AN INDIVIDUAL PERSPECTIVE
Perceptually Realistic Depiction of Human Figures

Martin Zavesky, Jan Wojdziak, Kerstin Kusch, Daniel Wuttig, Ingmar S. Franke and Rainer Groh
Institute of Software and Multimedia Technology, Department of Computer Science
Technische Universität Dresden, Dresden, Germany

Keywords: Multi-perspective, Spatial perception, Non-photorealistic rendering, Perceptually-based rendering.

Abstract: Projection of three-dimensional space onto a two-dimensional surface relies on the computer graphics camera based in design on the camera obscura. Geometrical limitations of this model lead to perspective distortions in wide-angle projections. Including the camera model, our approach is to involve the human perception in order to create a realistic spatial impression by a two-dimensional image. The aim is to provide human-centered interfaces for an efficient and coherent communication of spatial information in virtual worlds to support avatar-mediated interaction with its need for correct depiction of human figures concerning proportion and orientation. To this end, we explain an object-based and introduce a camera-based computer graphics procedure to prevent projective distortions and misalignments.

1 INTRODUCTION

Representations of individuals have become an essential feature in a wide range of applications. Users represented by their avatars are connected by broadband networks in virtual environments primarily in the field of entertainment such as three-dimensional communication tools and massively multiplayer online role-playing game (MMORPG). Avatars are graphical equivalents of users (Capin et al., 1999). They are realistic-looking antropomorphic shapes that pose communicational and social intentions of the users (Fu et al., 2008). Avatars do not only have to be realistic due to the appearance of the user but also due to his behavior (Bailenson et al., 2006).

Our consideration is focused on providing interfaces that assist the user in perceiving spatial information as effectively and efficiently as possible (Jokela et al., 2003). A necessary requirement to achieve this aim is a perceptually realistic projection of three-dimensional scenes. *Perceptual realism* means an integration of characteristics of human perception into the imaging process and the image (Groh et al., 2006). A common way of projecting objects on an image plane is the perspective projection using the computer graphics camera. In this model the observer and its perception are disregarded. A projection that rather relies on geometrical principles than on visual perception might prevent ambiguity and misinterpretation (Franke et al., 2008).

2 PERSPECTIVE PROJECTION IN COMPUTER GRAPHICS

The computer graphics camera model maps virtual three-dimensional space onto a two-dimensional surface. The transformation of virtual space onto the image plane of the camera follows the laws of perspective projection (Foley, 1999; Angel, 1997). The computer graphics camera model (figure 1) contains a view frustum as well as the center of projection [C], the viewing direction [V] and the up-vector [U]. The view frustum is defined by the aspect ratio of the projection plane and also near [N] and far [F] clipping plane. The intersection of the optical axis (viewing direction) and the image plane is the principle vanishing point of the image and corresponds to the intersection of the horizon and the sagittal line in the image (Franke et al., 2006).

The projection yields a perspective view along the viewing direction of the camera and obtains an uniform treatment of all objects in the image. As a direct consequence, objects located near the edges of the view frustum and close to the image plane are streched and misaligned. Especially, if objects are covered by a closed curved (convex) surface and are

Zavesky M., Wojdziak J., Kusch K., Wuttig D., S. Franke I. and Groh R..
AN INDIVIDUAL PERSPECTIVE - Perceptually Realistic Depiction of Human Figures.
DOI: 10.5220/0003317103130319
In *Proceedings of the International Conference on Computer Vision Theory and Applications* (VISAPP-2011), pages 313-319
ISBN: 978-989-8425-47-8

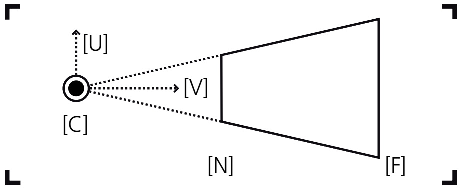

Figure 1: The common computer graphics camera model consists of center of projection [C], viewing direction [V] and up-vector [U] and also view frustum with near clipping plane [N] and far clipping plane [F].

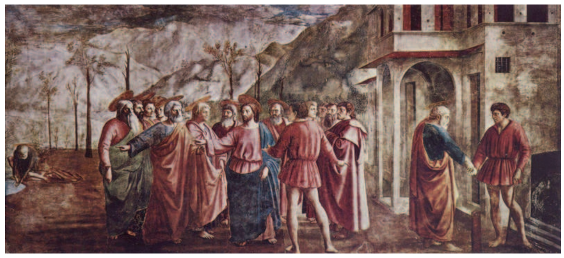

Figure 2: The Tribute Money, Tommaso di Ser Cassai (Masaccio), 1425-1428.

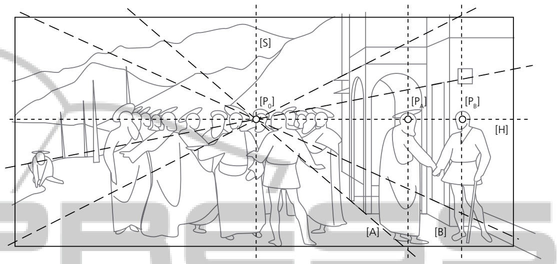

Figure 3: Sketch of 'The Tribute Money', Tommaso di Ser Cassai (Masaccio) with visualized horizontal [H] and sagittal [S] line as well as the principle vanishing point [P_0] and the additional principle vanishing points [P_A] and [P_B] of persons [A] and [B].

projected with a wide aperture angle, this stretching is perceived as unnatural and distorted (Franke et al., 2008; Yankova and Franke, 2008). This effect can be clearly observed at spheres, columns or human bodies as shown in figure 4 a (Groh, 2005; Ware, 1900). In addition, these shown human figures seem to possess an orientation which does not correspond with the expectation of the observer. The issues arise from the limitation of the standard camera model. There is only one single mathematical center of projection defined and therefore only one principal vanishing point exists. Accordingly, computer generated images are typically mono-perspective.

3 ARTISTIC APPROACH

The depiction of human figures has a long tradition in fine arts. Especially the artists of the Renaissance exerted themselves to analyze the proportions and orientations of human shapes in order to create rules for an aesthetic depiction.

Analyzing Renaissance paintings reveals that artists were faced with the same challenges as current computer graphics: to map a three-dimensional scene realistically onto a two-dimensional plane (Groh, 2005; Hockney, 2006; Ware, 1900). But there is a difference between paintings and computer generated images: The "device" for creating projections. A computer system calculates an image on geometrical rules of projection. Artists integrate a "human factor" in these rules because they construct paintings on their own (human) visual impression (Yankova and Franke, 2008). For analyzing the "human factor" and its applicability in computer generated images, the picture "The Tribute Money" (figure 2) will be discussed.

As seen in figure 3, the scene is constructed by the rules of perspective projection. The investigation of the perspective structure reveals that the construction design of the building in the right part of the painting is developed dependent on the principal vanishing point [P_0] of the whole scene. The shapes of all figures appear undistorted whereby figures proportions are kept constant. Each figure is shown in an individual perspective with a corresponding additional principal vanishing point (Franke et al., 2007). As an example the additional principle vanishing points [P_A] and [P_B] of persons [A] and [B] are shown in figure 3. In contrast to images rendered by the computer graphics camera, their proportions are regular regardless of their distance to the principal vanishing point [P_0] as a result of the multi-perspective visualization (Groh et al., 2006; Hockney, 2006). A virtual reconstruction with an angle of view of the computer graphics camera of 120 degrees, which is comparable to the painting "The Tribute Money" of Masaccio illustrates the differences between mono-perspective and multi-perspective images. To ensure the comparability of the human shapes, all figures base upon the same three-dimensional model. Especially the lateral figures [A] and [B] in figure 4 a are affected by distortions, that result from the projection process.

Hence, an essential feature of Renaissance fine arts is the use of multi-perspective projection methods in paintings to depict perceptually realistic human figures.

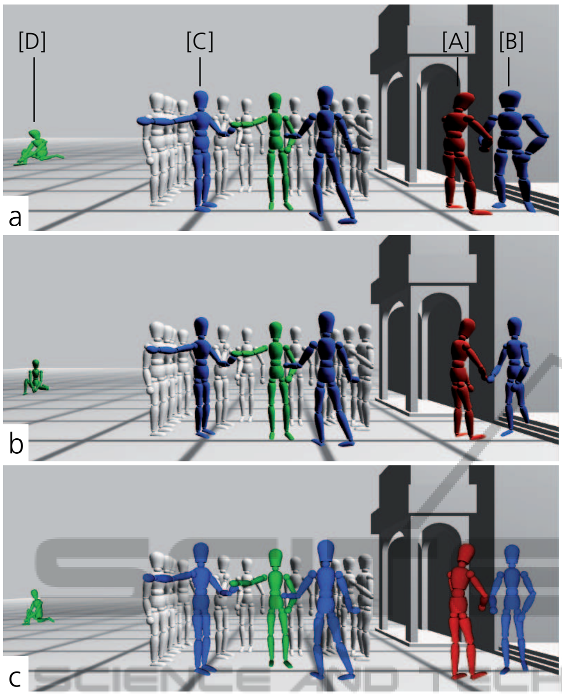

Figure 4: Virtual reconstruction of "The Tribute Money" (a) rendered with standard computer graphics camera (b) using OBPC on the colored human figures (c) using CBPC on the colored human figures.

4 MULTI-PERSPECTIVE IMAGING

The last section emphasized multi-perspective as a solution to create perceptually realistic images. The applicability in computer graphics implies a modification of the standard computer graphics camera model because a single viewing transformation cannot fulfill the requirements of multi-perspective imaging (cp. 2). Accordingly, individual objects of a computer graphics scene have to be visualized in their individual perspective.

4.1 Related Work

In multi-perspective images, the spatiality of a three-dimensional scene should be preserved concerning the proportions of objects. Spherical objects have to be visualized circularly and lines of cubic objects should be straight in multi-perspective images. The approach of (Zorin and Barr, 1995) enables to reduce geometric distortions in computer-generated and photographic images by constructing viewing transformations in a post processing stage of the rendering process. However, this procedure does not enable to visualize single objects in its individual perspective. Manipulations on the image plane are used by (Carroll et al., 2009) to preserve shapes and maintain straight lines of a scene that are marked by the user. The approach has to be supported manually and cannot be used in virtual worlds. (Zelnik-Manor et al., 2005) developed a procedure to construct panoramas from images taken from a single view point with no noticeable distortions in background and objects, but it is not applicable on discret objects of a scene. Another approach to achieve multi-perspective is to extend the camera model by using a multi-projection rendering algorithm with multiple cameras (Singh, 2002; Yu and McMillan, 2004). The final image is a composition of these camera views. The main focus of the procedures of (Singh, 2002) and (Yu and McMillan, 2004) are multi-perspetive panoramas and artistic multi-perspective rendering and not perceptually realistic imaging. (Agrawala et al., 2000) developed a framework for rendering multi-perspective images from three-dimensional models based on spatially varying projections. This approach is, similar to (Carroll et al., 2009), not completely self-acting. The concept of (Coleman and Singh, 2004) also operates with different cameras like the previous one. In contrast, the image is rendered by a single boss camera. The other cameras also called lackeys define a deformation on the scene objects. That approach addresses spatial scene coherence, shadows and illumination but no self-contained objects. We show in (Franke et al., 2007) one approach to visualize objects in its own perspective. A direct manipulation of the geometry is used to visualize objects in its individual perspective with the standard computer graphics camera model.

4.2 Object-based Perspective Correction (OBPC)

In (Franke et al., 2007) we presented an algorithmic solution implementing our concept of multi-perspective imaging. The approach belongs to the class of geometrical manipulations and has been already applied on abstract objects such as spheres and columns. In the following, the algorithm will be used on anthropomorphic shapes in the context of communication in virtual worlds. It will be referred to as *Object-Based Perspective Correction* to distinguish this algorithm from the *Camera-Based Perspective Correction* which will be described in the next subsection. The object-based correction algorithm performs affine transformations to modify the geometry of objects directly. These objects (avatars) are visualized in a way that their anthropomorphic shapes appear to be undistorted on the image plane. This procedure allows to influence only those objects selectively affected by a perspective distortion. A significant advantage of this approach is, that the standard computer graphics camera model can be used to create images. It is not necessary to render multiple views as it is done in image based solutions (Agrawala et al., 2000; Zorin and Barr, 1995). To imitate the multi-perspective approach of painters, transformations of rotation and shear are applied to the object geometry [O_A] (cp. figure 5). This results in a modification of the object on the image plane (cp. [D_A], [D_B] and [D_C]). It is a pre-rendering concept depending on the spatial constellation of camera and object that counteracts the perspective distortions. Consequently, the camera system requires only one center of projection [C] and one viewing direction [V] for multi-perspective imaging. The main steps of the algorithm are summed up as follows (cp. (Franke et al., 2007)):

1. Specify the pivot point of the object in the local camera coordinate system.
2. Compute the shear factors from that relative position
3. Compute the rotation angles
4. Rotate the object around x- and y-axis according to the rotation angles
5. Shear the object with a shear matrix based on the computed shear factors

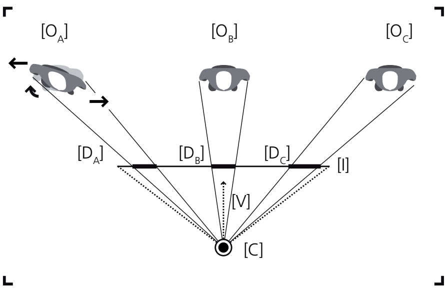

Figure 5: Transformations of rotation and shear by the OBPC. The depiction [D_A] of object [O_A] has the same size on the image plane [I] as the depiction [D_B] of object [O_B] on the viewing direction [V] originating from the center of projection [C] while the depiction [D_C] of object [O_C] covers a bigger region on the image plane.

A multi-perspective image of the reference scene based on the OBPC is shown in figure 4 b. The colored human figures were treated by the mentioned algorithm. The visualized peripheral avatars are undistorted and aligned correctly caused by their individual additional principle vanishing points. The resulting image "imitates" the perceived outcome of the natural human viewing behaviour. Correcting single objects embeds further perspectives in a original mono-perspective scene view.

4.3 Camera-based Perspective Correction (CBPC)

The rendering process allows the manipulation of the projection plane as another method to generate perception-adapted images. The aim is to offer an alternative procedure to generate perceptually realistic images based on multi-perspective. This is realized by the usage of several cameras combined according to predefined rules. Multiple cameras at the same position with different orientations simulate the successive shift of human visual attention. The algorithm creates a camera-framework consisting of a system camera and several object cameras as shown in figure 6.

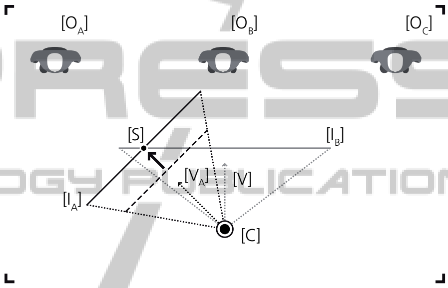

Figure 6: Principle of the CBPC. The system camera renders the whole scene ([O_B], [O_C]) excluding the object [O_A] attached to the object camera. The viewing direction [V_A] of the object camera is aligned towards its associated object. For size constancy the image plane [I_A] of the object camera is shifted towards the object right up to the intersection [S] of viewing direction [V_A] and image plane of the system camera [I_B].

For the creation of images attuned to visual perception it is necessary to synchronize the properties of the object cameras and the system camera. While translating the system camera in the three-dimensional scene the position of the object cameras are shifted corresponding to the position of the system camera. Contrary to this, the viewing direction of the object camera stays aligned to its assigned object. The final image is composed of the rendered frames of each camera ordered by the scene depth of the corresponding objects. As shown in figure 7, the image of the object camera is placed on the final image plane at the position it would have on the image plane of the system camera. Due to the different viewing directions the distance between the object and the image plane of the object camera is larger than the distance of the object and the image plane of the system camera. Therefore, the projected object covers a smaller region on the image plane of the object camera than on the respective plane of the system camera. To preserve the real size of the imaged object the projection

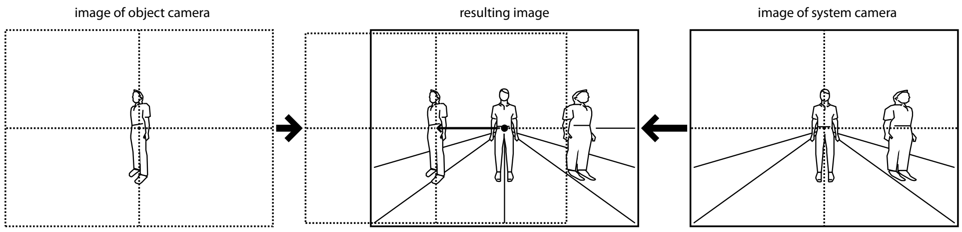

Figure 7: Composition of the image planes of system and object cameras.

plane of the attached camera is shifted towards the object right up to the intersection [S] of the optical axis and image plane of the system camera [I_B] (see figure 6). While translating and rotating the system camera, every object camera must be updated related to the position of the system camera in virtual space, its viewing direction and its order and position of the rendered frame on the final image. The main steps of the algorithm are summed up as follows:

1. Create an object camera and assign an object
2. Align viewing direction of the object camera to the pivot point of the object
3. Translate image plane towards the object
4. Calculate original position of the object on system image plane in screen coordinate system
5. Translate the image to the position on the final image
6. Sort image planes by depth

A result of this multi-perspective rendering process is visualized in figure 4 c. Each colored human shape is rendered with its individual object camera while the scene itself is visualized by the system camera. The proportions of the human shapes are visualized in correct relations. Accordingly, these computer generated images are multi-perspective in order to create a perceptually realistic spatial impression.

4.4 Comparison of Approaches

The conjoint aim of both procedures is providing human-centered interfaces for an efficient and coherent communication of spatial information (Jokela et al., 2003). With two ways for creating perceptually realistic multi-perspective images the purpose of reducing perspective distortion can be achieved. The proportions and the orientations of human figures are preserved (see figure 4 b and 4 c). We conducted a study to identify the benefit of multi-perspective images. The focus was set on the difference between the original orientation of a figure in three-dimensional scenes and its orientation perceived by the viewer.

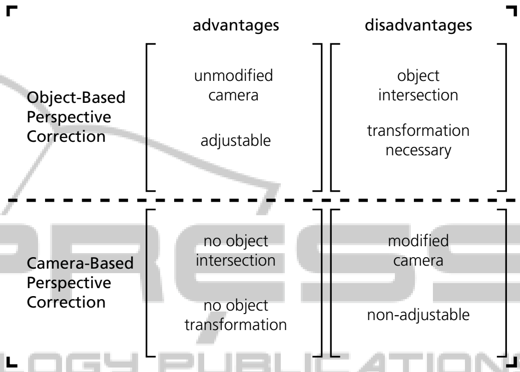

Figure 8: Differences between Perspective Correction procedures.

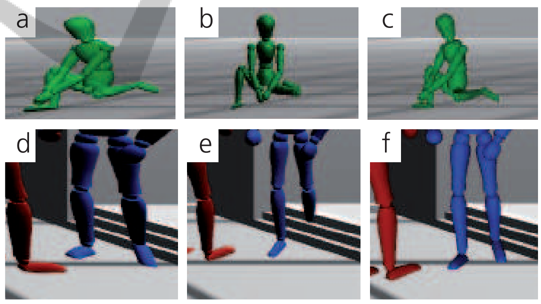

Figure 9: Details of virtual reconstruction of “The Tribute Money” (a+d) rendered with standard computer graphics camera b+e) using OBPC (c+f) using CBPC.

First results show a divergence. This difference depends on the figures original position and orientation. Figure 10 shows a lateral located figure which is rotated by -30 degrees. The perceived orientation in a mono-perspective image differs significantly from the orientation of the figure in the scene. In contrast, the perceived orientation in a multi-perspective image adapts the factual orientation approximately.

The following section discusses each Perspective Correction procedure related to the their application in interface design. To investigate both procedures in a realtime application the algorithms were realized in our framework for experimental three-dimensional computer graphics, called *Bildsprache*

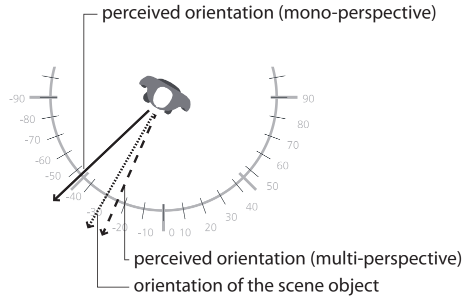

Figure 10: Divergence in orientation perception of figures in mono- as well as multi-perspective images from the orientation of the figure in the three-dimensional scene.

LiveLab, based on C++ and OpenGL. The OBPC algorithm uses an unmodified camera to create multi-perspective images. The geometric transformations are executed solely on scene objects. The applicability is also ensured in realtime applications. Another advantage is the possibility of an unrestricted geometrical transformation of objects. It is possible to adjust the algorithm steps shear and rotation independently. The adjustability of this transformation steps can be useful for adapting the proportion and orientation of avatars in respect of the context. For example some scene configurations require to counteract unwanted intersections of objects. This is illustrated by the right human figure's leg [B] that intersects the stair in figure 9 e.

A further aspect is the orientation of a human figure towards the image plane and other human figures. Examining the left person [C] in the original painting (figure 2) compared to its perspective reconstruction (figure 4 a) reveals that beside the distortions the orientation is falsified by the perspective projection. A second analysis of figure 4 b (in detail 9 b) shows that the adjustability of the OBPC allows a viewer-orientated depiction of person [C] just like in the painting by Masaccio (see figure 2).

In contrast, the CBPC is more restrictive; its parameters are less flexible. The functional dependency on the camera system only permits an absolute adjustment. As a constraint, this procedure is not sensitive to the context. An adjustable manipulation of proportion and orientation of human figures is not possible. Nevertheless, this method enables to correct the proportions of human shapes in a perceptually realistic manner (Yankova and Franke, 2008).

Compared with the object-based approach the camera-based algorithm causes considerably less object intersection. The reason is a geometrically unmodified scene. For example, there is no intersection between the leg of person [B] and the stair in the scene (cp. figure 9 f). This is achieved by arranging several image layers. The differences between the perspective correction procedures are summarized in figure 8. The comparison shows that both algorithms are appropriate for creating perceptually realistic images similar to Renaissance painting like "The Tribute Money". Both procedures complement each other in essential parts. They can be used on different objects in virtual scenes simultaneously without any interference.

5 CONCLUSIONS AND DISCUSSION

In the present paper computer graphics procedures to generate perceptually realistic images of three-dimensional scenes were presented and compared. Perceptual realism is achieved by creating multi-perspective images based on the rules of perspective projection enhanced by characteristics of visual perception and techniques of Renaissance painting.

The scenario of visualization of humanoid figures in which an appropriate depiction of proportions and orientation are of particular importance was chosen to underline the benefits of perceptually realistic images. Multi-perspective in the image was identified to emulate the human viewing behavior. A comparison of Renaissance painting and computer generated images revealed a insufficiency in depiction of humanoid figures in computer graphics. Subsequently, two methods to create multi-perspective images were introduced: an Object-Based and a Camera-Based Perspective Correction. Although these procedures differ in their way of proceeding to solve distortion and misalignment. They can be used in virtual worlds simultaneously without any interference to obtain a more perceptually realistic result.

6 FUTURE WORK

The general concept of Perspective Correction (Object- and Camera-Based) can enhance all kinds of visualization systems displaying human shapes and closed curved surfaces in general. It is suitable for contexts that require a wide visual range and a high amount of comparability of all objects simultaneously. This is found, for example, in ergonomic simulations, video games or in virtual training simulations. Yet, the context of the scene is not considered automatically. Our next step will be the adaptation of the algorithms based on the scene. This can be made by the semi-automatic adjustment of the parameters of the OBPC. This may conduce to perceptually realistic visualized objects which are tightly fitting in the context of the scene. For the CBPC it is conceivable to shift the object cameras. Thereby, a customizability similar to the object-based procedure can be achieved. As a result, the camera-based procedure would be sensitive to the context.

ACKNOWLEDGEMENTS

Jan Wojdziak and Martin Zavesky thank *European Social Fund (ESF) / the European Union and the Free State of Saxony*. Rainer Groh, Ingmar S. Franke and Kerstin Kusch thank *Deutsche Forschungsgemeinschaft (DFG)* Wahrnehmungsrealistische Projektion von dreidimensionalen Szenen (WaRP, DFG-GZ:GR 3417/1-1).

REFERENCES

Agrawala, M., Zorin, D., and Munzner, T. (2000). Artistic multiprojection rendering. In *Proceedings of the Eurographics Workshop on Rendering Techniques 2000*, pages 125–136.

Angel, E. (1997). *Interactive Computer Graphics - A top-down approach with OpenGL*. Addison-Wesley.

Bailenson, J. N., Yee, N., Merget, D., and Schroeder, R. (2006). The effect of behavioral realism and form realism of Real-Time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *Presence: Teleoperators & Virtual Environments*, 15(4):359–372.

Capin, T., Pandzic, L., Thalmann, N. M., and Thalmann, D. (1999). *Avatars in Networked Virtual Environments*. John Wiley and Sons Ltd.

Carroll, R., Agrawala, M., and Agarwala, A. (2009). Optimizing content-preserving projections for wide-angle images. In *SIGGRAPH'09:ACM SIGGRAPH 2009 papers*, pages 1–9.

Coleman, P. and Singh, K. (2004). Ryan: rendering your animation nonlinearly projected. In *NPAR '04: Proceedings of the 3rd international symposium on Non-photorealistic animation and rendering*, pages 129–156.

Foley, J. D. (1999). *Computer Graphics-Principles and Practice*. Addison-Wesley.

Franke, I. S., Pannasch, S., Helmert, J. R., Rieger, R., Groh, R., and Velichkovsky, B. M. (2008). Towards attention-centered interfaces: An aesthetic evaluation of perspective with eye tracking. *ACM Trans. Multimedia Comput. Commun. Appl.*, 4:1–13.

Franke, I. S., Zavesky, M., and Dachselt, R. (2007). Learning from painting: Perspective-dependent geometry deformation for perceptual realism. In *13th Eurographics Symposium on Virtual Environments (EGVE)*.

Franke, I. S., Zavesky, M., and Rieger, R. (2006). The power of frustum - die macht der geometrischen mitte. In *NMI 2006 "Film, Fernsehen und Computer" Neue Medien der Informationsgesellschaft*.

Fu, Y., Li, R., Huang, T. S., and Danielsen, M. (2008). Real time multimodal human avatar interaction. *IEEE Trans. CSVT*, 18(4):467–477.

Groh, R. (2005). *Das Interaktions-Bild - Theorie und Methodik der Interfacegestaltung*. TUDpress Verlag der Wissenschaft.

Groh, R., Franke, I. S., and Zavesky, M. (2006). With a painter's eye: An approach to an intelligent camera. In *Proceedings of The Virtual 2006*.

Hockney, D. (2006). *Secret Knowledge:Rediscovering the Lost Techniques of the Old Masters*. Thames & Hudson Ltd.

Jokela, T., Irvari, N., Matero, J., and Karukka, M. (2003). The standard of user-centered design and the standard definition of usability: Analyzing iso 13407 against iso 9241-11. In *CLIHC 03: Proceedings of the Latin American Conference on Human- Computer Interaction*, pages 53–60.

Singh, K. (2002). A fresh perspective. In *Proceedings of Graphics Interface 2002*, pages 17–24.

Ware, W. R. (1900). *Modern perspective: A treatise upon the principles and practice on plane and cylindrical perspective*. MacMillian and Co.

Yankova, A. and Franke, I. S. (2008). Angle of view vs. perspective distortion: a psychological evaluation of perspective projection for achieving perceptual realism in computer graphics. In *APGV'08:Proceedings of the 5th symposium on Applied perception in graphics and visualization*, pages 204–204.

Yu, J. and McMillan, L. (2004). A Framework for Multiperspective Rendering. In *Rendering Techniques 2004, Proceedings of Eurographics Symposium on Rendering*, pages 61–68, 408.

Zelnik-Manor, L., Peters, G., and Perona, P. (2005). Squaring the circles in panoramas. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, pages 1292–1299.

Zorin, D. and Barr, A. H. (1995). Correction of geometric perceptual distortions in pictures. In *SIGGRAPH '95: Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 257–264.