

Perspective picture from Visual Sphere: a new approach to image rasterization

Jakub Maximilian Fober*

January 7, 2020

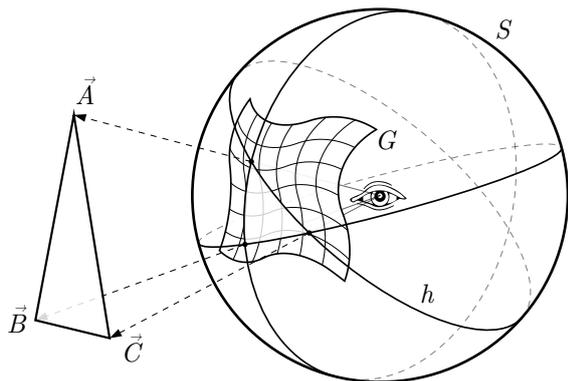


Figure 1: Projection of \overline{ABC} triangle onto image grid G of a visual unit sphere S . Projected edge \overline{CA} is formed by the great circle h .

Abstract

There is great demand for perspective projection model able to produce computer-generated (CG) image up to 360° of view, with lens-distortions, directly from three-dimensional space, to pixel data. Currently there is no practical direct method for rasterization of real-time graphics in curvilinear perspective. Every real-time perspective imagery incorporates *Pinhole Camera* model as a base, some with additional layers of distortion on top. Also to note, knowledge about relationship between motion and perspective has not been properly formulated, leaving a void in that field of image science.

This paper aims at solving those issues. Study involves exploring history of perspective picture, redefining abstract theorem of image (as recorded in common-knowledge), establishing rules of perspective image and presenting new, universal model for perspective projection and rasterization in CG graphics.

Keywords: perspective, non-linear projections, spherical perspective, graphics hardware, mathematics of art

* talk@maxfober.space & http://maxfober.space/

1 Introduction

This paper presents new perspective model. One based on a Gaussian Sphere, able to reproduce image of any shape, perspective geometry and angle of view (AOV). A model capable of combining pictures from sources previously incompatible, like fish-eye camera view and wide-angle lens picture. Based on this model, new rasterization method is presented, one able to render real-time curvilinear image directly from third-dimension.

This work is divided into nine sections, gradually unfolding the topic, from philosophical standpoint, to technical specifics. In following section, the history of perspective picture is introduced, establishing grounds for later discussion. Sections three and four overview perception of perspective picture, its geometry and sensation of motion. Following sections five and six present *Universal Perspective* model. Fifth section relates to mathematics of perspective; present equations of transformation and projection of 2D/3D data and lens distortions. Sixth section presents rasterization process for 3D surfaces. Section number seven refers to generating barycentric coordinates of a triangle in the *Universal Perspective* model. Section eight incorporates *Universal Perspective* model in measurement and simulation of real optical systems, with variable no-parallax point. Section nine followed by references, conclude all information and present direction for further studies. All support graphical material is presented on last five pages.



Figure 2: Rasterized polygon with texture (Paolo Uccello, 15th century) from a perspective map, where $\Omega_d = 270^\circ$, $k = 0.32$, $l = 62\%$, $s = 86\%$.

2 Previous work and history of the topic

Current image abstract theorem was established in 15th century, in a book *De Pictura*, by L. B. ALBERTI. Based on invention of F. BRUNELLESCHI, ALBERTI defined geometrical and theoretical rules for designing perspective projections.² These rules are currently used in CG polygon-based graphics.

Major theoretical statement that laid foundation for image projection technologies and current understanding of image nature can be found in ALBERTI abstract definition of image. He would describe a painting to be like a window in a wall.¹

“First of all, on the surface on which I am going to paint, I draw a rectangle of whatever size I want, which I regard as an open window through which the subject to be painted is seen.”¹

But in times of its discovery, as for now, linear perspective introduced itself with several issues. When there is a need for a wide-angle view, one close to the human visual field, geometrical distortions appear to dominate visual aspect of the picture. Those issues were noticed by Renaissance artists, like L. DA VINCI. He put to test the Alberti theorem and produced paintings of accurate-perspective.⁶ In his *Treatise on Painting*, DA VINCI notes that picture conforms to the idea of a window only when viewed from one specific point in space.⁷ Stating that seen otherwise, objects appear distorted, especially in the periphery. Picture then, viewed from a different point ceases to be like a window in a wall and becomes a visual symbol of an abstract point of view.^a

Some 18th century late Baroque and Neo-classical artists, when encountered those issues, introduced derivative projections. Like G. P. PANNINI with later re-discovered *Panini Projection*,¹⁷ or R. BARKER, who established the term *Panorama*.²⁵ It was a new type of perspective, in a form of cylindrical projection, where abstract window frame became horizontally curved, reducing deformation artifacts in a wide, panoramic depictions of architecture.

Invention of motion picture followed by the rise of film industry, resulted in demand for a new image geometry. Previously still, now pictures had to be pleasing to the eye, in motion. 1950s brought anamorphic cinematography to the wider audience. Lenses like CINEMASCOPE and later PANAVISION¹⁴ became a standard in film production. Figure 3 on page 11 shows example

^aEffect also referred to *Zeeman Paradox*.⁸

of mixed spherical and cylindrical projection in anamorphic lens, with perspective preservation.

Definition. *Conservation of perspective* - lines converging at the optical-axis vanishing-point remain straight (*also see perspective picture definition on the next page*).

CG image technology did not follow film industry progress in that field. Still based on ALBERTI theorem, computer graphics became incompatible with film, generating great costs, when two had to be joined together.²⁰ Which part took lens aberration rotoscopy, where geometry correction has to be performed manually at each frame.

Currently in computer-games industry, CG imagery is practically unable to produce realistic, curvilinear simulation of visual space (VS), or even simulate anamorphic lens geometry, due to limits of linear perspective and resource costs of overcoming those. Some hybrid solutions for real-time graphics were proposed,^{11,12} that combine rasterization with ray-tracing, or tessellation. Such approach allows for a semi-practical and limited production of real-time pictures in a non-linear perspective.

3 On visual space geometry and image perception

Perspective picture visible inside the visual space gives some sense of immersion (e.g. picture, film, computer game) even without visual illusion.⁸ It is perceived as a visual symbol of an abstract point of view. Picture immersion does not break, as long as object’s appearance do not exhibit too much deformation.

Representing point of view invokes separation from the surrounding. To enhance this effect, environment stimuli is being reduced, like in case of movie theater. To uphold the immersion lights are turned off and silence is expected. Or in case of horror-game session played at night, to separate from safe-space of home.

Remark. On the opposite side, picture which is an integral part of the surrounding, can be categorized under the *Trompe l’œil* technique.²³

Since most of the time picture presents point of view (e.g. film, video game, visualization), it’s wise to consider subject’s properties of vision when designing picture’s perspective. But instead of producing mechanical simulation, perspective should symbolize total sensory experience.²

Theorem. *To create immersive visual symbol of a visual space, it is necessary to use the curvilinear perspective instead of linear.*

Proof. Geometry of human visual space contradicts linear perspective principle, as visual field extends beyond linear perspective angle of view. Linear perspective, based on a tangent of an angle, exhibits limit of 179.9° of view. While visual field extends horizontally up to 220° for binocular vision.¹³ ■

At narrow AOV both types of perspective are suitable for immersive picture. In such case differences are negligible. Sub-figure 4a on page 11 presents these differences in comparison to five major perspective projections. Same differences seem exaggerated at higher AOV values (see Sub-figure 4b).

Corollary. *Practical limit for immersive picture in linear perspective is between 60° and 110° AOV. Wider-angles exhibit deformations known as the Leonardo Paradox,⁸ which are then dominant in the image perception and break picture's immersion.* ■

To show wider-angle picture it is necessary to use curvilinear projection. But there is a tendency to see the world not through sight but the understanding. We understand that the wall is flat, therefore we see it that way. Picture projected into the eye is just a visual symbol of a physical wall and has its own physical properties (e.g. perspective and shape). Therefore its visual representation is curvilinear, where curvature symbolizes wider field of view.

Exercise. Reader can validate the curvilinear nature of human visual space,^{4,10} by following A. RADLEY experiment:

*“Also when you have a moment, get a 30 cm ruler (...), and whilst looking forward bring it close to the bottom of your nose, and notice how its shape at the outer edges curves upwards and forwards. It may take you a few minutes to be able to see this effect, because you are so accustomed to not noticing it ! But once you do you will be amazed to see your curved field of view as it really is for the first time”.*¹⁸

Visual space symbolic picture

Problem. Which curvilinear perspective would be best for a visual symbol of visual space?

Proposition. *A model based on the anamorphic lens geometry; a mix between fish-eye, panini and anamorphic projection.*

fish-eye, as it can represent wider AOV, than linear perspective (e.g. π) and conforms to the curvilinear nature of VS.

panini, to symbolize binocular vision; two spherical projections combined into one panoramic image.^b

anamorphic, as cylindrical projections, like *Panini*, tend to elongate proportions vertically; there is a need for correction.

Visual sphere as a whole image

Common idea of an image is limited to a finite, two-dimensional plane. Which is subjective, due to constraints of human visual field and up-front placement of eyes. One can construct a rectangular frame, which at certain distance from the eyes will cover full visual field (VF). In case of some animals (e.g. horse, rabbit), visual space confines much wider VF.^{3,16} With only few blind spots, spanning to almost 360° . Such field cannot be enclosed by a single rectangular frame. Thus image nature is not of a frame. Another model has to be chosen instead. One able to cover full $\Omega = 360^\circ$ is a sphere.

Remark. Cylindrical projections cannot cover full 360° of view, in all directions. They are a hybrid between frame and spherical model. When vertically-oriented, full $\Omega_v < 180^\circ$.

All three-dimensional space around given observation point, can be projected onto a sphere, with given observation point as origin. Even the sphere itself is a 3D object, its surface (as well as image nature)¹⁹ is two-dimensional.

Therefore creating perspective picture, is matter of representing portion of the visual sphere on a flat surface, a fundamental topic in cartography.

Definition. Let us define perspective picture as the azimuthal projection, where lines converging at the optical axis vanishing point remain straight. Such that *conservation of perspective* occurs (see definition on the previous page).

Remark. Each projection of sphere on a flat surface is a compromise and can preserve only some properties (e.g. shape, area, distance or direction), which in case of perspective picture relates to some symbolic information about physical space.

Physical space properties preserved in azimuthal projections

Gnomonic (rectilinear) projects all great circles as straight lines, thus preserving directions. For 3D projection, straight-line in object-space remain straight.

^bEffect also referred as *stereopsis*.

It does not preserve proportions, angles nor area or distances (see figure 5a). Extreme distortion occur away from the center, in form of a radial stretch (see *Leonardo Paradox*)⁸. AOV $\Omega \in (0, \pi)$.

Example. Most common perspective type in painting, 3D graphics and architectural visualization. Sometimes it is used to overemphasize building's appearance by leveraging *Leonardo Paradox*.⁸ Wide AOV combined with lowered optical center creates an effect of acute corners, giving extraordinary look. This technique may confuse the public, as symbolic picture experience won't match building's visual-space appearance.

Stereographic (conformal) preserves angles (at line intersection point). There is no perceivable radial compression, thus smaller figures retain their shape. It does not preserve distances (non-isometric), nor angular surface area. For 3D projection, most important factor is the conservation of proportions (see figure 5b). AOV $\Omega \in (0, 2\pi)$.

Example. In a picture with stereographic projection, actor's face keeps its shape and proportions, even at wide AOV. This projection also gives best spatial sensation (where visual cues are available). Good use case is navigation through tight spaces and obstacles.

Equidistant preserves angular distance from the center point (see figure 5c). For 3D projection, angular speed of motion is preserved. Radial compression remains low-to-moderate at extreme Ω angles. AOV $\Omega \in (0, 2\pi]$.

Example. This projection is recommended for target aiming or radar map navigation, where all targets are projected onto a Gaussian Sphere.

Equisolid preserves angular area. Gives good sensation of distance (see figure 5d). Radial compression is moderate up to π . Near the maximum Ω , compression is high. AOV $\Omega \in (0, 2\pi]$.

Example. When there are no spatial cues, this is best projection to put emphasis on a distance to the viewer.¹² Good use case is flight simulation, where only sky and other aircraft are in-view.

Orthographic-azimuthal preserves planar illuminance. It is a parallel projection of a visual hemisphere. Has extreme radial compression, especially near π (see figure 5e). AOV $\Omega \in (0, \pi]$.

Example. Most commonly found in very cheap lenses, like the peephole door viewer. Thanks to illuminance preservation, it's commonly used in sky photography.²²

4 Image geometry and sensation of motion

Image perspective affects the way motion picture is perceived. It can enhance certain features, like proportions and shapes, movement or spatial sensation. It can also guide viewer's attention to a specific region of image (e.g. center or periphery). Knowledge about these properties is essential for conscious image design.

Attention focusing

Remark. In a film design, there are several techniques to focus viewer's attention on a specific portion of the picture, like *motion*, *color*, *light* and *composition*. Attention focusing through composition and motion is related to image perspective, as its geometry can compress and stretch the image. In composition, *rule of thirds* states that viewer's attention focuses on four corners of a rectangle produced by division of the image into three, equal-size rows and columns. In motion, attention generally drives towards objects approaching the camera or those growing in scale. Attention also focuses on objects entering the picture frame. Same rules apply loosely in reverse, as attention suspense.

Filmmakers tend to frame the image so that region of interest lays in accordance to the *rule of thirds*. In case of computer games, region of interest is usually located at the center, thus viewer must overcome the *principle of thirds* and some properties of *linear perspective* in order to switch attention on that region. In order to focus on the center, games usually incorporate some non-diegetic elements, like crosshair. Such approach may lower immersiveness of symbolic picture.⁵

Attention focusing motion of perspective

Radial stretching and compression are the main attention focusing factors of perspective projection. They give subconscious sensation of movement towards camera, and can amplify figures screen-relative speed of motion.

Gnomonic (rectilinear), due to extreme radial stretch, attention drives towards periphery. When approaching image bounds figures grow in scale and speed (see figure 5a). This combined with motion-sensitive peripheral vision adds to the effect. At wider AOV

amplified motion breaks immersion of symbolic picture.

Stereographic also draws attention towards periphery. Figures grow in scale near bounds, but immersion does not break, as proportions are preserved, even at wide AOV (see figure 5b).

Equidistant drives attention towards the center, as figures in periphery are radially compressed (see figure 5c). This projection preserves screen-relative, radial speed of motion, making it uniform and representative across the picture.

Equisolid also drives attention towards the center, as radial compression is even greater (see figure 5d). Figure speed of motion in screen-space slightly declines towards periphery.

Orthographic has extreme radial compression that breaks immersion of symbolic picture (see figure 5e). When in motion, image seems to be imposed on an artificial sphere.

Gnomonic and Orthographic projections are the two extremes of azimuthal spectrum. They are both least suited for an immersive picture.

Cylindrical perspective type, while symbolizing binocular vision, also gives visual cues for vertical axis orientation. Such cue is undesirable in case of camera roll motion, or when the view is pointing up or down, as image vertical axis will not align with depicted space orientation. In such case perspective geometry should transition from panini to spherical projection (see figure 6 on page 12).

5 Perspective picture transformations

Below are presented algorithms for producing custom perspective picture, from 3D and 2D data. For a proper transformation, 2D coordinates must be normalized for a given AOV type (e.g. vertical, horizontal or diagonal).

Example. For a pixel i in a picture of aspect-ratio 16:9 and AOV measured horizontally, coordinates (i_x, i_y) must be centered and horizontally normalized, so that $i_x \in [-1, 1]$ and $i_y \in [-\frac{9}{16}, \frac{9}{16}]$.

Universal perspective 3D→2D coordinates transformation

$$\begin{aligned}\hat{v} &= (\hat{v}_x, \hat{v}_y, \hat{v}_z) \\ \theta &= \arccos(\hat{v}_z \div \sqrt{\hat{v}_x^2 + l\hat{v}_y^2 + \hat{v}_z^2}) \\ R &= \begin{cases} \tan(k\theta) \div \tan(k\frac{\Omega}{2}), & 0 < k \leq 1 \\ \theta \div \frac{\Omega}{2}, & k = 0 \\ \sin(k\theta) \div \sin(k\frac{\Omega}{2}), & 0 > k \geq -1 \end{cases}\end{aligned}$$

$$\begin{bmatrix} \vec{f}_x \\ \vec{f}_y \end{bmatrix} = \begin{bmatrix} \hat{v}_x \\ \hat{v}_y \end{bmatrix} \frac{R}{\sqrt{\hat{v}_x^2 + l\hat{v}_y^2}} \begin{bmatrix} 1 \\ l(1-s) + s \end{bmatrix}$$

3D coordinates are represented by a normalized vector $\hat{v} \in [-1, 1]^3$, where view origin is at position $(0, 0, 0)$. Transformed picture coordinates are represented by vector $\vec{f} \in [-1, 1]^2$, where image center is at position $(0, 0)$.

Angle θ is between vector \hat{v} and the Z axis. R is the normalized distance between projected vector \vec{f} and the image center, where $\vec{f} \leftarrow \hat{v}$. Angle Ω is equal to AOV, where $\Omega_{\max} \in [\pi, 2\pi]$. Scalar k represents various projection types:

Gnomonic (rectilinear) $k = 1$
Stereographic..... $k = 0.5$
Equidistant..... $k = 0$
Equisolid..... $k = -0.5$
Orthographic..... $k = -1$

Scalar $l \in (0, 1]$ is the spherical projection factor, where $l \approx 0$ represents cylindrical projection and $l = 1$ spherical projection.

Scalar $s \in [4/5, 1]$ describes anamorphic correction of non-spherical image. For $s = 1$ or $l = 1$ there is no anamorphic correction.

Universal reverse 2D→3D coordinates transformation

$$\begin{aligned}\vec{f} &= (\vec{f}_x, \vec{f}_y) \\ R &= \sqrt{\vec{f}_x^2 + l\vec{f}_y^2} \\ \theta &= \begin{cases} \arctan(\tan(k\frac{\Omega}{2})R) \div k, & 0 < k \leq 1 \\ \frac{\Omega}{2}R, & k = 0 \\ \arcsin(\sin(k\frac{\Omega}{2})R) \div k, & 0 > k \geq -1 \end{cases}\end{aligned}$$

$$\begin{bmatrix} \hat{v}_x \\ \hat{v}_y \\ \hat{v}_z \end{bmatrix} = \left\| \begin{bmatrix} \vec{f}_x \\ \vec{f}_y \\ 1 \end{bmatrix} \begin{bmatrix} \sin(\theta)/R \\ \sin(\theta)/R \\ \cos(\theta) \end{bmatrix} \begin{bmatrix} 1 \\ \frac{1}{l(1-s)+s} \\ 1 \end{bmatrix} \right\|$$

Picture coordinates are represented by vector $\vec{f} \in [-1, 1]^2$, where image center is at position $(0, 0)$.

Transformed 3D coordinates are represented by normalized vector $\hat{v} \in [-1, 1]^3$. This transformation is a reverse of the Universal 3D→2D transform on the previous page.

Both transforms produce *perspective picture* (see definition on page 3). Such formula allows for a smooth adjustment of image geometry in accordance to visible content. Base projection type is adjusted by the k component. It manipulates image perception. Cylindrical projection, is adjusted by the l component. Low l values should represent view at level (see figure 6a). For roll motion, recommended value l is 100% (see figure 6b). Anamorphic correction of non-spherical image, driven by the s component, depends on the subject in view. Purpose of the s scalar is to adjust figure's proportions.

Combination of 3D and 2D transformation can be used to map between two different projections, for example *Stereographic* ↔ *Equidistant*, using two separate projection components, k_i and k_o for input and output image.

Lens distortion of perspective picture

Creating perspective picture of a real optical system may require additional deformation of the vector data. Most commonly used is the *Brown-Conrady* lens distortion model.²⁴ Lens distortion vector is added to the frame coordinates \vec{f} vector.

$$\begin{aligned} r &= \vec{f} \cdot \vec{f} \\ \vec{f}' &= \begin{bmatrix} \vec{f}_x \\ \vec{f}_y \end{bmatrix} + \begin{bmatrix} \vec{f}_x \\ \vec{f}_y \end{bmatrix} \begin{bmatrix} k_1 r \\ k_1 r \end{bmatrix} \\ &+ \begin{bmatrix} \vec{f}_x \\ \vec{f}_y \end{bmatrix} \begin{bmatrix} k_2 r^2 \\ k_2 r^2 \end{bmatrix} + \begin{bmatrix} q_1 r \\ q_2 r \end{bmatrix} \\ &+ \left(\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \cdot \begin{bmatrix} \vec{f}_x \\ \vec{f}_y \end{bmatrix} \right) \begin{bmatrix} \vec{f}_x \\ \vec{f}_y \end{bmatrix} \end{aligned}$$

Where r is the dot product of two \vec{f} vectors. k_1 and k_2 are the radial distortion coefficients. q_1 and q_2 are the prism aberration coefficients. p_1 and p_2 are the misalignment coefficients.

6 From visual pyramid, to visual sphere

Visual pyramid of ALBERTI theorem is by definition limited to acute angles, which is restricting in terms of projections it can describe. This property makes stitching or layering multiple pictures defined in such space a problematic task.

In standard perspective model 3D point position is transformed into 2D screen coordinates. But in a case of some curvilinear projections,

points get stretched into lines, like in equidistant projection, where at $\Omega = 360^\circ$ point opposite to the camera direction forms a ring around picture bounds.

In proposed visual sphere model, every perspective picture has its own spherical coordinates, at each pixel. Thus single point can occupy multiple places, conforming to the principles of curvilinear perspective. Such format of perspective allows for stitching and layering images of any perspective geometry.

Remark. Visual sphere image can be reconstructed from six visual-pyramid pictures,^c each covering $\Omega = 90^\circ$, with three mutually perpendicular and three adjacent camera view directions.

Points in spherical projection model are no longer transformed into screen space. Rather lines are calculated and combined to form a polygon image (see figure 7a on page 13). This process involves rotating perspective vector map data. The goal is to align one of the axis with a great circle of the polygon edge (see figure 7b). One way to rotate one axis component is to calculate dot product between the perspective map vector and unit vector perpendicular to the two edge points.

Rasterization of the \overline{ABC} polygon triangle using vector \hat{G} from perspective map

Projected polygon geometry is always part of a great circle. The goal of the algorithm is to rasterize polygon shape formed by those spherical lines. Rasterization process involves determining orientation of the great-circle. Then rotating perspective map, so that the vertical axis aligns with this great-circle. Next, the step-function is performed on the \hat{G}'_x component of the rotated vector \hat{G}' . Full polygon picture is defined by intersection of three such operations.

Rotation of \hat{G}_x axis component to a great circle of two triangle points \vec{A} and \vec{B}

$$\begin{aligned} \hat{X} &= \|\vec{A} \times \vec{B}\| \\ \hat{G}'_x &= \hat{G} \cdot \hat{X} \end{aligned}$$

Where $\hat{X} \in [-1, 1]^3$ is a normalized cross product of triangle vectors \vec{A} and \vec{B} , where $\{\vec{A}, \vec{B}, \vec{C}\} \in \mathbb{R}^3$. Rotated perspective map \hat{G}'_x component is derived from a dot product of the rotated \hat{X} axis vector and perspective map \hat{G} vector. At position $\hat{G}'_x = 0$ lays the great circle crossing points \hat{A} and \hat{B} .

^cProcess known as *cube mapping*

$$\text{pstep}(\hat{G}'_x) = \left(\frac{\hat{G}'_x}{|\Delta\hat{G}'_x|} + \frac{1}{2} \right) \cap [0, 1]$$

Function $\text{pstep}(x) \in [0, 1]$ is a variation of a simple $\text{step}(x)$ function, but with anti-aliasing. It rasterizes hemisphere on which the opposite point \hat{C} lays. $|\Delta\hat{G}'_x|$ is equivalent to the $\text{fwidth}(x)$ function, which returns pixel width of x . If back-facing polygons are to be rendered, the \hat{G}'_x must be flipped ($-\hat{G}'_x$) when polygon is facing back.

$$\begin{aligned} \vec{N} &= (\vec{A} - \vec{B}) \times (\vec{C} - \vec{B}) \\ \text{sign}(x) &= \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x = 0 \\ -1, & \text{if } x < 0 \end{cases} \\ f &= \text{sign}(\vec{N} \cdot \vec{A}) \end{aligned}$$

If the triangle is back-facing, sign of a dot product between triangle normal \vec{N} and one of the triangle points will be equal to -1 . If it's front-facing the result will be 1 . Therefore multiplying \hat{G}'_x component by f will properly render back-facing polygons.

Rasterization of triangle with matrix

Since triangle edge rasterization is performed three times, full triangle rasterization can be represented as a 3×3 matrix multiplication.

$$\begin{aligned} \vec{T} &= \begin{bmatrix} \hat{G}'_x \\ \hat{G}'_y \\ \hat{G}'_z \end{bmatrix} \begin{bmatrix} \|\vec{A} \times \vec{B}\| \\ \|\vec{B} \times \vec{C}\| \\ \|\vec{C} \times \vec{A}\| \end{bmatrix} \\ t &= \text{pstep}(\vec{T}_1) \wedge \text{pstep}(\vec{T}_2) \wedge \text{pstep}(\vec{T}_3) \\ &= [0, 1] \cap \sum_i^n t_i \end{aligned}$$

Rasterising \overline{ABC} triangle image t of perspective vector map \hat{G} is equivalent to the intersection of $\text{pstep}(x)$ function of each component of $\vec{T} \in [-1, 1]^3$ vector. \vec{T} vector is derived from multiplication of perspective vector map \hat{G} vector and a rasterization matrix. Each row of the rasterization matrix represents great-circle tangent vector, which is a normalized cross product between points forming the triangle edge. When combining multiple triangle masks it's important to sum each t mask, otherwise gaps in-between polygons will occur.

Miter of the anti-aliased triangle outline

In special case, when projected polygon edges meet at very narrow angle, its corners will extend

beyond the outline (due to half-pixel offset in the $\text{pstep}(x)$ function). It can be corrected by a miter mask. There are many ways to form such mask, one is to define the smallest circle over projected \overline{ABC} triangle. Furthermore rasterization matrix can be extended to 4×3 , which product is four-component mask vector.

$$\begin{aligned} \vec{T} &= \begin{bmatrix} \hat{G}'_x \\ \hat{G}'_y \\ \hat{G}'_z \end{bmatrix} \begin{bmatrix} \|\vec{A} \times \vec{B}\| \\ \|\vec{B} \times \vec{C}\| \\ \|\vec{C} \times \vec{A}\| \\ \hat{S} \end{bmatrix} \\ t &= \text{pstep}(\vec{T}_1) \wedge \text{pstep}(\vec{T}_2) \wedge \text{pstep}(\vec{T}_3) \\ &\quad \wedge \text{pstep}(\vec{T}_4 - S) \end{aligned}$$

Where $\vec{T} \in [0, 1]^4$ represents source for triangle masks of each edge ($\vec{T}_1, \vec{T}_2, \vec{T}_3$) and miter \vec{T}_4 . Vector \hat{S} is the smallest-circle center vector and S is a step function threshold point, equal to cosine of an angle between \hat{S} and circle edge. To determine smallest circle center and step threshold, based on C. ERICSON solution,⁹ following algorithm can be defined.

$$\begin{aligned} \hat{B}' &= \hat{B} - \hat{A} \\ \hat{C}' &= \hat{C} - \hat{A} \\ d &= (\hat{B}' \cdot \hat{B}')(\hat{C}' \cdot \hat{C}') - (\hat{B}' \cdot \hat{C}')^2 \\ \vec{b} &= \begin{bmatrix} \frac{(\hat{B}' \cdot \hat{B}')(\hat{C}' \cdot \hat{C}') - (\hat{B}' \cdot \hat{C}')(\hat{C}' \cdot \hat{C}')}{(\hat{B}' \cdot \hat{B}')(\hat{C}' \cdot \hat{C}') - (\hat{B}' \cdot \hat{C}')(\hat{B}' \cdot \hat{B}')} \\ \frac{2d}{2d} \end{bmatrix} \\ s &= \vec{b}_1 \\ t &= \vec{b}_2 \\ \vec{S} &= \begin{cases} \frac{\hat{A} + \hat{C}}{2}, & \text{if } s \leq 0 \\ \frac{\hat{A} + \hat{B}}{2}, & \text{else if } t \leq 0 \\ \frac{\hat{B} + \hat{C}}{2}, & \text{else if } s + t \geq 1 \\ \hat{A} + s\hat{B}' + t\hat{C}', & \text{otherwise} \end{cases} \\ S &= |\vec{S}| \end{aligned}$$

Where $\vec{b} \in \mathbb{R}^2$ is the two-component barycentric position of the circumcenter. If $d = 0$, all projected $\hat{A}, \hat{B}, \hat{C}$ points lay in line,⁹ therefore triangle is degenerate and miter mask can be omitted. Vector \vec{S} represents center of the smallest circle. Threshold for $\text{pstep}(x)$ function is denoted by S as cosine of the angle between smallest circle edge and the center vector, which is equivalent to the length of \vec{S} .

Wire-frame \overline{AB} line rasterization

Following algorithm will produce wire-frame image of projected line segment.

$$\begin{aligned}
\text{lpstep}(\hat{G}'_x) &= 1 - \left(\frac{|\hat{G}'_x|}{|\Delta\hat{G}'_x|} \cap [0, 1] \right) \\
\hat{L} &= \frac{\hat{A} + \hat{B}}{2} \\
m &= \text{pstep}(\hat{G} \cdot \hat{L} - \hat{L} \cdot \hat{A}) \\
l &= m \wedge \text{lpstep}(\hat{G}'_x)
\end{aligned}$$

Where \hat{L} is the \overline{AB} line-middle vector, $\text{lpstep}(\hat{G}'_x)$ function rasterizes great-circle line-image, crossing points \hat{A} and \hat{B} . Radial mask m is formed by a $\text{pstep}(x)$ function of dot product between perspective map vector \hat{G} and line-middle vector \hat{L} , minus dot product of \hat{L} and one of the line points (here \hat{A}). Radial mask m combined with a great circle image forms \overline{AB} line segment image l .

Simple-particle procedural rasterization

Following algorithm will produce approximate gradient image of a procedural spherical particle for given position and radius.

$$\begin{aligned}
\text{scr}(x) &= x(2 - x) \\
\text{parl}(\hat{G}, \vec{P}, r) &= \text{scr} \left(\frac{\hat{G} \cdot \hat{P} - \cos \theta}{1 - \cos \theta} \right) \cap [0, 1] \\
\cos \theta &= \sqrt{1 - \sin^2 \theta} \\
&\equiv \sqrt{1 - \frac{r^2}{\vec{P} \cdot \vec{P}}}
\end{aligned}$$

Where $\vec{P} \in \mathbb{R}^3$ is the particle position, r is representing particle radius and α is the angle between \vec{P} and a ray from observation point O , tangent to the particle surface, so that $\angle PrO = 90^\circ$.

7 Fragment data interpolation from barycentric coordinates

Rendering realistic polygon graphics involves shading and texture mapping. Values of normal, depth and UV coordinates associated to each vertex are interpolated across polygon surface using barycentric coordinates of the fragment point.

$$\vec{N} = (\vec{A} - \vec{B}) \times (\vec{C} - \vec{B})$$

Normal vector \vec{N} of the triangle plane \overline{ABC} is derived from cross product of two triangle edges, where $\vec{N} \in \mathbb{R}^3$. Length of this vector is equal to the area of a parallelogram formed by those two edges.

$$\begin{aligned}
u &= \frac{\vec{N} \cdot \vec{A}}{\vec{N} \cdot \hat{G}} \\
&\equiv |\hat{G} \rightarrow \overline{ABC}|
\end{aligned}$$

Distance u represents multiplier of the visual sphere vector \hat{G} , to intersection point on the \overline{ABC} triangle plane. Since \hat{G} is a unit vector, value u can be exported as depth. Here, vector \vec{A} can be replaced by any point on the triangle plane (e.g. \vec{A} , \vec{B} and \vec{C}).

$$\begin{bmatrix} \vec{b}_1 \\ \vec{b}_2 \\ \vec{b}_3 \end{bmatrix} = \begin{bmatrix} |(\vec{B} - u\hat{G}) \times (\vec{C} - u\hat{G})| \\ |(\vec{C} - u\hat{G}) \times (\vec{A} - u\hat{G})| \\ |(\vec{A} - u\hat{G}) \times (\vec{B} - u\hat{G})| \end{bmatrix} \frac{1}{|\vec{N}|}$$

Barycentric vector \vec{b} is a proportion of surface area. From vector \vec{b} , various vertex properties can be interpolated (e.g. depth, normal direction and UV coordinates), given each vertex A , B and C has associated value.

$$\text{frag}(\vec{b}, A_p, B_p, C_p) = \vec{b}_1 A_p + \vec{b}_2 B_p + \vec{b}_3 C_p$$

Interpolated triangle property p is a product of a function $\text{frag}(\vec{b}, A_p, B_p, C_p)$. Which is equivalent to the dot-product of barycentric vector \vec{b} and p value associated to each triangle vertex.

8 No-parallax point mapping

Real optical systems exhibit phenomenon known as the floating no-parallax point¹⁵, where each portion of the picture represent different projection origin. To simulate such perspective, vector length can encode z position offset, such that $\vec{G} = g\hat{G}$, $g = |\vec{G}|$ and $\hat{G} = \|\vec{G}\|$.

In spherical lens, z offset can be described as a product of function $P(\theta)$. Offset value can be approximated by an optical measurement of the parallax alignment (see Figure 9 on page 15).

To measure origin offset, first static NPP picture must be produced. If camera lens does not produce such image, it can be derived from a sequence of images, each taken at different z position (see Subfigure 9b). Offset map value is then equivalent to the source image z position.

To reproduce picture with a floating NPP, each 3D point must be transformed prior to projection, accordingly to perspective map position and associated offset value.

9 Conclusion and future work

I proved that visual sphere model is superior to the visual pyramid of linear perspective. Later

I described design rules for perspective picture which give symbolic meaning to geometric attributes of perspective projection. Also presented mathematical equations for producing universal perspective vector maps, combined with the algorithm for rendering 3D polygon graphics directly from three-dimensional space. This new solution fits well into current graphical pipeline, replacing only low-level rasterization processes.

Presented visual sphere model unites all type of perspective projections under one technical solution, making perspective geometry a fluid construct, as perspective maps can easily be combined and interpolated.

Picture geometry can now be designed to smoothly adapt to the visual story, giving new dimension of control over mental perception of image. Presented concepts and equations may also find its use in other fields, not imagery-related.

This study did not fully explore the process of rendering floating NPP images. Further studies will include research over calibration and simulation of real optical systems with floating NPP. Also performance tests, comparison to current solutions should be performed in future research. Psychological evaluation of perspective geometry magnitude of influence on depicted space perception, performed on a large sample data, could be an interesting field of study.

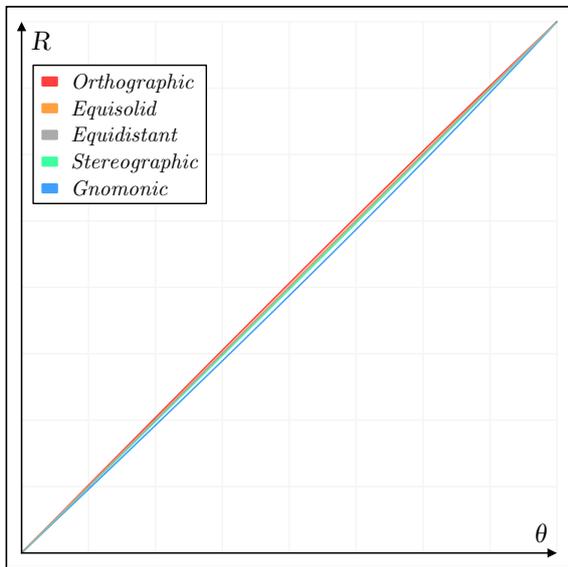
References

1. Alberti, L. B. *On Painting* in. *On painting and on sculpture* trans. by Grayson, C., 55 (Phaidon, London, England, 1967). http://as.vanderbilt.edu/koepnick/Modernism_f05/materials/images/alberti_window.htm.
2. Argan, G. C. & Robb, N. A. The Architecture of Brunelleschi and the Origins of Perspective Theory in the Fifteenth Century. *Journal of the Warburg and Courtauld Institutes* **9**, 96 (1946).
3. Bagley, L. H. & Lavach, D. in Varga, M. *Textbook of Rabbit Medicine E-Book* chap. 9.1 Ocular anatomy and physiology (Butterworth-Heinemann, 2013).
4. Baldwin, J., Burleigh, A. & Pepperell, R. Comparing Artistic and Geometrical Perspective Depictions of Space in the Visual Field. *i-Perception* **5**, 536–547. <http://journals.sagepub.com/doi/pdf/10.1068/i0668> (2014).
5. Casamassina, M. *King Kong's Immersive Style* IGN. <http://ign.com/articles/2005/07/18/king-kongs-immersive-style> (2019).
6. Da Vinci, L. *Annunciation* oil and tempera on panel. Painting. Uffizi Gallery, Florence, Italy, 1475-1480. http://en.wikipedia.org/wiki/De_pictura#/media/File:Leonardo_Da_Vinci_-_Annunciazione.jpeg.
7. Da Vinci, L. *A Picture is to be viewed from one Point only in A treatise on painting* trans. by Rigaud, J. F., 52 (Project Gutenberg, Salt Lake City, USA, 2014). <http://gutenberg.org/ebooks/46915>.
8. Dixon, R. *Perspective Drawings in Mathographics* 82–83 (Dover Publications, Great Britain, 1987).
9. Ericson, C. *Minimum bounding circle (sphere) for a triangle (tetrahedron)* <http://realtimecollisiondetection.net/blog/?p=20> (2020).
10. Erkelens, C. J. The Perspective Structure of Visual Space. *i-Perception* **6**, 204166951561367. <http://journals.sagepub.com/doi/pdf/10.1177/2041669515613672> (2015).
11. Gascuel, J.-D., Holzschuch, N., Fournier, G. & Péroche, B. *Fast Non-linear Projections Using Graphics Hardware* in *ACM Symposium on Interactive 3D Graphics and Games* (ACM, Redwood City, California, 2008), 107–114. <http://maverick.inria.fr/Publications/2008/GHFP08/i3d2007.pdf>.
12. Glaeser, G. & Gröller, E. Fast generation of curved perspectives for ultra-wide-angle lenses in VR applications. *The Visual Computer* **15**, 365–376 (1999).
13. Hueck, A. *Von den Gränzen des Sehvermögens* ger, 84. http://hans-strasburger.userweb.mwn.de/materials/hueck_1840_ocr.pdf (Müller's Archiv, Tartu, 1840).
14. Königsberg, I. *The complete film dictionary* 12. <http://archive.org/details/complectefilmdict00koni> (Nal Penguin INC, New York, USA, 1987).
15. Littlefield, R. Theory of the No-Parallax Point in Panorama Photography. <http://janrik.net/PanoPostings/NoParallaxPoint/TheoryOfTheNoParallaxPoint.pdf> (2006).
16. Miller, P. E. & Murphy, C. J. *Equine Vision in Equine Ophthalmology* (ed Gilger, B. C.) chap. Visual perspective and field of view (Elsevier Health Sciences, St. Louis, USA, 2010).

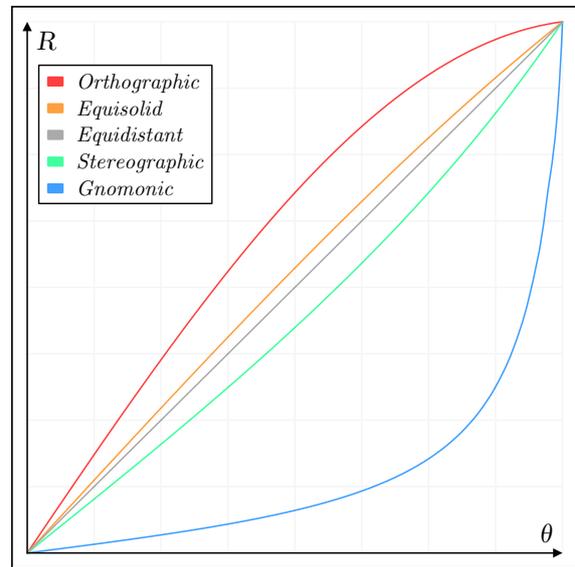
17. Pannini, G. P. *Interior of St. Peter's* oil on canvas. Painting. National Gallery of Art, Washington DC, USA, 1754. <http://web.archive.org/web/20130327080328/vedutismo.net/Pannini/>.
18. Radley, A. *Lookable User Interfaces and 3D* in Cipolla-Ficarra, F. V. *Handbook of Research on Interactive Information Quality in Expanding Social Network Communications* 46 (IGI Global, 2015). <http://books.google.com/books?id=90yfBwAAQBAJ>.
19. Rybczyński, Z. *a Treatise on the Visual Image* (ed Oszmiańska, H.) chap. 1.2.2. <http://books.google.com/books?id=mUHNAQAACAAJ> (Art Stations Foundation, Poznań, Poland, 2009).
20. Rydzak, J. *Zbig Rybczyński, Oscar-Winning Filmmaker* interview. Arizona, USA, 2014. http://youtube.com/watch?v=071_XgZ4H2I.
21. ShareGrid, I. *Todd AO High-Speed Anamorphic 35mm T1.4 at T2.0_log* from The Ultimate Anamorphic Lens Test. <https://learn.sharegrid.com/sharegrid-lens-test-todd-ao-anamorphic> (2020).
22. Thoby, M. *About the various projections of the photographic objective lenses. Orthographic fisheye projection* http://michel.thoby.free.fr/Fisheye_history_short/Projections/Various_lens_projection.html (2019).
23. Wade, N. J. & Hughes, P. Fooling the Eyes: Trompe L'Oeil and Reverse Perspective. *Perception* **28**, 1115–1119. [http://wexler.free.fr/library/files/wade%20\(1999\)%20fooling%20the%20eyes.%20trompe%20l'oeil%20and%20reverse%20perspective.pdf](http://wexler.free.fr/library/files/wade%20(1999)%20fooling%20the%20eyes.%20trompe%20l'oeil%20and%20reverse%20perspective.pdf) (1999).
24. Wang, J., Shi, F., Zhang, J. & Liu, Y. A new calibration model of camera lens distortion. *Pattern Recognition* **41**, 607–615. <http://dia.fi.upm.es/~lbaumela/Vision13/PapersCalibracion/wang-PR208.pdf> (2008).
25. Wikipedia, c. *Robert Barker (painter)*. *Biography* Wikipedia, The Free Encyclopedia. [http://en.wikipedia.org/w/index.php?title=Robert_Barker_\(painter\)&oldid=907715733#Biography](http://en.wikipedia.org/w/index.php?title=Robert_Barker_(painter)&oldid=907715733#Biography) (2019).



Figure 3: Still from TODD-AO HIGH-SPEED ANAMORPHIC lens (35mm T1.4) with visible curvilinear perspective. This type of lens was featured in films like *Conan the Barbarian*, *Dune* and *Mad Max*.²¹ © 2017 ShareGrid, Inc.

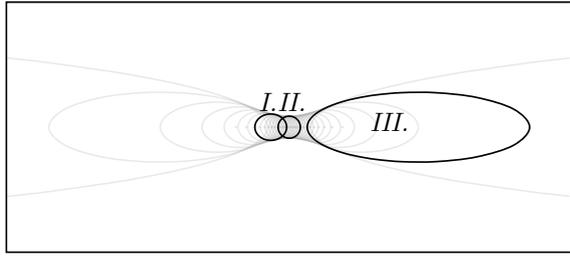


(a) Linear-scale graph plotting ray of angle $\theta \in [0^\circ, \Omega]$, as the horizontal axis, and screen-position radius $R \in [0, 1]$, as the vertical axis, where $\Omega = 40^\circ$ (equivalent of $R = 1$).

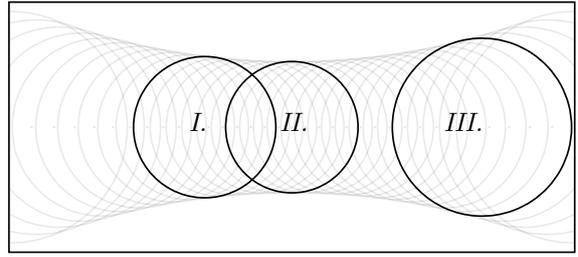


(b) Linear-scale graph plotting ray of angle $\theta \in [0^\circ, \Omega]$, as the horizontal axis, and screen-position radius $R \in [0, 1]$, as the vertical axis, where $\Omega = 170^\circ$ (equivalent of $R = 1$).

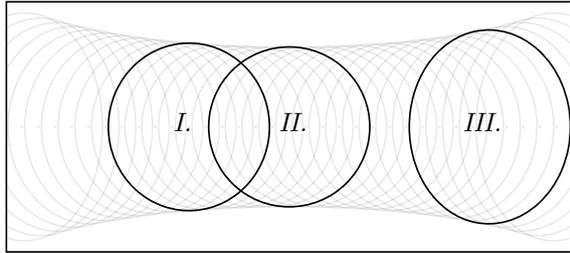
Figure 4: Chart comparison of radial compression in five major azimuthal projections, across two different AOV values (represented by Ω); narrow (4a) and wide (4b).



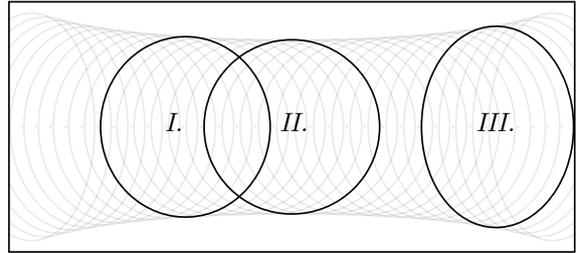
(a) Rectilinear (Gnomonic) projection.



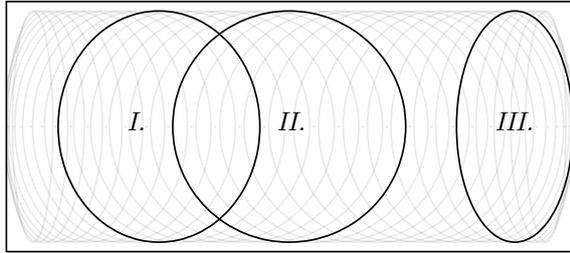
(b) Stereographic projection.



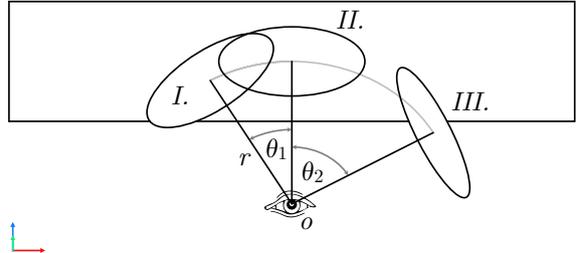
(c) Equidistant projection.



(d) Equisolid projection.

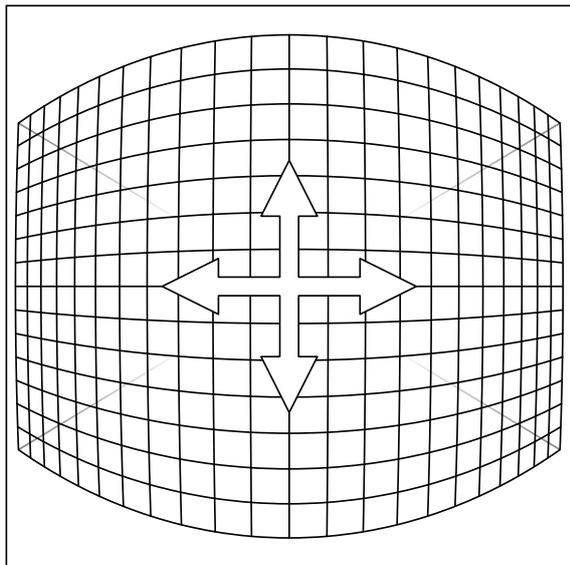


(e) Orthographic-azimuthal projection.

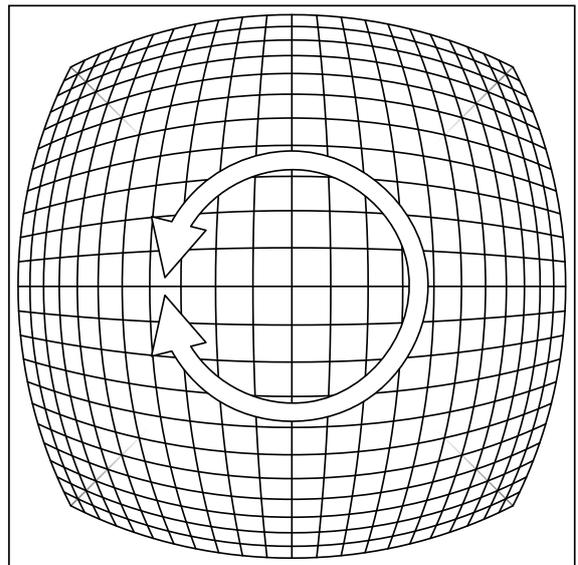


(f) Examples model of equal-size discs (*I*, *II*, *III*) at equal distance r , facing observation point o , where $\Omega_h = 170^\circ$, $\theta_1 = 30^\circ$ and $\theta_2 = 60^\circ$.

Figure 5: Example of motion in perspective picture in various azimuthal projections, where Subfigure 5f presents viewed elements layout.

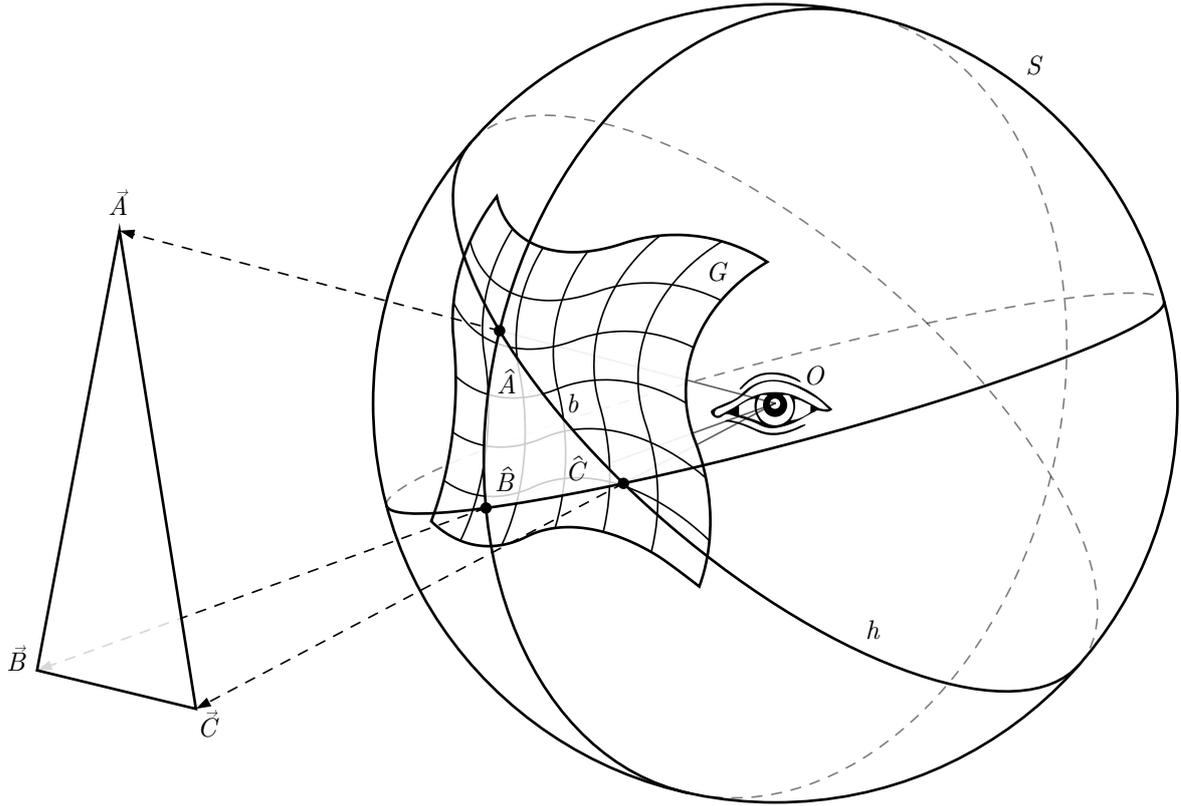


(a) Example of image geometry for pitch and yaw motion (arrows), where $\Omega_h = 120^\circ$, $k = 0$, $l = 10\%$ and $s = 95\%$.

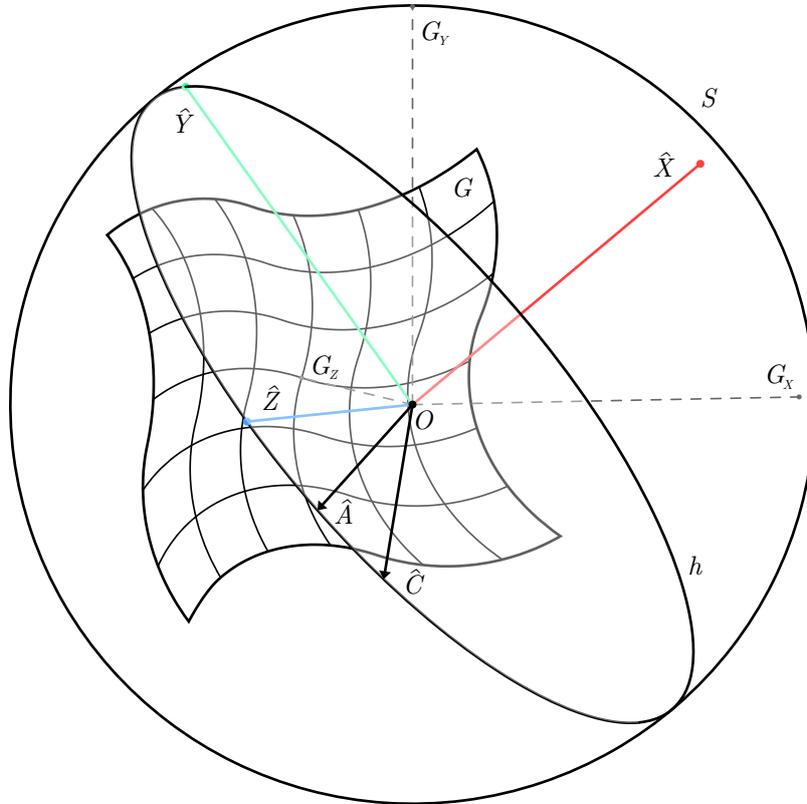


(b) Example of image geometry for roll motion (arrows), where $\Omega_h = 120^\circ$, $k = 0$, $l = 100\%$ and $s = 95\%$.

Figure 6: Examples of image geometry for given type of view motion.



(a) Projection of polygon triangle \overline{ABC} onto an image grid G imposed over the visual sphere S , where projected triangle line b belongs to a great circle h .

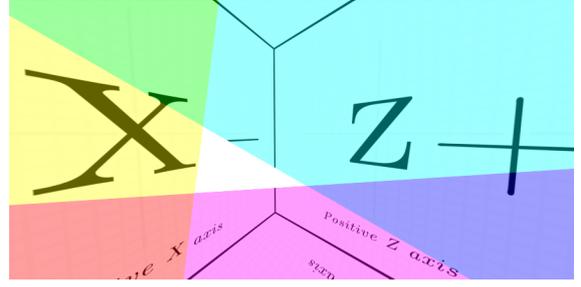


(b) Rotation of image grid \hat{G}_x component, to unit vector \hat{X} , where $\hat{X}^\perp = h$ and $X \perp Y \perp Z$, so that $\{\overline{AC}, \hat{Y}, \hat{Z}\} \in h$. Required components can be calculated from cross and dot product; $\hat{X} = \|\vec{A} \times \vec{C}\|$ and $\hat{G}'_x = \hat{X} \cdot \hat{G}$.

Figure 7: Projection of triangle \overline{ABC} onto visual unit sphere S , where projection origin O is at the sphere center. Edge of the projected triangle is always produced by an arc of a great circle, here $\overline{AC} \in h$. Image grid G represents visual sphere vector map, where each pixel color is a spherical unit vector.



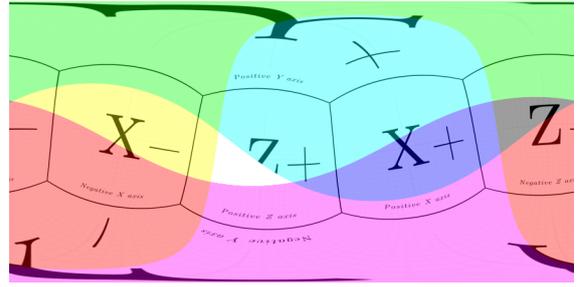
(a) Rectilinear perspective vector map, where $\Omega_d = 140^\circ$, $k = 1$, $l = 1$.



(b) Rasterized triangle in rectilinear perspective using vector map, where $\Omega_d = 140^\circ$, $k = 1$, $l = 1$.



(c) Equirectangular projection vector map of whole sphere, where $\Omega_h = 360^\circ$ and $\Omega_v = 180^\circ$.



(d) Rasterized triangle in equirectangular projection, where $\Omega_h = 360^\circ$ and $\Omega_v = 180^\circ$.



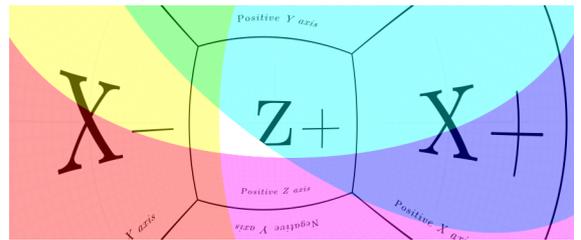
(e) Mustache style lens distortion vector map, where $\Omega_d = 131^\circ$, $k = 0.32$, $l = 62\%$, $s = 86\%$, $k_1 = -0.6$, $k_2 = 0.4$.



(f) Mustache style lens distortion triangle rasterization, where $\Omega_d = 131^\circ$, $k = 0.32$, $l = 62\%$, $s = 86\%$, $k_1 = -0.6$, $k_2 = 0.4$.



(g) Curvilinear perspective vector map, where $\Omega_d = 270^\circ$, $k = 0.32$, $l = 62\%$, $s = 86\%$.



(h) Rasterized triangle in curvilinear perspective, where $\Omega_d = 270^\circ$, $k = 0.32$, $l = 62\%$, $s = 86\%$.

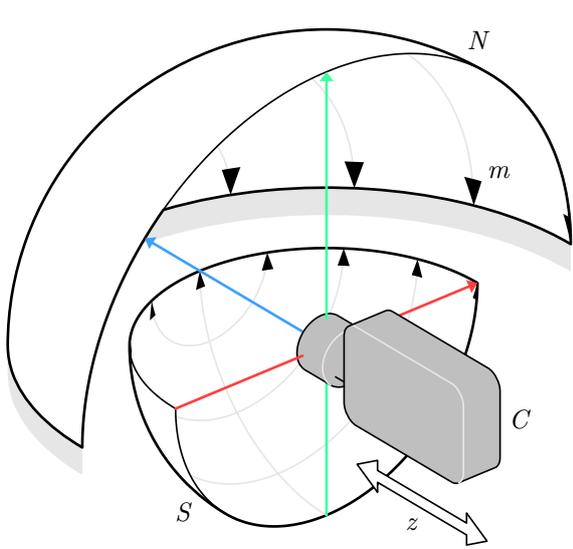


(i) Perspective vector map of five-screen horizontal array in rectilinear projection, where single screen $\Omega_{h,v} = 60^\circ$, $k = 1$ and $l = 1$, giving a total $5\Omega_h = 300^\circ$.

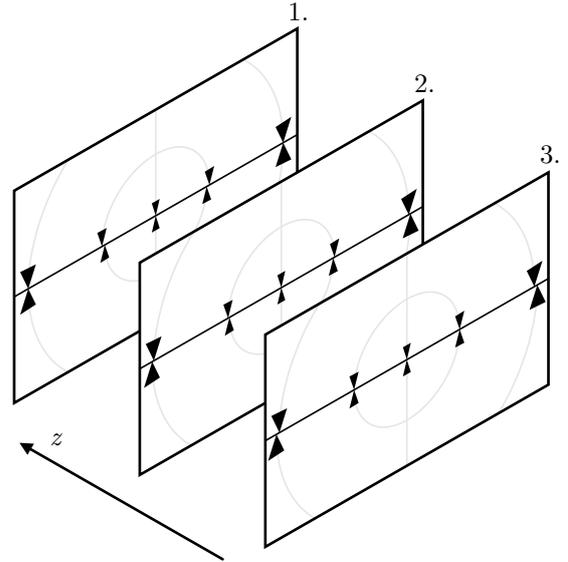


(j) Rasterized triangle of five-screen horizontal array, where single screen $\Omega_{h,v} = 60^\circ$, $k = 1$ and $l = 1$, giving a total $5\Omega_h = 300^\circ$.

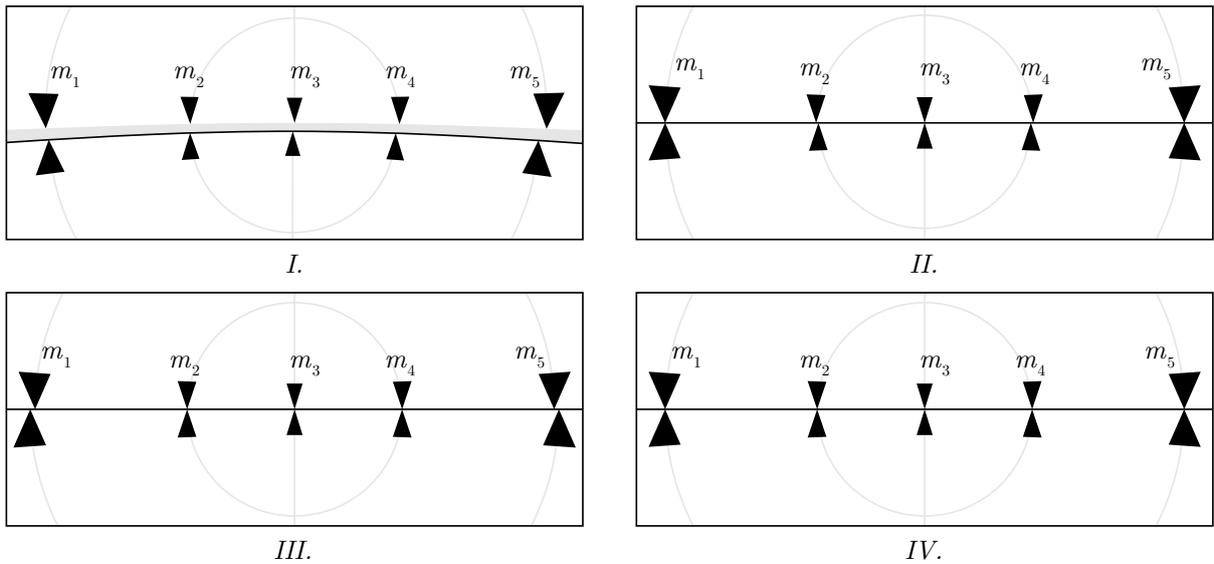
Figure 8: Examples of polygon triangle rasterization directly from three-dimensional space to the image, using visual sphere vector map G , where $G_n \in [0, 1]^3 \leftrightarrow [-1, 1]^3$.



(a) Model of no-parallax point (NPP) calibration rig. Which measures misalignment of markers (m) between northern (N) and southern (S) hemisphere as seen through camera (C).



(b) Example of a variable camera position z encoded in an image sequence (1, 2, 3). Element 1 presents alignment of the peripheral markers, while element 3 presents alignment of the side markers, element number 2 presents position in-between.



(c) Visualization of the NPP alignment using calibration rig (9a). Example *I* presents misalignment of the camera in all three axes. Examples *II*, *III* and *IV* present alignment in X, Y axes. In example *II* peripheral marker pairs m_1 , m_5 present alignment near horizontal AOV, while pairs m_2 , m_4 are misaligned due to floating NPP. In example *III* markers m_2 and m_4 are aligned, while peripheral markers m_1 , m_5 are not. Position z in example *III* is less than *II*. Example *IV* presents “slit-scan” composite of variable z position, where all markers are aligned.

Figure 9: Floating no-parallax point calibration process.

© 2020 Jakub Maksymilian Fober (the Author).

The Author provides this document (the Work) under the Creative Commons CC BY-ND 3.0 license available online. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nd/3.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA. The Author further grants permission for reuse of images and text from the first page of the Work, provided that the reuse is for the purpose of promoting and/or summarizing the Work in scholarly venues and that any reuse is accompanied by a scientific citation to the Work.

