

Pictorial relief for equiluminant images

Andrea J. van Doorn^{*a}, Huib de Ridder^a, Jan J. Koenderink^b

^aDelft University of Technology, Landbergstraat 15, 2628 CE Delft, The Netherlands

^bUniversiteit Utrecht, PO Box 80 000, 3508 TA Utrecht, The Netherlands

ABSTRACT

Pictorial relief depends strongly on “cues” in the image. For isoluminant renderings some cues are missing, namely all information that is related to luminance contrast (*e.g.*, shading, atmospheric perspective). It has been suggested that spatial discrimination and especially pictorial space suffer badly in isoluminant conditions. We have investigated the issue through quantitative measurement of pictorial depth-structure under normal and isoluminant conditions. As stimuli we used monochrome halftone photographs, either as such, or “transposed” to Red/Green or Green/Red hue modulations. We used two distinct methods, one to probe pictorial pose (by way of correspondences settings between pictures of an object in different poses), the other to probe pictorial depth (by way of attitude settings of a gauge figure to a perceptual “fit”). In both experiments the depth reconstructions for Red/Green, Green/Red and monochrome conditions were very similar. Moreover, observers performed equally well in Red/Green, Green/Red and monochrome conditions. Thus, the general conclusion is that observers did not do markedly worse with the isoluminant Red/Green and Green/Red transposed images. Whereas the transposed images certainly looked weird, they were easily interpreted. Much of the structure of pictorial space was apparently preserved. Thus the notion that spatial representations are not sustained under isoluminant conditions should be applied with caution.

Keywords: pictorial space, isoluminant images, pictorial relief, picture perception, pictorial cues

1. INTRODUCTION

When a human observer looks at a picture, the observer is aware of the picture as a flat sheet covered with pigments in a certain spatial order. Moreover, the observer typically “looks into” the picture and perceives a three dimensional “pictorial space”. Whereas the picture is perceived as a physical object in the scene of which the observer’s body is part, the pictorial space is perceived as “virtual” in the sense that it is not made up of physical bodies that might be touched because the observer’s body is not part of that space. Nevertheless the pictorial space contains objects, many of which are apparently opaque bodies, bounded by surfaces, so called “pictorial relief”. The surfaces are seen as being illuminated by the “pictorial light field”, pictorial objects are perceived to occlude each other, and so forth. Pictorial space is due to the creative imagination of the observer, checked by the image structure. In this context the image structure is usually described in terms of “pictorial cues”, although the cues should not be understood as mere patterns of pigmentation on the picture surface, but as structures defined by such patterns that derive their significance (that is their “cue”- nature) due to the perceptual abilities of the observer. It is generally believed that humans are unique in the animal kingdom (with the possible exception of some primates) in experiencing “pictorial scenes”.

The “pictorial cues” are necessarily “monocular cues”, for the very structure of the picture does not allow for pictorial accommodation, vergence, disparity, or parallax cues. Such cues are indeed present, but they relate not to the pictorial scene, but rather to the picture as a physical object. Thus such cues are conflicting cues with respect to pictorial space. Indeed, it has been shown that such cues tend to counteract the production of pictorial space¹.

All cues specify the scene only modulo extensive groups of ambiguities. Even taking all cues into account there remains a large group of ambiguities. This leaves an important role for the “beholder’s share”² in pictorial vision. We have shown that human observers exploit this freedom to perform “mental changes of perspective”³.

A monochrome picture, i.e., a photograph, may be “transposed” into an isoluminant color image. This is a rather artificial transformation, never encountered under natural circumstances. For instance, the grayscale may be mapped on a red-green scale such that green corresponds to white, red to black, and yellow to average gray, all mixtures of red and green being of equal luminance. Many different “transpositions” are possible of course. It has been suggested that under such circumstances pictorial space would effectively vanish⁴, although the issue is not undisputed⁵. This is the topic of the present research.

Why would pictorial space deteriorate under isoluminant conditions? For one thing, for isoluminant renderings some cues are expected to weaken considerably. An example that immediately comes to mind is the “shading cue”. Other cues might be expected to remain potentially effective. Cues of the latter are for instance :

- contour,
- contourform,
- occlusion,
- texture gradient,
- foreshortening,

whereas cues that are expected to diminish in importance are those based on information that is related to luminance contrast such as:

- shading,
- atmospheric perspective,
- figure-ground.

It is common knowledge that spatial discrimination of borders and gradients almost vanishes in the isoluminant condition. It would be expected that (3D) pictorial space will suffer because of this and perhaps might almost vanish in the isoluminant condition. In this study we have tried to find out if this is indeed the case. We investigated the issue through the quantitative measurement of pictorial depth structure under normal and isoluminant conditions.

As stimuli we used monochrome halftone photographs, either used as such, or “transposed” to Red/Green or Green/Red hue modulations. The photographs were of one and the same human torso (a plastic mannequin used for shop window advertising) but exposed with the torso in three different poses while keeping the vantage point of the camera and the position of the lightsource unchanged (Fig. 1). Poses differed by 45 degrees of angle rotation about the vertical. (In the following they are called respectively Pose0, Pose45, Pose90.) Notice that the shading was different for the three poses because the light field rotated relative to the object. We used all three halftone renderings as well as the Red/Green and Green/Red transposed versions of them. To have a convenient data structure for our quantitative measurements we computed triangulations of the renderings. These triangulations were not revealed to the observers.

For our measurements we used two distinct methods, one to probe pictorial “pose” (Experiment 1), the other to probe pictorial “depth” (Experiment 2). In both experiments the viewing conditions were the same. Viewing distance was 50 cm and the center of the monitor screen was in the straight ahead direction. Images subtended 16 x 24 degrees of visual angle. The monitor was linearized (gamma of one). Viewing was monocular and a chin rest was used for head fixation. The room was dark. In Experiment 1 six observers completed the experiment. In Experiment 2 two observers participated.

We determined the equiluminant R/G ratio for each observer by means of the motion stand-still method⁶. This is necessary because there is quite a wide range of the “equiluminant” R/G ratio for different observers. We thus made sure that we measured at each observer’s individual isoluminant setting, thus avoiding possible luminance artifacts.

2. EXPERIMENTS

2.1 Experiment 1: image correspondences

2.1.1 Procedure

In this experiment the observers were confronted with a pair of pictures at a time, presented side by side⁷. On the fiducial (left) picture a dot was superimposed. The observers were required to place a dot on the test (right) picture at a

position in pictorial space corresponding with the position of the fiducial dot (also in pictorial space). An example is shown in Fig. 2.

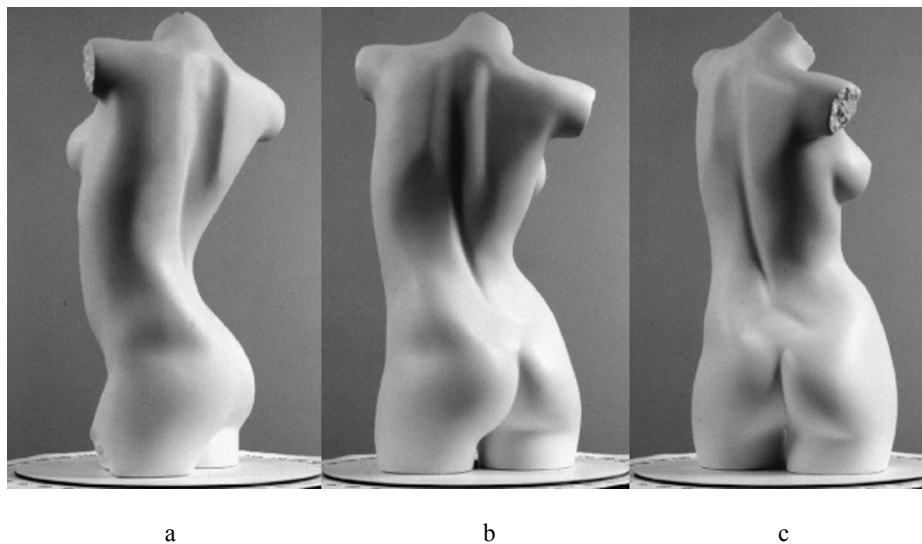


Figure 1: The three (dorsal) views of the torso. The viewing directions differ by 45 degrees of angle. Camera and light source were fixed, whereas the object was rotated between exposures. a: Pose0, b: Pose45, c: Pose90.

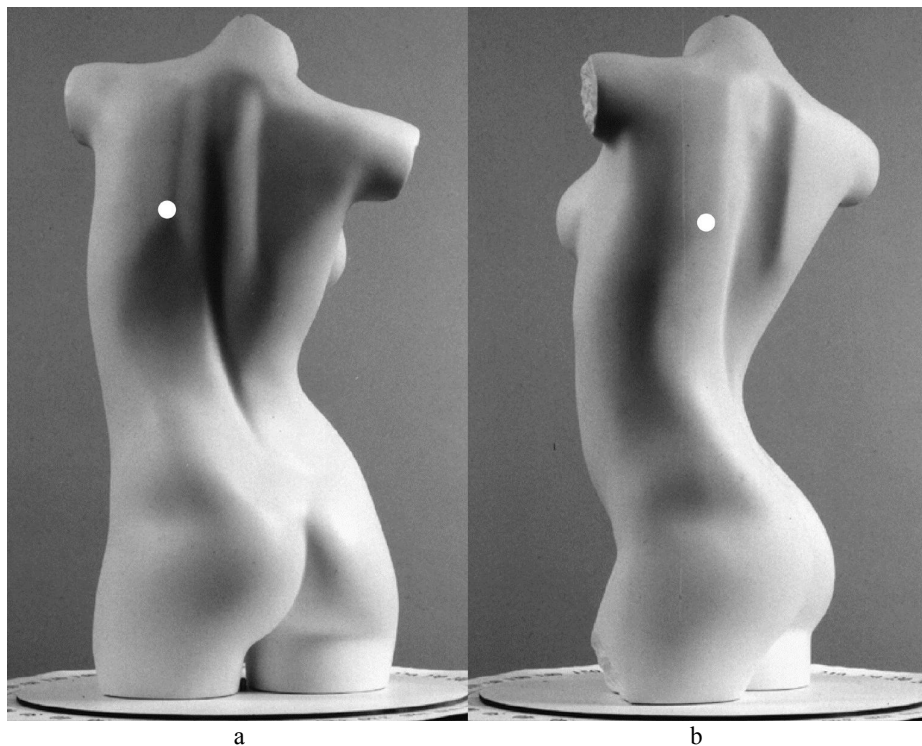


Figure 2: The correspondences task. a: fiducial picture with fixed marker. b: test image with a marker position adjusted by the observer.

Observers needed only a few seconds per setting. We gathered settings for all vertices in the triangulation of the fiducial image, one by one, in random order. The set of stimulus pairs consisted of the following combinations:

Red/Green transposed Pose45 versus Green/Red transposed Pose0	(pair 1)
Red/Green transposed Pose45 versus Green/Red transposed Pose90	(pair 2)
Red/Green transposed Pose45 versus Red/Green transposed Pose0	(pair 3)
Red/Green transposed Pose45 versus Red/Green transposed Pose90	(pair 4)
Monochrome Pose45 versus Green/Red transposed Pose0	(pair 5)
Monochrome Pose45 versus Green/Red transposed Pose90	(pair 6)
Monochrome Pose45 versus Red/Green transposed Pose0	(pair 7)
Monochrome Pose45 versus Red/Green transposed Pose90	(pair 8)
Monochrome Pose45 versus Monochrome Pose0	(pair 9)
Monochrome Pose45 versus Monochrome Pose90	(pair 10).

For each stimulus pair (that means 10 pairs) the observers repeated the correspondences settings three times. This allows us to estimate the spread in their (repeated) settings.

2.1.2 Analysis of the data

From these settings the average deviations in the horizontal and vertical direction were calculated for each observer and for all pairs. The scatter in the deviations appeared to be roughly equal in the horizontal and vertical directions. This is an interesting result in itself, because the object was rotated about the vertical axis, that is to say that all changes due to the relief were in the horizontal direction. Apparently the observers did not use the technique of trying to find corresponding heights but really used “landmarks” on their 3D pictorial reliefs to find the correspondences. From the combination of horizontal deviations in pair 1 and horizontal deviations in pair 2 we could find a surface in pictorial space that explains the results best in the least squares sense for the comparison-condition Red/Green transposed versus Green/Red transposed. The same could be done by combining horizontal deviations:

- in pair 3 and 4 for the comparison-condition Red/Green transposed versus Red/Green transposed,
- in pair 5 and 6 for the comparison-condition Monochrome versus Green/Red transposed,
- in pair 7 and 8 for the comparison-condition Monochrome versus Red/Green transposed,
- in pair 9 and 10 for the comparison-condition Monochrome versus Monochrome.

Fig. 3 shows these calculated depth profiles of one observer (AV) in the five different comparison-conditions. Other observers have comparable results.

For the same observer (AV) we show in Fig. 4 straight scatterplots of depth values extracted from his depth profiles (Fig. 3) for different comparison-conditions. (We found comparable results for all other observers.)

In Fig. 5 we show (again for the same observer AV) the complete set of R-squared values of such scatterplots of depths over all conditions. (Again, all observers have the same quantitative behavior).

From Figs. 4 and 5 it is evident that the results for the various conditions are very similar indeed. We do not deem it sufficiently interesting to pursue possible minor differences. Notice especially that the isoluminant conditions do not lead to results that are markedly different from those obtained for the natural gray scale images.

2.1.3 Conclusions

Apparently the observers were not seriously handicapped by the transposition into isoluminant renderings. Despite the fact that such renderings did indeed look weird to them, they were nevertheless able to perform tasks in pictorial space. Thus the conclusion from this image correspondences experiment (Experiment 1) is (perhaps unexpectedly) that observers performed about equally well in Monochrome versus Monochrome (thus NOT isoluminant), in Monochrome

versus Red/Green transposed, in Monochrome versus Green/Red transposed, in Red/Green transposed versus Green/Red transposed, and in Red/Green transposed versus Red/Green transposed isoluminant conditions.

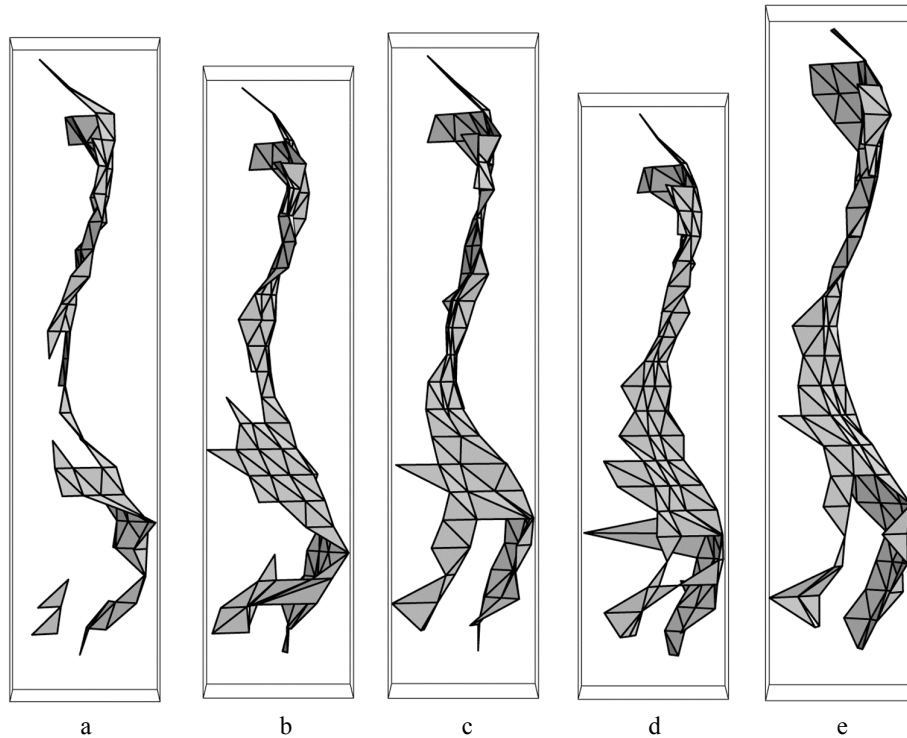


Figure 3: Depth profiles for observer AV. The comparison-conditions are respectively: a: Red/Green transposed versus Green/Red transposed, b: Red/Green transposed versus Red/Green transposed, c: Monochrome versus Green/Red transposed, d: Monochrome versus Red/Green transposed, e: Monochrome versus Monochrome.

2.2 Experiment 2: surface attitudes

2.2.1 Procedure

In this experiment the observer was confronted with only one picture at any time. We used a method of "fit" to measure the spatial attitude of pictorial surface elements at many points (five to six hundred) in the image. The observer adjusted the shape of a gauge figure, superimposed upon the picture, such as to produce a perfect fit. The gauge figure was an ellipse (in general orientation and with variable eccentricity). The "fit" was achieved if the gauge figure looked to the observer as a circle painted on the pictorial surface⁸. Observers needed only a few seconds for such a setting. We gathered settings for all barycentra in the triangulation of the presented picture. The settings were done in random order, one at a time. The same stimuli were used as in the previous experiment. In Fig. 6a,b we show an example of the procedure (this picture was not among the stimuli).

The set of stimuli consisted of all three monochrome renderings (Pose0 Monochrome, Pose45 Monochrome, Pose90 Monochrome), all three Green/Red transposed renderings (Pose0 Green/Red, Pose45 Green/Red, Pose90 Green/Red), and all three Red/Green transposed renderings (Pose0 Red/Green, Pose45 Red/Green, Pose90 Red/Green).

For each stimulus (in total 9 different conditions) the observers completed three independent sessions of gauge figure settings.

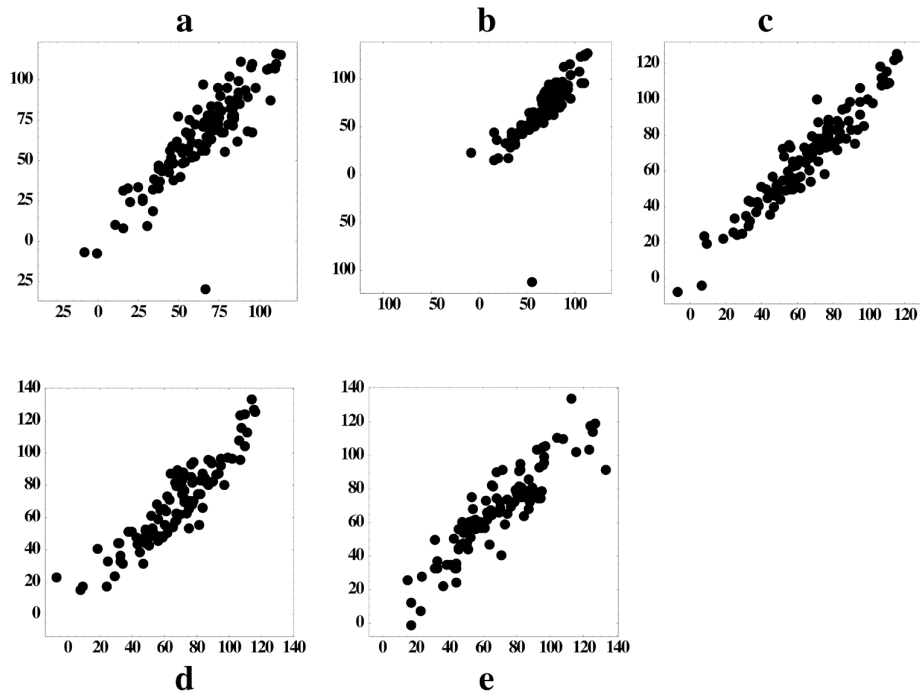


Figure 4: Scatterplots of depth values for observer AV for the following cases:
 a: horizontal axis depth values for comparison-condition Monochrome versus Monochrome, vertical axis depth values for comparison-condition Monochrome versus Red/Green transposed.
 b: horizontal axis depth values for comparison-condition Monochrome versus Monochrome, vertical axis depth values for comparison-condition Red/Green transposed versus Green/Red transposed.
 c: horizontal axis depth values for comparison-condition Monochrome versus Red/Green transposed, vertical axis depth values for comparison-condition Monochrome versus Green/Red transposed.
 d: horizontal axis depth values for comparison-condition Monochrome versus Red/Green transposed, vertical axis depth values for comparison-condition Red/Green transposed versus Green/Red transposed.
 e: horizontal axis depth values for comparison-condition Red/Green transposed versus Green/Red transposed, vertical axis depth values for comparison-condition Red/Green transposed versus Red/Green transposed.

2.2.2 Analysis of the data

As one interprets the elliptical gauge figure as a foreshortened circle the ellipse parameters can be converted into the attitude (slant and tilt values) of the local tangent plane of the pictorial relief (see Fig. 6c). After completion of all sessions we averaged the attitude samples for each stimulus and found the set of depth values on the vertices of its triangulation that best explained the average attitude samples in the least squares sense. From this we constructed the operationally defined "pictorial reliefs" for each stimulus. The observers produced consistent pictorial reliefs which is a confirmation of previous findings⁸. The scatter in repeated data at any location on the reliefs was of the order of 20-30 % of the gradient magnitude.

The results of these depth reconstructions for the nine different stimulus conditions are shown in Figs. 7 and 8 for one of the observers (AD). The results for another observer were comparable. Fig. 7 shows the observer's pictorial reliefs as calculated from the average attitude settings and Fig. 8 shows the isodepth plots of these pictorial reliefs.

We quantitatively compared the reconstructed pictorial reliefs by making scatterplots of the depth values that were computed for the different stimulus conditions. In Fig. 9 the complete set of scatterplots of depths is shown for one observer (again AD).

For the same observer (AD) we have plotted the slopes of these scatterplots in Fig. 10a. Although some of the slopes may be (just) significantly different, there does not seem to be any obvious pattern in the results.

In Fig. 10b we have plotted for the same observer (AD) the R-squared values of the scatterplots. The figure shows that there is not really any difference between the Monochrome and isoluminant Red/Green or Green/Red transpositions.

Another observer showed the same quantitative behavior.

Table of R^2 values over all conditions for AV

AV	Mon/Mon	Mon/RG	Mon/GR	RG/GR	RG/RG
Mon/Mon	*				
Mon/RG	0.743	*			
Mon/GR	0.839	0.911	*		
RG/GR	0.590	0.836	0.871	*	
RG/RG	0.806	0.861	0.857	0.813	*

Figure 5: Table of R-squared values for observer AV.

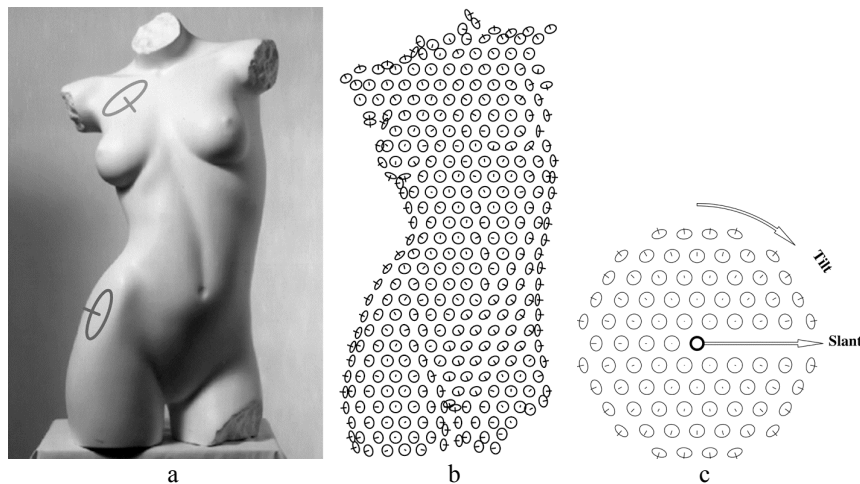


Figure 6: a: two gauge figures, superimposed upon the pictorial object. One fits, the other does not fit. b: result of a measurement (one setting at a time, in random order), a field of local surface attitudes. c: the attitude of the local tangent plane can be defined by two parameters, namely the slant, that is the angle between the normal on the local tangent plane and the viewing direction, and the tilt, that is the direction in which the local tangent plane is slanted. The figure is a polar plot of all possible slants and tilts. Slant has values between 0 and 90 degrees of angle, tilt between 0 and 360 degrees of angle.

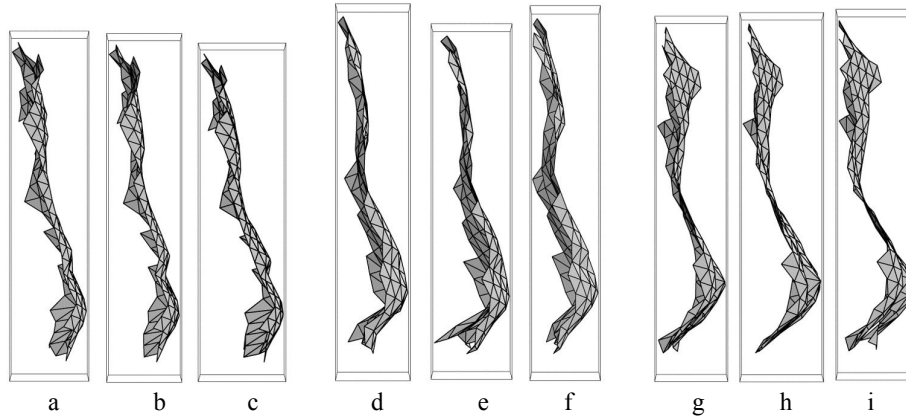


Figure 7: Pictorial reliefs for observer AD. The conditions are respectively:

a: Pose0 Green/Red, b: Pose0, Red/Green, c: Pose0 Monochrome, d: Pose45 Green/Red, e: Pose45 Red/Green, f: Pose45 Monochrome, g: Pose90 Green/Red, h: Pose90 Red/Green, i: Pose90 Monochrome.

2.2.3 Conclusions

The conclusions from this surface attitude experiment (Experiment 2) are:

- The depth reconstructions for the Green/Red, Red/Green and Monochrome conditions appeared to be very similar for each of the three poses.
- Observers performed about equally well in the Green/Red, Red/Green and Monochrome conditions for each of the three poses.

Thus (again, perhaps unexpectedly) there appears to be no empirical evidence for the (not uncommon) expectation that pictorial space would suffer badly under isoluminant conditions.

3. GENERAL CONCLUSIONS

The results from the two experiments are in perfect agreement, and the major conclusion seems to be clearcut: perhaps remarkably so, and certainly against common expectations (but see⁵), observers do *not* do markedly worse with the isoluminant Red/Green and Green/Red transposed images. Whereas the transposed images certainly looked weird to our observers, they were easily interpreted in terms of pictorial structure. Much of the structure of pictorial space is apparently preserved, despite the fact that one may expect various pictorial cues to be significantly degraded.

Apparently the notion that spatial representations are not sustained under isoluminant conditions should be applied with caution. This is perhaps somewhat understandable in view of the fact that human observers are generally able to generate rich pictorial scenery on the basis of very scant image structure. Examples are abundant, think for instance of cartoon renderings as compared with full scale photographic prints. Indeed, in the visual arts it is common practice to produce renderings with greatly diminished (or even conflicting or arbitrary) cue structure for the sake of “visual interest”. Observers appreciate such renderings for their “artistic value” and perhaps enjoy the greater effort of their creative imagination involved in parsing such productions in terms of pictorial content. If this is indeed the explanation one would expect a much larger interobserver variation though. In our experiments we find hardly any signs of greatly increased idiosyncrasy.

Because some cues are doubtless considerably affected by transposition of the gray scale to isoluminant renderings (for instance shading), the question is what are the cues that are especially resistant against such transformations? In future research we hope to address this problem via progressive, controlled deterioration of monochrome images.

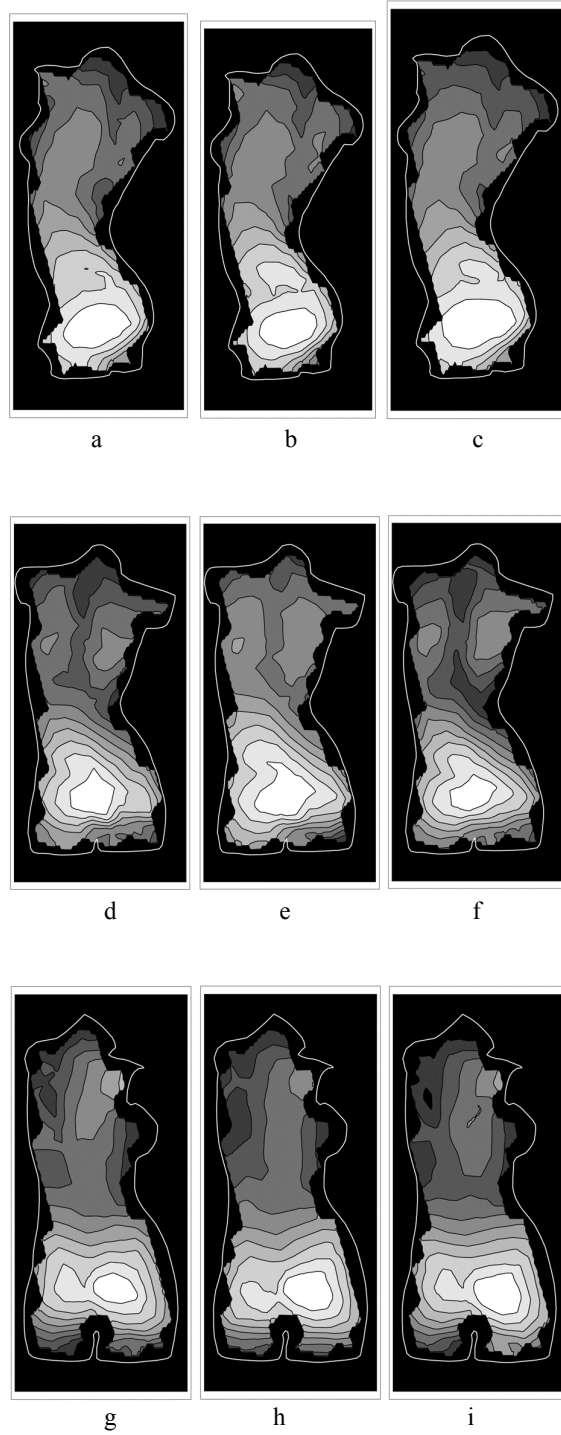


Figure 8: Isodepth plots of the pictorial reliefs (see Fig. 7) for observer AD. The conditions are respectively: a: Pose0 Green/Red, b: Pose0 Red/Green, c: Pose0 Monochrome, d: Pose45 Green/Red, e: Pose45 Red/Green, f: Pose45 Monochrome, g: Pose90 Green/Red, h: Pose90 Red/Green, i: Pose90 Monochrome.

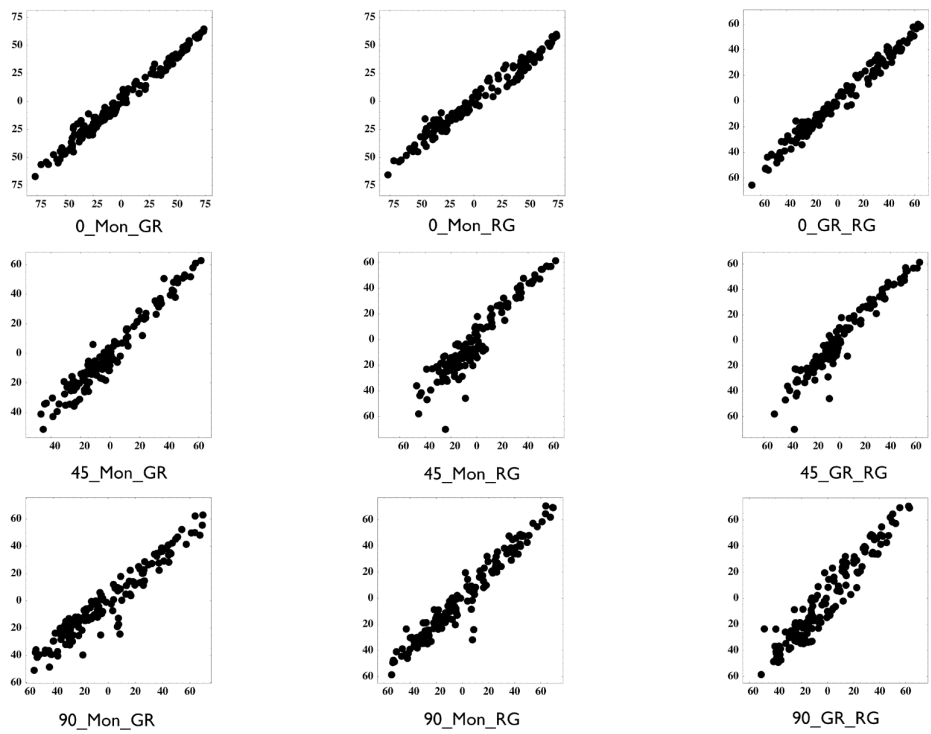


Figure 9: Scatterplots of depth values for observer AD. The plotted depth values in each figure are, respectively from left to right and from top row to bottom row, for the following conditions:
 horizontal axis Pose0 Monochrome, vertical axis Pose0 Green/Red,
 horizontal axis Pose0 Monochrome, vertical axis Pose0 Red/Green,
 horizontal axis Pose0 Green/Red, vertical axis Pose0 Red/Green,
 horizontal axis Pose45 Monochrome, vertical axis Pose45 Green/Red,
 horizontal axis Pose45 Monochrome, vertical axis Pose45 Red/Green,
 horizontal axis Pose45 Green/Red, vertical axis Pose45 Red/Green,
 horizontal axis Pose90 Monochrome, vertical axis Pose90 Green/Red,
 horizontal axis Pose90 Monochrome, vertical axis Pose90 Red/Green,
 horizontal axis Pose90 Green/Red, vertical axis Pose90 Red/Green.

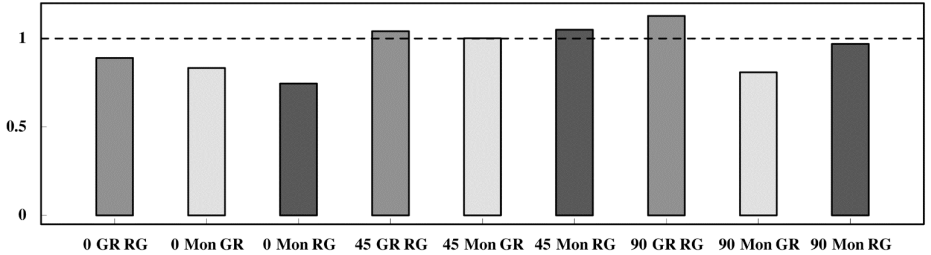


Figure 10a: Slopes of the scatterplots of depth values (Fig. 9) for observer AD. The bars correspond from left to right with the scatterplots of respectively:
 0 GR RG: depth values of Pose0 Green/Red against depth values of Pose0 Red/Green,
 0 Mon GR: depth values of Pose0 Monochrome against depth values of Pose0 Green/Red,
 0 Mon RG: depth values of Pose0 Monochrome against depth values of Pose0 Red/Green,

45 GR RG: depth values of Pose45 Green/Red against depth values of Pose45 Red/Green,
 45 Mon GR: depth values of Pose45 Monochrome against depth values of Pose45 Green/Red,
 45 Mon RG: depth values of Pose45 Monochrome against depth values of Pose45 Red/Green,
 90 GR RG: depth values of Pose90 Green/Red against depth values of Pose90 Red/Green,
 90 Mon GR: depth values of Pose90 Monochrome against depth values of Pose90 Green/Red,
 90 Mon RG: depth values of Pose90 Monochrome against depth values of Pose90 Red/Green.

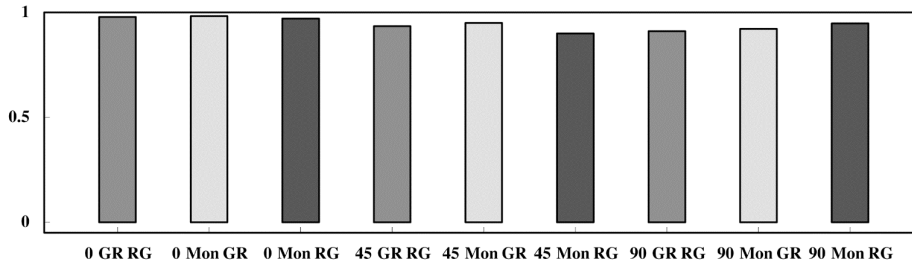


Figure 10b: R-squared values of the scatterplots of depth values (Fig. 9) for observer AD. The bars correspond from left to right with the scatterplots of respectively:

0 GR RG: depth values of Pose0 Green/Red against depth values of Pose0 Red/Green,
 0 Mon GR: depth values of Pose0 Monochrome against depth values of Pose0 Green/Red,
 0 Mon RG: depth values of Pose0 Monochrome against depth values of Pose0 Red/Green,
 45 GR RG: depth values of Pose45 Green/Red against depth values of Pose45 Red/Green,
 45 Mon GR: depth values of Pose45 Monochrome against depth values of Pose45 Green/Red,
 45 Mon RG: depth values of Pose45 Monochrome against depth values of Pose45 Red/Green,
 90 GR RG: depth values of Pose90 Green/Red against depth values of Pose90 Red/Green,
 90 Mon GR: depth values of Pose90 Monochrome against depth values of Pose90 Green/Red,
 90 Mon RG: depth values of Pose90 Monochrome against depth values of Pose90 Red/Green

REFERENCES

1. J. J. Koenderink, A. J. van Doorn, A. M. L. Kappers, "On so-called paradoxical monocular stereoscopy", *Perception*, Vol. 23, pp. 583-594, 1994.
2. E. H. Gombrich, *Art and Illusion: A study in the psychology of pictorial representation*, Phaidon Press, London, 1959.
3. J. J. Koenderink, A. J. van Doorn, A. M. L. Kappers, J. T. Todd, "Ambiguity and the "mental" eye in pictorial relief", *Perception*, Vol. 30, pp. 431-448, 2001.
4. M. Livingstone, D. Wiesel, "Segregation of form, color, movement, and depth: anatomy, physiology, and perception", *Science, New Series*, Vol. 240, nr. 4853, pp. 740-749, 1988.
5. P. Cavanagh, "Vision at Equiluminance", In: *Vision and Visual Dysfunction: Limits of Vision*, eds. Cronley-Dillon, Kulikowski, Walsh, Murray, Chapter 18, pp. 234-250, MacMillan, 1991.
6. Zhong-Lin Lu, L. A. Lesmes, G. Sperling, "Perceptual motion standstill in rapidly moving chromatic displays", *PNAS*, Vol. 96, pp. 15374-15379, 1999.
7. J. J. Koenderink, A. M. L. Kappers, F. E. Pollick, M. Kawato, "Correspondence in pictorial space", *Perception & Psychophysics*, Vol. 59, pp. 813-827, 1997.
8. J. J. Koenderink, A. J. van Doorn, A. M. L. Kappers, "Surface perception in pictures", *Perception & Psychophysics*, Vol. 52, pp. 487-496, 1992.

*a.j.koenderink-vandoorn@io.tudelft.nl