Advance Access publication on 19 April 2013
doi:10.1093/iwc/iwt027

# Dynamic Anamorphosis as a Special, Computer-Generated User Interface

ROBERT RAVNIK, BORUT BATAGELJ, BOJAN KVERH AND FRANC SOLINA*

*Faculty of Computer and Information Science, University of Ljubljana,*
*Tržaška cesta 25, 1000 Ljubljana, Slovenia*
**Corresponding author: franc.solina@fri.uni-lj.si*

**A classical or static anamorphic image requires a specific, usually a highly oblique view direction, from which the observer can see the anamorphosis in its correct form. This paper explains dynamic anamorphosis which adapts itself to the changing position of the observer so that wherever the observer moves, he sees the same undeformed image. This dynamic changing of the anamorphic deformation in concert with the movement of the observer requires from the system to track the 3D position of the observer's eyes and the re-computation of the anamorphic deformation in real time. This is achieved using computer vision methods which consist of face detection and tracking the 3D position of the selected observer. An application of this system of dynamic anamorphosis in the context of an interactive art installation is described. We show that anamorphic deformation is also useful for improving eye contact in videoconferencing. Other possible applications involve novel user interfaces where the user can freely move and observe perspectively undeformed images.**

RESEARCH HIGHLIGHTS

- We explain how the classical concept of anamorphosis can be extended to computer-generated images so that the images adapt dynamically in real time to a moving observer.
- A simple method for observer localization is proposed.
- The use of the concept was demonstrated in the context of an art installation.
- With the help of two experiments we show that the concept might improve eye contact in videoconferencing.

*Keywords: intelligent user interfaces; computer vision; interactive systems and tools; human computer interaction (HCI)*

*Editorial Board Member: Gitte Lindgaard*

*Received 22 November 2012; Revised 19 March 2013; Accepted 25 March 2013*

## 1. INTRODUCTION

Anamorphosis is a distorted projection requiring the viewer to occupy a specific view point or to use special devices to see the undistorted image. The first type of anamorphosis or perspectival anamorphosis which was developed during Renaissance (fifteenth century) requires just a proper positioning of the observer to gain the correct viewpoint from which he can see the undeformed image (Seckel, 2004). The other type of anamorphosis called a mirror or catoptric anamorphosis was developed about two centuries later and requires a conical or cylindrical mirror to be placed on a precise position on the flat distorted image so that the reflection from the mirror shows the image undeformed. The word anamorphosis comes from the Greek words *ana* (back, again) + *morphoun* (to shape).

In this paper, we show how the principle of perspectival anamorphosis can be extended to dynamic anamorphosis which can adapt itself to the changing position of the observer in such a way that he always sees the image in its correct undeformed form. Dynamic anamorphosis as a new concept was first described in 2007 (Nacenta *et al.*, 2007; Solina and Batagelj, 2007). First, we discuss more thoroughly the classical principle of anamorphosis in Section 2. Next, in Section 3 we describe the principle of dynamic anamorphosis. A mathematical formulation of the anamorphic deformation is given and methods for the localization of the observer are discussed. The anamorphic deformation of the displayed image is computed as a planar homography. The most unobtrusive methods for localization of the observer use computer vision methods that find human faces in images. Section 4 describes the implementation of our system for dynamic anamorphosis. In Section 5, we describe the interactive art installation which gave the motivation to develop the system for dynamic anamorphosis and discuss other possible applications of dynamic anamorphosis, in particular how anamorphosis can help in maintaining better eye contact between teleconferencing partners. Conclusions are provided in Section 6.

## 2. ANAMORPHOSIS

Our mind is constantly interpreting and giving structure to the raw visual input from our eyes. We prefer an ordered world, familiar shapes and regular patterns. One of such features of human perception is that our brain tends to order visual features in a regular, orderly, symmetric and simple manner, as formulated by the Gestalt school in psychology (Koffka, 1935). Therefore where possible, we see stable rectangular forms although these forms appear most of the time distorted due to perspective projection and are also constantly changing due to our movement. This principle is called shape constancy (Pizlo, 1994).

Perspectival anamorphosis or anamorphic projection was discovered in art in the late fifteenth century both as a challenge and as a confirmation of the rules of linear perspective which were discovered at the same time (Collins, 1992a,b). Classical linear perspective is based upon the Euclidean paradigm that light travels in straight lines and when light reflected from an object intersects a planar surface an accurate representation of the original object is reflected on that surface. While we normally look at images frontally from a limited range of viewing angles, the viewer of an anamorphic image must usually be at a radically oblique angle to the picture plane to see the anamorphic image undistorted. The anamorphic image looked at up front is in such cases usually so distorted as to be unrecognizable.

Probably the most famous example of anamorphosis in art history is the 1533 painting *The Ambassadors*, by Hans Holbein (Fig. 1). On the bottom of this painting appears a diagonal blur which appears as a human skull when viewed from the upper right (Topper, 2000).

Perspectival anamorphosis is also closely related to illusionistic or trompe l'oeil painting. Perspective construction is used in both cases to create an image which is seen correctly just from a particular viewpoint. Perspective anamorphosis is usually seen correctly from an unconventional viewpoint and is from a standard viewpoint usually so distorted to be almost unrecognizable as in the example in Fig. 1. An illusionistic painting, on the other hand, presents an invented image, which from a standard viewpoint looks as if it were reality. The best-known examples are Baroque ceiling paintings that from a particular viewpoint make a flat ceiling look like extending into imaginary architecture of domes, towers or the heaven itself. A 3D illusion based on perspective construction is the Ames room (Gregory, 1970).

**Figure 1.** The *Ambassadors* by Hans Holbein, 1533, Oil on oak, 207 × 209 cm, National Gallery, London. The diagonal blur on the bottom appears as a human skull when viewed from the upper right.

Special examples of anamorphic projection are also deformed images or signage which after being projected on slanted surfaces (typically on pavements in front of stores) appear undeformed. Traffic signs painted on the road surface are also often in reality elongated so that from the perspective of a traffic user approaching the sign they more readily appear in the right proportion. Nowadays, anamorphic chalk images are even produced as pavement or sidewalk art, often as part of numerous street painting festivals or advertising and publicity campaigns (Beever, 2010; Stader, 2010; Wenner, 2010). Owing to the ease of producing anamorphic images using computer graphics they appear now often in newspapers and magazines.

Virtual advertising inserts images such as signs, brand logos or even product packages into live or previously taped television programs. Live sports, in particular, are surrounded with multiple billboards that can be overlaid with new images. The overlaid images must be first deformed so that they can be precisely aligned to the actual billboards which are perspectively deformed when seen from a general viewpoint. In this way, the overlaid new images appear as if they are actually present in the scene. Virtual advertisements can readily be changed and different advertisements can be played for different markets across the globe.

Since the appreciation of anamorphic images requires an *eccentric* viewing point as opposed to a *normal* or orthogonal viewing point, anamorphosis is a term popular with many postmodern theorists used mainly as a metaphor for the relativity of vision or the subjectivity of human experience (Topper, 2000). Anamorphosis serves as a model for the concept of the gaze, which suggests that visual appreciation rather than passive *looking* requires active *observing* (Lacan, 1978; Žižek, 1989). To appreciate an anamorphic image requires indeed from the observer that he positions himself precisely in the right spot and directs his gaze in the right direction as opposed from the *normal* or *centric* vision (Arnheim, 1984) where the viewer sees himself at the center of the world and as he moves, the center of the world moves with him and the world surrounding him stays coherent. Further philosophical considerations of anamorphosis are nicely summed up by Massey (1997).

To view an anamorphic image one has to transform an oblique and non-uniform focal plane into a coherent, 2D image which is sometimes facilitated by viewing it with one eye or with half-closed eyes (Collins, 1992b). This enables the dissociation of the image from the screen or the supporting surface and the anamorphosis re-forms itself (Topper, 2000). Viewing *normal* pictures from an oblique angle does not result in a distorted picture since human perception can automatically compensate for the distortion using the principle of shape constancy (Cutting, 1987; Todorović, 2008). Straying away from the right viewpoint of an anamorphic image, on the other hand, can quickly deteriorate the effect. A person viewing a *normal* picture from an arbitrary viewpoint must treat separately two perspective effects on the image that he perceives: the perspective deformation due to oblique viewing *and* the perspective deformation that is due to the content of the picture. Namely, if the pictorial perception would depend only on the geometry of the projected retinal image, the perception of the depicted space would be deformed in comparison to the actual depicted pictorial space (Goldstein, 1987).

The exact mechanism that supports correct space perception from deformed retinal images is still disputed in the human perception research community. Cutting (1987), for example, explains that the visual system corrects the distortions based on the assumption that objects are rigid. Sedgwick (1993) gives a theoretical analysis based on the concept of available visual information. More recent research in human perception has shown that the adjustment to oblique viewing is achieved before the contents of the image are interpreted (Vishwanath *et al.*, 2005). The adjustment to oblique viewing is based on the local slant of the picture which can be estimated from binocular disparity and the perspective of the picture frame. When viewing at the picture's surface from a very oblique position so that the image slant is larger than 60°, the estimation of the local slant becomes uncertain and the adjustment for oblique viewing does not occur (Vishwanath *et al.*, 2005). This explains why the dissociation of the supporting surface and the image is so important for the anamorphic effect to materialize. If there is a frame around the anamorphic image it should also be deformed so that it supports the anamorphic effect.

In perception of pictures viewed at an angle it is important to distinguish two different perceptual phenomena (Goldstein, 1993): the perception of the layout in 3D space of objects represented in the pictures and the direction a pictured object appears to point when extended out of the picture into observer's space. The perception of spatial layout depicted in a picture remains relatively constant with changes in the viewing angle of the observer. The second phenomenon, the perceived orientation of objects pointing out more or less perpendicularly from the picture plane, undergoes large changes with changes in viewing angle. Objects on pictures that point directly out of the picture appear to rotate so that when the observer moves relative to the picture they keep pointing directly at the observer. Objects on pictures that point to the side, however, rotate less, and they do not maintain a constant direction relative to a moving observer. This is a manifestation of the differential rotation effect (DRE) (Goldstein, 1987).

## 3. DYNAMIC ANAMORPHOSIS

To see a static anamorphic image one has to position oneself on the right spot and then view the image in the right direction. Nowadays, computer technology is often used for display of images, especially moving images. Since one can project the anamorphic image using a video projector which is connected to a computer or to use a large computer monitor, we can reshape the anamorphic image whenever the observer moves in such a way that the re-formed image stays the same for the observer. Nacenta *et al.* (2007) named this new capacity perspective-aware interface while we referred to it independently and at about the same time as dynamic anamorphosis (Solina and Batagelj, 2007). To achieve this constancy of the re-formed anamorphic image one has to track the position of the observer in real time and then according to the established position, pre-deform the projected anamorphic image in real time so that it appears un-deformed from that particular view point. Dynamic anamorphosis or perspective-aware interface is therefore a combination of observer localization and image warping that adapts itself to the changing observer location.

The methodology and technology to achieve anamorphic deformation is known and straightforward. Anamorphic deformation of an image is in fact just an application of image warping (Wolberg, 1998). Owing to recent advances in computer technology real time dynamic anamorphosis of video imagery can be achieved now also in practice using standard computer equipment. Tracking of the observer's head can be achieved reliably by wearing passive or active sensors on the head (Nacenta *et al.*, 2007). We decided to use in lieu of sensors a face detection method that can reliably detect faces of observers in images. Based on the position and size of the face in the image the corresponding 3D position of the observer can be determined. Owing to the fact that most computer monitors are now equipped with built-in cameras this method is unobtrusive and much more affordable.

### 3.1. Related work

A somewhat similar concept involving imagery that adapts to the position of the viewer is described by Mann (2001). The observer wears special eyeglasses that track where the person is and then the system generates stabilized images on displays to sustain the illusion of a transparent window showing the subject matter behind the display in exact image registration with what one would see it if the display were not present.

Nacenta *et al.* (2007) developed E-conic, a new perspective-aware interface intended particularly for multi-display environments. In a multi-display environment it is difficult to assure that the user or that all users in a multi-user scenario are positioned perpendicular to all display surfaces. Even when a single user is looking at a large flat monitor from close proximity he is not able to see all parts of the screen from a perpendicular direction. Looking at an oblique angle makes viewing, reading and manipulating information due to perspective distortion more difficult. Nacenta *et al.* (2007) therefore used anamorphic deformation to address this difficulty. They conducted a systematic empirical study where they compared the performance of human subjects at typical computer interface tasks such as targeting, steering, copying/aligning, pattern matching and reading from an oblique viewpoint first on undeformed displays and then on anamorphically deformed displays. The experiment showed that the anamorphic deformation improves the user performance by 8% to 60% depending on the task. To track the heads of the users E-conic used an ultrasonic 3D tracker with the sensor placed in a baseball cap.

A related group of authors (Hancock *et al.*, 2009) also studied the effects of changing projection geometry on the interpretation of 3D visualizations on tabletop displays. Since users of a tabletop display can look at the displayed 3D information from all sides of the table, it is for the perception of the scene very important if the point of view and the center of projection are aligned or not. An empirical experiment showed that errors in judging an object's orientation were increasing with the increasing discrepancy between the point of view and the center of projection. Errors were even bigger if the displayed 3D objects were perpendicular to the tabletop display giving rise to the DRE effect.

Elvira Vreeswijk (2010) developed a concept for a 3D display based on catoptric anamorphosis for her graduation project at Leiden university in 2007. Anamorphically deformed imagery is projected by a video projector on a flat surface in the middle of which stands a cylindrical mirror. The mirror resolves the anamorphic deformation so that the image on the mirror is seen correctly. A camera-based motion tracking system determines the position of the single user who can move freely around the display so that the projected anamorphic image can be rotated into a corresponding position and that the user can see in the mirror all the time the same undeformed image.

Sander ter Braak (2010), a digital designer, created for the final project at the Utrecht School of the Arts a project that he referred to as augmented anamorphosis. He projects an image of a 3D cube on the floor. The actual projected shape changes depending on the position of the observer so that the perceived 3D shape of the cube remains stable when the observer walks around the cube. This is a nice example of augmented reality which works in actual 3D space but is observable only from a singular view point and therefore in practice useful just for a single user. Details on how the position of the observer is determined have not been published.

Recent computer-controlled video projection systems have one or more built-in cameras to provide a visual feedback that can automatically compensate for the so called *keystone* deformation. The keystone deformation can be represented in the most general way as a planar homography mapping points in the projector plane onto the screen plane, corresponding to a 3-degrees of freedom (DF) alignment (pan, tilt and screw) (Brazzini and Colombo, 2005). To eliminate the effect of the keystone, its associated homography can be estimated and used to suitably pre-deform the image being displayed. The same homography can be used to make a *virtual anamorphosis* so that the image is seen undeformed only for observers looking at the screen from a particular viewpoint (Brazzini and Colombo, 2005). The authors call this functionality directional vision and compare it to directional audio. We use this homography to deform the projected image in such a way that it looks undeformed from the viewpoint of the observer.

To enable dynamic anamorphosis we therefore need two crucial components, anamorphic deformation of the image and localization of the observer's eyes in space in real time.

### 3.2. Anamorphic deformation

Let us assume that the real wall, on which the image of size $w \times h$ pixels will be projected, lies on the plane $z = 0$. Further, let us assume that the origin of our coordinate system is in the center of the computer screen, which is projected to the wall. Before the transformation, the projected image is assumed to have its center aligned with the coordinate system origin and to extend from $-0.5$ to 0.5 (or less if we want to preserve the image width and height ratio) in both $x$ and $y$ directions.

We want to find a transformation that will transform the image plane ($z = 0$) in such a way, that the vector from the center of the observer's eyes to the center of the projected image on the imaginary wall will be perpendicular to the transformed plane (see Fig. 2). Computation of the anamorphic deformation of an image proceeds in two steps. The image is first transformed to the imaginary wall with the normal vector parallel to vector from the observer's eyes. The image from the imaginary wall is then projected back (bold line) to the real wall in order to make it look like it was being displayed on the imaginary wall.

The described anamorphic transformation can now be formalized as follows:

(1) We take the vector from the center of the displayed image $I$ (the origin) to the center of observer's eyes,

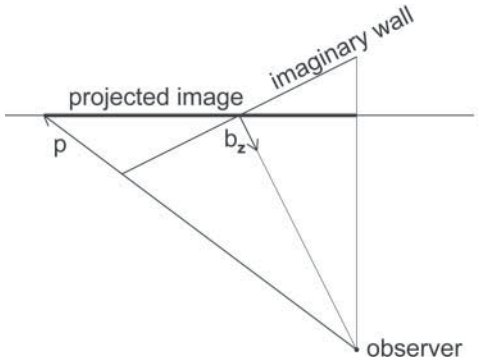


**Figure 2.** Computation of anamorphic deformation.

which is $(v_x, v_y, v_z)$, and normalize it. Let us denote this vector as $\mathbf{b_z}$ (Fig. 2).

(2) We calculate $\mathbf{b_z} \times (0, 0, 1)$ (where $\times$ stands for cross product between two 3D vectors) and normalize it. Let us denote this vector as $\mathbf{b_y}$. This vector is actually the axis around which the image has to be rotated.

(3) We calculate vector $\mathbf{b_x} = \mathbf{b_y} \times \mathbf{b_z}$.

(4) Now we can construct a matrix $\mathbf{R}$, that will align vector $\mathbf{b_y}$ with $(0, 1, 0)$ as follows:

$$\mathbf{R} = \begin{bmatrix} \mathbf{b_x} \\ \mathbf{b_y} \\ \mathbf{b_z} \end{bmatrix}. \quad (1)$$

Think of $\mathbf{b_x}$, $\mathbf{b_y}$ and $\mathbf{b_z}$ as $1 \times 3$ submatrices in this case.

(5) We construct a matrix, which will rotate the coordinate system around axis $\mathbf{b_y}$ for the same angle as it is between vectors $(0, 0, 1)$ and $\mathbf{b_z}$ as follows:

$$\mathbf{P} = \begin{bmatrix} c & 0 & -s \\ 0 & 1 & 0 \\ s & 0 & c \end{bmatrix}, \quad (2)$$

where $c = \mathbf{b_z} \cdot (0, 0, 1)$ and $s = (0, 0, 1) \cdot \mathbf{b_x}$ ($\cdot$ stands for dot product),

(6) finally, we need a matrix that will align vector $(1, 0, 0)$ back with $\mathbf{b_x}$. The inverse of $\mathbf{R}$ will perform this operation, but since $\mathbf{R}$ is a rotation matrix, we can calculate $\mathbf{R}^{-1}$as $\mathbf{R}^{\mathrm{T}}$.

(7) multiplying all three matrices together, namely $\mathbf{T} = \mathbf{R}^{\mathrm{T}} \cdot \mathbf{P} \cdot \mathbf{R}$, we obtain the matrix, which transforms the original image to the imaginary wall.

For each point of the image $I$ (let us denote it with $(i_x, i_y, i_z)$), we can now calculate its position on the imaginary wall, by multiplying its vector $\mathbf{i} = (i_x, i_y, i_z)$ by matrix $\mathbf{T}$. To find out where this point should be projected back at the original wall in order to make an image look like it was shown on the imaginary wall, we need to take a vector which goes from the observer's eyes to $\mathbf{T} \cdot (i_x, i_y, i_z)$ and extend it up to the original wall. This can be achieved with the similar triangles principle as

$$\mathbf{p} = (\mathbf{T} * \mathbf{i} - \mathbf{v}) \cdot \frac{v_z}{v_z - t_z}, \quad (3)$$

where $\mathbf{v} = (\mathbf{v_x}, \mathbf{v_y}, \mathbf{v_z})$ and $t_z$ is the $z$ component of vector $\mathbf{T} * \mathbf{i}$. Vector $\mathbf{p}$ is the same as shown in Fig. 2. Coordinates of projected points in world coordinate system can thus be obtained from vector $\mathbf{v} + \mathbf{p}$.

If we would like the observer to see a picture of the same size, regardless of his distance from camera, we can use the built-in features of graphical libraries. They support showing the scene in perspective projection, which requires setting a few parameters to specify the type of projection precisely. We can specify the distance of each graphical element from the projection plane and the objects which are far away, will appear smaller than those, which are close to the projection plane. In our case, we should ensure that $d_{\mathrm{o}} + d_{\mathrm{p}}$ stays constant, where $d_{\mathrm{o}}$ is the distance of the observer from the camera and $d_{\mathrm{p}}$ is the distance of the plane, containing the picture, to the projection plane. So when the observer comes closer to the camera ($d_{\mathrm{o}}$ becomes smaller) we need to increase $d_{\mathrm{p}}$, meaning pushing the plane, containing the picture, further away from the projection plane and vice versa.

### 3.3. Localization of the observer

To drive the dynamic anamorphic projection we need to know the position of the observer's eyes so that if the observer looks at the anamorphic projection he or she sees the image on the anamorphic projection un-deformed. Several techniques can be employed for tracking the location of a human head from which the position of eyes can be reliably estimated. The person can wear some active or passive devices which are used to determine location of the observer as was done in the E-conic system (Nacenta *et al.*, 2007). Typically, special markers are worn by the observer on predefined body parts. Considering the orientation and position of detected markers the observer's location can be calculated. Passive markers facilitate solving the correspondence problem if stereo computer vision methods are used. For this task, the observer could wear eyeglasses equipped with markers so that the location of the eyes could be determined very accurately. Such approach gives accurate results but it is rather impractical for everyday usage.

Less obtrusive methods to determine the position of objects in a given scene can also be provided by computer vision. A general system setup for the described problem could consist of two or even more cameras which eye the observer. By using the principle of stereo reconstruction of distances we can further determine the position of the user's head in 3D space. The most difficult problem in stereo reconstruction is the correspondence problem—to find for a given point in the first image the corresponding point in the second image (Faugeras, 1993). Since the number of possible matches goes into thousands of points this is a computationally intensive task.

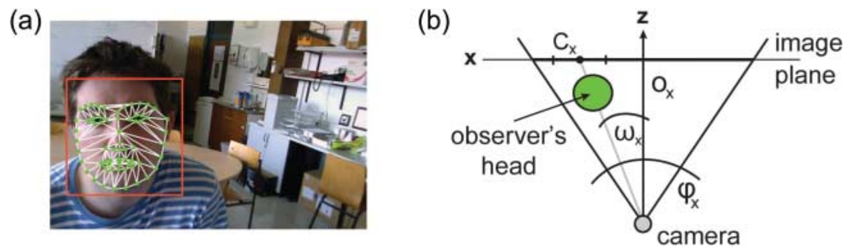


**Figure 3.** Observer localization. (a) Face detection, tracking and registration is performed on the camera image. (b) The horizontal plane $(x, z)$ is shown, where the horizontal angle $\omega_x$ between the camera's optical axis $z$ and the line through $C_x$, the projection of the face center onto axis $x$ is computed. $\varphi_x$ is the camera's horizontal angular field of view.

Since we would like to use the proposed system also on a standard desktop or laptop computer which usually has a single built-in camera we propose instead a simple yet effective system for localization of the observer's face using just a single camera. We employ face detection and tracking algorithms to estimate the observer's position relative to the camera. Facial movement tracking is an important technology in the study of human–computer interaction in general (Ward, 2004). The acquired facial image is then registered in order to obtain facial feature points, such as nose and eyes coordinates (Fig. 3a). Since adult faces are size-wise uniform we can estimate the distance between an observer and the camera.

This approach to face localization can be used at short distances when a user sits behind a computer as well as for larger distances when a user stands on the far side of a room. The maximum distance up to which the method works depends on the camera's field of view and image resolution.

### 3.3.1. *Face detection*

We use a face detection method to determine the position of the user's face in the pictorial plane. Face detection is now a mature technology and methods such as the one developed by Viola and Jones (2004) can run in real time. The processing is done as follows: an input image is scanned at all possible locations and scales by a sub-window. Face detection is posed as classifying the pattern in the sub-window either as a face or a non-face. The face/non-face classifier is learned from face and non-face training examples using statistical learning methods. For our purpose we used the AdaBoost learning-based method because it is so far the most successful in terms of detection accuracy and speed. The hit rate of this face detection method is reported to be 98% (Lienhart and Maydt, 2002). With the proposed face detection method we obtain the location in the image plane of all the present faces regardless of their position and scale down to the size of $20 \times 20$ pixels (Fig. 3a).

### 3.3.2. *Face tracking and registration*

Face tracking and registration of observer's face is performed using an active appearance model (AAM) method. The AAM simultaneously models the intrinsic variation in shape and texture of a deformable visual objects as a linear combination of basis modes of variation. Although linear in both shape and appearance, overall, AAMs are nonlinear parametric models in terms of the pixel intensities. Fitting an AAM to an image consists of minimizing the error between the input image and the closest model instance; i.e. solving a nonlinear optimization problem (Matthews and Baker, 2004; Saragih and Göcke, 2009). A registered facial image of an observer is shown in Fig. 3a.

### 3.3.3. *Estimating observer's 3D position*

First, we determine the view direction towards the registered face relative to the camera coordinate system. The view direction is determined by two angles, $\omega_x$ in the horizontal plane and $\omega_y$ in the vertical plane.

After detection of face $F$ in the image plane $(x, y)$ and its registration we obtain the position of 66 facial feature points. Each eye is described with six feature points that form a convex polygon around the eye orbit. We calculate the centroid of this polygon in order to estimate the location of observer's eye center. We denote centroid points of left and right eye as $E1$ and $E2$. Now we can calculate the center point $C$ of face $F$ as midpoint between $E1$ and $E2$.

We introduce two more parameters: camera's horizontal angular field of view $\varphi_x$ and camera's vertical angular field $\varphi_y$ whose values are camera-specific and can be obtained from technical specifications.

The horizontal angle $\omega_x$ in the horizontal plane $(x, z)$ and the vertical angle $\omega_y$ in the vertical plane $(y, z)$ that determine the direction of the detected face from the camera's optical axis $z$ can now be calculated as

$$\omega_x = \varphi_x * (C_x/I_{\text{width}} - 0.5), \quad (4)$$

$$\omega_y = \varphi_y * (C_y/I_{\text{height}} - 0.5), \quad (5)$$

where $I_{\text{width}}$ and $I_{\text{height}}$ denote width and height of camera image $I$ in pixels, respectively. The geometrical scheme of the described calculus in the horizontal plane is presented in Fig. 3b.

After obtaining angles $\omega_x$ and $\omega_y$ we try to estimate the distance between the observer and the camera. Since most adult human faces share a similar or almost the same interpupillary distance (IPD) (Dodgson, 2004), we presume for this estimation of observer's distance that the IPD is a constant. According to

Dodgson (2004) more than 90% of adults have IPD between 57 and 69 mm, with a mean around 63 mm. We can calculate IPD as Euclidean distance between $E1$ and $E2$:

$$\text{IPD} = \|E1 - E2\|. \tag{6}$$

Distance between the face and the camera is inversely proportional to the observer's IPD. Based on this consideration, we propose the following distance estimation function $\hat{F}_{\text{dist}}$:

$$\hat{F}_{\text{dist}}(x) = Ax^{-1} + B, \tag{7}$$

where $A$ and $B$ are camera-specific constants that depend on the field of view of the camera lens and the resolution of the camera picture sensor and $x$ is the estimated IPD in pixels. The camera that we used throughout all of our experiments reported in this article is a Logitech WebCam Pro 9000 with a horizontal field of view 63.1°, a vertical field of view 49.4° and an image resolution of 1600 × 1200 pixels. We experimentally set $A$ to 21528.8 and $B$ to −7.78 using standard numerical packages.

The accuracy of the distance estimation function $\hat{F}_{\text{dist}}$ was evaluated by comparing the computed distances with the actual measured distances between the observer and the camera using a tape meter at several distances between 0.3 and 10 m. Two participants were asked to stand in front of the camera and consecutively move away from it. After each move, the distance between the camera and a participant increased by 10 cm and a picture of the scene was taken. Fig. 4 presents experimentally measured distances and distances computed using $\hat{F}_{\text{dist}}$ based on IPD.

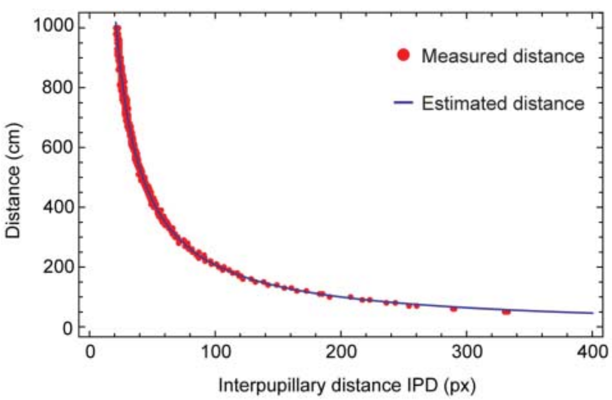


**Figure 4.** Experimental evaluation of distance estimation function.

If the observer of the image needs to keep eye contact with the person on the displayed anamorphic image, it is important that the generated anamorphic image resolves into the correct direction and that the directional error is smaller than 5° (Stokes, 1969). For the camera configuration that we used, if the face of the observer is 0.5 m away from the camera, a 4-cm shift in lateral direction is already enough to reach the 5° limit for losing eye contact. At a distance of 3 m the corresponding shift grows to 26 cm and at 10 m the permissible lateral error of position can be up to 87 cm before we lose the eye contact. The error in direction $z$ influences only the scale of the displayed image which does not have any effect on the possibility of establishing eye contact. Fortunately, while the error of position estimation grows with the distance from the camera, the need for lateral accuracy decreases, both due to the same geometrical consideration about the diverging spatial angles and image quantization. A change of position of the observer closer to the camera requires larger anamorphic deformations than when the same change happens farther away.

After estimating the distance between the observer and the camera as well as angles $\omega_x$ and $\omega_y$, we have all the information needed to estimate the observer's position in 3D space. The described method operates in real time and achieves sufficient accuracy for the given problem.

## 4. IMPLEMENTATION

Presented concepts were implemented in a standalone application. System uses face detection to determine the approximate 3D position of the observer and uses it to calculate the anamorphic deformation of the projected image thus making it look undistorted to the observer regardless of his position (see Fig. 5).

The face detection part uses OpenCV library (Baggio *et al.*, 2012), while the image is rendered with the help of Mesa library (Surhone *et al.*, 2010). The application runs on a machine with Intel Core2 Quad Q9000, 2.00 GHz with 4 GB of RAM using NVidia Quadro FX 2700M graphic card. The application is suitable for real-time environments.


**Figure 5.** (a) Undistorted predetermined image. (b) Anamorphic projected image corresponding to viewer position (10, 0, 2). To see the anamorphic image undeformed, look at it from the right side with the face just above the plane of the paper.

The dynamic nature of our application requires constant face tracking, observer localization and appropriate anamorphic deformation of the image. Even though the face tracking algorithm is reliable, there are situations where it can lose track of the person for several reasons. The observer can move quickly to another position or even outside the area covered by the camera or he can just turn his face away from the camera and thus become invisible to the face-tracking algorithm. To prevent rapid large changes of the anamorphically deformed image and to smoothen the changes of the anamorphic deformations we employ the following rules that try to cover such cases.

When the position of the observer changes significantly, we use an interpolation between the previous and currently detected position, so that the anamorphically deformed image for the latest detected position reaches its final form in a few iterations. If the position of the face changes again during this accommodation sequence, the system redirects the interpolation towards the new position. This handles the cases of quick moves and possible false localizations of the face. If the observer moves outside of the camera view or turns away from the camera, the application interpolates towards the initial position $(0, 0, z)$. Similarly, when the observer appears in the camera view and his face is detected, the application performs the interpolation between the initial position and the observer's position.

In practice, another important constraint is also the medium for displaying the anamorphically deformed image. If a computer display is used then the maximum angle under which one can still observe the image on the display depends largely on the display technology. This viewing angle constraint is not as severe if we use projected images.

## 5. APPLICATION OF DYNAMIC ANAMORPHOSIS

Our interest in anamorphosis started with images of faces. We have also designed the described system for applying dynamic anamorphosis based on face detection of the observer. If a face is displayed on the observed image, the question of its gaze direction and the possible eye contact with the observer follows quite naturally. People are very sensitive to the direction of the eye gaze. Our eyes express our emotions and intentions and they help us direct attention (Langton, 2000). Cultural norms dictate when, for how long and in what situations it is appropriate to gaze into another person's eyes. We can determine very accurately if somebody is actually looking at us. Eye gaze is important in conversations and the lack of proper eye contact, for example, in videoconferencing systems, is a serious limitation of such systems (Colburn *et al.*, 2000).

Gaze direction is a vector pointing from the fovea of the looker through the center of the pupil to the gazed-at spot. An intuitively obvious and simple cue to determine eye gaze direction is the dark-white configuration or the position of the pupil in the visible part of the eye. However, Todorović (2006) demonstrated clearly that the perception of gaze direction depends not only on the eye turn but also on the head turn of the looker. The observer-related gaze direction is according to Todorović simply an additive combination of observer-related head orientation and looker-related gaze direction. With this geometrical basis for gaze perception he explained two well-known phenomena of gaze direction: the Mona Lisa effect and the Wollaston effect.

The Mona Lisa effect, named after Leonardo's famous painting, is actually an instance of the DRE. The DRE occurs not only on objects that extend in depth from the picture plane towards the observer but also on perceived gaze direction. Goldstein (1987) demonstrated the well-known observation, known already in ancient times, that 'when a straight-on face is looking directly at an observer, its eyes will rotate to follow the observer so that they appear to be looking directly at the observer no matter where he or she is relative to the picture'. However, Goldstein also demonstrated that, in accordance to DRE, the more the perceived gaze is looking to the left or to the right, the less it follows the observer when he moves around.

The Wollaston effect demonstrates that the perceived gaze direction of a portrait depends not only on the position of the irises but also on the orientation of the head (Todorović, 2006). For example, given an image of a face with the face turned leftwards and the irises rightwards, the resulting gaze is directed towards the observer. If you make a mirror duplicate of that image, turning the face consequently rightwards, and then pasting into the mirrored face in place of the mirrored irises the original rightwards looking irises, the resulting gaze also shifts rightwards. The Wollaston effect thereby indicates that it is possible to change the perceived gaze direction without any manipulation of the irises but only by turning the head of the looker.

### 5.1. Interactive art installation

The Computer Vision Laboratory where we developed the described system for dynamic anamorphosis has a long tradition of using the latest information technology in fine arts, in particular, we are interested in using computer vision methods in the context of interactive art installations (Peer and Batagelj, 2009). We started to develop the concept of dynamic anamorphosis also in the context of an art installation (Solina and Batagelj, 2007). The subject of the projected image in this installation is a human face looking straight ahead (Fig. 6). We used a short excerpt of the Big Brother from the film *Nineteen Eighty-Four* (1956) after George Orwell's novel of the same name published in 1949. The Big Brother personifies complete surveillance by the authorities of all members of the society, mainly using television screens with the image of the Big Brother. The anamorphic deformation on Fig. 6 resolves at the viewer position (5, 1.5, 2)—move your head to the right and up so that the upper-right corner of the image is the closest to your eyes. The installation was housed in a dark room where

**Figure 6.** Anamorphic deformation of the Big Brother.

the only illumination comes from video projection. The face of a single observer is illuminated by the projection itself and stands out clearly from the dark background.

The resolution of the camera for face detection and localization of the observer should be adjusted to the size of the room as described in Section 3.3.3. When setting up the installation, all dimensions of the room must also be taken into consideration, for example, what is the maximum size of the surface available for the projection. Oblique viewing requires that very deformed and elongated images must be projected. It may be necessary to restrict the movement of the observers only to those parts of the room where consistent effects can be achieved. Since the installation can truly be experienced only by a single user the entrance to the room with the installation should be controlled. Another possible scenario for the exhibition would allow several people in the audience. Of course, only one of them must somehow be selected for the virtual anamorphic experience using either face recognition or some other distinguishing feature used for identification. Other people in the room could enjoy the projection which will be deformed from their viewpoint or they will try to move into the position from which the anamorphosis will re-form itself.

In contrast to traditional perspectival anamorphosis which requires an accurate, often *eccentric* viewpoint, this installation uses anamorphosis to separate the human spatial orientation from the visual cues and can thus provoke a crisis in the visual faculty—wherever the observer moves in space, he sees the same re-formed image.

In the installation the projected face with the eye gaze turned directly ahead will meet the eyes of the installation user. Owing to the viewpoint-sensitive anamorphic deformation, the projected face in the installation will stare at the installation user wherever he moves. There will be no way to escape its gaze. This should be for most installation users a rather unnerving situation. On a symbolic level the installation epitomizes the personification of ubiquitous video surveillance systems (Levin *et al.*, 2002).

According to the Mona Lisa effect, a user of the installation should encounter the same experience of the Big Brother following him or her with his gaze everywhere even without any anamorphic deformation of his image. However, the installation users related that their experience was somehow more direct, especially since the frame of the projection also remains stable during their movement so that they could not deduce their position in the room from the perspective deformations of the picture frame.

### 5.2. Enabling better eye contact in videoconferencing

Video-teleconferencing enables people to communicate face-to-face over remote distances. One of the major unresolved issues of teleconferencing concerning the user experience is the loss of eye contact between participants of a teleconferencing session (Colburn *et al.*, 2000). Eye contact seems to be so important especially in individual-to-individual communications that poor eye contact may have hurt a wider adoption of videoconferencing technology since people associate poor eye contact with deception (Bekkering and Shim, 2006). Gaze patterns in general provide an extremely important and rich set of serial signals which should be taken into account for videophone design (Bruce, 1996). While a person's eyes are usually directed at the center of the computer screen where the teleconferencing partner's face is displayed, the cameras are usually placed above the display. Studies have shown that a video link where the camera is mounted above the computer monitor is less trusted than a centrally mounted camera but even less trusted than just a voice connection or email (Bekkering and Shim, 2006). It has been shown, that if the vertical distortion angle between the line from the camera to the eyes and the line from the eyes to the screen is more than 5° the loss of eye contact is noticeable (Stokes, 1969). Using average-sized desktop computer displays at a normal viewing distances, this angle is usually between 15° and 20° (Yip and Jin, 2004), which results in an inevitable loss of eye contact. An example is shown in Figs 7 and 8. The same problem also arises if the face of a video conference correspondent is displayed in a smaller window on a large screen if the angle between the correspondent's face in the window and the position of the camera is larger than 5°. Users of small handheld devices do not experience these problems since the angle between the center of the display and the camera measured from the usual viewing distance of the user's face is usually smaller than 5°.

This problem of gaze awareness has been known and addressed for many years. Initially, hardware-based solutions were proposed ranging from using half mirrors, beamsplitters and integrating the camera into the center of the computer screen. However, these solutions to align the camera axis with the center axis of the computer monitor require specialized systems which are expensive and generally not accessible. With increased computer power, software solutions of the problem involving the manipulation of the image itself were proposed (Gemmell *et al.*, 2000). To generate a virtual camera view from the middle of the computer screen anywhere from two (Ott *et al.*,

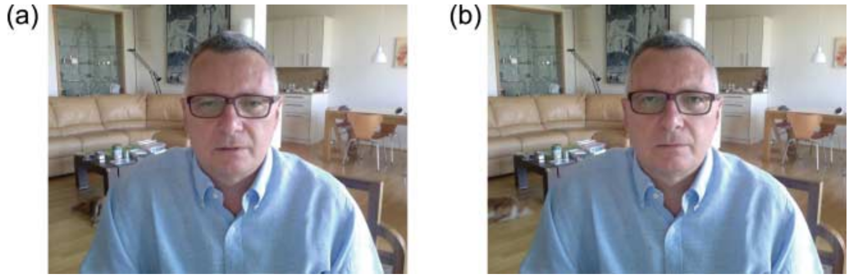


**Figure 7.** Images of a user captured by the camera mounted above a 27 inch computer display at a normal working distance: (a) in the left image the user is looking into the middle of the display, (b) in the right image the user is looking straight into the camera. Image (a) on the left illustrates the problem of the missing eye contact in teleconferencing: when the user is looking at the middle of the display, where the eyes of his teleconferencing partner are displayed, his partner on the other side sees image (a) where the user appears to be looking downwards and not directly in the partner's eyes such as in image (b).

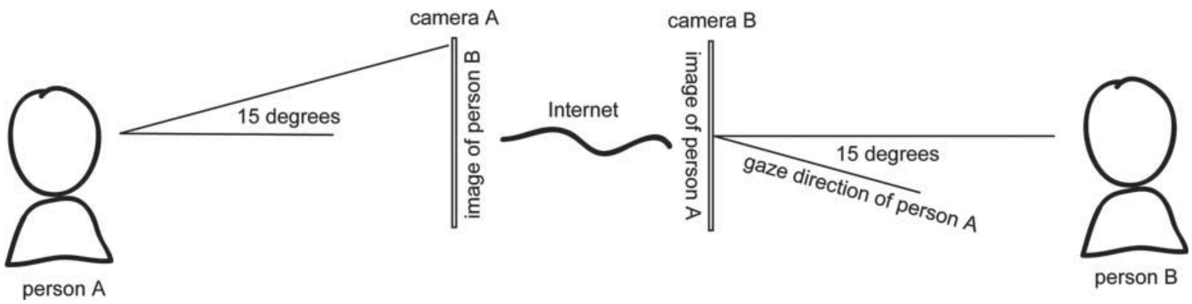


**Figure 8.** Person A is looking at person B on screen in front of him (left). Because camera A is mounted above the computer screen, person B, when looking at image of person A, cannot meet his eye gaze (right).

1993) to eight cameras have been proposed (Dumont *et al.*, 2008). These systems are based on stereo matching and image morphing methods. Proposed was even a single camera system which requires separate rectification of the face and eyes with affine transformation (Yip, 2005). Since users now expect to use videoconferencing on their own desktop or portable computers all these solutions seem to be overly complicated.

Microsoft researchers first reported that informal observation suggests that it might be possible to change the eye gaze by rotating the image if the face is initially not looking at the viewer, but concluded that more work needs to be done to figure out the exact relationship between rotating an image and eye gaze direction (Zitnick *et al.*, 1999). We elaborated on this idea further. Based on our work involving dynamic anamorphosis we proposed a very simple way of providing a better eye-contact experience in videoconferencing (Solina and Ravnik, 2011). It should be stressed that all methods that try to alleviate the problem of missing eye contact, including our proposed method, do not force teleconferencing partners into eye-contact but just make it possible when and if the partners want to engage into eye contact.

If we rotate the image of our videoconferencing partner for a moderate angle around axis $x$ with the top of the picture moving away from us we still perceive the partner as before since our human perception estimated the amount of the rotation to correct for the perspectival deformation (Fig. 9). We observed that the eye gaze of the video partner which in the original, not rotated image, is directed below our face (as in Fig. 7a) seems also to rotate. When the amount of the rotation of the image plane is appropriate, observers report a better eye contact (Fig. 10).

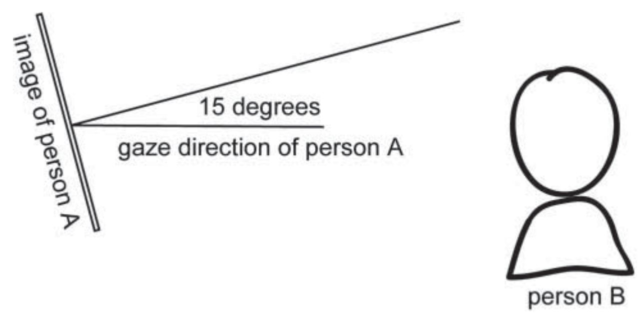


**Figure 9.** By rotating the image of person A, A's perceived eye gaze direction also seems to rotate so that person B now has a better subjective experience of having eye contact with person A.

### 5.3. Experiment

To test our hypothesis we performed the following experiment. We took still pictures of four different people who were sitting

**Figure 10.** The image of the videoconference participant in Fig. 7a was anamorphically transformed so that the image plane was rotated around axis $x$ for 15°. According to the experiments that we made, this may improve the subjective eye contact experience in comparison to the original non-rotated image.

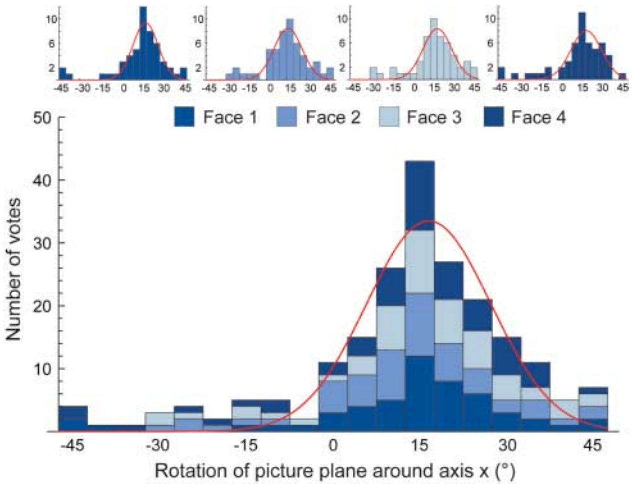


**Figure 11.** Results of testing if the rotation of the image of an assumed videoconferencing partner around axis $x$ can improve the subjective perception of eye contact.

in front of a 27 inch monitor and looking towards the center of the monitor with a camera mounted above the monitor, as was the case in Fig. 7a. We devised a web application which applies an anamorphic deformation of the selected picture around axis $x$ which is for small angles similar to the rotation of the image around axis $x$.

We asked a group of 54 mostly undergraduate and graduate students to visit the experiment's web page http://papagena.fri.uni-lj.si/dyana/ to rotate the pictures in such a position where they experience the best eye contact with the person on the image. The time that the subjects needed to perform the orientation was not limited. Two of the subjects responded that no amount of rotation could improve eye contact. The results of the other 52 subjects are summarized in Fig. 11.

The histogram of votes suggests that an increase of eye contact satisfaction can be achieved by rotating the image of the videoconferencing partner so that the top of the image moves away from us. 80% of the votes selected this approach to improve eye contact. The peak of the histogram shows that the required amount of rotation is around 15°. When we sit in front of a 27 inch monitor 60 cm away our eye gaze directed to the center of the screen also changes the angle of direction for about 15° when we look at the camera mounted above the display. Forty-two percent of votes selected angles between 10° and 20° and 60% of votes selected angles between 5° and 25° (Fig. 11).

### 5.4. Follow-up experiment

To better assess whether the simple action of rotating the image of the videoconferencing partner for 15° around axis $x$ really improves the eye contact we performed the following additional experiment. Using the same equipment (27 inch monitor, built-in camera above the screen) we took still images of nine people (five females and four males) looking into the center of the screen where the face of their video conferencing partner would normally appear and a second image when they were looking into the camera above the screen—similar to Figs 7a and b.

By means of a web application we were asking participants in the test group to select among two displayed images of the same person the image that gave him or her a better subjective perception of eye contact with the assumed teleconferencing partner. There were four types of images that we were evaluating for eye-contact in this experiment (Fig. 12): (a) image of a person looking into the center of the screen, (b) image of a person looking into the camera above the screen, (c) image of a person looking into the center of the screen, rotated for 15° around axis $x$ and (d) image of a person looking into the center of the screen, rotated for 15° around axis $x$ and with vertically chopped off left and right edges. Since we compared pairwise all four types of images, there were six image combinations for each of the nine persons in the test group. Images of persons looking straight into the camera (b) were included as a reference, since these images should offer the best eye-contact. Images of a person looking into the center of the screen rotated for 15° around axis $x$ (c) should, according to our hypothesis and the results of the initial experiment (Section 5.3), offer better eye-contact than images of type (a). Finally, to better understand the role of the perspectively deformed picture frame in the perception of the rotated pictures we included in the comparison also images rotated for 15° around axis $x$ where the left and right perspectively deformed edges were then vertically chopped off (Fig. 12d).

The experiment was conducted over the World Wide Web on the following way. Each participant of the test was asked to open the web application http://papagena.fri.uni-lj.si/dyana2/ and to switch to a full screen mode during the test. Participants were then guided through 18 image pairs, each pair representing one

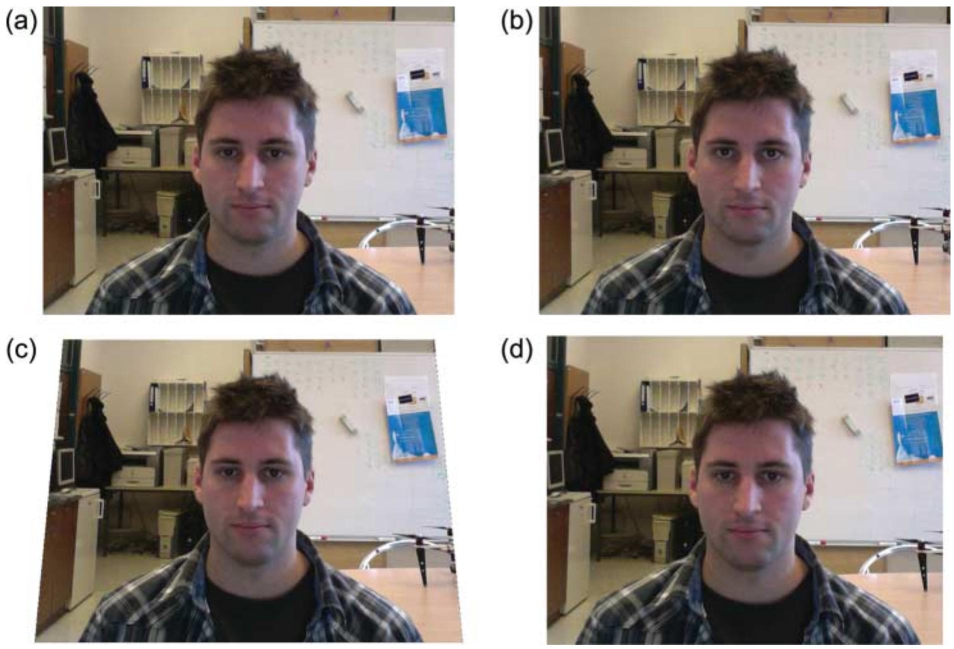


**Figure 12.** Four types of images used in the follow-up experiment: (a) image of a person looking into the center of the screen, (b) image of a person looking into the camera above the screen, (c) image of a person looking into the center of the screen, rotated for 15° around axis $x$, (d) image of a person looking into the center of the screen, rotated for 15° around axis $x$ with vertically chopped off left and right edges. Images of nine persons (4 males and 5 females) were used in the experiment.

of the six possible combinations of the four images of the same person (Fig. 12). For each image pair, the participant was asked to select among the two images the one which gave him or her a better experience of eye-contact. If a participant could not decide on any of the two images because the difference was so small, he could select the third option—cannot decide. To remove any bias, the sequence of 18 image pairs was generated for each test participant individually in the following way: the sequence of persons appearing on test images was selected randomly and each among the nine imaged persons appeared in an image pair exactly two times, among the 18 image pairs each type of image pair combination out of the six possible combinations appeared three times. The left/right ordering of each image pair was also generated randomly. Therefore, in none of the testing sequences the same image pair could appear twice.

In this follow-up experiment 229 participants voted on the experiment's web page in a single day. The participants of this experiment did not overlap with the participants of the first experiment and a large majority of them was also not familiar with our hypothesis on how to improve eye-contact. Descriptive results of voting are summarized in Table 1. The columns show: the comparison number (Cmp), percentage of votes for designative answers $(i)$ and $(j)$ and percentage of undecided votes (Undecided).

Statistical analysis was performed separately for each of the six comparisons. The data of each comparison were collected in a $229 \times 3$ table, where rows represent participants and columns denote the cumulative number of votes for each possible answer. Each cell therefore contained a numerical value between 0 and 3 (inclusive), giving the row sum of 3. A new data column $\Delta_{i-j}$ was defined as a difference between cumulative votes of two non-undecided answers $(i)$ and $(j)$. We calculated the values of $\Delta_{i-j}$ by subtracting from the number of cumulative votes of answer $(i)$ the number of cumulative votes of answer $(j)$. Values of $\Delta_{i-j}$ are between $-3$ and 3 (inclusive) and express the inclination of each participant towards one of the two non-undecided answers. The closer the $\Delta_{i-j}$ value is to zero, the less considerable is the difference between answers.

The detailed results of the statistical analysis of the follow-up experiment are presented in Table 2. Note that the order of comparisons is the same as in Table 1. The columns present: the comparison (Cmp), the difference formula ($\Delta_{i-j}$), the number of participants ($N$), average $\Delta_{i-j}$ value (mean), and standard deviation of the mean (SD). Results of the one sample $t$-test ($\alpha = 0.05$, two-tailed) are presented with $t$-value, DF and the calculated $p$-value.

In the case of comparison 1 where participants were comparing an image of a person looking into the center of the screen (a) and an image of a person looking into the camera (b), we observe a high mean bias of 2.40 towards option (b) (see Table 2, comparison 1, column 4). The bias significance was also confirmed using the one sample two-tailed $t$-test when

**Table 1.** Voting results of the follow-up experiment where 229 participants selected which image in a given image pair offered a better subjective feeling of eye-contact

| Cmp | Which image gives a subjectively better feeling of eye-contact? Answer ($i$) | Undecided | Answer ($j$) |
|---|---|---|---|
| 1 | (a): image of a person looking into the center of the screen<br>8% | Undecided<br>3% | (b): image of a person looking into the camera above the screen<br>89% |
| 2 | (a): image of a person looking into the center of the screen<br>19% | Undecided<br>45% | (c): image (a) rotated for 15° around axis $x$<br>36% |
| 3 | (a): image of a person looking into the center of the screen<br>21% | Undecided<br>54% | (d): image (a) rotated for 15° around axis $x$ with vertically chopped off edges<br>25% |
| 4 | (c): image (a) rotated for 15° around axis $x$<br>12% | Undecided<br>2% | (b): image of a person looking into the camera above the screen<br>86% |
| 5 | (d): image (a) rotated for 15° around axis $x$ with vertically chopped off edges<br>10% | Undecided<br>3% | (b): image of a person looking into the camera above the screen<br>87% |
| 6 | (d): image (a) rotated for 15° around axis $x$ with vertically chopped off edges<br>21% | Undecided<br>48% | (c): image (a) rotated for 15° around axis $x$<br>31% |

Since there were for each person, whose images were used in the experiment, four different image types (Figs 12a–d), six possible image pair combinations for comparison were possible. Images of nine people were used in the experiment.

comparing the mean of $\Delta_{b-a}$ to 0 (see Table 2, comparison 1, columns 6 and 8).

The second comparison addresses our basic hypothesis that rotating the image around axis $x$ (c) should improve the experienced eye-contact over an un-rotated image (a). Although almost a half of the participants could not decide which image is better (see Table 1, comparison 2), the rest of the participants prefer the rotated image almost twice as often (36%) as the original image (19%). Statistical significance in means was also confirmed with the one sample two-tailed $t$-test (see Table 2, comparison 2, columns 6 and 8) indicating $t$-value and $p$-value of 4.66 and $p < 0.0001$, respectively.

If the perspectively deformed left and right edges of the rotated image are vertically chopped off (d), the same statistical test shows no significant difference between the selection of the original image (a) and the rotated chopped off image (d) (see Table 2, comparison 3).

Comparisons 4 and 5 that involve images of persons looking into the camera (b) show, as expected, a high bias towards (b) and a statistically significant preference in eye-contact perception in relation to (c) and (d) images (see Table 2, comparison 5 and 6, column 8).

Analysis of comparison 6 between rotated images (c) and cropped rotated images (d) indicates a trend in favor of (c).

**Table 2.** Statistical analysis of the follow-up experiment.

| Cmp | $\Delta_{i-j}$ | $N$ | Mean | SD | $t$-value | DF | $p$-value |
|---|---|---|---|---|---|---|---|
| 1 | b–a | 229 | 2.40 | 1.02 | 35.80 | 228 | <0.0001 |
| 2 | c–a | 229 | 0.49 | 1.59 | 4.66 | 228 | <0.0001 |
| 3 | d–a | 229 | 0.12 | 1.37 | 1.35 | 228 | 0.1774 |
| 4 | b–c | 229 | 2.22 | 1.16 | 29.05 | 228 | <0.0001 |
| 5 | b–d | 229 | 2.31 | 1.13 | 30.88 | 228 | <0.0001 |
| 6 | c–d | 229 | 0.29 | 1.67 | 2.61 | 228 | 0.0096 |

One sample $t$-test (two-tailed) shows significant mean difference between designated answers in comparisons 1, 2, 4, 5 and 6.

Calculated $p$-value ($p = 0.0096$) is not as distinct as in other comparisons but still satisfies significance criterion (see Table 2, comparison 6, column 8).

### 5.5. Discussion

Viewing pictures from a reasonable oblique angle does not result in a distorted picture since human perception can automatically compensate for the distortion using the principle of shape constancy (Vishwanath *et al.*, 2005). When a picture is viewed from its center of projection it generates the same retinal image as the original scene, so the viewer perceives the scene correctly. When a picture is viewed from other directions, the retinal image changes, but we normally do not notice the change. This invariance or shape constancy is in human perception achieved through estimation of the local surface orientation (Vishwanath *et al.*, 2005). The DRE should be noted again since it makes objects pointing out of the picture, also including gaze directions, follow the moving observer (Goldstein, 1987). The Wollaston effect is also important for the discussion because it demonstrates that gaze direction can be manipulated by changing either the position of the iris or the position of the head of the looker (Todorović, 2006).

Most studies of eye gaze perception experimented with left and right movements of the eyes and not with the up and down movements that we are addressing. According to Sedgwick (1993) moving up-down is from a geometrical point of view the same as moving left-right relative to the observed picture. However, due to the elongated shape of the eye, the possible up-down movement of the iris is much smaller than the left-right movement.

The results of our experiment are in the light of published psychophysical studies on perception of gaze direction and of slanted pictures somewhat surprising. According to Todorović (2006), correcting the eye gaze just by rotating an image should not be possible: ‘if the looker’s gaze misses me from one vantage point by a particular angle, specified by a particular combination of head and eyes cues, then it will miss me from most any vantage point by a similar angle.’

The observer-related gaze direction is an additive combination of observer-related head orientation and looker-related gaze direction. A possible underlying mechanism for gaze determination is that the visual system first independently extracts iris eccentricity and head orientation information from the looker’s head, and then combines these two measures to form the gaze direction judgment (Todorović, 2006). This explanation is also supported by the fact that reaction times for judgements of gaze direction are shorter when the eyes and the head are turned in the same direction than when they are turned in different directions (Langton, 2000). This means that the gaze direction is determined independently and in parallel from the eyes and the head and then combined. These two sources may either be congruent or incongruent, leading to corresponding acceleration or de-acceleration of gaze processing (Todorović, 2006). The Wollaston effect tells us that it is possible to manipulate with the perceived gaze direction by changing either the position of the iris or the head individually, or both at the same time, but in different amounts so that the relation between the iris and the head are changed.

Based on the above deliberation, our hypothesis for the mechanism behind the results of our experiment is as follows: by rotating the picture, the DRE is employed for the perceived gaze direction which is estimated only from the iris eccentricity, and the face/head of the looker. Since the perceived gaze direction “extends” out of the picture more than the head, the gaze seems to turn according to DRE more than the face itself. This difference in perceived rotation introduces a perceived change in the head/eyes relation, causing in line with the Wollaston effect a change in the perceived gaze direction of the looker. The perceived gaze direction of the looker in Figs 10 and 12d which was determined only on the looker’s eyes therefore rotates according to DRE much more than the face of the looker. The final perceived gaze direction is then assembled after the DRE effect out of the rotated eye-based gaze direction and the head-based gaze direction. Since in our case the head of the looker is in frontal orientation, the head orientation does not have a large influence on determining the perceived gaze direction anyway.

There is even an earlier psychophysical experiment that can be interpreted in support of our findings. Goldstein (1987) performed an experiment (no. 3 in the article) to study how the viewing angle affects the perceived gaze direction of portraits looking directly at the observer and for six frontally faced portraits that appear due to iris eccentricity to be looking to the left or right of the observer. Goldstein used drawings of faces in the experiment. For the portrait looking directly at the observer the results are exactly as predicted by DRE, for all viewing angles the perceived gaze direction is the same as the viewing angle, meaning the portrait looks directly into the observer. The faces with other gaze directions, however, rotate less for the same change of viewing angle. For example, for the portrait that is looking 15° to the left and is observed frontally (from 0°), the perceived gaze direction is 30° to the left, if the same portrait is observed from 20° to the left, the perceived gaze direction is 45° to the left, if the portrait is observed from 40° to the left, the perceived gaze direction is 60° and if the portrait is observed from 70° to the left, the perceived gaze direction is 75°. Because the perceived gaze direction rotates slower than the direction of observation, the original offset between the looking direction and the perceived gaze direction at the frontal position is decreasing when the observer moves in the direction of the perceived gaze. When the direction of observation is the same as the perceived angle direction, ideal eye contact could be established. The experiment demonstrates that if the observer tries to reach the perceived gaze direction of the looker on a picture by moving or rotating his gaze into the direction of the looker's gaze, the difference between the two gaze directions decreases. This strategy was also used in our experiment. Note again that Goldstein performed his experiment for left-right movements of the head and the eyes, while the experiment described in this article treats up and down movements.

There are still many open questions whether this strategy of following the perceived gaze of the videoconferencing partner by rotating his image can be really successful. Goldstein's experiment described above maybe hint that the difference in view directions can be in this way only decreased but never completely eliminated. Since both experiments described in this article were conducted over the Internet, we could not control the type of device (desk-top or hand-held), the size of the computer monitor that the participants in the experiment were using, nor the position of the participants relative to the monitor. People whose images were used in the experiments were also not very strictly guided how to sit behind the monitor and how to look at it. These inter-personal differences also contributed to differences in results. The initial experiment allowed that the participants rotated the image themselves into a position that they preferred. In the follow-up experiment where the participants were given a fixed pre-rotated image that they could perceive as better, equal or worse in regard to perceived eye-contact the observed effect was not so strong as in the first experiment but still statistically significant. Therefore, any user interface that would use this effect should preferably offer a rotationally adjustable image. What perceptual mechanisms really underlay the observations that we made with our two experiments should be explored in the future by carefully designed psychophysical experiments.

### 5.6. Other applications of dynamic anamorphosis

The above problem of missing eye contact in two-way telecommunication can be alleviated by dynamic anamorphosis. Dynamic anamorphosis, however, can improve the one-way communication in the same way as it works in the described art installation. When a user who needs to constantly move in his work place is being addressed by somebody by means of his video image on a computer screen, anamorphic deformation can take care that this person is always *turned* towards the user. Although due to the principle of shape constancy we easily understand images which are perspectively deformed because their image plane is rotated out of alignment with our gaze direction, it would still be beneficial if any crucial information on the computer screen, pictorial or alphanumeric, could be *turned* towards the user. Nacenta *et al.* (2007) have clearly demonstrated that an anamorphic deformation improves targeting, steering, aligning, copying and reading from oblique screens.

Such dynamic 3D alignment of the virtual image plane towards the position of the user is especially appealing for advertising which is already demonstrated by the recurrent use of 3D pavement art for commercial purposes. Gaming is another application area which could benefit from such adaptive viewing experience.

## 6. CONCLUSIONS

We described a system that implements the principle of dynamic anamorphosis as an extension of perspectival anamorphosis. Perspective-aware interface or dynamic anamorphosis was first defined in 2007 (Nacenta *et al.*, 2007; Solina and Batagelj, 2007). The concept can be implemented by combining known and standard methods and techniques. A perspectival anamorphosis can be seen in its true or intended shape only from a particular viewpoint, from other viewpoints it looks deformed or even not discernible. A dynamic anamorphosis adapts itself to the position of the observer so that it can be seen in its undeformed form from almost anywhere in front of the projected image. Classical anamorphosis requires precise positioning of the observer in space while dynamic anamorphosis disassociates the geometric space in which the user moves from the visual cues he sees, since wherever the observer moves, he sees the same image.

Tracking of the observer can be achieved by different means. In the described system we used computer vision techniques to locate the user's face in 3D space and in real time. The anamorphic deformation of the still or video images must also be computed in real time.

The described system was initially developed for an art installation where a human face was projected with its eye gaze directed straight ahead to meet the eyes of the installation user. Due to the dynamic anamorphic projection the gaze of the Big Brother is always frontally oriented towards the viewer giving a stronger impression that it is not possible to escape his view.

The problem of missing eye-contact in video conferencing systems is due to the fact that when people want to engage in eye-contact with the speaking partner they cannot achieve it due to technical limitations of the teleconferencing systems. The normal way of establishing eye contact is to mutually look into the partner's eyes. Since the viewing axes of the cameras and displayed face images on videoconferencing systems are not aligned, this simple strategy does not work and at least subconsciously causes frustration. To solve the problem of gaze awareness we performed two experiments on how anamorphic deformation could be used in videoconference user interfaces. The two experiments demonstrated that by rotating the image of a videoconferencing partner around axis $x$ of the image so that the top of the image moves away from us, the subjective experience of a better eye contact between the videoconferencing partners could be improved. We also offer a possible psychophysical explanation for this effect but these are very preliminary findings that need further investigation. We expect that other possible uses of dynamic anamorphosis based on face tracking for observer localization, especially in multi-display environments will be further explored, especially in advertising and gaming applications.

## ACKNOWLEDGEMENTS

We thank anonymous reviewers for valuable comments and suggestions. We would also like to thank Professor Anuška Ferligoj from University of Ljubljana for her advice on statistical analysis.

## FUNDING

All authors acknowledge support by the Slovenian Research Agency, research program Computer Vision (P2-0214).

## REFERENCES

Arnheim, R. (1984) The Power of the Center: A Study of Composition in the Visual Arts. University of California Press.

Baggio, D.L., Emami, S., Escriva, D.M., Ievgen, K., Mahmood, N., Saragih, J. and Shilkrot, R. (2012) Mastering OpenCV with Practical Computer Vision Projects. Packt Publishing.

Beever, J. (2010) Pavement drawings. Online: http://users.skynet.be/J.Beever/pave.htm. Accessed 23 July 2010.

Bekkering, E. and Shim, J. (2006) Trust in videoconferencing. Commun. ACM, 49, 103–107.

Braak, S. (2010) Augmented anamorphosis. Online: http://vimeo.com/hautetechnique. Accessed 23 July 2010.

Brazzini, A. and Colombo, C. (2005) Computer Vision for Interactive Skewed Video Projection. In Roli, F. and Vitulano, S. (eds.) Image Analysis and Processing - ICIAP 2005. Lecture Notes in Computer Science, Vol. 3617, pp. 139–146. Springer.

Bruce, V. (1996) The role of the face in communication: Implications for videophone design. Interact. Comput., 8, 166–176.

Colburn, R.A., Cohen, M.F. and Drucker, S.M. (2000) The Role of Eye Gaze in Avatar Mediated Conversational Interfaces. Technical Report MSR-TR-2000-81. Microsoft Research, Microsoft Corporation.

Collins, D.L. (1992a) Anamorphosis and the eccentric observer: History, technique, and current practice. Leonardo, 25, 179–187.

Collins, D.L. (1992b) Anamorphosis and the eccentric observer: Inverted perspective and construction of the gaze. Leonardo, 25, 72–82.

Cutting, J.E. (1987) Rigidity in cinema seen from the front row, side aisle. J. Exp. Psychol. Human Percept. Perform., 13, 323–334.

Dodgson, N.A. (2004) Variation and Extrema of Human Interpupillary Distance. In Woods, A.J., Merritt, J., Benton, S. and Bolas, M.T. (eds.) Proc. SPIE, pp. 36–46. SPIE, San Jose, California, USA.

Dumont, M., Maesen, S., Rogmans, S. and Bekaert, P. (2008) A Prototype for Practical Eye-Gaze Corrected Video Chat on Graphics Hardware. In Assunção, P.A.A. and de Faria, S.M.M. (eds.) SIGMAP 2008, pp. 236–243. INSTICC Press.

Faugeras, O. (1993) Three-Dimensional Computer Vision: A Geometric Viewpoint. MIT Press.

Gemmell, J., Toyama, K., Zitnick, C.L., Kang, T. and Seitz, S.M. (2000) Gaze awareness for video-conferencing: a software approach. IEEE MultiMedia, 7, 26–35.

Goldstein, E.B. (1987) Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. J. Exp. Psychol.: Human Percept. Perform., 13, 256–266.

Goldstein, E.B. (1993) Perceived Orientation, Spatial Layout and the Geometry of Pictures. In Ellis, S.R., Kaiser, M.K. and Grunwald, A.J. (eds.) Pictorial Communication in Virtual and Real Environments (2nd edn), pp. 480–485. Taylor & Francis, Inc., Bristol, PA.

Gregory, R.L. (1970) The Intelligent Eye. McGraw-Hill.

Hancock, M., Nacenta, M., Gutwin, C. and Carpendale, S. (2009) The Effects of Changing Projection Geometry on the Interpretation of 3D Orientation on Tabletops. In Proc. ACM Int. Conf. on Interactive Tabletops and Surfaces, pp. 157–164. ACM, New York, NY.

Koffka, K. (1935) Principles of Gestalt Psychology. Harcourt, Brace and World, New York, NY.

Lacan, J. (1978) Four Fundamental Concepts of Psycho-Analysis. W.W. Norton.

Langton, S. (2000) The mutual influence of gaze and head orientation in the analysis of social attention direction. Q. J. Exp. Psychol., 53, 825–845.

Levin, T.Y., Frohne, U. and Weibel, P. (2002) CTRL [Space]: Rhetorics of Surveillance from Bentham to Big Brother. MIT Press, Cambridge, MA.

Lienhart, R. and Maydt, J. (2002) An Extended Set of Haar-Like Features for Rapid Object Detection. IEEE Int. Conf. Image Processing 2002, pp. 900–903. IEEE, Rochester, New York, USA.

Mann, S. (2001) Intelligent Image Processing. Wiley.

Massey, L. (1997) Anamorphosis through descartes or perspective gone awry. Renaissance Q., 50, 1148–1189.

Matthews, I. and Baker, S. 2004. Active appearance models revisited. Int. J. Comput. Vis., 60, 135–164.

Nacenta, M.A., Sakurai, S., Yamaguchi, T., Miki, Y., Itoh, Y., Kitamura, Y., Subramanian, S. and Gutwin, C. (2007) E-conic: a Perspective-Aware Interface for Multi-Display Environments. Proc. 20th Annu. ACM Symp. User Interface Software and Technology, pp. 279–288. ACM, New York, NY.

Ott, M., Lewis, J.P. and Cox, I. (1993) Teleconferencing Eye Contract using a Virtual Camera. INTERACT '93 and CHI '93 Conf.

Companion on Human Factors in Computing Systems, pp. 109–110. ACM, New York, NY.

Peer, P. and Batagelj, B. (2009) Art—a Perfect Testbed for Computer Vision Related Research. In Grgic, M., Delac, K. and Ghanbari, M. (eds.), Recent Advances in Multimedia Signal Processing and Communications. Studies in Computational Intelligence, Vol. 231, pp. 611–629. Springer, Berlin.

Pizlo, Z. (1994) A theory of shape constancy based on perspective invariants. Vis. Res., 34, 1637–1658.

Saragih, J. and Göcke, R. (2009) Learning AAM fitting through simulation. Pattern Recognition 42, 2628–2636.

Seckel, A. (2004) Masters of Deception: Escher, Dalí & the Artists of Optical Illusion. Sterling Publishers.

Sedgwick, H.A. (1993) The Effects of Viewpoint on the Virtual Space of Pictures. In Ellis, S.R., Kaiser, M.K. and Grunwald, A.J. (eds.) Pictorial Communication in Virtual and Real Environments (2nd edn), pp. 460–479. Taylor & Francis, Inc., Bristol, PA.

Solina, F. and Batagelj, B. (2007) Dynamic Anamorphosis. In Luciani, A. and Cadoz, C. (Eds.) Proc. 4th Int. Conf. Enactive Interfaces (ENACTIVE 07), pp. 267–270. ACM, Grenoble, France.

Solina, F. and Ravnik, R. (2011) Fixing Missing Eye-Contact in Video Conferencing Systems. In V. Luzar-Stiler, I. Jarec, Z.B. (Ed.), Proceedings of the ITI 2011 33rd International Conference on Information Technology Interfaces (ITI), pp. 233–236.

Stader, M. (2010) Master of 3D-street painting. Online: http://www.3d-street-art.com. Accessed 23 July 2010.

Stokes, R. (1969) Human factors and appearance design considerations of the mod ii picturephone station set. IEEE Trans. Commun. Technol., 17, 318–323.

Surhone, L.M., Tennoe, M.T. and Henssonow, S.F. (2010) Mesa 3D (OpenGL). VDM Publishing.

Todorović, D. (2006) Geometrical basis of perception of gaze direction. Vis. Res., 46, 3549–3562.

Todorović, D. (2008) Is pictorial perception robust? the effect of the observer vantage point on the perceived depth structure of linear-perspective images. Perception, 37, 106–125.

Topper, D. (2000) On anamorphosis: Setting some things straight. Leonardo, 33, 115–124.

Viola, P. and Jones, M. (2004) Robust real-time face detection. Int. J. Comput. Vis., 57, 137–154.

Vishwanath, D., Girshick, A. and Banks, M. 2005. Why pictures look right when viewed from the wrong place. Nat. Neurosci., 8, 1401–1410.

Vreeswijk, E. (2010) Anamorviesis: An anamorphic 3D solution. Online http://anamoviesis.elviertje.nl. Accessed 23 July 2010.

Ward, R. (2004) An analysis of facial movement tracking in ordinary human-computer interaction. Interact. Comput., 16, 879–896.

Wenner, K. (2010) Artist and architect. Online: http://www.kurtwenner.com. Accessed 23 July 2010.

Wolberg, G. (1998) Image morphing: a survey. Visual Comput., 14, 360–372.

Yip, B. (2005) Face and Eye Rectification in Video Conference using Affine Transform. In ICIP 2005, pp. 513–516. IEEE, Genoa, Italy.

Yip, B. and Jin, J.S. (2004) Pose Determination and Viewpoint Determination of Human Head in Video Conferencing based on Head Movement. In Chen, Y.P.P. (ed.) 10th Int. Multimedia Modeling Conf. pp. 130–135. IEEE Computer Society, Brisbane, Australia.

Zitnick, C.L., Gemmell, J. and Toyama, K. (1999) Manipulation of Video Eye Gaze and Head Orientation for Video Teleconferencing. Technical Report MSR-TR-99-46. Microsoft Research, Microsoft Corporation.

Žižek, S. (1989) The Sublime Object of Ideology. Verso.