
James E. Cutting
Cornell University

Goldstein (1987) studied the perception of pictures seen from the front and the side. Several distinctions arose from his results and analysis, but only one is central to the reanalysis presented here: The perceived orientation of objects within a picture with respect to the external world is a function of viewer position in front of the picture. For example, the eyes of a portrait subject appear to follow an observer who moves around a gallery. Viewed from many positions, such objects can be said to rotate, following a mobile viewer. Goldstein called this the differential rotation effect because those objects that point directly out of the picture (at 90°) rotate most; those pointing at other angles rotate in decreasing amounts. Goldstein offered no theoretical model and little in the way of explanation for this effect. This Observation offers a model based on the affine geometry and the analyses of La Gournerie (1859). This analysis transforms pictorial space (the space behind a photograph or representational picture) by shears, compressions, and dilations according to the viewpoint of the observer in relation to the composition point of the picture. These effects account for Goldstein's differential rotation effect quite well.

Representational pictures, particularly photographs, have a dual character. On the one hand, they portray objects in a particular environment; on the other, they, as pictures, are dual character. On the one hand, they portray objects in a particular environment; on the other, they, as pictures, are dual character. Consider an interesting fact that stems from this dual, nested character: Certain perceived relations about objects in a picture, which are aspects of the first characteristic, are a function of the relation of the viewer to the picture surface, which is an aspect of the second.

The most memorable examples of these relations concern the eyes and arms of portrait subjects. If the portrayed individual appears to look at the viewer when the latter is standing directly in front of the picture, the portrait appears to follow the viewer as he or she moves about the gallery hall (see, for example, Anstis, Mayhew, & Morley, 1969; Brewster, 1842; Goldstein, 1979; Pirenne, 1970; Wallach, 1976). Moreover, if the portrayed subject points at the viewer standing in front of the picture, that subject points at all who look at the portrait, regardless of their viewpoint. The most famous example is Alfred Leete's 1914 British recruiting poster of the Secretary of State for War Kitchener ("Your country needs you"), copied many times, including the 1917 Army recruiting poster of Uncle Sam ("I want YOU for the U.S. Army"; see Thompson & Davenport, 1980). Kitchener's and Uncle Sam's arm and forefinger follow a mobile viewer. This phenomenon is striking and well worth sustained interest.

Recently Goldstein (1987) explored this phenomenon empirically. He presented slanted pictures to observers and had them make judgments about spatial relations in the picture with respect to an external referent. In two experiments pursuing the portraiture example above, he had observers assess where the portrait was looking. In two others the pictorial relations concerned dowel rods, and he asked observers to imagine a plane that passed through these rods and then through the picture surface into the surrounding world. Goldstein then asked them to assess the angle at which this plane intersected the picture surface. Results of all four studies showed systematic changes with the viewing angle of the observer to the picture plane. Goldstein called it the differential rotation effect because those objects pointing directly out at the observer appear to move with the viewer more than do those pointing obliquely in other directions.

Goldstein (1987) offered no theoretical model for differential rotation; instead he offered only the idea and the supporting evidence that this type of pictorial space was different from others and that confusions concerning picture perception resulted from mixture of these properties. I will return to this idea in my conclusion, but here let me say that the notion seems generally apt and well motivated. This Observation offers a model for the differential rotation data, based on analyses from affine geometry, first sketched by La Gournerie (1859). Various treatments and discussions of pictorial distortions of slanted pictures can be found in Farber and Rosinski (1976), Kubovy (1986), Lumsden (1980), Pirenne (1970), Rosinski, Mulholland, Degelman, and Farber (1980), and Sedgwick (1986). None of these works, however, had the opportunity to explore data as rich and controlled as Goldstein's (1987).
Reconstructing Pictorial Space

To proceed, we must reconstruct the space behind slanted and unslanted pictures. First, I will consider a picture as a planar cross section of the optic array of a possible (as in a perspective painting or drawing) or actual (as in a photograph) environment projected to a particular observation point, which I call the composition point (Cutting, 1986b, 1987), indicated as point c in left panel of Figure 1. The original spatial relations are shown schematically from above in that panel. Within that panel there are two points, a and b, forming a slanted line, and to the right of the line is a partial checkerboard of squares in the horizontal (xz) plane.

La Gournerie (1859) reasoned that viewing a perspective picture from a point other than its composition point (assuming the composition point to be in the center of the picture) should create systematic distortions in pictorial space. In particular, and as shown in the middle panel of Figure 1, when a viewer moves to one side, the imaginary planes in depth parallel to and behind the picture plane (xy planes stacked in depth) slide sideways past one another. Moreover, imaginary vertical planes initially perpendicular to the picture plane (yz planes) follow the angle of the observer to a particular point on the picture’s surface. This second point has threefold import: It is usually the center of the picture itself (and is the center of all standard photographs), it is the point on the picture typically nearest the composition point, and it is often the vanishing point in one-point, or Albertian, perspective (e.g., Carlbom & Paciorek, 1978; Cutting, 1986a). I call it the principal point, or p, for the construction of perspective, and following convention (e.g., Kubovy, 1986), the ray connecting the composition point to the principal point on the image surface and extending into the depth of pictorial space is called the principal ray.

The change in layout of pictorial space from a grid of squares to a grid of parallelograms is called shearing and is an affine transformation. For example, all objects and parts of objects in a particular square of the checkerboard can be found in the corresponding parallelogram after the transformation. Operationally, this means that the ratio of distances D/d remains constant for all new viewpoints, where D is the distance from the composition point (or new viewpoint) to a particular location in the virtual space behind the picture and where d is that from the composition point (or new viewpoint) to the picture surface. Because all points are anchored on the picture surface, one can now begin to see why the eyes of a portrait subject might follow a viewer around a gallery hall: The principal ray for all affine reconstructions is anchored to the observer’s eye position (with binocular differences ignored).

The other affine transformations are the compression or dilation of one dimension, here always the z axis, as compared with the others, x and y. This compression (or dilation) occurs when the viewer moves closer to (or farther away from) the picture or when camera lenses of greater (or lesser) focal length are used. Compression due to more proximal viewing by the observer, or to a longer focal-length lens, is shown in the left panel of Figure 1. Before trying to account for Goldstein’s differential rotation effect, it will serve us to reconstruct the proper depth relations in the photographs that he used.

These belong under the rubric of spatial relations rather than perceived orientations, which Goldstein argues are different. Nonetheless, I will apply this affine analysis throughout.

Lens Length and Compressed Pictorial Depth

The lenses used on everyday cameras typically compress the depth in pictures. Cameras are fitted with such lenses because most individuals use cameras to photograph people. And a person, when photographed from relatively close range with a standard lens, will appear to have a bulbous nose. A long lens counteracts this effect because is presents a more nearly parallel projection (see, for example, Hochberg, 1986; Kraft & Green, in press). Thus, to consider the reconstruction of pictorial depth, we must consider such lenses.

A long lens is one that has a focal length greater than the film size. Although effective focal length is determined by the distance between the nodal point in the lens to the photographic plate, it also determines the size of the film on which an image can be projected. Interestingly, these two numbers are the same. Thus, a 35-mm camera often has a 50-mm lens because the diagonal of the film image is about 50 mm. The diagonal is important here because the lens creates a circular image and, because a photograph has a rectangular frame, the diagonal of the film must fit within the image circle. Such a lens is regarded as a “normal” lens (see, for example, Cox, 1971; or Swedlund, 1974); it neither compresses nor expands pictorial depth.

A 35-mm camera often has a longer lens—85-mm and 135-mm lenses are common. These lenses expand the image sizes to 85 and 135 mm, but, of course, the camera cannot be fitted for such film. Thus, only the central 50-mm region of the potential 85- or 135-mm image is exposed. This region, however, is an expanded version of what was seen with the shorter lens. This expansion of the image, when viewed from the same distance, compresses depth. The manner in which depth is compressed is the same as outlined in the right panel of Figure 1 for an observer closer than normal to a photograph. The amount of depth compression is a simple ratio of lens length to film size. Thus, the 85- and 135-mm lenses compress depth by factors of 1.7 (85/50) and 2.7 (135/50), respectively. These facts are relevant to Goldstein (1987), particularly to a result in Experiment 1.

Experiment 1: Mislocation of Rods in Pictorial Space

Depth compression is pertinent to this Observation because Goldstein (1987) used as lightly magnified lens. His 3.5-in. × 4.25-in. (90 × 105 mm) format Polaroid (with an effective

---

1 Here I will discuss only the affine distortion of virtual space behind a picture, not the perspective distortions that accompany picture viewing from the side (Cutting, 1986a, 1986b). I return to perspective distortions in my Conclusions.

2 Hagen and her colleagues (Hagen & Elliott, 1976; Hagen, Elliott, & Jones, 1978; but see Kubovy, 1986) have shown that viewers prefer to look at pictures shot with a longer than normal lens, a phenomenon they call the “zoom effect.” Telephoto lenses compress depth and yield a photograph that is more nearly a parallel projection.

3 Goldstein kindly provided me with information and photocopies
OBSERVATIONS

Figure 1. Affine distortions of pictorial space. (The left panel shows the original composition point \(c\), the principal point \(p\), picture plane [the thick horizontal lines], and several arrangements in pictorial depth. To the left are two points, \(a\) and \(b\), and a line of orientation drawn between them, and to the right is a partial checkerboard. The line length \(D\) indicates the distance from the composition point to one corner of the checkerboard grid, and the length \(d\) indicates that from the composition point to the point on the image representation that corner. Radios \(D/d\) are constant in reconstruction of all points in pictorial space from all viewpoints. The middle panel shows the affine shear in an observer's viewpoint moved to the side. Both the slant of the line and the shape of the checkerboard change. The right panel shows the affine compression due to an observer moved closer to the image [or to an image blown up with a long lens], and again the slants and shapes change.)

diagonal of roughly 140 mm) was fitted with a 215 mm lens. The magnification factor (and depth compression factor) is 215/140 or 1.53. Thus, although Goldstein (1987) took care to recreate the angular subtense of his dowel rods and faces, the pictorial depth was nonetheless compressed because of the lens length.

This compression has several affects. One is that most of Goldstein's (1987) pictorial angles are slightly different than he indicates. Consider his arrangement of dowels for his Experiments 1 and 2, shown in my Figure 2. In the upper panel are given the results for his observers' reconstruction of rods in space. In a real situation, the orientation of Rods B to A was 100° as measured on the upper panel of Goldstein's Figure 2. But with depth compressed, the distance between Rods A and B is shortened, and the reconstructed angle becomes 104°, as shown in the lower panel of my Figure 2. This effect goes some distance in accounting for the misplacements of dowel positions in his Experiment 1.

To go further in accounting for the misplacements, an additional factor must be considered—the perceived line of sight to particular points on the image. That is, the line of sight between the nodal point of the camera lens (and hence the viewer's eye) and Rod B will be taken by observers to be 90°. In fact, it is 81.5°, because the principal point on the image is considerably to the right, between Rods B and C. These two angles are shown in the bottom panel of my Figure 2. If one adds the 8.5° difference error to the 104° projected angle, the resulting orientation of the line formed by Rods BA is 112.5°. Measuring the top panel of Goldstein's (1987) Figure 2 yields an angle of 115°, a fine fit. The value of 112.5° for the angle of projected line BA is used in the analyses below; similar analyses were done to derive the angle for the projected line CA.

Reconstructing Projected Pictorial Slants

Having adjusted the proper depth relations for compression, we can now consider the differential rotation effect. Affine geometry, stemming from insights of La Gournerie (1859), should indicate how the shears and compressions (or dilations) occur and how differential rotation of objects with respect to viewer position in front of the picture can be predicted. Derivations can proceed in many ways with many equivalent results, but the simplest formulation I have seen, and one that is suitable for the conditions of Goldstein (1987), is given by Rosinski et al. (1980).

Given that the viewing is done from the station point, two angles are of central interest: \(\alpha\), the slant of the picture's surface with respect to the viewpoint (where 90° is the correct viewing angle) and \(\theta\), the slant of the pertinent pictured line (where 0° is parallel to the picture plane). When both \(\alpha\) and \(\theta\) can vary, the projected slant of the pictured line can be determined by the following equation:

\[
\theta' = \arcsin\left[\sin \theta \sin \alpha / (\sqrt{1 + \cos \alpha \sin^2 \theta})\right]
\]

so long as the observer remains at the same distance from the principal point. And, of course, the projected slant, \(\theta'\), is the

4 This analysis is related to and inspired by the discussion of slant misestimates by Perrone (1980). See also Hagen and Elliott (1976) and Hagen et al. (1978) for their descriptions of the zoom effect.

of the original photographic stimuli on which my analysis is based. I could not extend this analysis to his earlier work (Goldstein, 1979) because certain critical information about photographic distances is unavailable.
Experiment 2: Perceived Orientations of Rods

Using the scheme and formula given above, I show the predictions from an affine compression and shear analysis of pictorial space based on the original insights of La Gournerie (1859) for Goldstein's (1987) Experiment 2 in my Figure 3. The fit is reasonably good, with a high correlation, \( r = .97, t(19) = 17.6, p < .001 \). Of course, correlations are a good measure of trend but not necessarily a good measure of fit; the average discrepancy is about 11°, with a standard deviation of about 15°. Two salient aspects of the predictions do not fit the data: first, the difference between the 20° and 45° data for the orientation of the line formed by Rods AC; and second, the lack of rotation in the predictions for the line formed by Rods BC. Goldstein's (1987) results here are in tune with other data and with my own experience with slanted pictures, and I have no principled explanation for the mismatch between prediction and data. Other factors may be involved, particularly that of a visible frame (see Goldstein, 1987, Experiment 5). At any rate, large deviations are generally evident only at extreme slants and should, I think, be regarded as a subsidiary effect.

Experiment 3: Perceived Orientation of Eye Glances

The same kind of analysis and prediction can proceed with Goldstein's (1987, Experiment 3) investigation of eye glances. Here the perceived direction of gaze replaces perceived slant, \( \theta' \), but all else is the same. Given that the same camera was used as in his Experiment 2 to make stimuli, his viewing angles of 45°, 60°, 75°, 90°, 105°, 120°, and 135° are compressed to 33°, 48.5°, 68°, 90°, 112°, 131.5°, and 147°, respectively. Table 1 shows the various predictions, retaining Goldstein's angles for purposes of nomenclature. There are a few widely discrepant points generally at extreme slants, but as in the previous analysis, the fit of predictions is adequate (mean...
Table 1
Predictions of Perceived Orientation Compared With Goldstein’s (1987, Experiment 4) Data

<table>
<thead>
<tr>
<th>Gaze direction and orientation</th>
<th>20</th>
<th>45</th>
<th>70</th>
<th>90</th>
<th>110</th>
<th>135</th>
<th>160</th>
<th>M diff.</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted 45°</td>
<td>8</td>
<td>17</td>
<td>27</td>
<td>33</td>
<td>38</td>
<td>40</td>
<td>30</td>
<td>23</td>
<td>8</td>
</tr>
<tr>
<td>Observed 45°</td>
<td>-16</td>
<td>-12</td>
<td>-2</td>
<td>2</td>
<td>13</td>
<td>22</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted 60°</td>
<td>11</td>
<td>24</td>
<td>37</td>
<td>49</td>
<td>60</td>
<td>76</td>
<td>80</td>
<td>31</td>
<td>8</td>
</tr>
<tr>
<td>Observed 60°</td>
<td>-5</td>
<td>-4</td>
<td>7</td>
<td>21</td>
<td>26</td>
<td>37</td>
<td>39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted 75°</td>
<td>14</td>
<td>32</td>
<td>51</td>
<td>68</td>
<td>86</td>
<td>113</td>
<td>147</td>
<td>26</td>
<td>19</td>
</tr>
<tr>
<td>Observed 75°</td>
<td>13</td>
<td>19</td>
<td>33</td>
<td>49</td>
<td>55</td>
<td>68</td>
<td>88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted 90°</td>
<td>20</td>
<td>45</td>
<td>70</td>
<td>90</td>
<td>110</td>
<td>135</td>
<td>160</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Observed 90°</td>
<td>22</td>
<td>47</td>
<td>72</td>
<td>91</td>
<td>110</td>
<td>135</td>
<td>159</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted 105°</td>
<td>33</td>
<td>63</td>
<td>94</td>
<td>112</td>
<td>129</td>
<td>148</td>
<td>166</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>Observed 120°</td>
<td>95</td>
<td>105</td>
<td>114</td>
<td>123</td>
<td>137</td>
<td>160</td>
<td>171</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted 135°</td>
<td>100</td>
<td>104</td>
<td>120</td>
<td>131</td>
<td>143</td>
<td>156</td>
<td>169</td>
<td>23</td>
<td>9</td>
</tr>
<tr>
<td>Observed 135°</td>
<td>132</td>
<td>136</td>
<td>142</td>
<td>152</td>
<td>171</td>
<td>173</td>
<td>180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted 150°</td>
<td>150</td>
<td>140</td>
<td>142</td>
<td>147</td>
<td>153</td>
<td>163</td>
<td>172</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td>Observed 150°</td>
<td>151</td>
<td>151</td>
<td>162</td>
<td>169</td>
<td>174</td>
<td>187</td>
<td>190</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M diff.</td>
<td>20</td>
<td>22</td>
<td>20</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>19</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>SD</td>
<td>22</td>
<td>15</td>
<td>9</td>
<td>10</td>
<td>13</td>
<td>16</td>
<td>22</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. M diff. = mean difference.

difference = 20°, \( \sigma = 15° \) and the correlation high, \( r = .94, t(47) = 18.9, p < .001 \).

Goldstein (1987) took the difference between extreme values of perceived orientation as a measure of differential rotation. Figure 4 shows the values he obtained and those predicted by an analysis of affine distortions. Although the general shape of the function is different, the fit is fairly good (mean discrepancy = 29°, \( \sigma = 21° \) and correlation reasonable, \( r = .80, t(5) = 4.97, p < .05 \).

The Sphere Demonstration

As a demonstration of the changes in perceived orientation, Goldstein also presented to the subjects of his Experiment 3 a picture of a sphere with a mark on it. In one condition the mark was at 90°, closest to the observer, and in another it was at 150°, toward the far left edge. He had observers look at these marks from a steep angle, 20°, and had them make judgment of orientation of the surface normal (the line perpendicular to the surface at a given point). His observers determined that the surface normals at these marks were at 25° and 153°, respectively. If we assume no compression (that is, that the lens was appropriate for the film), the prediction from an analysis of affine distortions is that they should have appeared at 20° and 157°. This too is a fine fit.

Goldstein (1987) conducted his Experiment 3 and his sphere demonstration to contrast with his Experiment 2. In particular, he thought that perspective information might be used in judging orientation of rods, but such information could not be used with judging glances or with judging the orientation of surface normals to a sphere. Yet the reconstruction of depth through affine compressions and shears here appears to fit the data in the two situations about equally well. That is, true lines of oriented rods that can be drawn in a picture, and imagined lines of sight from a portrait can be dealt with, and accounted for, in a similar fashion. Thus, insofar as the reconstructions carried out here are based on a geometric analysis of pictorial space and insofar as perspective is a related geometric analysis, I claim that Goldstein (1987) did not rule out the contribution of perspective information to judgments of perceived orientation.

Conclusions

Affine geometry can be used to reconstruct the virtual space behind a slanted photograph. Moreover, its application goes quite some distance in explaining Goldstein’s (1987) perceived orientation data, as it did in the analyses of Farber and Rosinski (1978). Contrary to what Goldstein (1987) might have predicted, this affine reconstruction adequately predicts...
perceived orientations in photographs with and without straight-line content—that is, with and without “perspective information.” More particularly, it does about equally well for the orientation of lines through rods, which can be easily imagined, and the orientation of an eye glance, which is less easily discerned (Brewster, 1842; Gibson & Pick, 1963; Noll, 1976).

On the basis of the differential rotation effect and other data, Goldstein (1987) argued for three different kinds of pictorial/spatial attributes—perceived orientation, projection, and spatial layout. Let me consider briefly the latter two, in turn.

Projection is not discussed in detail by Goldstein (1987). It is the effect of relatively squashed width in the optics of pictures seen from the side. The squashed width is a corollary of the affine reconstructions presented in this Observation, but there is another important effect of projection that I have discussed elsewhere (Cutting, 1986b, 1987): There is a relative lengthening and shortening of the near and far edges of the slanted picture. These create another set of distortions overlaid on the affine distortions of pictorial space. Goldstein (1987) notes that these would be easily discriminated if a slanted picture were compared with an identical one that is unslanted. My point here is that what Goldstein calls projection distortions occurs both in optics of pictorial space (that is, behind the picture surface) and in the optics of the picture surface (which is trapezoidal to the eye rather than rectangular). I claim that we rarely notice the former, but we can easily notice the latter.

Spatial layout concerns judgments of spatial relations of objects portrayed in a picture, with reference to the original layout. As Goldstein (1987) noted, observers can physically reconstruct the geometric layout of a scene fairly accurately, and these reconstructions are remarkably unchanged with viewing angle. (This is true despite the fact that my affine compression analysis of his Experiment 1 data fits the 90° condition so well.) I claim that spatial layout judgments are those which an observer makes by taking into account the picture slant and by attempting to reconvert the affine space into Euclidean space (see Kubovy, 1986; and Pirenne, 1970).

Elsewhere, I have called this process Euclidean rectification (Cutting, 1986b, 1987), and I have argued in the context of cinema that the process is unnecessary for the perception of certain attributes of objects, particularly rigidity. Here I claim that Euclidean rectification explains Goldstein’s spatial layout effect (Goldstein, 1987, Experiment 1). Thus, contrary to my conclusion concerning cinema (Cutting, 1987), I think that Euclidean rectification is possible and even likely with still pictures. The rationale for this difference is that, unlike cinema, pictures are typically viewed from relatively close up, with a smaller image and with strong cues about slant from frame shape, specularities (bright spots due to light reflections), graded binocular disparities, and even accommodation.

Finally, Goldstein (1987) argued that to confuse the three properties that he discusses—perceived orientation, spatial layout, and projection—is to confuse the various aspects of picture perception. He claimed that Pirenne (1970) and others may have done just this, and here I think Goldstein is on solid ground. I would add to this a set of reflections. First, the affine transformed orientation of objects in a picture is information in the optics of pictures as presented to the eye. Thus,

---

5 Only frame shape and possible specularities were available to Goldstein’s (1987) viewers. In his Experiment 5, he removed the possibility of the latter by presenting back-illuminated line drawings. These increased the rotation effect, but he did not measure observers’ abilities to reconstruct spatial layout.
perceived orientation is not a derived percept; it stems directly from visual information. Second, projection information is also in the optics of slanted pictures but may be noticed only under conditions of discrimination. And third, spatial layout is a derived percept from information about surface slant or from the noticed optical changes in pictures seen at extreme slants.

References


Received June 5, 1987
Revision received July 13, 1987
Accepted July 14, 1987