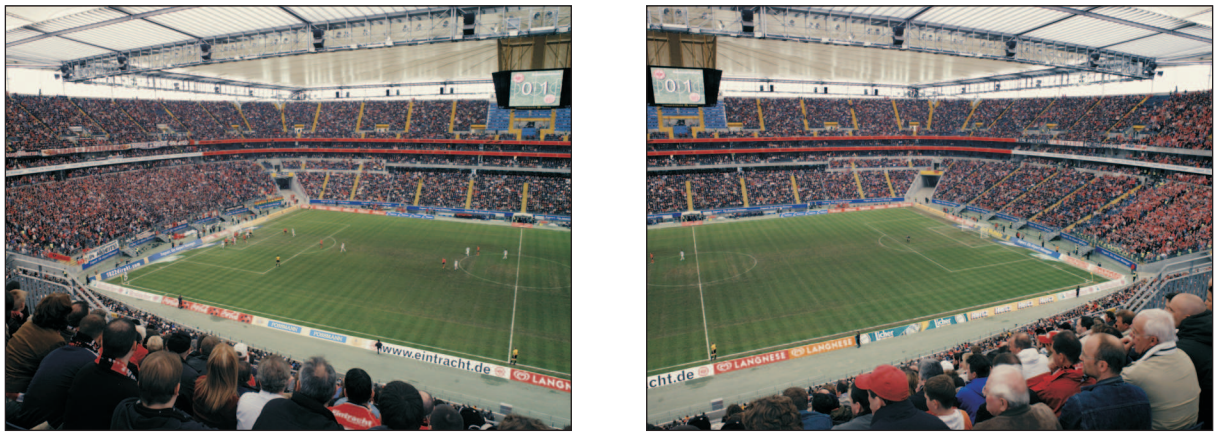


Figure 3. The left and right view images (after color grading) with a spatial resolution of 2880x2160 pel.

4.1. Photometric Adjustments

Image formation is related to several factors that influence the scene radiance recorded in a camera [14]:

- **Color response.** Due to production tolerances the spectral sensitivity curves of the individual cameras' imaging sensors might be slightly different leading to slight differences in color rendition.
- **Vignetting.** The term *vignetting* refers to a gradual reduction of brightness in the image periphery compared to the image center. In a high quality camera, this "illumination fall-off" is mostly caused by the combined effect of *natural* and *optical* vignetting, which are – to different degrees – inherent to every lens design [10, 14].
- **White balancing.** This term refers to a chromatic adjustment that can be performed by a digital camera to ensure that the object colors remain invariant under different lighting conditions.

During our shootings we found that our two ARRIFLEX D-20 "film-style" digital cameras produced sufficiently photoconsistent imagery and that also vignetting wasn't a problem with the chosen aperture sizes and lighting conditions. Thus, the only photometric operation that was required was a digital color grading (using a 3D look-up table (3D-LUT) from ARRI) to create the desired 'look' of the footage.

4.2. Geometric Adjustments

The spatial alignment of the left and right view images comprised the following processing steps:

Right view. In order to register both views into a joint panoramic image plane $\mathcal{I}_p$, we first defined a "virtual" panorama camera that has the same center of projection (nodal point) C_p as the right camera, but that is set-up parallel to the football field with it's optical axis pointing down the mid-line (see Fig. 5). The relationship between the captured right view and the "virtual" panoramic image can then be described by an 8-parameter planar perspective motion model (or *homography*) [27] that is defined by the 3x3 matrix:

$$\mathbf{H}_{r\to p} = \mathbf{A}_p \mathbf{R}_{r\to p} \mathbf{A}_r^{-1} \ , \tag{1}$$

where $\mathbf{R}_{r\to p}$ is the 3x3 rotation matrix that results from the different viewing directions of the two cameras and $\mathbf{A}_r$, resp. $\mathbf{A}_p$, are the intrinsic parameters of the right, resp. the "virtual" panorama camera [33].[1]

To vertically center the scenary in the panoramic image plane $\mathcal{I}_p$, we applied an additional, "virtual" shift s_v to the camera's sensor target. With this, the total transform of the right view image results to:

$$\mathbf{H}_{r\to p}^{total} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & s_v \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{H}_{r\to p} \ . \tag{2}$$

Compared to a conventional camera tilt-down, the "virtual" sensor shift has the advantage that vertical scene structures, such as the goalposts, also appear vertical in the transformed panorama images (see Fig. 4).

Left view. In principle, we could adjust the left view in a way similar to the right view by "rotating" it into

[1] In general, the intrinsic camera parameters $\mathbf{A}_r$ could be estimated using a suitable calibration technique (see e.g. [30, 34]). We simply approximated them from the cameras' sensor chip specifications and the lens parameters (focal length f) assuming that the optical axis intersects the image plane $\mathcal{I}$ in the center of the sensor target and that lens distortions are neglectible [33].

(a) (b)

Figure 4. Perspective appearance of the right goal in the panorama: (a) With a conventional camera tilt-down; (b) With a "virtual" shift s_v of the camera's sensor target.

the desired viewing direction, or – as it is mostly done in state-of-the-art panorama creation software (see e.g. [22, 29]) – by estimating a global transform that minimizes a suitable error metric (e.g., the per-pixel color difference) within the overlap area of the two images. However, due to the fact that the nodal points C_l and C_r of our two capturing cameras are not perfectly aligned in a common 3D point (see again Section 3), we would experience parallax artifacts that would result in the perception of double images (ghosting) and blurred details in the overlapping scene parts.

While such parallax artifacts cannot be avoided completely with our non-perfect acquisition setup, we can, however, ensure that no errors are created in the plane Π defined by the football field, i.e., the area that is of most interest for our application. For doing so, we first try to determine a homography $\mathbf{H}_{l\to r}$ that exactly transforms all field points $\mathbf{p}_l$ in the left view into their corresponding images $\mathbf{p}_r$ in the right view, i.e.:

$$\tilde{\mathbf{p}}_r \cong \mathbf{H}_{l\to r}\tilde{\mathbf{p}}_l \ , \tag{3}$$

where the tilde denotes the use of projective coordinates and the symbol $\cong$ means "equal up to scale" [9].

From Eq. (3), $\mathbf{H}_{l\to r}$ can be estimated from four or more point correspondences using Hartley's well-known 8-point algorithm [8]. In addition, we can use the fact that the sidelines $\mathbf{l}_l$ and $\mathbf{l}_r$ in the left and right views must be aligned in the joint panoramic view to yield the additional constraints [9]:

$$\tilde{\mathbf{l}}_r \cong \mathbf{H}_{l\to r}^{-1}\tilde{\mathbf{l}}_l \ . \tag{4}$$

The resulting total transformation from the left view to the joint panoramic view can then be written as:

$$\mathbf{H}_{l\to p}^{total} = \mathbf{H}_{r\to p}^{total}\mathbf{H}_{l\to r} \ . \tag{5}$$

According to "plane + parallax" theory [15, 25] the "residual parallax" $\Delta\mathbf{p}$ for a 3D point P outside the plane Π is then directly proportional to the height $P^{\perp}$ of the point above the plane and inversely proportional to its depth Z from the right view camera [15], i.e.:

$$\Delta\mathbf{p} = \mathbf{H}_{r\to p}^{total}\left(\mathbf{p}_r - \mathbf{H}_{l\to r}\mathbf{p}_l\right) = \alpha_u \cdot \frac{bP^{\perp}}{ZC_l^{\perp}} \ , \tag{6}$$

where α_u is the focal length (in pixel units) and b the baseline between the left and right camera (see Fig. 5).

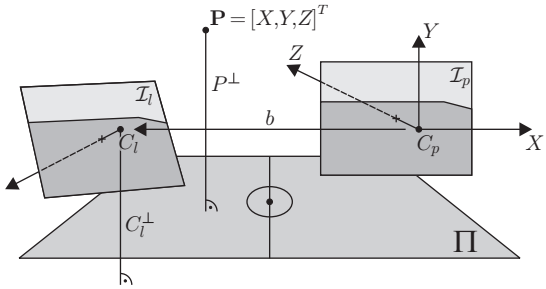


Figure 5. "Plane + parallax" configuration.

Parallax Error Correction. After the left and right views have both been transformed into the joint panoramic image plane $\mathcal{I}_p$, the remaining parallax errors $\Delta\mathbf{p}$ have to be corrected in the left image before blending. In principle, this correction would be required for all 3D points P that do not belong to the plane Π defined by the football field, i.e., for the stands, for the roof of the stadium, for the video cube, etc. However, because the parallax errors are rather small and, thus, do not lead to perceptible global perspective distortions, we only need to ensure that corresponding scene parts match in the overlapping left and right image portions (i.e., in the later blending area).

We achieve this by defining a large rectangular patch that covers the whole grandstand (from down to the field up to the stadium roof) in the left view and by correcting this image section by means of a *bilinear transformation* or *bilinear mapping* [7, 11, 32]. Such mappings are most commonly used in the field of computer graphics (CG) for warping rectangles into general quadrilaterals. As shown in Fig. 6, they can be computed by first linearly interpolating by fraction $\hat{u} \in [0..1]$ along the top and bottom edges of the quadrilateral, and then by linearly interpolating by fraction $\hat{v} \in [0..1]$ between the two interpolated points to yield the destination point [11]:

$$\mathbf{p} = (1-\hat{u})(1-\hat{v})\mathbf{p}_1 + \hat{u}(1-\hat{v})\mathbf{p}_2 + \hat{u}\hat{v}\mathbf{p}_3 + (1-\hat{u})\hat{v}\mathbf{p}_4 \ .$$

In the more general matrix notation this can also be written as:

$$\mathbf{p} = \begin{bmatrix} u \\ v \end{bmatrix} = [\hat{u}\hat{v},\ \hat{u},\ \hat{v},\ 1] \begin{bmatrix} a & e \\ b & f \\ c & g \\ d & h \end{bmatrix}, \qquad (7)$$

where the eight matrix coefficients a to h are calculated from the four corner point correspondences of Fig. 6.

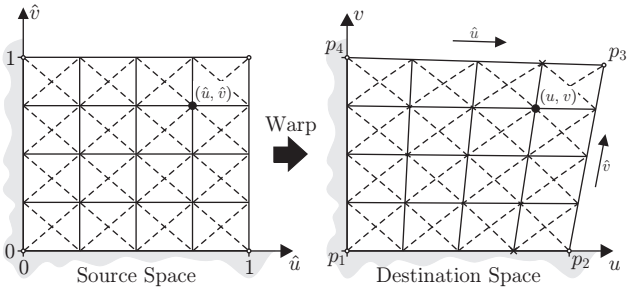
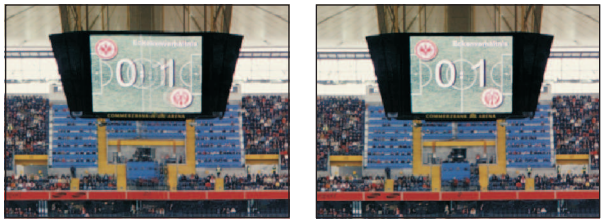


Figure 6. Bilinear warp from source to destination space. Equispaced points are preserved on lines that are horizontal or vertical in the source space (after [11]).

In contrast to the well-known perspective mappings (i.e., the projective transformations that are defined by a homography), bilinear warps have the interesting property that the forward transform from source to destination space preserves lines that are horizontal or vertical in the source space – this follows from the bilinear interpolation that is used to realize the transformation [32] –, and that it preserves equidistantly spaced points along such lines. (On the other hand, a bilinear warp *does not* preserve diagonal lines which are mapped to quadratic curves in the destination space [11].) By slightly displacing only the upper right corner of the rectangle, we can thus correct for the vertically increasing parallax in the stands without creating perceptible discontinuities at the bottom and the left border of the source patch.

Because the video cube, which is centered above the kick-off point in the middle of the stadium, is closer to the capturing cameras and, thus, leads to larger parallax errors than the parts of the grandstand that are equally far above the football field, we need an additional, second bilinear transformation to also get rid of the remaining small ghosting effects.

The results of our parallax error correction procedure can be seen in Fig. 7. Without any correction (a), the blending of the left and right views leads to annoying double contours (ghosting) in the grandstand and, even more conspicuous, in the video cube. By correcting the left image with the described bilinear mappings (b), these artifacts can be eliminated as far as possible.



(a) (b)

Figure 7. Close-ups of the upper part of the grandstand and the video cube: (a) Ghosting resulting from parallax errors; (b) After correction of parallax errors using a bilinear mapping.

4.3. Results

The final 5016x1400 pel (5k) football panorama is shown in the upper part of Figure 8. The reproduced perspective corresponds to what a spectator would see from a good seat on the height of the midline. In addition to the complete field, the stands are fully visible on three sides of the ground, yielding a total field-of-view (FOV) of almost 100°. This very large viewing angle as well as the possibility to see the crowds (in combination with the original stadium sound) significantly contribute to the feeling of being actually 'present' at the match. The three close-ups on the bottom of the figure give an idea of the high amount of spatial detail that is available in the panoramic imagery.

5. 5k Projection

The display of 5016x1400 pel (5k) imagery is currently not possible with even the most advanced high-end digital projectors. (The Sony SXR-R110CE SXRD projector offers a horizontal resolution of 4096 pel (4k) [28], state-of-the-art DLP [5] projectors provide a resolution of 2048 pel (2k) horizontally). Such ultra-high resolutions can, however, be achieved by stitching together a number of lower resolution video frames in order to yield a single image of ultra-high definition. For this purpose, a scalable and modular multi-projection system has been developed at Fraunhofer HHI [31]. The core of this system is our so-called CineCard (see Fig. 9), a PCI plug-in card that contains dedicated hardware (e.g., for alpha blending, black level adjustment, and colormetric corrections) for providing seamless transitions between the individual sub-images. In contrast to other multi-projection concepts (e.g., [16, 19]), which often use clusters of PCs, where each PC serves a single projector only, the CineCard is able to control up to four separate projectors (with

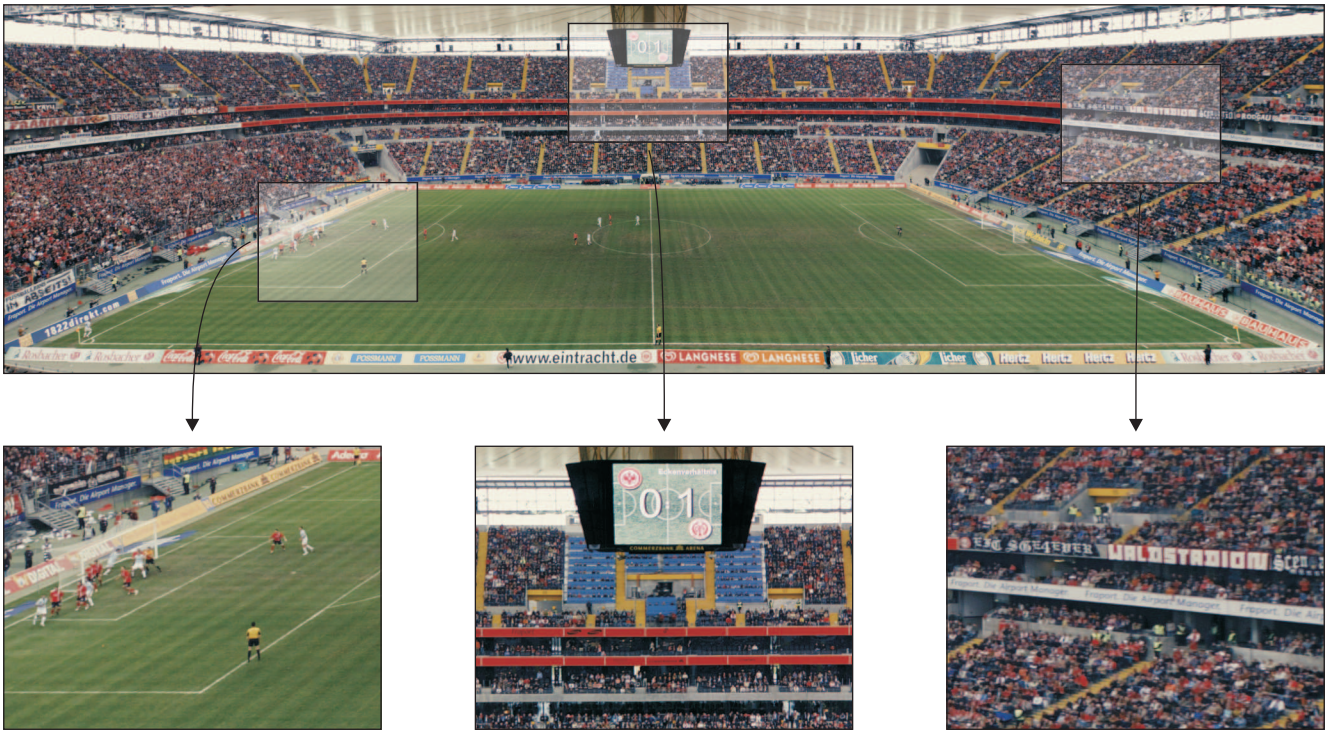
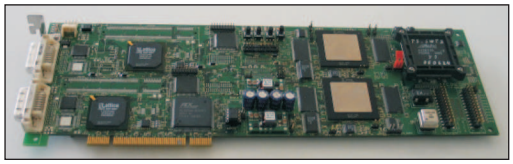
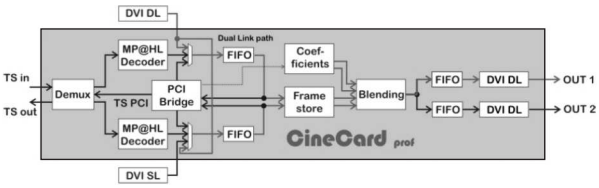


Figure 8. The final football panorama with a spatial resolution of 5016x1400 pel.

either four times XGA or two times HDTV (SXGA or SXGA+) resolution) simultaneously yielding a highly efficient and cost-effective solution.



(a)



(b)

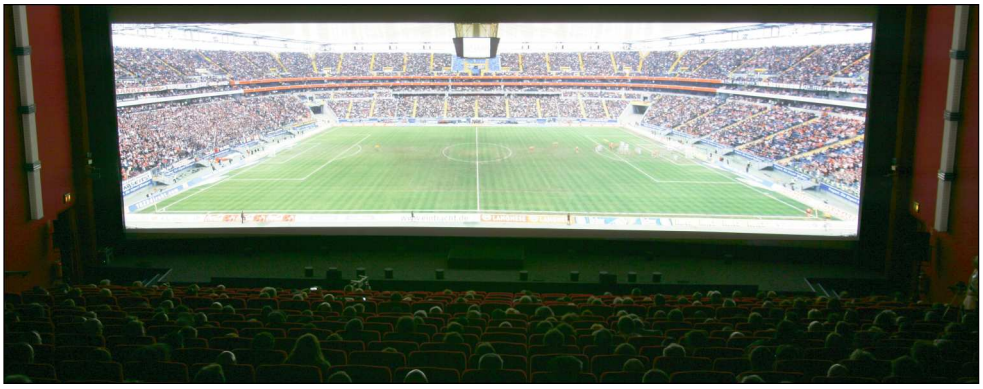
Figure 9. CineCard developed at Fraunhofer HHI: (a) PCI plug-in board; (b) System design.

The CineCard can be driven with either compressed or uncompressed motion imagery. Uncompressed videos are grabbed from a Dual-Link DVI connector (either two SXGA, resp. SXGA+ streams with a bit depth of 8bpp (bits/pel) or one stream with XGA resolution and a bit depth of 16bpp), whereas compressed videos are taken from an MPEG-2 transport stream (either loaded from local disk via the PCI bus or streamed from a LVDS SPI input) [12]. In the latter case, the TS is demultiplexed and the resulting MPEG-2 encoded video streams are decompressed using two onboard MP@HL (Main Profile@High Layer) decoders [13]. Projector arrays of arbitrary size and configuration can be realized by cascading an unlimited number of CineCards. In this case, precise synchronization between the individual video streams is achieved by means of a patented recovery of sync data from the MPEG-2 transport layer.

For displaying our 5k football panoramas, we used a configuration with five Christie DS+8K projectors, each having a resolution of 1400x1050 pel (SXGA+) [3], that were driven from three CineCards operated in a single PC. The original 5016x1400 pel panoramic views were split into five overlapping sub-images (in landscape mode) and encoded with MPEG-2 MP@HL at a bitrate of 25 Mbps each. The resulting MPEG-2 elementary streams were then multiplexed into three synchronized MPEG-2 transport streams (two TS with two videos and one with only a single video and the accompanying six-channel surround sound).

The multi-projection system was presented to a selected audience of journalists, cinematographers, and

(a)



(b)

Figure 10. The presentation of the multi-projection system during the FIFA World Cup 2006 in Berlin: (a) Projector setup (the mirrors enabled us to project in landscape mode); (b) Spectators view on the screen.

industry professionals during the FIFA World Cup 2006 in a 600 seat CinemaxX movie theater in Berlin. Pictures of the projector setup and the spectators view on the screen are shown in Fig. 10. The overall size of the panoramic 5k renditions was 20 m x 5.6 m, the total light intensity was equal to 42,500 ANSI Lumens (8,500 ANSI Lumens per individual projector).

6. Conclusions & Future Work

This paper described an approach for creating high-resolution panoramas of sport events. In an examplary football scenario, we used a dual-camera setup consisting of two ARRIFLEX D-20 "film-style" digital cameras to capture both sides of the playing field as well as large parts of the stands. From these left and right view images, we created a joint panoramic view with a resolution of 5016x1400 pel (5k) using sophisticated image processing algorithms. The generated football panoramas were screened with our modular, high-resolution multi-projection system in a local cinema in Berlin during the FIFA World Cup 2006.

For the future, we plan to automate some of the image processing steps that currently have to be initialized manually (e.g., the selection of point and line correspondences within the area of overlap or the correction of the residual parallax errors in the stands). In addition, we would like to combine our system with state-of-the-art methods for sports analysis (e.g., for ball and/or player tracking [6, 23]) in order to further enhance the immersive viewing experience with supplementary information such as ball possession per team or individual player statistics.

7. Acknowledgments

This work has been sponsored by the Federal Ministry of Economics and Technology (BMWi) within the German research project CineVision 2006. The authors would like to thank all project partners (Fraunhofer IIS, ARRI, Flying Eye, Fraunhofer HHI, Kinoton, and Deutsche Telekom) for their support and for their

input to this publication. In addition, we would like to thank Lars Plath from the Deutsche Fußball Liga (DFL), Carsten Knoop from Eintracht Frankfurt Fussball AG, as well as Marcel Boyé from Stadion Frankfurt Management GmbH for the permission to record a DFL football match in the Commerzbank Arena Frankfurt.

References

[1] ARRI [Online] (2006). http://www.arri.de.

[2] S. E. Chen. QuickTime VR – An Image-Based Approach to Virtual Environment Navigation. In *Proceedings of ACM SIGGRAPH*, pages 29–38, Los Angeles, CA, USA, 1995.

[3] Christie [Online] (2006). http://www.christiedigital.com.

[4] Digital Cinema Initiatives, LLC. Digital Cinema System Specification. V1.0, July 2005.

[5] DLP Texas Instruments [Online] (2006). http://www.dlp.com.

[6] E. Guizzo. Bugged Balls for Tough Calls (Ball Tracking Systems). *IEEE Spectrum*, 42(6):14–15, June 2005.

[7] H. Hægland. Streamline Tracing on Irregular Grids. Master's thesis, University of Bergen, Norway, Dec. 2003.

[8] R. I. Hartley. In Defence of the 8-Point Algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(6):580–593, June 1997.

[9] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, 2000.

[10] D. Hasler and S. E. Susstrunk. Color Handling in Panoramic Photography. In *Proceedings of SPIE Videometrics and Optical Methods for 3D Shape Measurements*, pages 62–72, San Jose, CA, USA, Jan. 2001.

[11] P. S. Heckbert. Fundamentals of Texture Mapping and Image Warping. Master's thesis, University of California, Berkeley, CA, USA, June 1989.

[12] ISO/IEC JTC 1/SC 29/WG 11. Information Technology – Generic Coding of Moving Pictures and Audio: Systems. ISO/IEC 13818-1:1996, Apr. 1996.

[13] ISO/IEC JTC 1/SC 29/WG 11. Information Technology – Generic Coding of Moving Pictures and Audio: Video. ISO/IEC 13818-2:1996, Apr. 1996.

[14] J. Jia and C.-K. Tang. Tensor Voting for Image Correction by Global and Local Intensity Alignment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(1):36–50, Jan. 2005.

[15] R. Kumar, P. Anandan, and K. Hanna. Direct Recovery Of Shape From Multiple Views: A Parallax Based Approach. In *Proceedings of International Conference on Pattern Recognition*, pages 685–688, Jerusalem, Israel, Oct. 1994.

[16] K. Li and Y. Chen. Optical Blending for Multi-Projector Display Wall System. In *Proceedings of* 12th *Lasers and Electro-Optics Society Annual Meeting*, pages 281–282, San Jose, CA, USA, Jan. 1999.

[17] N. Lodge. Being Part of the Fun – Immersive Broadcasting for the New Millenium. In *Proceedings of* 7th *International Conference on Image Processing and its Applications*, pages 1–7, Manchester, UK, July 1999.

[18] M. Lombard and T. Ditton. At the Heart of it all: The Concept of Presense. *Journal of Computer-Mediated Communication*, 3(2), Sept. 1997.

[19] P. Lyon. Edge Blending Multiple Projection Displays on a Dome Surface to form Continuous Wide Angle Field-of-View. In *Proceedings of* 7th *I/ITEC*, pages 203–209, 1985.

[20] Mega Vision [Online] (2006). http://www.megavision.co.jp.

[21] T. Pintaric, U. Neumann, and A. Rizzo. Immersive Panoramic Video. In *Proceedings of International Multimedia Conference*, pages 493–494, Los Angeles, CA, USA, Oct. 2000.

[22] Realviz [Online] (2006). http://www.realviz.com.

[23] J. Ren, J. Orwell, G. A. Jones, and M. Xu. A General Framework for 3D Soccer Ball Estimation and Tracking. In *Proceedings of International Conference on Image Processing*, pages 1935–1938, Singapore, Oct. 2004.

[24] C. S. Schwartz, editor. *Understanding Digital Cinema*. Focal Press, 2004.

[25] A. Shashua and N. Navab. Relative Affine Structure: Canonical Model for 3D from 2D Geometry and Applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):873–883, Sept. 1996.

[26] M. Shipman-Mueller. D-20 – Let's Debayer a Color Filter Array, Shall We? ARRI Newsletter, No. 9, Nov. 2003.

[27] H.-Y. Shum and R. Szeliski. Panoramic Image Mosaics. Technical Report MSR-TR-97-23, Microsoft Research, 1997.

[28] Sony SXRD 4k Projector [Online] (2006). http://www.sonybiz.net/4k.

[29] The Panorama Factory [Online] (2006). http://www.panoramafactory.com.

[30] R. Y. Tsai. A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-The-Shelf TV Cameras and Lenses. *IEEE Journal of Robotics and Automation*, 3(4):323–344, Aug. 1987.

[31] C. Weissig, I. Feldmann, J. Schüssler, U. Höfker, P. Eisert, and P. Kauff. A Modular High-Resolution Multi-Projection System. In *Proceedings of* 2nd *Workshop on Immersive Communication and Broadcast Systems*, Berlin, Germany, Oct. 2005.

[32] G. Wolberg. *Digital Image Warping*. IEEE Computer Society Press, Los Alamitos, CA, USA, 1990.

[33] G. Xu and Z. Zhang. *Epipolar Geometry in Stereo, Motion and Object Recognition*. Kluwer Academic Publishers, Dordrecht, NL, 1996.

[34] Z. Zhang. A Flexible New Technique for Camera Calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, Nov. 2000.