Perspective Projection: The Wrong Imaging Model[1]

Margaret M. Fleck
Department of Computer Science
University of Iowa

Abstract:

Perspective projection is generally accepted as the ideal model of image formation. Many recent algorithms, and many recent judgements about the relative merits of different algorithms, depend on this assumption. However, perspective projection represents only the front half of the viewing sphere and it distorts the shape and intensity of objects unless they lie near the optical axis. It is only one of several projections used in lens design and it does not accurately model the behavior of many real lenses. It works well only for narrow-angle images.

This paper surveys the properties of several alternative models of image formation. A model based on stereographic projection of the viewing sphere is shown to be a better general-purpose imaging model than perspective projection. The new model can represent wider fields of view and more closely approximates real wide-angle lenses. It preserves a suitable range of shape properties, including local symmetries. It approximates narrow-angle perspective locally, enabling local affine features to be used throughout the field of view.

1 Introduction

Perspective projection is generally accepted as the ideal model of image formation. Considerable effort has been spent deriving mathematical results (e.g. invariant properties) for perspective projection. But is perspective the right model for imaging?

Ideally, a camera model should:

- Be a close match to the behavior of real camera systems, and
- Have convenient mathematical properties.

For narrow-angle lenses, perspective seems to satisfy both of these criteria.

However, perspective projection is not a convenient representation for wide-angle images, i.e. images covering a field of view over 90 degrees (Kingslake 1992). A perspective image cannot cover

[1]This research was carried out at the Department of Computer Science, University of Iowa. It was supported in part by Research Initiation grant IRI-9209728 and Research Instrumentation grant CDA-9121985 from the NSF.

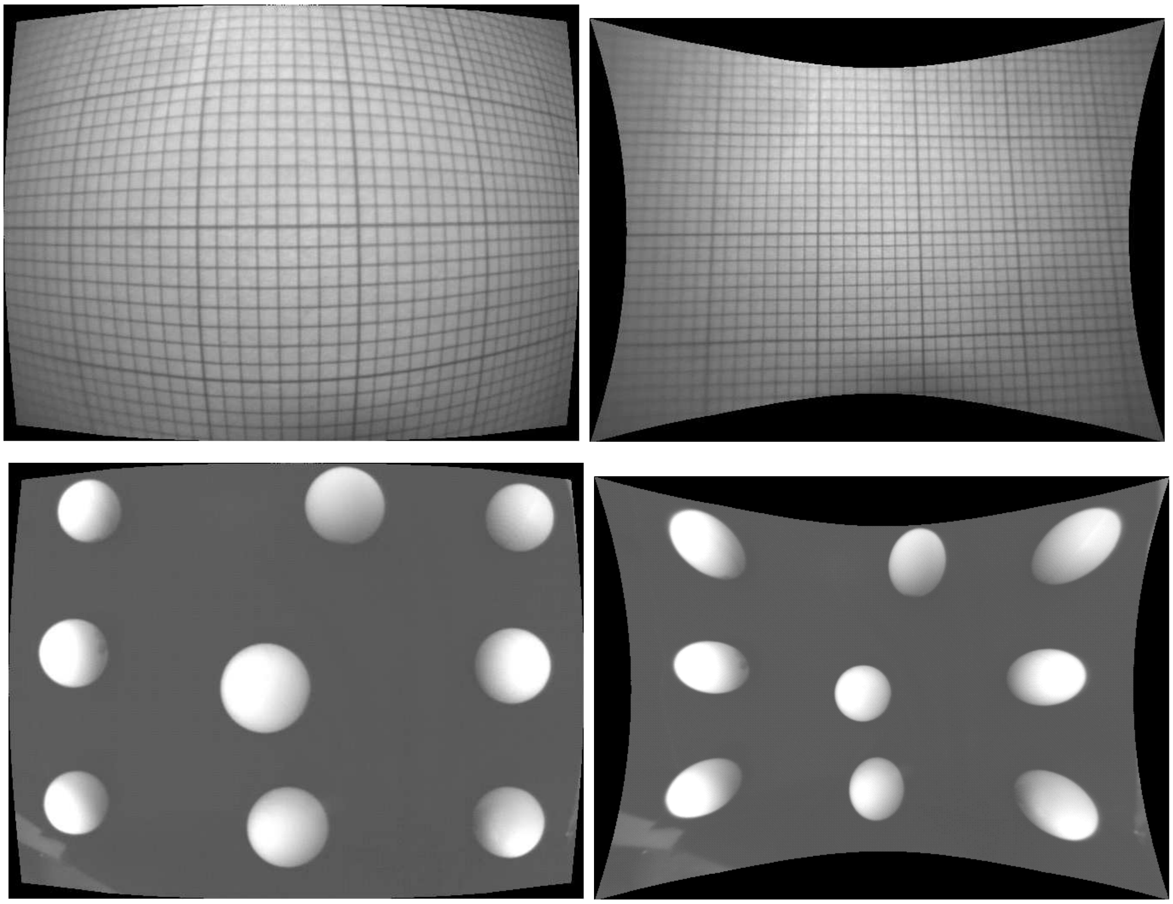

Figure 1: 3D lines look straight in perspective projection (top right) but curved in stereographic projection (top left). Spheres look circular in stereographic projection (bottom left) and elliptical in perspective projection (bottom right).

more than 180 degrees, less than the field of view of some lenses. In fact, real perspective lenses are limited to a field of view less than 140 degrees because insufficient light reaches locations further from the optical axis. In an ideal perspective image, objects change shape as they move across the field of view. In particular, symmetric objects appear asymmetrical when they are in the periphery of the image (figures 1 and 2), (Liu et al. 1993). In practice, real wide-angle lenses have substantial "barrel distortion," deviating from ideal perspective projection in a way that decreases the distortion of symmetric objects but makes straight lines appear bent (figure 3).

These problems have received little attention in computer vision, because researchers, with limited exceptions (Kato et al. 1994, Oh and Hall 1989, Stevenson and Fleck 1994), have avoided wide-angle lenses. Object recognition and "self-calibration" algorithms use normal to narrow-angle lenses (field of view less than 55 degrees), because they assume that perspective projection is approximately orthographic (Huttenlocher 1988, Liu et al. 1993, Mukherjee 1994, Pillow et al. 1994, Zerroug and Nevatia 1993) or because they assume no radial distortion is present (Dron 1993, Faugeras 1992, Forsyth et al. 1992, Forsyth et al. 1994, Hartley 1992, Liu et al. 1993,

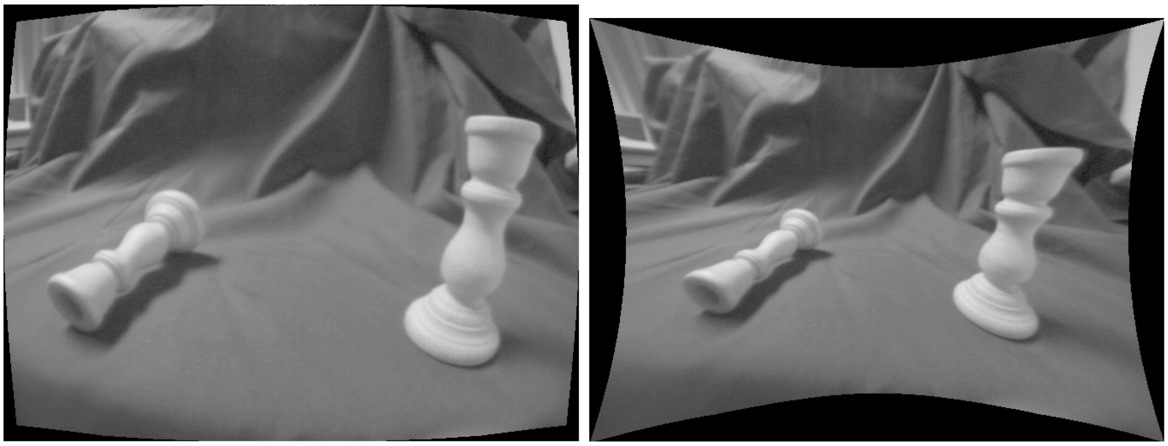

Figure 2: Rotationally symmetric objects appear locally symmetric in stereographic projection (left), but not in perspective projection (right).

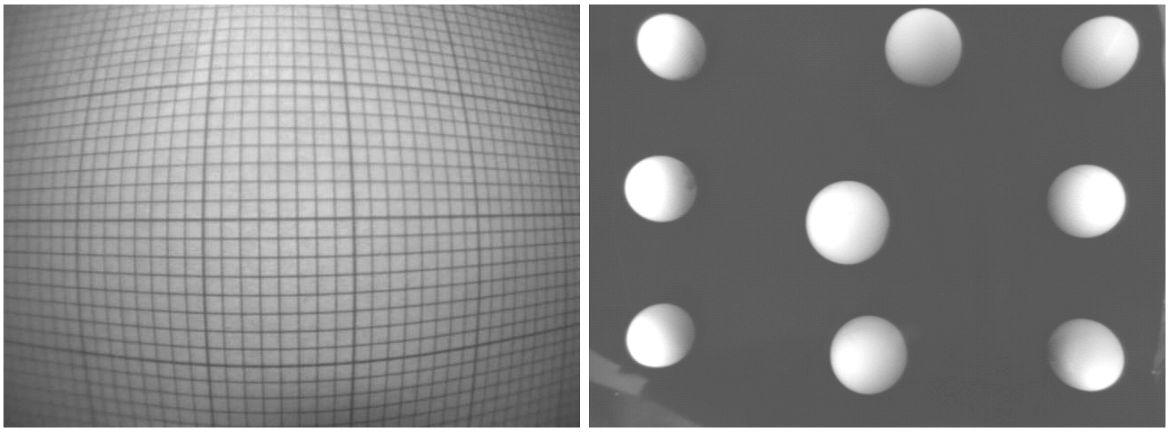

Figure 3: The output of our Computar 2.6 mm lens (NEC NX18A Camera) is midway between stereographic and perspective projection. Spheres are slightly elliptical and lines bend by ± 10 degrees.

Rothwell 1992). Calibration algorithms incorporating models of radial distortion have been tested on lenses with fields of view less than 78 degrees, typically less than 65 degrees (Du and Brady 1993, Fiala 1994, Tsai 1986, Weng et al 1992). Our experiments (Stevenson and Fleck 1994) suggest that their low-order polynomial models of the radial distortion function are inadequate for wide-angle lenses.

It is time to re-examine whether perspective projection is the most suitable imaging model. Perspective projection is only one of several mathematically convenient projections consistent with the optics of real camera systems. This paper will show that one of these alternative models, based on stereographic projection of the viewing sphere, is a better choice for general-purpose imaging. The stereographic projection model represents almost the entire field of view, reduces intensity drop-off in the periphery of the image, and more closely approximates the geometry of real wide-angle lenses.

Stereographic projection preserves circularity and, thus, projects 3D local symmetries onto 2D local symmetries. Perspective properties such as straightness and bitangency are not preserved globally. However, for small regions anywhere in a stereographic image, the projection from 3D approximates perspective projection near the optical axis. Therefore, perspective properties are preserved approximately over small regions. Furthermore, affine (scaled orthographic) representations for small compact features can be used throughout the field of view.

This paper will begin with three sections motivating the usefulness of wide-angle images and introducing the alternative imaging models. Sections 5–8 will discuss properties of the imaging models in detail. Section 9 explains why calibration information is useful in interpreting images and why strange-looking images are not necessarily bad. Finally, the conclusions are presented in section 10.

2 Why bother with wide-angle lenses?

Although neglected in computer vision, wide-angle images are widely used in photography and in biological vision. A wide field of view simplifies four types of tasks:

- Mapping the local environment for visual search (Wixson and Ballard 1994), planning actions, navigation (Kato et al. 1994, Oh and Hall 1989), and detection of hazards (e.g. oncoming moving objects),
- Obtaining a representative sample of colors for color constancy (Forsyth 1990) or a large set of features for identifying one's current location,
- Photographing large objects (e.g. buildings, landscapes), nearby objects, and objects in confined spaces, and
- Robust analysis of egomotion.

The first requirement explains why surveillance cameras typically use wide-angle lenses, why prey animals may have horizontal fields of view as large as 360 degrees, and why a 140 degree field of

Figure 4: Perspective is exaggerated in wide-angle images: the head looks too large compared to the hand, the forward parts of the face are enlarged. Straight lines (e.g. the edges of the monitor, the power conduit on the far right) may be bent.

view is required for an Iowa driver's license (Iowa 1990). (The normal human binocular field of view is about 180 degrees horizontally and 120 degrees vertically.)

Peripheral areas of the visual field play an important role in determining egomotion, because objects near the center of the field of view move in the same direction under camera pan (or tilt) as they do when the camera is translated horizontally (or vertically). It is easy to tell when there is a translational component to the motion—objects at different depths move different distances—but it is difficult to reliably estimate how much of the motion is translational. In peripheral areas, objects move in different directions under the two components of camera motion, making it easier to separate translation from pan and tilt.

Wide-angle lenses are frequently used in TV news photography, particularly in photographing indoor scenes. These lenses are also used in some TV ad campaigns to make the actors seem "in your face." The output of a wide-angle lens has three distinctive features:

- Background objects seem implausibly far away or implausibly small, compared to foreground objects,
- The front parts of faces (e.g. noses) seem unnaturally prominent, and
- Straight 3D lines near the edges of the image may appear curved.

Figure 4 shows extreme examples of these effects.

The images in this paper were produced by an ordinary, inexpensive C-mount lens, whose field of view is 116.5 degrees horizontally. We calibrated its radial distortion using techniques described in (Stevenson and Fleck 1994). This field of view is not at the limit of what is commercially available: C-mount lenses have fields of up to 138 degrees, 35mm lenses up to 220 degrees (Hedgecode 1993, Kingslake 1989, Peterson 1991). Wide-angle images can also be assembled from two or more images (Hartley 1994).

Figure 5: A point X on the viewing sphere can be represented by its angular distance α from the optical axis and its angle β around the optical axis. Projection onto the plane preserves β and puts the point at distance $p(\alpha)$ from the image center.

The theoretical imaging models used in computer vision must be able to handle wide-angle images, so that we can emulate biological systems and so that we can process the full range of available images.

3 Generalized imaging model

The imaging geometry of most cameras used in computer vision and photography, though not aerial slit cameras (Hartley and Gupta 1994, Kingslake 1992), can be described as a two-stage process: perspective projection of the 3D world onto a unit sphere (the *viewing sphere*) followed by a projection π of the sphere onto the image plane. The spherical image is convenient for theoretical analysis, because it represents the entire field of view uniformly (Cipolla and Blake 1992, Horn 1986, Malik and Rosenholtz 1994). However, a flat image simplifies storage of digitized images, low-level image processing (e.g. convolution), geometry, and film manufacture. Although it is not generally known in computer vision, several choices for π are used in optics and in mathematics.

This paper will consider only projections rotationally symmetric about the optical axis (see appendix A for panoramic cameras). Assume that the viewing sphere is a unit sphere centered at the origin and that $C = (0, 0, 1)$ is the direction of gaze. A point $X = (x, y, z)$ on the viewing sphere can be represented by two angles (figure 5). The angular distance from C is $\alpha = \mathrm{atan}(q, z)$, where $q = \sqrt{x^2 + y^2}$. The angular position around C is $\beta = \mathrm{atan}(y, x)$. Standard camera systems then map X onto the point at distance $p(\alpha)$ from the origin and angle β from horizontal.

There are at least five reasonable choices for the radial projection function p:

$p(\alpha) = k \tan \alpha$ (perspective projection)

$p(\alpha) = k \tan(\frac{\alpha}{2})$ (stereographic projection)

$p(\alpha) = k\alpha$ (equidistant projection)

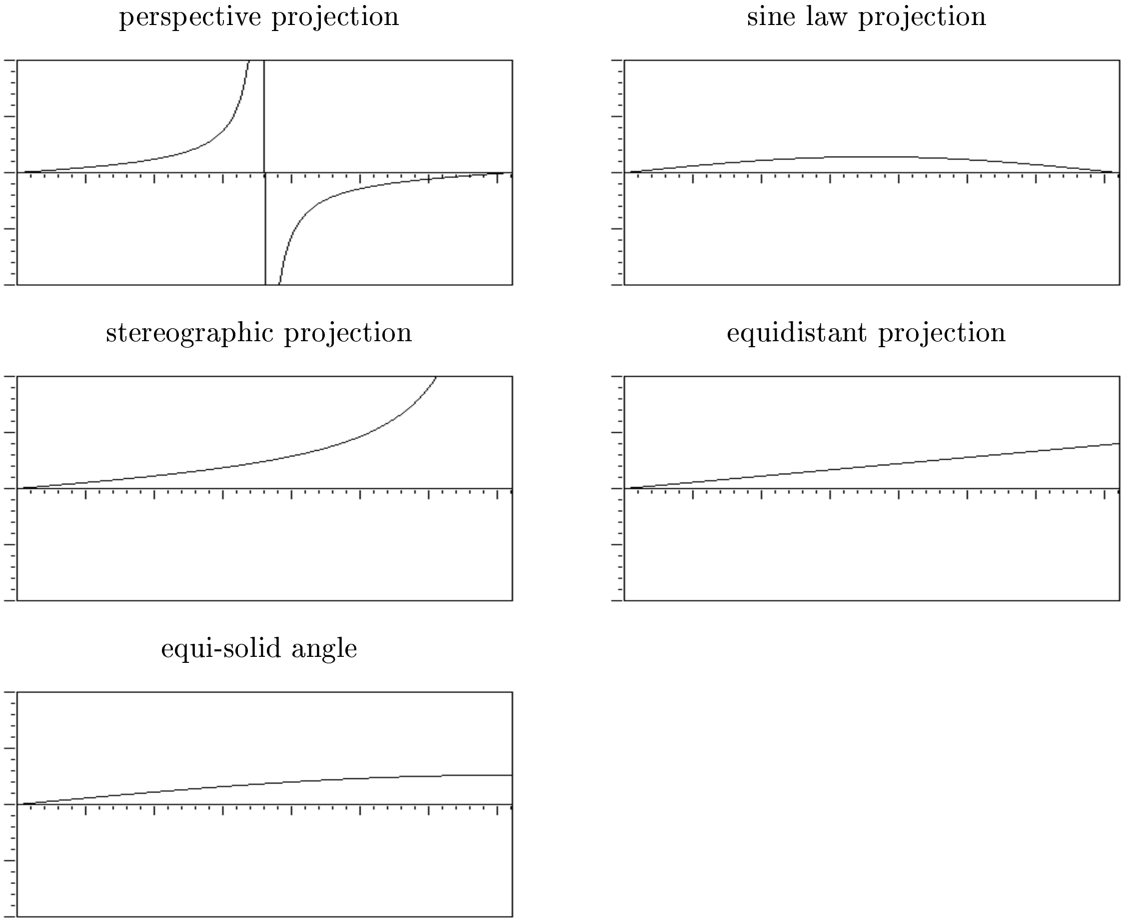

Figure 6: Five radial projection functions, mapping the angles $[0, 180]$ degrees onto distances from the image center, normalized to have the same value at 45 degrees.

$p(\alpha) = k \sin \alpha$ (sine-law projection)

$p(\alpha) = k \sin(\frac{\alpha}{2})$ (equi-solid angle projection)

where k is a scaling constant. Graphs of these functions are shown in figure 6. Geometrical constructions for three of them are shown in figure 7.

In theory, narrow angle and some wide-angle lenses are designed to emulate perspective projection. Some wide-angle and all extremely wide-angle ($\geq 140°$ field of view) lenses are designed to emulate equidistant, sine law, or equi-solid angle projection (Kingslake 1989, Ray 1994, Smith 1992). In practice, designing a lens involves choosing a compromise between several competing goals: projection geometry, size, even illumination, sharp focus, and cost. Therefore, the radial projection function for a real lens may lie almost anywhere in the range between perspective and sine-law projection (Smith 1992), as illustrated in figure 8 from (Stevenson and Fleck 1994). Thus, wide-angle lenses must be calibrated for precision applications (Smith 1992, Stevenson and Fleck 1994,

Figure 7: Geometrical constructions of the image of a point X under sine law projection (left), perspective projection (middle), and stereographic projection (right).

Tsai 1986).

Stereographic projection is widely used in mathematics (Berger 1987, Hahn 1994, Hilbert and Cohn-Vossen 1932) and occasionally in graphics (Hill 1990), shape from shading (Horn 1986), and cartography (Monmonier 1991). Though known in optics, it is rarely used in lens design (Ray 1994). However, there do not seem to be any technical obstacles preventing construction of a stereographic lens, if there were a market for one. Furthermore, in the front half of the field of view (up to 90 degrees from the optical axis), the differences between stereographic, equidistant, and equi-solid angle projection are small, so equidistant or equi-solid angle lenses can be used as approximate stereographic lenses. Therefore, any of these five projections is a possible choice for our ideal camera model.

4 Choosing an imaging model

In narrow-angle imaging, images from all five models approximate spherical images, so the choice of ideal projection function has little practical consequence. Away from the optical axis, however, the different models become significantly different, in some cases sometimes infinitely different. This paper will evaluate each projection using three criteria:

- How well it preserves metric properties of the spherical image (angular position and area),
- How wide a field of view it can represent, and
- What shape features it preserves, both globally and locally.

These properties are summarized in table 1.

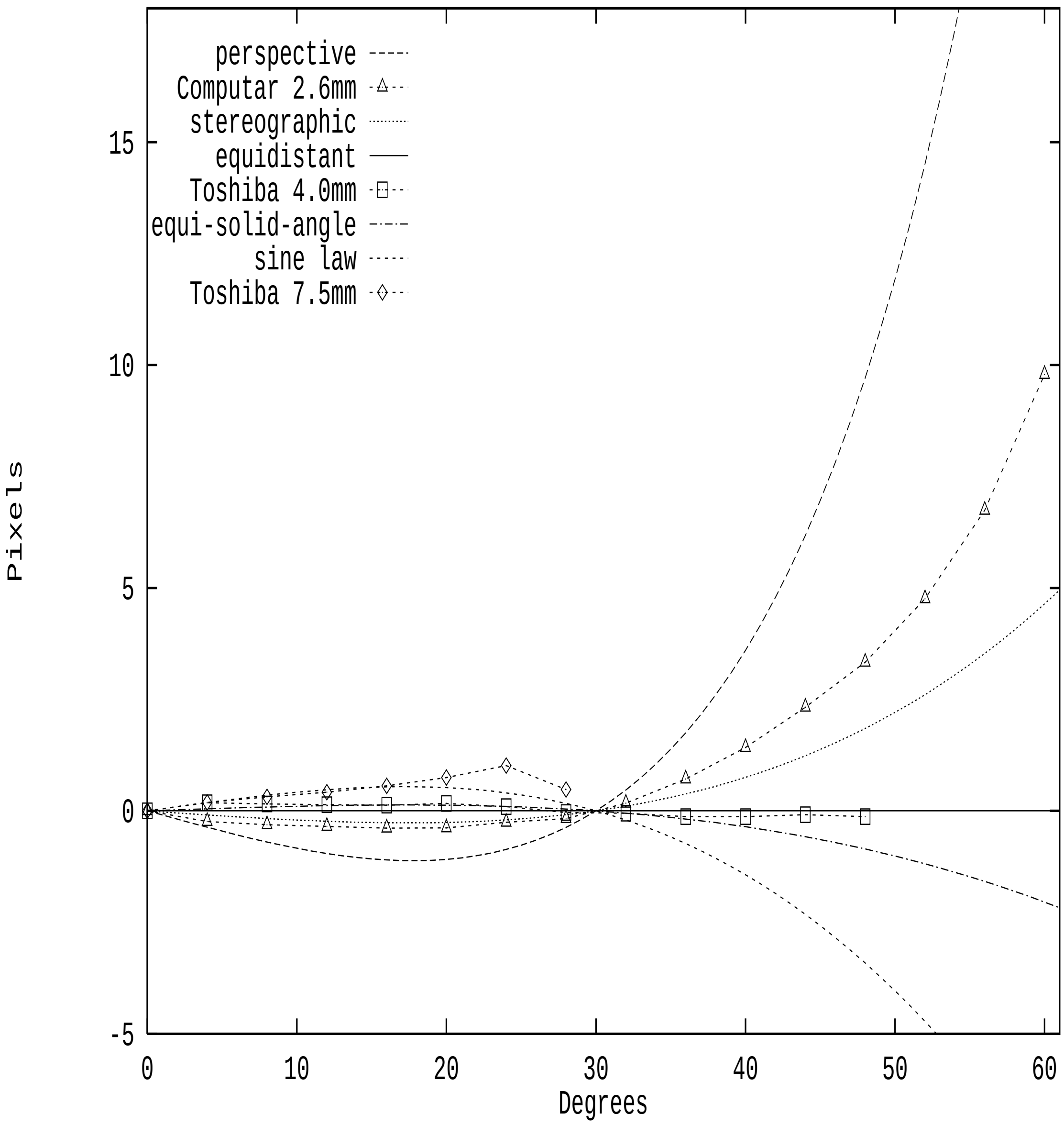

Figure 8: Radial projection functions for three lenses (2.6mm C-mount on 1/2" format camera, 4mm and 7.5mm lenses on a Toshiba miniature camera) compared to four idealized projection models. Each was normalized to have the value 30 at 30 degrees and the line $y = x$ was subtracted.

Table 1: Properties of projections (* = approximate)			
	metric properties	max field of view	shapes preserved
perspective		$140°/180°$	straightness
stereographic		$250°/360°$	circles intersection angles local shape (*)
equidistant	angular position angular area (*)	$345°/360°$	
equi-solid angle	angular area angular position (*)	$360°$	
sine law		$180°$	

This analysis serves two purposes. First, it will be used to argue that stereographic projection is the best choice for a single, general-purpose projection model. Specifically, we will see that

- It can represent the full 360 degree field of view; metric distortion and intensity drop-off are modest within the front half,
- It preserves a range of global shape properties analogous to those preserved by perspective projection, and
- The shape of small regions does not depend on their location in the field of view.

Perspective and sine law projection, by contrast, cannot handle very wide fields of view. Equidistant, equi-solid angle, and sine law projection do not preserve powerful shape properties. And, for all the non-stereographic projections, the shape of a small region changes as it moves across the field of view.

This analysis can also be used to choose a projection for a particular application. For example, time to contact equations for round objects have a simpler form in equi-solid angle projection, symmetry testing is best done in stereographic projection, and it is easiest to analyze curvature sign in perspective projection. If calibration information is available, an algorithm may be significantly simplified by first transforming the image or contours into the appropriate projection. The transformation between perspective and stereographic coordinates is particularly simple (appendix B). In a specialized application, it may be convenient to choose a lens implementing the projection most appropriate to the algorithms.

5 Preservation of angular position and area

Equidistant and equi-solid projections are useful in applications which measure angular positions or angular areas. Traditional applications include maritime naviation (Monmonier 1991), astrophotography, and mapping the human visual field (as in opthalmology and psychophysics). In computer vision, these projections might be useful for robot navigation (Oh and Hall 1989) or in time to contact algorithms.

All five projections preserve angular position about the viewing direction (β). Thus, great circles through $C = (0, 0, 1)$ on the viewing sphere project to straight lines in the image. In addition, equidistant projection preserves the radial angle (α). Equidistant and equi-solid angle projection closely approximate one another (figure 6), particularly in the front half of the visual field, and so equi-solid angle projection approximately preserves radial angle.

Equi-solid angle projection preserves the area of regions on the spherical image, up to a scaling constant (Ray 1994). To show this, it is sufficient to consider filled circular regions centered about the optical axis. Preservation of area for other regions can then be derived by adding and subtracting filled circular regions, and by using the fact that the projection is rotationally symmetric about the optical axis.

Consider a filled region R of angular radius α on the (unit) viewing sphere. By standard geometrical formulas (Beyer 1984), the surface area of R is $2\pi(1 - \cos\alpha)$. R projects onto a circle of radius $k\sin(\frac{\alpha}{2})$ on the image plane. The area of this circle is $\pi k^2 \sin^2(\frac{\alpha}{2})$. But $\sin^2(\frac{\alpha}{2}) = \frac{1-\cos\alpha}{2}$ (a standard trignometric identity). Therefore, the two areas differ by a constant factor of $\frac{k^2}{4}$.

In other projections, regions near the optical axis appear larger or smaller than those in the periphery. To calculate the change in area, consider a thin annulus on the viewing sphere, of angular radius α and width 2β, centered around the optical axis. Its projection on the image has area proportional to $p^2(\alpha+\beta) - p^2(\alpha-\beta)$ where p is the radial projection function. In the limit, as β goes to zero, ignoring scaling constants, this is simply the derivative of $p^2(\alpha)$.

For equi-solid angle projection $p^2(\alpha) = \sin^2(\frac{\alpha}{2})$, which is equal to $\frac{1}{2}(1 - \cos\alpha)$. Therefore, its derivative is proportional to $\sin\alpha$. To calculate the area of a small region seen at angle α, relative to its area if seen at the optical axis, we compare $\sin\alpha$ to the value of $p^2(\alpha)$ for other projections, ignoring constant factors,

For equidistant projection, the derivative of $p^2(\alpha)$ is proportional to α. So the change in area is proportional to $\frac{\alpha}{\sin\alpha}$. Similarly, for sine-law projection, the change in area is proportional to $\cos\alpha$. And, for perspective projection, the change in area is proportional to $\frac{1}{\cos^3\alpha}$. For stereographic projection, the derivative of $p^2(\alpha)$ is proportional to $\frac{\sin(\frac{\alpha}{2})}{\cos^3(\frac{\alpha}{2})}$. So the change in area from the viewing sphere to the image is proportional to $\frac{\sin(\frac{\alpha}{2})}{\sin\alpha\cos^3(\frac{\alpha}{2})}$. By trigonometric identities, this is proportional to $\frac{1}{\cos^4(\frac{\alpha}{2})}$ which is proportional to $\frac{1}{(1+\cos\alpha)^2}$.

These formulas reveal the extent of area distortion in the other projections. Equidistant projection approximately preserves area over much of the field of view because $\sin\alpha$ is approximately proportional to α. Stereographic projection changes area by only a factor of 4 in the front half of the field of view, but the areas are infinitely expanded at 180 degrees from the viewing direction. Perspective and sine-law projection have unbounded distortions of area at 90 degrees from the viewing direction: perspective expands area infinitely and sine-law collapses regions to zero area.

6 Field of view

The maximum field of view represented by an imaging model is determined by two factors. First, the mathematical projection must be one-to-one and continuous throughout the field of view. Second, it must be possible to construct a physical lens implementing the projection. Lens construction is primarily limited by the need to prevent the peripheral areas of the image from becoming much darker than the central areas. Small changes in illumination can be removed by calibration or tolerated by making vision algorithms insensitive to them. However, if large differences are present (e.g. a factor of 20), features in the periphery cannot be made visible without saturating the center of the image.

The five projections have different topological properties. Stereographic, equidistant, and equisolid angle projections map nearly all the viewing sphere one-to-one onto the plane (see figure 6). Thus, they can be used to construct extremely wide-angle images for applications such as egomotion analysis or navigation. Perspective and sine-law projection are one-to-one only on the front half of the viewing sphere, so their field of view is limited to 180 degrees. This limitation applies whether the image is produced by a real lens or synthesized (e.g. by pasting together several smaller images).

In addition, in a real camera system, the image of a patch of Lambertian surface typically appears brighter if it lies near the optical axis and less bright when it is in the periphery of the field of view. This intensity drop-off is due to the fact that

- the lens transmits less of the light from solid angles in the periphery, and
- solid angles in the periphery cover more image area than solid angles near the optical axis.

Minimizing such variation in intensity is an important goal in lens design.

Consider a small patch P of 3D Lambertian surface lying at angle α from the optical axis of the camera. If the lens projects the world onto a spherical image surface, locations in the image of P have intensity values proportional to $\cos\alpha$. This drop-off occurs because the lens aperture, as seen from P, is tilted at angle α. Thus, its apparent area is reduced (foreshortened) and fewer of the rays coming from P enter the lens (more miss the lens and go off into space). Vignetting by other components of the lens may further alter the amount of light transmitted (Kingslake 1989).

Conservation of energy implies that intensity values drop when light from a given solid angle is spread out over a larger image area. So, the intensity drop-off for flat images is calculated by dividing the intensity values for a spherical image plane by the change in area equations derived in section 5. For example, the change in area for perspective projection is proportional to $\frac{1}{\cos^3\alpha}$, so the intensity values are proportional to $\cos^4\alpha$. This formula is well-known, though usually derived in a different way (Horn 1986, Kingslake 1989, Kingslake 1992, Ray 1994, Smith 1992). This steep drop-off in intensities limits the field of view of ideal perspective lenses to about 120-125 degrees, corresponding to a reduction in intensity by a factor of 16–20 (Ray 1994).

By adjusting the details of the optics, the amount of light transmitted by the lens can be made nearly uniform. For example, the front elements of wide-angle lenses are curved and coma may be deliberately introduced, so that more of the light from peripheral objects can enter and exit the lens (Kingslake 1989, Ray 1994, Smith 1992). The illumination in the center of the image can be reduced using graduated neutral density filters (Ray 1994). Thus, the practical limit for perspective lenses is a $\cos^3 \alpha$ drop-off in illumination, which allows a field of view up to 140 degrees (Ray 1994).

The other four projections reduce the image area covered by peripheral regions. Therefore, lenses implementing these projections have a slower drop in intensities and larger maximum fields of view. Depending on whether the amount of light transmitted by the lens is uniform across the field of view or varies as $\cos \alpha$, perfectly even illumination is produced by either equi-solid angle projection or sine-law projection (Kingslake 1989, Ray 1994). Equidistant projection has a very slow intensity drop-off, between $\frac{\sin \alpha}{\alpha}$ and $\frac{\cos \alpha \sin \alpha}{\alpha} = \frac{\sin(2\alpha)}{2\alpha}$. For stereographic lenses, the drop-off is between $(1 + \cos \alpha)^2$ and $\cos \alpha (1 + \cos \alpha)^2$. This imposes a maximum field of view of 250 degrees, larger than that of existing lenses.

Thus, three projections—equidistant, stereographic, and equi-solid angle—support very wide fields of view. Sine-law projection is limited to a 180 degree field of view. Perspective projection is limited to a 180 degree field of view for synthetic images, and a 140 degree field of view for real images.

7 Global shape

Because they are continuous, all five projections preserve basic curve features such as curve intersections, tangency, and sharp corners. However, only two projections preserve important aspects of the shape of regions. Perspective projection preserves straightness: it maps arcs of great circles on the viewing sphere (e.g. the images of 3D lines) onto straight lines, line segments, and rays in the image. Stereographic projection preserves circularity: it maps circles on the viewing sphere (e.g. images of 3D spheres) onto circles and lines on the image plane (appendix C). Stereographic projection is also *conformal*: it preserves the angle at which two curves intersect (Berger 1987, Hahn 1994, Hilbert and Cohn-Vossen 1932).

Shape analysis is typically based on the outline of an object (possibly including internal boundaries). In order to be useful in most computer vision applications, particularly object recognition, a feature of the outline's shape must have three properties:

- Frequent visibility: the feature must appear on the outline from many viewpoints.
- Good projection: whenever the feature is visible, it must appear in similar form on the viewing sphere.
- 2D invariance: the feature must appear in similar form on the flat image, regardless of how the viewing sphere is projected onto the image.

Clearly, 2D invariance is only possible within a restricted class of projections, such as all stereographic projections of the viewing sphere or all perspective projections.

To avoid special cases in the following discussion, lines and planes will be considered degenerate types of circles and spheres. When convenient, flat 3D objects will be treated as having a nonzero thickness and sharp corners will be treated as slightly rounded. Finally, the discussion will consider only idealized, exact features, with the understanding that the results apply approximately to features approximately matching these models.

7.1 Perspective

Because perspective projection preserves straightness, three basic features are used for recognition in perspective images (Forsyth et al. 1992, Huttenlocher 1988, Mokhtarian and Mackworth 1986, Rothwell 1992, Kriegman 1993, Liu et al. 1993): lines, curvature zero-crossings, and bitangents. Specifically, a 3D line projects onto a line in the image. If a 3D surface is tangent to a plane at two points, the images of these points will be tangent to a common line whenever they both appear on the outline. If the tangent plane at some point A crosses the surface, the tangent line at the image of A will cross the outline whenever A lies on the outline (except for an infinitely small set of degenerate viewpoints).

For a point A to appear on the outline, it must meet the following *basic visibility conditions*:

(1) The ray from the lens center to A is tangent to the surface,

(2) The curvature of the surface along this ray is positive at A (so it is not occluded locally),

(3) No other part of the surface lies in front of A, and

(4) The angular position of A lies within the camera's field of view.

It is relatively easy to understand when conditions (2)-(4) will be satisfied: interesting questions about how frequently a feature will be visible typically concern condition (1).

A single curvature zero-crossing or bitangent in 3D is frequently visible only if it lies on a sharp edge. In order for a smooth 3D surface to frequently generate a particular 2D feature, it must contain an infinite family of 3D features. For example, if a rotationally symmetric object has a concavity in its outline, it will have a rotationally symmetric family of bitangent planes "capping off" the concavity. From any particular viewpoint, at most two representatives of the family generate bitangents of the outline. The family is visible from many viewpoints, in the sense that two of its members are visible.

Two representatives from a rotationally symmetric family of bitangents are sufficient for reconstructing the projection of the axis of symmetry (Liu et al. 1993). Each pair also defines two viewpoint-invariant points along the axis. The cross-ratio of four such points remains constant for all views of the object and can be used for recognition (Forsyth et al. 1992, Liu et al. 1993). The weak link in this method is that existing algorithms for pairing corresponding bitangents (Liu et al. 1993) assume the 2D outline is approximately symmetric, which is true only in narrow-angle images.

7.2 Stereographic

Because stereographic projection preserves circularity, it preserves local symmetry. A group of points on a planar or spherical image are said to have a *local symmetry* or be a *symmetry group* if there is a circle (the *symmetry circle*) tangent to the contour at all points in the set (Blum and Nagel 1978, Brady 1983, Brady and Asada 1984, Bruce et al. 1985, Giblin and Brassett 1985). Similarly, a set of points on a 3D surface have a local symmetry if there is a sphere (the *symmetry sphere*) tangent to the surface at all points in the set (Bruce et al. 1985, Giblin and O'Shea 1990). For example, the points in any cross-section of a rotationally symmetric object form a symmetry group.

A symmetry group of a smooth 3D surface appears as 2D symmetry (both on the viewing sphere and in the stereographic image) whenever two of its points meet the basic visibility conditions from section 7.1. Symmetry groups containing only two points (e.g. most of those in flat objects) rarely generate 2D symmetries, because they rarely satisfy condition (1). A larger 3D symmetry group, however, may generate a 2D symmetry visible in many viewpoints. For example, a cross-section of a surface of rotation is frequently visible as a symmetry (appendix D).

Objects containing extended families of local symmetries visible from many viewpoints include:

- A sphere, partial sphere, or ellipsoid,
- A surface of rotation,
- A surface formed by sweeping a sphere along a space curve, with or without smooth changes in the sphere's radius (Pillow et al. 1994),
- A partial surface of rotation containing sections on opposite sides of the axis of rotation,
- Complex objects containing the above components, and
- Objects approximating the above models (e.g. fingers, hexagonal pencils, straight homogenous generalized cylinders with fat elliptical cross-sections).

When a 3D symmetry generates a 2D symmetry of the outline, the 2D symmetry circle is the projection of the 3D symmetry sphere. However, the center of the 2D circle is not, in general, the projection of the center of the 3D circle. For example, for a circle of angular radius 20 degrees, located at 60 degrees from the optical axis, the projection of the 3D center is displaced from the 2D center by 10% of the 2D radius. In fact, none of the alternative symmetry centers proposed by previous authors (Blum and Nagel 1978, Brady 1983, Brady and Asada 1984, Bruce et al. 1985, Giblin and Brassett 1985, Leyton 1988) is preserved under the projection from 2D to 3D.

Viewpoint invariant construction of the axis is possible, however, for an elongated region with a straight axis. If the axis of the 3D region is straight, it projects into the stereographic image as a circle C. For each local symmetry along the axis, C must intersect the 2D symmetry circle at a right angle. This is clearly true if C projects onto a straight line through the camera center. The general result follows from the fact that stereographic projection preserves the angle at which two curves intersect on the viewing sphere.

Now, choose three symmetry groups belonging to the region. (The choice does not need to be viewpoint invariant.) Except for degenerate cases, there is a 1D family of circles perpendicular to a fixed pair of circles (Hahn 1994) and exactly one circle intersecting all three symmetry circles at right angles. This circle must be the image of the 3D symmetry axis. This construction should work approximately for short sections of regions whose axes curve only gently.

Suppose that we identify four viewpoint invariant features which lie along a line L in 3D. The cross-ratio of their locations in the stereographic image is invariant under changes in viewpoint.[2] If L projects onto a straight line through the image center, the result for stereographic projection can be derived from the result for perspective projection using the transformation in appendix B. The general result follows from the fact that cross-ratios are preserved under transformations between two stereographic projections of the viewing sphere (the *Möbius transformations*) (Hahn 1994). The weak link in this method is detecting a wide enough class of viewpoint-invariant points.

7.3 Wrap-up

Thus, stereographic projection preserves a set of global shape features similar to those preserved by perspective projection, with a similar range of viewpoint-invariant properties. Perspective projection seems to offer more promising features for recognizing flat objects (Forsyth et al. 1991, Forsyth et al. 1992, Liu et al. 1993, Rothwell 1992). On the other hand, a typical rotationally symmetric object contains extended sequences of local symmetries. Thus, aspect ratio constraints can be used to prune symmetries created by random alignments and detect prominent symmetries in linear time (Brady and Asada 1984, Connell and Brady 1987, Fleck 1986, Fleck 1994). This approach cannot be used to filter bitangents, because they involve small discrete sets of points and, thus, all known algorithms for detecting bitangents are $O(n^2)$ (Kriegman 1993, Liu et al. 1993, Ponce 1989).

The two projections seem to provide complementary types of shape information. This suggests that better shape descriptions could be produced by using calibration information to transform between the two projections (see appendix B). Alternatively, for relatively small objects, perspective features can be detected approximately in stereographic images, as we will see in the next section.

8 Local shape

Many objects, features, and surface patches encountered in computer vision subtend only small visual angles. For example, the terminal screen in the upper left of figure 9 has an angular radius of perhaps 10 degrees. Because stereographic projection is conformal, it approximately preserves the shape of small regions on the viewing sphere (Berger 1987, Hahn 1994, Hilbert and Cohn-Vossen 1932). Therefore, small regions look approximately the same no matter where they lie in the field of view: their projection approximates perspective projection near the optical axis.

This narrow-angle approximation can be made as accurate as required by considering a small

[2] This cross-ratio should be computed by representing the 2D feature points as complex numbers as in (Hahn 1994), not by computing distances along the curve which is the image of L.

Figure 9: In stereographic projection (left), the shape of an object is approximately the same, regardless of where it occurs in the field of view, unless it subtends a large visual angle. In perspective projection (right), objects near the periphery (e.g. the phone) are visibly distorted.

Table 2: Error in circle centers as a percentage of circle radius

angular	angular radius			
position	5°	10°	15°	20°
20°	0.8	1.5	2.3	3.1
40°	1.6	3.2	4.8	6.4
60°	2.5	5.1	7.6	10.2
80°	3.7	7.3	11.0	14.8
100°	5.2	10.4	15.7	21.0

enough region. For example, table 2 shows the distance between the center of a circle in image coordinates and the projection of its center in spherical coordinates. These are the maximum errors that would occur if centers of 2D symmetry circles were used to locate the 3D symmetry axis, or if the ratio of two distances in the stereographic image was used to estimate the ratio of the corresponding angular distances on the viewing sphere. The errors are substantial for large peripheral circles, but modest for small circles within a wide field of view.

In stereographic projection, the sides of polygonal regions are curved. However, this distortion is bounded. The images of straight 3D lines are great circles on the viewing sphere. Of these, the equator of the viewing sphere ($\alpha = 90°$) projects onto the smallest circle in the stereographic image, with radius k (where k is the scaling constant for stereographic projection). This sets a global upper bound on the curvature of the images of straight lines. If a contour has higher curvature that this, its pre-image in 3D cannot have been straight.

Because the sides curve in stereographic projection, the angles in an n-gon are larger than the $180(n-2)$ degrees expected for a planar n-gon. Because stereographic projection preserves intersection angles, the excess in a stereographic image is exactly that for an n-gon on the unit sphere: its angular area. If the angular radius of the n-gon is α, its area on the viewing sphere is smaller than the area covered by a circle of radius α: $360(1 - \cos\alpha)$. So, for n-gons with radii 5, 10, 15, and 20 degrees, the error in the sum of the angles is only 1.4, 5.5, 12.3, and 21.7 degrees.

Thus, we can compute perspective features (e.g. bitangent lines, curvature zero crossings) approximately from a stereographic image, so long as the points belonging to each feature all lie within a small neighborhood. Bounds on the errors in this approximation can be derived from bounds on the field of view of the camera. If no better information is available, worst-case bounds can be derived from the fact that no commercial lens has a field of view greater than 220 degrees.

If an object spans a small range of depths relative to its distance from the camera, as well as subtending a small visual angle, its projection onto the stereographic image approximates scaled orthographic projection. Images of the object from different viewpoints are then approximately affine transformations of one another. A large number of texture (Malik and Rosenholtz 1994), stereo (Otto and Chau 1989), motion (Cipolla and Blake 1992), and shape algorithms (Huttenlocher 1988, Lamdan 1988, Mukherjee 1994, Pillow et al. 1994, Zerroug and Nevatia 1993) have been developed using affine matching and affine invariants. When they are applicable, these techniques are simpler and more powerful than general perspective methods (Forsyth et al. 1991, Forsyth et al. 1992, Liu et al. 1993, Rothwell 1992).

In the four projections other than stereographic, small regions change shape as they move across the field of view. For example, in perspective projection, a small spherical object at angle α from the optical axis appears as an ellipse of elongation $\frac{1}{\cos\alpha}$. The distortion is small within about 25 degrees of the optical axis, leading some authors (Malik and Rosenholtz 1994, Pillow et al. 1994) to dismiss this effect as small. However, the ellipse's elongation increases rapidly as the object's position moves past 45 degrees: 1.1 at 20 degrees, 1.3 at 40 degrees, but 2.0 at 60 degrees and 5.8 at 80 degrees. The distortion of shape is obvious in figures 1, 2, and 9.

This distortion of peripheral regions occurs no matter how small an angle the regions subtend. As a result, stereographic features (e.g. local symmetry) cannot be computed from wide-angle perspective images without large errors. Similarly, affine algorithms can be applied only to objects lying near the optical axis. The full extent of the problem with perspective projection cannot be observed empirically without calibrating one's camera. Barrel distortion is common even in lenses with modest fields of view—a slight bending of straight lines may not be obvious to the casual observer—and it will tend to move the projection towards stereographic and improve the quality of affine approximations.

Thus, in wide-angle perspective projection, we can use only perspective features. In stereographic projection, however, we can use not only stereographic features, but also (locally) approximate perspective and affine features. This asymmetry is a strong argument in favor of stereographic projection.

9 Why do wide-angle images look strange?

No matter which projection is used, wide-angle images look strange. Either straight lines appear curved or spheres appear elliptical. Exaggerated perspective makes objects seem misshapen:

Figure 10: If a sphere lies at distance d from the center of the viewing sphere and it subtends an angle 2α, then its 3D radius is $d \sin \alpha$.

Figure 4 is not a flattering portrait of the author. Manuals of photography (Hedgecode 1993, Kingslake 1992, Ray 1994) often contain advice about how to avoid or conceal these artifacts: e.g. the horizon should be kept straight by making it pass through the center of the image, portraits should be taken with narrow-angle lenses. Furthermore, despite acknowledging some of the arguments given above, photographic texts present perspective projection as the ideal model.

The apparent discrepancy between the claims made in this paper and standard photographic practice is due to differences in goals between standard photography and imaging for computer vision. In computer vision, images produced by a video camera are transmitted directly to algorithms for interpreting them. In standard photography, images are displayed as prints or on a TV or projector screen. This flat 3D image is then projected through the human eye to produce the input to human visual processing.

When you look at a photographic print or a TV image, your natural tendency is to interpret it as if the image on your retina was generated by 3D objects in your environment. Therefore, objects in photographs look distorted unless the camera's field of view and the viewing distance of the print are chosen so that the images of objects in the print subtend the same angles that the objects would subtend if they were physically present (Kingslake 1992). Furthermore, because the first step in human vision is perspective projection (onto a spherical retina), the print must be created using perspective projection.

To understand why perspective seems exaggerated in wide-angle images, suppose that we see two spheres whose relative sizes we know. Let their 3D radii be r and lr, where l is known and r is unknown. Suppose that the images of the spheres subtend (observed) angles 2α and $2\alpha'$, with centers separated by (observed) angle β. The distances to the two spheres are then $\frac{r}{\sin \alpha}$ and $\frac{lr}{\sin \alpha'}$ (figure 10).

The distance between the centers of the spheres is then rk, where k is given by:

$$k^2 = \frac{1}{\sin^2(\alpha)} + \frac{l^2}{\sin^2(\alpha')} - \frac{2l \cos \beta}{\sin(\alpha)\sin(\alpha')} \tag{1}$$

(figure 11). This equation can be used to estimate the distance between two people, relative to the size of the people, because all human heads are approximately the same size. It also implies that a

Figure 11: The distance between two spheres of known relative sizes, can be calculated (relative to their radii) from their images on the viewing sphere.

surface formed by sweeping a sphere along a space curve (Pillow et al. 1994) can be reconstructed, except its overall size, from a single calibrated image.

For this calculation, the camera's field of view must be known, in order to convert image distances into angular distances. Human observers use calibration information for their eyes in interpreting 3D scenes. The human eye has a single, fixed field of view. Therefore, human perception algorithms need not be invariant to changes in field of view, i.e. zooming, because they do not occur. So, a human observing figure 4 at normal reading distance will interpret it using an erroneous estimate of the effective field of view. The constraints imposed by equation 1 will then imply that my hand is unnaturally small relative to my head, or my arm is unnaturally long.

More generally, a photographer can make two objects appear to have any desired separation by adjusting the camera's field of view. Consider two spheres lying at different distances from the camera. If we change the field of view of the lens and adjust the viewing distance so that the front sphere stays a constant size, the size of the back sphere will change. This mimics a change in the distance between the objects. Similarly, zooming a video camera increases or decreases the apparent separation between objects in an unnatural way. Photographers avoid zooming when this artifact would be significant, except as a deliberate special effect (Kingslake 1992).

Photographers favor narrow-angle lenses because standard-size prints subtend small angles at normal reading distance. Wide-angle images look strange because we typically see them from the wrong distance. Prints taken with non-perspective lenses look strange because this is the wrong way to produce prints to be shown to human observers. In computer vision, as in other types of scientific photography (e.g. astrophotography), these considerations are not important because the images are passed directly to processing algorithms.

10 Conclusions

In this paper, we have seen five alternative models for wide-angle imaging, all of which have special properties which might be useful in particular applications. Of these, only three have low variation in intensity values and support very wide-angle images: stereographic, equi-solid angle, and equidistant projection. Only stereographic and perspective projection preserve powerful shape properties useful in object recognition. Finally, because stereographic projection is conformal, it locally approximates narrow-angle perspective, so the projection of small, compact objects is approximately orthographic. This enables local affine techniques to be used anywhere in the field of view. Based on a combination of these properties, stereographic projection is the best choice for a general-purpose imaging model.

Acknowledgements

Dan Stevenson's calibration work (Stevenson and Fleck 1994) provided important data (e.g. figure 8) to back up my theoretical arguments. I would also like to thank Terry Boult, Michael Covington, David Forsyth, Brian Madden, Howard Moraff, Steve Osborne (Electronic Video Systems), Mike Wall, and Andrew Zisserman for useful tips.

Appendix A: Panoramic cameras

This paper considers only projections that are rotationally symmetric about the optical axis. A second class of projections are also used in optics and in cartography. In *panoramic projections*, the horizontal direction is special. A point on the viewing sphere is represented by two angles: $\gamma = \text{atan}(x, z)$ and $\delta = \text{asin}(y)$. The projection function π maps this point onto the point $(\gamma, p(\delta))$.

To implement such projections, panoramic cameras focus light through any of the lenses described in section 3, but a slit mounted behind the lens allows only a vertical strip of light to reach the film. The lens is then rotated, sweeping out a panoramic image on the film (Kingslake 1992). The projection function p depends on the lens chosen. Panoramic cameras are primarily used for photographing landscapes and large groups of people.

Panoramic projections have been studied primarily because of their applications in cartography. Some of the discussion in this paper can be generalized to panoramic projections. For example, a conformal projection (the Mercator projection) can be produced by setting $p(\delta) = \log(\tan(\frac{(\delta+90)}{2}))$ (Monmonier 1991, Berger 1987). An area-preserving projection is produced if $p(\delta) = \sin\delta$ (Berger 1987, Hilbert and Cohn-Vossen 1932, Monmonier 1991).

Appendix B: How are stereographic and perspective coordinates related?

If the camera is calibrated, it is straightforward to transform image locations between any two of the projections defined in section 3. In particular, the stereographic projection (x_s, y_s) and the perspective projection (x_p, y_p) of a point on the viewing sphere are related via the following equations adapted from (Horn 1986):

$$(x_s, y_s) = \left(\frac{2x_p}{1+\sqrt{1+r_p^2}}, \frac{2y_p}{1+\sqrt{1+r_p^2}}\right) \tag{2}$$

where $r_p^2 = x_p^2 + y_p^2$ and

$$(x_p, y_p) = \left(\frac{4x_s}{4-r_s^2}, \frac{4y_s}{4-r_s^2}\right) \tag{3}$$

where $r_s^2 = x_s^2 + y_s^2$.

These formulas assume that the radial projection function for perspective projection is $\tan(\alpha)$ and that for stereographic projection is $2\tan(\frac{\alpha}{2})$. To convert pixel coordinates, the formulas must be adjusted to incorporate appropriate image scaling constants.

Appendix C: Why stereographic projection preserves circles

It is well-known in mathematics that stereographic projection preserves circles. A geometrical proof can be found in (Hilbert and Cohn-Vossen 1932); (Hahn 1994) contains a proof using complex numbers. To sketch the main idea, suppose that point P lies at angle α from the optical axis NS (figure 12). The stereographic image P' of P lies where the image plane intersects the line passing through P and the far pole N. The two line segments ON and OP have the same length, implying that the two angles marked β are the same. Since $\gamma = 180 - \alpha$ (sum angles to one side of line NS) and also $\gamma = 180 - 2\beta$ (sum angles within triangle), $\beta = \frac{\alpha}{2}$, confirming that the formula given in section 3 agrees with this geometrical construction.

The two angles labelled δ must be the same, because they are both equal to $90 - \beta$ (add up the angles to one side of the tangent at P; NSP is a right triangle). So the tangent plane at P and the image plane intersect the projection line NP at the same angle. Let C be a circle on the viewing sphere, centered at P. C must lie in a plane parallel to the tangent plane at P.

To project the circle C onto the image plane, we construct a cone of rays from N passing through the points of C. The cross-section of this cone is an ellipse. Its axis of elongation lies in the plane of the cross-section shown in figure 12 (left). Because the image plane and the plane containing C cut the cone at the equal (but opposite) angles in this cross-section, their intersections

Figure 12: The tangent plane to the sphere at P and the image plane intersect the stereographic projection line NP at the same angle δ (left). Any circle centered at P lies in a plane parallel to the tangent plane at P. These three planes cut an elliptical cone at the same angle, so their intersections with the cone must have the same shape (right).

with the cone must have the same shape (figure 12, right). So, since C is circular, its image is also circular. If the point N happens to lie on the circle C, this construction will not work; in this case, all the projection rays are coplanar and C's image is a line.

Appendix D: Visibility of symmetries in surfaces of rotation

Cross-sections of surfaces of rotation are examples of 3D symmetry sets which frequently generate 2D local symmetries. To estimate how often a 2D symmetry occurs, let C be such a cross-section. The angle 2α that C subtends on its symmetry sphere measures how quickly the cross-section radius is changing. The rays from the camera center are tangent to C's symmetry sphere along another circle Y. If the symmetry sphere subtends an angle of 2β on the viewing sphere, then the radius of Y is $90 - \beta$ degrees.

Condition (1) from section 7.1 will be satisfied if C and Y intersect in at least two points, i.e. if C's center (on the symmetry sphere, not in 3D) lies in the band of radius α around Y. The fraction of the sphere's area covered by this band is

$$\frac{1}{2}\min(1, \sin(\beta + \alpha)) - \frac{1}{2}\max(-1, \sin(\beta - \alpha))$$

So if α is large (cross-section radius changing slowly) and β is small (object not very close to the camera), a 2D symmetry will appear much of the time.

References

- Berger, Marcel (1987) *Geometry II*, Springer-Verlag, Berlin, trans. Michael Cole and Silvio Levy.
- Beyer, William H., ed. (1984) *CRC Standard Mathematical Tables*, 27th edition, CRC Press, Boca Raton, FL.
- Blum, Harry and Roger N. Nagel (1978) "Shape Description using Weighted Symmetric Axis Features," *Pattern Recognition* 10, pp. 167–180.
- Brady, J. Michael (1983) "Criteria for Representations of Shape," in Jacob Beck, Barbara Hope, and Azriel Rosenfeld, eds., *Human and Machine Vision,* Academic Press, New York, pp. 39–84.
- Brady, J. Michael and Haruo Asada (1984) "Smoothed Local Symmetries and Their Implementation," *Intern. J. of Robotics Res.* 3/3, 36–61.
- J. W. Bruce, Peter J. Giblin and C. G. Gibson, "Symmetry Sets," Proc. Roy. Soc. Edinburgh 101A (1985) 163–186.
- Cipolla, Roberto and Andrew Blake (1992) "Surface Orientation and Time to Contact from Image Divergence and Deformation," *European Conf. on Comp. Vision* 1992, 187–202.
- Connell, Jonathan H. and J. Michael Brady (1987) "Generating and Generalizing Models of Visual Objects," *Artificial Intelligence* 31/2, pp. 159–183.
- Dron, Lisa (1993) "Dynamic Camera Self-Calibration from Controlled Motion Sequences," *Proc. IEEE Conf. Comp. Vision and Pattern Recogn.* 1993, pp. 501–506.
- Du, Fenglei and Michael Brady (1993) "Self-Calibration of the Intrinsic Parameters of Cameras for Active Vision Systems," *Proc. IEEE Conf. Comp. Vision and Pattern Recogn.* 1993, pp. 477–482.
- Faugeras, Olivier D. (1992) "Camera Self-Calibration: Theory and Experiment," *European Conf. Comp. Vision* 1992, pp. 321–334.
- Fiala, John C., Ronald Lumia, Karen J. Roberts, and Albert J. Wavering (1994) "TRICLOPS: A Tool for Studying Active Vision," *Intern. J. of Comp. Vision* 12/2-3, pp. 231–250.
- Fleck, Margaret M. (1986) "Local Rotational Symmetries," *Proc. IEEE Conf. Comp. Vision and Pattern Recogn.* 1986, pp. 332–337.
- Fleck, Margaret M. (1994) "Shape and the Wide-Angle Image," technical report 94-04, Computer Science, University of Iowa.
- Forsyth, David A. (1990) "A Novel Algorithm for Colour Constancy," *Intern. J. of Comp. Vision* 5/1, pp. 5–36.
- Forsyth, David, Joseph Mundy, Andrew Zisserman, Chris Coelho, Aaron Heller, and Charles Rothwell (1991) "Invariant Descriptors for 3-D Object Recognition and Pose," *IEEE Tr. Pattern Anal. and Machine Int.* 13/10, pp. 971–991.

- Forsyth, David, Joseph Mundy, Andrew Zisserman, and Charles Rothwell (1992) "Recognizing Rotationally Symmetric Surfaces from the Outlines," *European Conf. on Comp. Vision* 1992, 639–647.

- Forsyth, David, Joseph Mundy, Andrew Zisserman, and Charles Rothwell (1994) "Using Global Consistency to Recognize Euclidean Objects with an Uncalibrated Camera," *Proc. IEEE Conf. Comp. Vision and Pattern Recogn.* 1994, pp. 502–507.

- Giblin, Peter J. and Donal B. O'Shea, "The Bitangent Sphere Problem," American Mathematical Monthly 97/1 (1990) 5–23.

- Giblin, Peter J. and S. A. Brassett, "Local Symmetry of Plane Curves," American Mathematical Monthly 92/10 (1985) 689–707.

- Hahn, Liang-shin, *Complex Numbers and Geometry*, Math. Assoc. of Amer., Washington DC, 1994.

- Hartley, Richard I. and Rajiv Gupta (1994) "Linear Pushbroom Cameras," *European Conf. Comp. Vision* 1994, pp.555–566.

- Hartley, Richard I. (1992) "Estimation of Relative Camera Positions for Uncalibrated Cameras," *European Conf. Comp. Vision* 1992, pp. 579–587.

- Hartley, Richard I. (1994) "Self-Calibration from Multiple Views with a Rotating Camera," *European Conf. Comp. Vision* 1994, vol. I., pp. 471–478.

- Hilbert, D. and S. Cohn-Vossen, *Geometry and the Imagination,* Chelsea Publishing Co., New York, second edition, 1990, translation by P. Nemenyi of *Anschauliche Geometrie*, Springer-Verlag, Berlin, 1932.

- Hedgecoe, John, *The Photographer's Handbook*, Alfred Knopf, New York, 1993

- Hill, Francis S., Jr., *Computer Graphics*, MacMillan, New York, 1990.

- Horn, Berthold K.P., *Robot Vision,* MIT Press, Cambridge MA, 1986.

- Huttenlocher, Daniel P. (1988) "Three-Dimensional Recognition of Solid Objects from a Two-Dimensional Image," Ph.D. thesis, Elec. Eng. and Comp. Sci., MIT.

- Iowa Department of Transportation, *1990/91 Iowa Driver's Manual*, 1990.

- Kato, Kouichi, Tadashi Nakanishi, Akio Shio, and Kenichiro Ishii (1994) "Structure from Image Sequences Caputured Through a Monocular Extra-Wide Angle Lens," *Proc. IEEE Conf. Comp. Vision and Pattern Recogn.* 1994, pp. 919–924.

- Kingslake, Rudolf, *A History of the Photographic Lens,* Academic Press, San Diego, 1989.

- Kingslake, Rudolf, *Optics in Photography,* SPIE, Bellingham, Washington, 1992.

- Kriegman, David J., B. Vijayakumar, and Jean Ponce (1993) "Reconstruction of HOT Curves from Image Sequences," *Proc. IEEE Conf. Comp. Vision and Pattern Recogn.* 1993, pp. 20–26.

- Lamdan, Yehezkel, Jacob T. Schwartz, and Haim J. Wolfson (1988) "Object Recognition by Affine Invariant Matching," *Proc. IEEE Conf. Comp. Vision and Pattern Recogn.* 1988, pp. 335–344.

- Michael Leyton, "A Process-Grammar for Shape," Artificial Intelligence 34/2 (1988) 213–247.

- Jane Liu, Joe Mundy, David Forsyth, Andrew Zisserman, and Charlie Rothwell (1993) "Efficient Recognition of Rotationally Symmetric Surfaces and Straight Homogeneous Generalized Cylinders," *Proc. IEEE Conf. Comp. Vision and Pattern Recogn.* 1993, pp. 123–128.

- Malik, Jitendra and Ruth Rosenholtz (1994) "Recovering Surface Curvature and Orientation from Texture Distortion: A Least Squares Algorithm and Sensitivity Analysis," *European Conf. Comp. Vision* 1994, vol I, pp. 353–364.

- Mokhtarian, Farzin and Alan Mackworth, "Scale-Based Description and Recognition of Planar Curves and Two-Dimensional Shapes," *IEEE Tr. Pattern Anal. and Machine Int.* 8/1 (1986) pp. 34–43.

- Monmonier, Mark (1991) *How to Lie With Maps,* U. Chicago Press, Chicago.

- Mukherjee, Dipti Prasad, Andrew Zisserman, and Michael Brady (1993) "Shape from Symmetry–Detecting and Exploiting Symmetry in Affine Images," *Proc. Roy. Soc.*.

- Oh, Sung Jun and Ernest L. Hall (1989) "Calibration of an Omnidirectional Vision Navigation System Using an Industrial Robot," *Optical Engineering* 28/9, pp. 955–962.

- Otto, G. P. and T. K. W. Chau (1989) " 'Region-Growing' algorithm for matching of terrain images," *Image and Vision Comp.* 7/2, pp. 83–94.

- Peterson, B. "Moose," *Nikon System Handbook*, Images Press, New York, 1991.

- Pillow, Nic, Sven Utcke, and Andrew Zisserman (1994) "Viewpoint-Invariant Representation of Generalized Cylinders using the Symmetry Set," *Proc. British Machine Vision Conference*, 1994, to appear *Image and Vision Computing.*

- Ponce, Jean, David Chelberg, and Wallace B. Mann (1989) "Invariant Properties of Straight Homogeneous Generalized Cylinders and Their Contours," *IEEE Tr. Pattern Anal. and Machine Int.* 11/9, pp. 951–966.

- Ray, Sidney F. (1994) *Applied Photographic Optics,* second edition, Focal Press, Oxford.

- Rothwell, Charles, Andrew Zisserman, David Forsyth, and Joseph Mundy (1992) "Canonical Frames for Planar Object Recognition," *European Conf. on Comp. Vision* 1992, 757–772.

- Smith, Warren J. (1992) *Modern Lens Design: A Resource Manual,* McGraw-Hill, New York.

- Stevenson, Daniel E. and Margaret Fleck (1994) "Robot Aerobics: Four Easy Steps to a More Flexible Calibration," technical report 94-09, Computer Science, University of Iowa.

- Tsai, Roger Y., "An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision," CVPR 1986, pp. 364–374.

- Ulupinar, Fatih and Ramakant Nevatia (1990) "Shape from Contour: Straight Homogeneous Generalized Cones," *Proc. Intern. Conf. Comp. Vision* 1990 pp. 582–586.

- Weng, Juyang, Paul Cohen, and Marc Herniou (1992) "Camera Calibration with Distortion Models and Accuracy Evaluation," *IEEE Tr. Pattern Anal. and Machine Int.* 14/10, pp. 965–980.

- Wixson, Lambert E. and Dana H. Ballard (1994) "Using Intermediate Objects to Improve the Efficiency of Visual Search," *Intern. J. of Comp. Vision* 12/2–3, pp. 137–172.

- Zerroug, Mourad and Ramakant Nevatia (1993) "Quasi-Invariant Properties and 3-D Shape Recovery of Non-Straight, Non-Constant Generalized Cylinders," *Proc. IEEE Conf. Comp. Vision and Pattern Recogn.* 1993, pp. 96–103.