Geometry of binocular imaging II : The augmented eye

Victor S. Grinberg
Gregg W. Podnar
M. W. Siegel

Robotics Institute, School of Computer Science, Carnegie Mellon University
5000 Forbes Ave., Pittsburgh, PA, 15213

ABSTRACT

We address the issue of creating imagery on a screen that, when viewed by naked human eyes, will be indistinguishable from the original scene as viewed through a visual accessory. Visual accessories of interest include, for example, binoculars, stereomicroscopes, and binocular periscopes. It is the nature of these magnifying optical devices that the transverse (normal) magnification and longitudinal (depth-wise) magnification are different. That is why an object viewed through magnifying optical devices looks different from the same object viewed with the naked eye from a closer distance – the object looks "squashed" (foreshortened) through telescopic instruments and the opposite through microscopic instruments. We rigorously describe the quantitative relationships that must exist when presenting a scene on a screen that stereoscopically simulates viewing through these visual accessories.

Keywords: binocular imaging, stereo imaging, binoculars, stereomicroscope, periscope, simulation, projective geometry

1. INTRODUCTION

Binocular (stereoscopic) imaging is one of the most powerful means to use 2D images for realistically simulating the appearance of 3D objects when viewed with two eyes. All binocular imaging systems must provide a means for presenting one image to the left eye and another to the right. We investigate the geometric issues related to binocular imaging. We consider a system where images for both the left and right eyes are created by appropriate software on one computer screen, and where a means to ensure that each eye sees only the intended image is employed. Our goal is to study the "correct geometry" of this process, i. e., how synthetic objects existing in the synthetic 3D world must be mapped onto the screen for the truest binocular perception. We concentrate on these pure geometric issues, putting aside other effects that can contribute to depth perception, such as color, lighting, texture, etc.

In our previous paper[4] we concentrated on the naked-eye case. We analysed the geometry that must be precisely implemented to satisfy the requirement that the imagery be geometrically indistinguishable from the reality that would be perceived by the "naked" human visual apparatus.

This paper uses the conceptual framework of [4]. Now we consider the augmented-eye case. Namely we investigate the geometry needed to create imagery on a screen that, when viewed by naked human eyes, will be indistinguishable from the original scene as viewed through a visual accessory (ignoring in this discussion issues such as apertures which restrict the field of view). We use the generic term *optical device* for such an accessory. Typical optical devices of interest are binoculars, stereomicroscopes, and binocular periscopes. Usually optical devices introduce certain geometric distortions in the viewed scene, and therefore what a viewer sees through the device looks different from the real scene. Typically the object looks "squashed" (foreshortened) through telescopic instruments and "stretched" (lengthened front-to-back) through microscopic instruments*. An example of a device that is free from this defect is a simple periscope that contains two mirrors and no lenses. However the presence of lenses in a device inevitably leads to the presence of this kind of distortion.

---

*Generally the stretching by microscopic instruments is not readily visible because to deliver enough light for the observer to see anything they must have large apertures, and hence they have small depths-of-field. However the foreshortening in telescopic instruments is well known; it is often used for certain dramatic effects in movies and television.

The correct geometry must simulate exactly what a viewer sees through an optical device, including the distortions.

Our approach derives from first principles a framework for understanding presentation of stereoscopic images for a pair of augmented eyes.

2. CONCEPTS OF BINOCULAR OPTICAL DEVICE SIMULATION

Naturally we shall not consider every effect that happens in real optical devices. First, remaining in the frame of geometric optics, we put aside all effects tied with diffraction and interference. In other words we assume that all optical elements are much larger than the wavelengths characteristics of visible light. Second, we shall neglect chromatic aberration and other dispersion effects, thus assuming that either the indices of refraction are wavelength independent or the light is monochromatic. Under these assumptions an optical device can be described as a collection of elements that consecutively change the direction of light rays. These changes in direction may be effected by reflection (mirrors) or by refraction (lenses, prisms, etc.).

Now we have to answer the crucial question: what will one see looking at an object through an optical device? We assume that two devices are used, one per eye (although they may share some elements, e.g., a stereomicroscope with two oculars but a single objective). Let us begin with the simplest object − a point source. First, we are interested only in rays that are emitted by the point source and that eventually enter an eye.

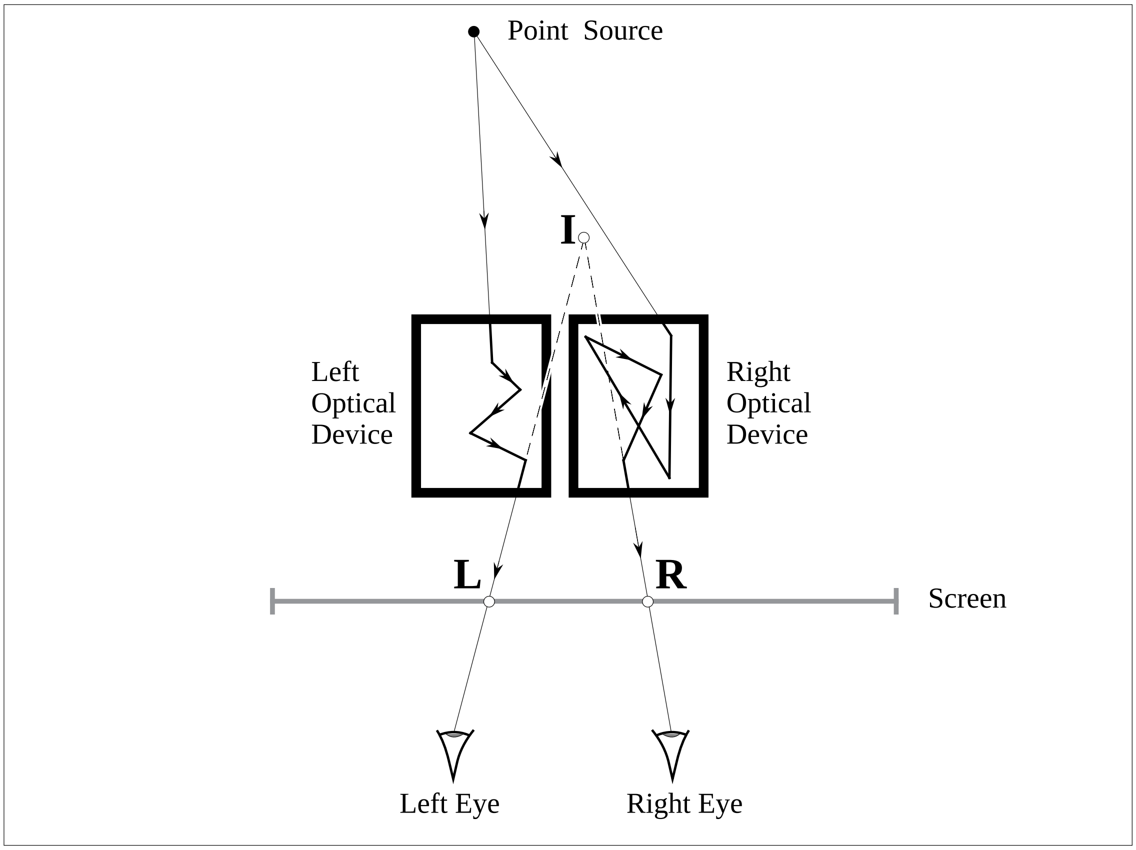

Figure 1

Figure 1 shows two such rays. One of them enters the left eye, the other the right eye. Before entering an eye, each ray passes through the corresponding optical device that somehow changes its direction. Each eye sees the direction from which the ray comes, and perceives the point source in this direction. So if these two rays intersect after extension, our binocular vision apparatus will perceive the point source exactly at the intersection point. If these rays of light do not intersect (which happens sometimes because of bad alignment) the viewer will see two separate points and will not typically be able to fuse them into one point at a well defined distance. Hereafter we assume that the device is properly aligned so that these two rays do intersect, and we denote the intersection point by **I**. Knowing the position of **I**, it is easy (see [4]) to find two points **L** and **R** on the display screen where left and right projections of **I** can be drawn to simulate point **I** at the correct depth. The preliminary algorithm for simulation is thus:

1) find the ray that is emitted by the point source and enters the left eye after passing through the left optical device;

2) find the ray that is emitted by the point source and enters the right eye after passing through the right optical device;

3) find the point **I**, the intersection point of these rays after extension (we assume the system is properly aligned so they do intersect);

4) find the point **L**, the intersection of the left ray and the surface of the screen, where the left image of the point source must be drawn on the screen;

5) find the point **R**, the intersection of the right ray and the surface of the screen, where the right image of the point source must be drawn on the screen;

6) repeat steps 1)-5) for enough points of the object to draw complete left and right images on the screen.

3. COMPUTATIONAL ASPECTS

The algorithm described in section 2 has one weak feature: how can steps 1) and 2) be performed? How can we find the rays that enter the eyes after passing through the optical devices? The problem in general is hard and computationally expensive. Fortunately all useful optical devices have a remarkable property that allows us to overcome this difficulty: focusing. That is, all rays emitted by a point, after passing through the device and suitable extension of the exit rays, intersect at another point, the image (real or virtual) of the first one. Real devices (except flat mirrors) do not satisfy this condition precisely, but they can satisfy it rather closely. We shall consider ideal optical devices, which by definition do have this property. An ideal lens, for example, is a device that is exactly described by the lens formula

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f} \quad .$$

Thus an ideal optical device maps each point into some other point – its image. In other words it performs the transformation of 3D space into itself. We shall call this transformation an *OD-transformation*. Knowing the OD-transformation that corresponds to a given optical device allows us to find particular rays rather straightforwardly. Namely, these rays are straight lines that join *the image* of the point with the left and right eyes.

Now we come to the problem how to find the OD-transformation of a given device, or in other words, how to find the image of a given object point. The obvious way to do this is to analyse sequentially each elementary device that constitutes the composite device. For each simple element, like any surface of a lens or a mirror, an interface between two media with different refractive indices, the corresponding OD-transformation is simple and well known. Evidently, the OD-transformation of a composite device is a composition of these elementary transformations in the order in which the rays transverse elementary devices. This solves the problem in principle.

4. PROJECTIVE TRANSFORMATIONS

The direct application of the described method for image calculation requires calculations of the chain of images for each point of the object: the image that the first optical element creates, then the image of this image that the second optical element creates and so on up to the final image that is projected onto the screen. One would achieve significant calculating speed-up if it were possible to avoid calculation of all the intermediate images in this chain of images. It turns out that this is possible. The key idea is the representation of OD-transformations as projective transformations. A description of projective transformations can be found in any book on projective geometry (see for example [6], [7]). We write below, in the form that is needed for our applications, the definition of projective transformations, and a summary of their properties (without proofs).

4.1. Definition

A *projective transformation* is a map of 3D space into itself that is defined by an arbitrary non-zero 4×4 matrix **T** in the following way: given a point **x** in 3D space with co-ordinates (x_1, x_2, x_3), form the quadruple $(1, x_1, x_2, x_3)$; then multiply **T** by this quadruple, getting the quadruple (z_0, z_1, z_2, z_3); and finally define $y_1=z_1/z_0$, $y_2=z_2/z_0$, $y_3=z_3/z_0$. The point **y** with co-ordinates (y_1, y_2, y_3) is the image of **x**.

4.2. Properties of projective transformations

a) the definition above looks co-ordinate dependent. Actually it is not: if a transformation is projective in one coordinate system (i. e., there exists a corresponding matrix), it is projective in any other co-ordinate system (with some other corresponding matrix);

b) a projective transformation is defined by its 4×4 matrix uniquely up to non-zero multiplicative factor; so the set of all projective transformations has 4×4−1=15 independent parameters;

c) any linear transformation is projective;

d) any translation is a projective transformation;

e) a projective transformation is a one-to-one correspondence, so for any projective transformation described by matrix **T** there is an inverse transformation, that maps the image space into the object space; this transformation is also projective, and furthermore, it is described by the matrix $\mathbf{T}^{-1}$ (in the same co-ordinate system);

f) a composition (product) of two projective transformations, described by matrices **S** and **T**, is also a projective transformation and it is described by the matrix $\mathbf{S}\times\mathbf{T}$ (in the same co-ordinate system);

g) projective transformations preserve straight lines: the image of a straight line is a straight line** ;

h) any OD-transformation is projective*** .

4.3. Examples: OD-transformation for elementary optical devices as projective transformations

4.3.1. Flat mirror

According to 4.2.h) a flat mirror performs a projective transformation. Using the law "the angle of incidence equals the angle of reflection" it is easy to calculate the corresponding matrix.

In the co-ordinate system shown in Figure 2 this matrix is

$$\left\|\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{array}\right\|$$

Any object point (x_1, x_2, x_3) is mapped into the corresponding image point $(x_1, x_2, -x_3)$.

4.3.2. Flat interface between two media with different refractive indices

According to 4.2.h) such interface performs a projective transformation. Using Snell's law, $n_1 \sin\theta_1 = n_2 \sin\theta_2$, it is

---

**Sometimes this property is sacrificed in order to obtain focusing, e. g., fish-eye lens. The image created by such a lens *is not* described by a projective transformation.

***The converse is not true, i. e., all focusing optical systems are described by projective transformations, but not all projective transformations can be realized via optical systems. This can be seen simply from the fact that the projective transformation has 15 free parameters, whereas any ideal optical system can be fully described by one or two focal lengths and the distance between two principal planes (see [1] pp.150-154, [2] pp. 141-149, [3] pp. 27-6-27-7, [5] pp. 17-21). Adding 6 parameters that define the position and orientation of the device, we conclude that the number of free parameters of the optical device is at most 9.

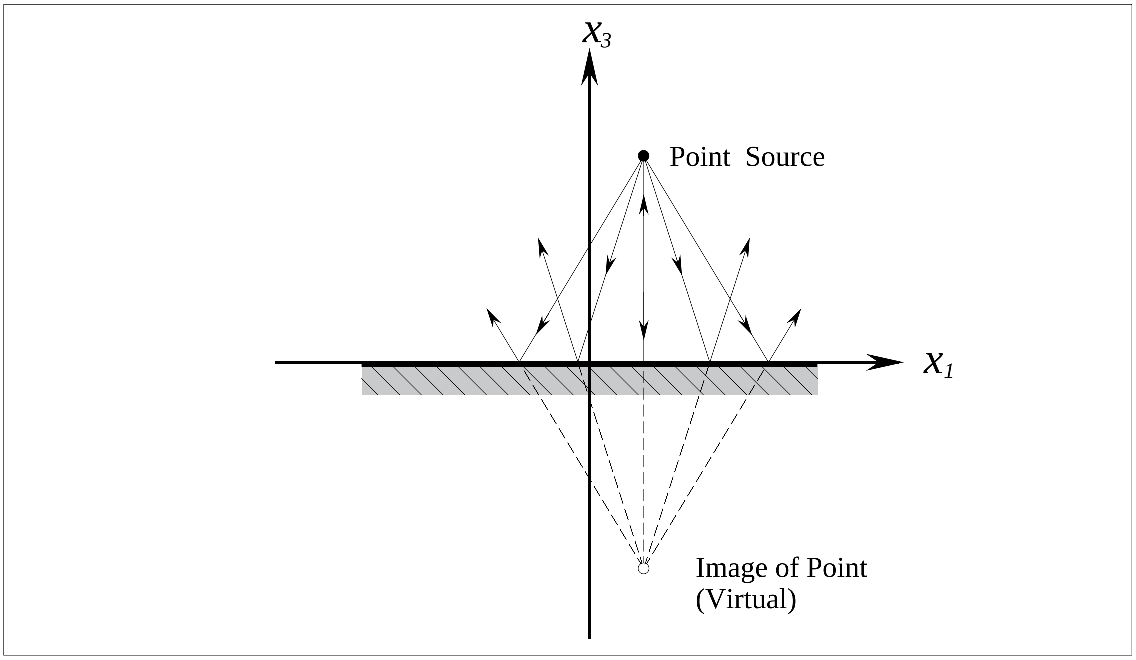

Figure 2. A flat mirror as an ideal optical device.

possible to calculate the corresponding matrix.

In the co-ordinate system shown in Figure 3 this matrix is

$$\begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & n_2/n_1 \end{Vmatrix},$$

where n_1 and n_2 are the refractive indices of two media. Note that this is true only approximately, because this device maps a point not onto a point but rather onto a short line segment. The approximation holds well when only rays not too far from the normal to the interface are allowed.

4.3.3. Thin lens, spherical mirror

According to 4.2.h) a thin lens performs a projective transformation. Using the lens equation it is possible to construct the corresponding matrix.

In the co-ordinate system shown in Figure 4 this matrix is

$$\begin{Vmatrix} f & 0 & 0 & -1 \\ 0 & f & 0 & 0 \\ 0 & 0 & f & 0 \\ 0 & 0 & 0 & f \end{Vmatrix},$$

where f denotes the focal length of the lens. With the appropriate sign conventions, a convex or concave mirror

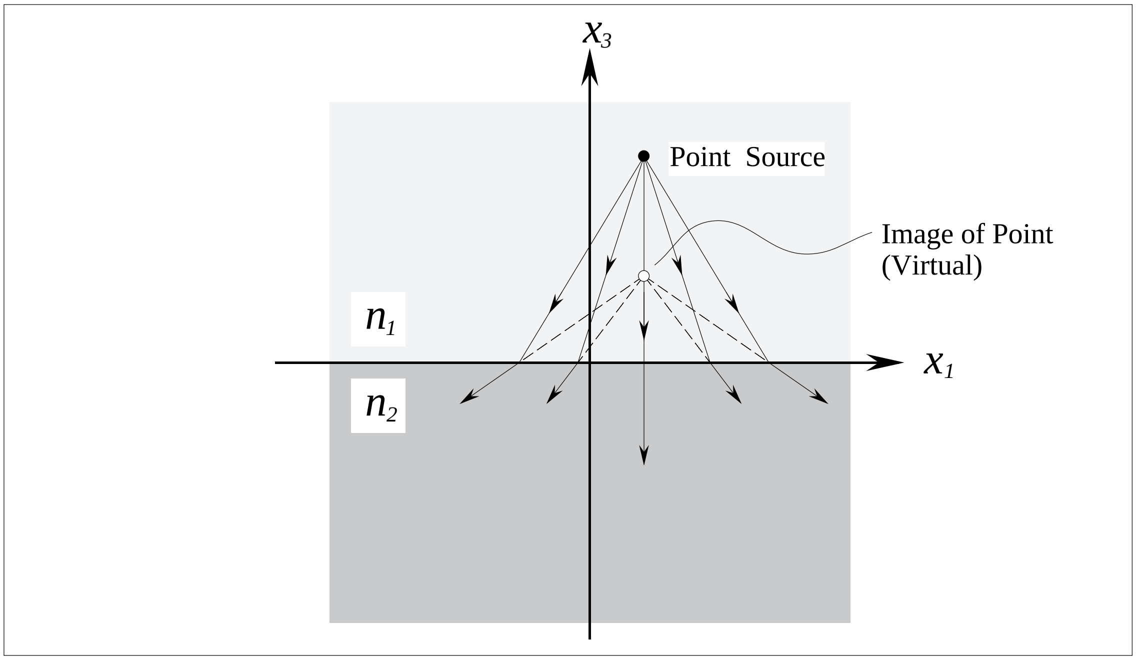

Figure 3. Flat interface between two media as an ideal optical device.

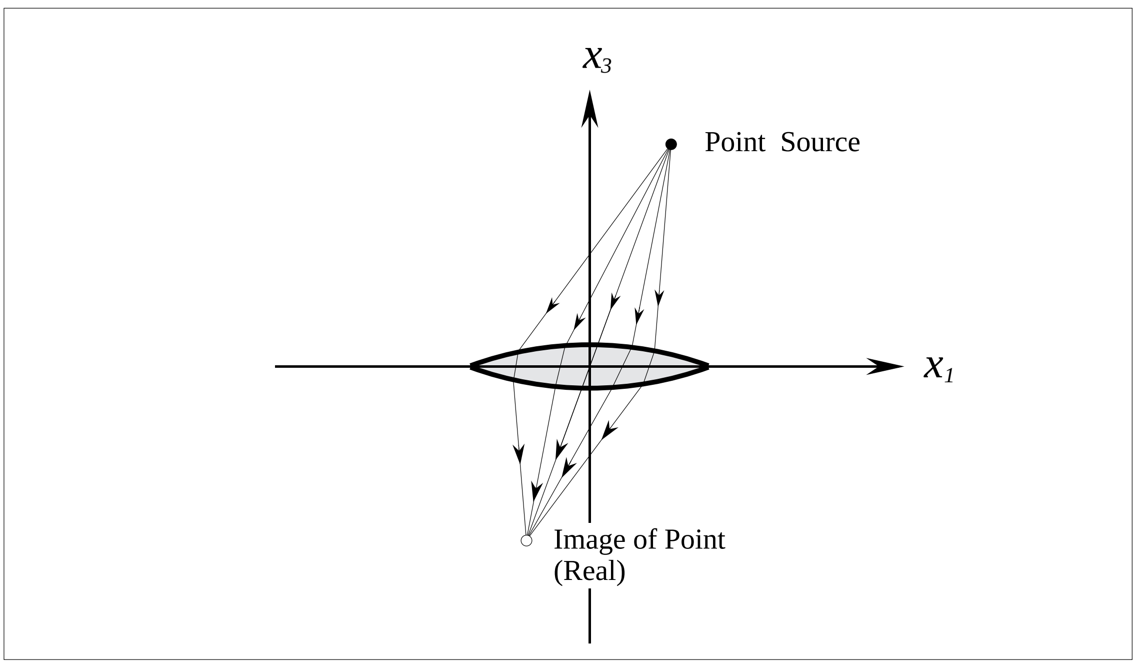

Figure 4. A lens as an ideal optical device.

satisfies the same equation (with proper positive or negative f), and therefore has the same corresponding matrix.

4.4. Use of projective representation of optical devices

The preceding representation of elementary optical devices allows us to define the transformation performed by a composite optical device that consists of several elementary ones in the following way:

1) using 4.2.h) describe the action of each elementary device by a corresponding matrix, introducing an appropriate co-ordinate system for each device;

2) choose and fix a global co-ordinate system;

3) using 4.2.a), 4.2.c), 4.2.d), for each elementary device construct the matrix of its transformation in the global co-ordinate system;

4) multiply all these matrices in the order in which the rays traverse the elementary devices.

According to 4.2.f) the result will describe the transformation performed by the composite optical device. The advantage of this approach is that as long as one has this compact description of the device, it can be used for calculation of all images. The preliminary work is done only once and then is used for all points of the viewed object. This saves a significant amount of calculation time in comparison with ray tracing through arbitrary media with refractive and reflecting properties but without focusing properties.

Another possibility to save calculation time arises from 4.2.g). To create the image of a straight line segment it is sufficient to calculate the images of its endpoints. 4.2.g) guarantees that the image of the segment will be a segment bounded by these two points.

5. APPLICATIONS

Stereoscopic optical device simulation can be useful in the design of binocular optical instruments. We think it can be especially helpful in stereomicroscope design. According to some sources, roughly half of the people using stereomicroscopes use it as a monoscopic tool. This neglect of the advantages of stereoimaging can in some cases be attributed to vision defects in the user, but in others perhaps discomfort is resulting from "unnatural" optical design. Computer simulation of stereomicroscopes may open the way to low cost human factors experiments aimed at developing more comfortable instruments.

6. ACKNOWLEDGMENTS

This work was supported by ARPA High Definition Systems Grant MDA972-92-J-1010.

7. REFERENCES

1. Max Born and Emil Wolf, Principles of Optics, Pergamon Press, Oxford, New York, 1969.
2. E. B. Brown, Modern Optics, Reinhold, New York, 1965.
3. R. P. Feynman, R. B. Leighton, M. Sands, The Feynman Lectures on Physics, Vol 1, Addison-Wesley, 1963.
4. V. S. Grinberg, G. W. Podnar, M. W. Siegel, "Geometry of Binocular Imaging", in *Stereoscopic Displays and Applications V*, Proceedings of the 1994 IS&T/SPIE Conference on Electronic Imaging: Science & Technology, San Jose, California, USA, SPIE, 8-10 Feb., 1994, pp. 56-65.
5. R. S. Longhurst, Geometrical and physical optics, Wiley, New York, 1967.
6. R. A. Rosenbaum, Introduction to Projective Geometry and modern algebra, Addison-Wesley, 1963.
7. P. Samuel, Projective Geometry, Springer-Verlag, New York, 1988.