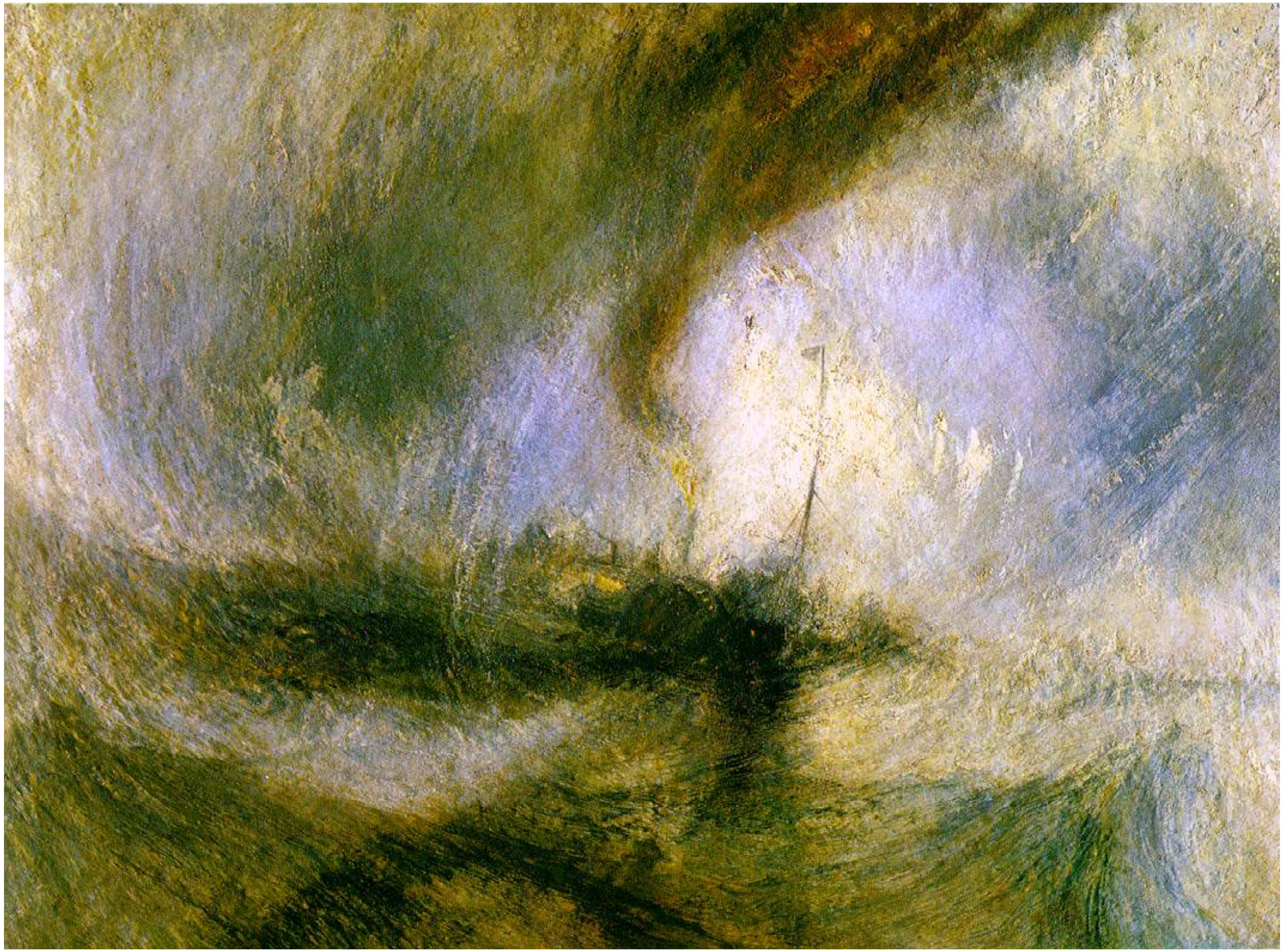
Pictorial Space

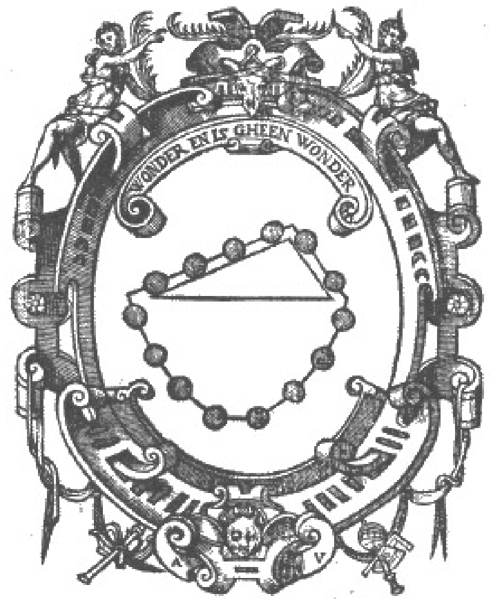
Jan Koenderink



DE CLOOTCRANS PRESS

Pictorial Space

Jan Koenderink



De Clootcrans Press, MMXII

Front cover: Joseph Mallord William Turner (1775–1851). Snow Storm – Steam-Boat off a Harbour's Mouth, exhibited 1842. Oil on canvas, 91.5 x 122 cm. Tate collection, London.

De Clootcrans Press
Utrecht The Netherlands
jan.koenderink@telfort.nl

pax / jan koenderink

First edition, 2012

10 9 8 7 6 5 4 3 2 1

Foreword

This short E–Book is intended as preliminary reading material for a summer course in visual perception. It is one of a series of short introductions.

The book was prepared in PDF–LaTeX, using the movie15 and hyperref packages. In order to make use of its structure use the (free!) Adobe Reader. Make sure that all options are enabled! Video and sound clips will start by clicking (or double clicking) the image, they are embedded in the file.

Make sure you are online. Clicking the light blue text will get you to Internet sites with additional material (try this page!). Be sure to read some of that, if good material is available elsewhere I'll skip it in the text. Most references occur only once, saving both you and me time. However, it means that you may have to backtrack at times.

All links were active when I checked last. I will check them again with the next reprint.

Utrecht, february 12, 2012 — Jan Koenderink

Jan Koenderink
Katholieke Universiteit Leuven
Laboratory of Experimental Psychology
Tiensestraat 102 – bus 3711
3000 Leuven
Belgium
jan.koenderink@ppw.kuleuven.be
Personal page

PICTORIAL SPACE

You become aware of "pictorial space" when you look into a picture. Pictorial space is completely disjunct from the space you move in, moreover, its geometrical structure is completely different. In some cases there appears to be a "wormhole" that connects these spaces, like Alice entering "through the looking glass". In such cases you are simultaneously aware of the picture as a physical object (say a planar surface covered with pigments in some configuration), and the pictorial space. However, neither the picture frame, nor the pictorial surface belong to pictorial space. Even more important, neither is the eye in pictorial space. This becomes obvious once you move with respect to the picture (as an object). Visual space and pictorial space are affected in very different ways. They are not really "connected" at all.

Looking into pictures

"Looking into pictures" sounds simple enough, but it really is an art. Some time ago science has decided that "stereopsis" (which originally meant "spatial (3D) vision") is the same as "binocular stereopsis". (You may try various modern dictionaries and the Internet to check this.) As a by-effect it has become "a fact" that monocular stereopsis is impossible. There are papers with "paradoxical monocular stereopsis" in the title. This is unfortunate, as virtually all painting before the twentieth century was supposed to show the observer 3D pictorial space. In Victorian times many families owned a stereoscope (as is well known), but—and this is less known—they probably also owned optical machines to view single pictures. Some view boxes were meant for monocular viewing, some for binocular viewing. The latter are often confused with stereoscopes. They are not. They merely (like the zograscope) presented identical images to the two eyes, thus minimizing the binocular cue that revealed the picture to be flat. In the early twentieth century Zeiss still designed (superior) view boxes like the Verant[1], as well as systems that essentially removed binocular disparity[2]. These machines were used to see pictorial space by looking into single pictures. These were doubtless instances of "stereopsis", albeit not *binocular* stereopsis. The modern usage of the word completely disregards the history of the topic!

Mixtures of pictorial and visual space are perfectly possibe. (Baroque palaces and churches are full of then.) They are best enjoyed from a single view point though, or by way of photographs. The moment you move the "tiger" breaks down and reduces to a flat array of colors on the pavement.

For pictures of a certain size (like most paintings) you don't need optical asistance at all, because the binocular disparity and physiological cues

[1]Designed by Moritz von Rohr (1868–1940)with important input from Allvar Gullstrand (1862–1930).

[2]This is the synopter, also designed by von Rohr.

(monocular parallax, accommodation, depth of field) are much weaker than the pictorial cues (often called monocular cues) used by the painter. When you stand in front of a nineteenth century landscape painting you should do the following

- stand in front of a painting at a distance of about two times the picture size;
- view the picture frontally, with the eye roughly centered;
- close one eye;
- wait.

After a short while you should become aware of pictorial space. It is every-bit as compelling as looking into a stereoscope. Many people even prefer monocular stereopsis over binocular stereopsis, because the latter often yields a depth segregation that is like a series of flat cutouts placed at various distances, whereas the former often shows pleasant plasticity. At the time this was common knowledge[3]. Nowadays it is largely forgotten.

Lord Kitchener is still looking at you straightly when his picture is seen oblique. He merely looks "thinner", apparently due to foreshortening. This clearly illustrates that the eye is not in pictorial space. Naive people think it "uncanny" that the eyes of a portrait en face *"follow you about the room". They ascribe it to almost supernatural powers of the artist.*

[3] You may want to read the paper (of 1941) by Harold Schlosberg (1904–1964).

This is just one of arbitrary many examples: what—do you think—is the preferred (some would say: uniquely implied) vantage point for this picture. The answer is: INDETERMINATE. *The point is that the eye is not in pictorial space. More precisely, neither the eye, nor the picture (as a physical object) are in pictorial space.*

A movement into depth. (video clip).

Modern people are sometimes amazed at the discovery that a frontally painted haunted portrait "follows you about the room with its eyes", a discovery of the difference between pictorial space and physical space.

Adolf Hildebrand

Adolf von Hildebrand (1847–1921) was a German sculptor, well known in his time. He was highly influenced by his contacts with the painter Hans von Marées (1837–1887) and the theoretician (and maecenas) Konrad Fiedler (1841–1895). In 1893 he published a book "*Das Problem der Form in der bildenden Kunst*" (The Problem of Form in the Visual Arts) that rightly became famous. Although highly influenced by his relation with Konrad Fiedler, the book is evidently Hildebrand's own. It is written from the perspective of an active visual artist.

The book attracted considerable criticism at the time, because it places pictorial space in a prominent position, which appears perhaps odd for a sculptor. I will not enter into that here, I'll merely mention Hildebrand's views as they immediately apply to pictorial space.

One technical distinction Hildebrand makes is between the *Fernbild* (far view) and the *Nahbild* (near view). These are modes of vision. The *Fernbild* corresponds to what I call pictorial vision, the *Nahbild* to vision in the space you move in. Thus the *Nahbild* is pieced together from multiple perspectives and multiple fixations, it is vision by active probing via movements of eye, head and body. The *Fernbild* is artistically more interesting, because more of a unitary entity. The *Fernbild* can be taken in by way of a single view, preferably monocularly, with stationary body. The *Fernbild* leads to visual awareness due to the artistic configuring of the artwork, whereas the *Nahbild* leads to the awareness of a dynamic exploration. Hildebrand considers the latter in artistically doubtful taste, the task of the creative artist is to build a *Fernbild* in such a way that the observer is unconsciously forced to accept the artist's intentions.

Adolf von Hildebrand (born at Marburg, 1847, died in München 1921). Worked mainly at Berlin, Florence, and later at Münich (e.g., the monumental Wittelsbacher Brunnen).

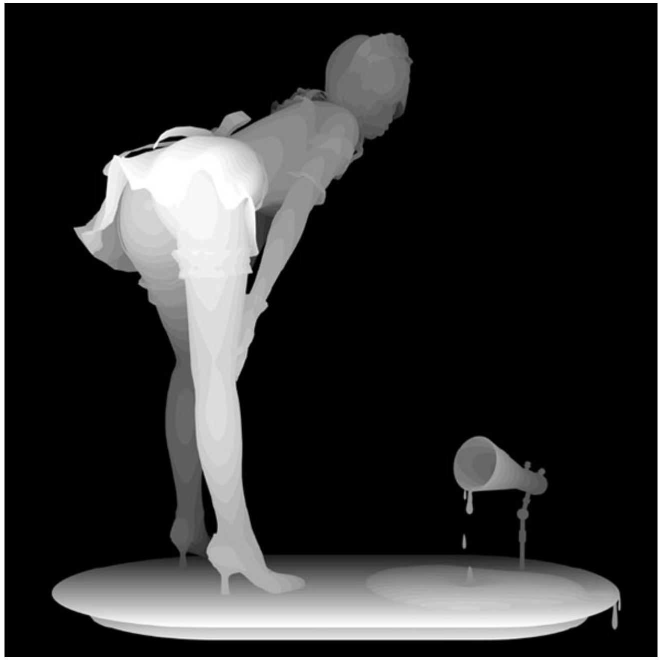
The notion of *Fernbild* is well known in the arts. Already Leonardo (1452–1519) tells the painter to prefer viewpoints at a distance of at least three times the size of the subject. This corresponds to a medium tele, say ninety millimeters, lens of a Leica camera[4].

As a sculptor Hildebrand is especially concerned with shape and space. He considers shape to reside in the depth articulation of the surfaces of opaque, rigid objects (like statues). These surfaces are articulated in *depth*, which is to be distinguished from "range", that is (Euclidean) distance reckoned from the eye. Depth is a sense of separateness from the self. It has no origin, nor a unit measure. This is why Hildebrand speaks of a *Reliefauffassung*, which I will translate as (pictorial) "relief".[5] Hildebrand illustrates the difference with

[4] The Leica was designed by Oskar Barnack (1879–1936) in the early nineteen-twenties. He combined two movie frames (the camera was to use 35mm perforated movie stock) into what became the world standard 24x36mm frame.

[5] At his time well known art historians wrote that the *Reliefauffassung* as related to sculpture is like "wooden iron". Hildebrand couldn't care less though. He meant it.

Euclidean distance with the fact (well known to sculptors) that observers from some distance fail to distinguish between flattish work (often called "relief" sculpture) and work "in the round". The depth dimension is volatile and often indeterminate. It is the task of the artist to create the relief for the observer in such a way that it easily "reads". In order for that to happen the artist must often use means that deviate from merely copying physical reality. What artistically counts is only the visual awareness of the observer, not veridicality as referred to physical reality.

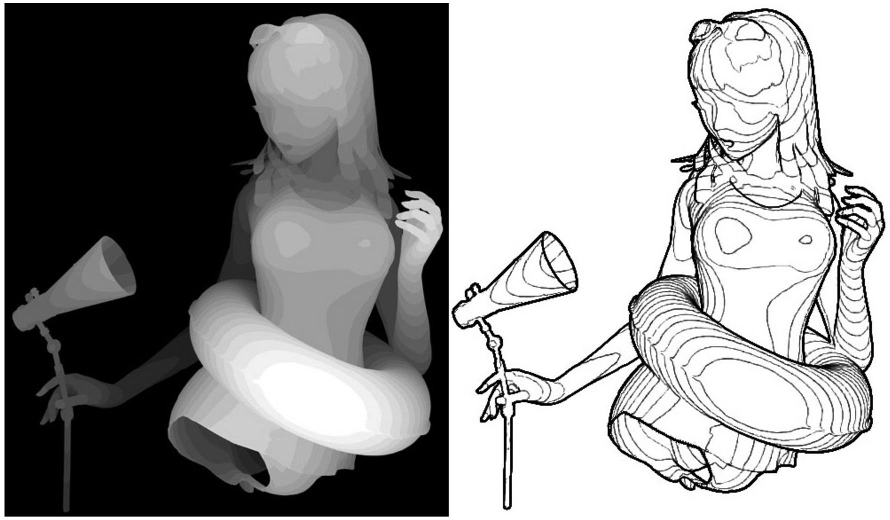


Kazuki Takamatsu paints "depth maps". Notice how the graytone is monotonically related to depth in pictorial space. It is a modern way to implement Hildebrand's Reliefauffassung.

In order to suggest how the *Relielfauffassung* may be reached Hildebrand suggests that you think of a form as existing between two frontoparallel panes of glass. You are to relate the shape to these planes. Artistic composition takes place in terms of this representation, it works by planes, even if the object is curved.

The two parallel planes of Hildebrand are an implementation of an affine gauge (although Hildebrand was unaware of that). It allows you to calibrate depth as the fraction of the depth slice defined by the pair of parallel planes (see below).

The depth slicing also allows you to describe the relief in a qualitative, topological manner. The pattern of equal depth curves through the singular points of the relief (minima, maxima and saddle points of depth) divide the relief into a juxtaposed and nested order of bulges and depressions.



I ran an edgefinder on one of Kazuki Takamatsu "depth maps". It allows you to study the topological pattern of juxtaposed and nested bulges and depressions.

In discussing visual space Hildebrand introduces a notion of "depth flow" (*Tiefenbewegung*). You sense the space by moving into it until the movement is stopped by some pictorial object. It is the artist's task to regulate this flow into well shaped patterns.

When the flow hits a relief it is diverted and takes the shortest path into depth along the surface. This reveals a topological pattern of hills and dales that is complementary to the pattern of bulges and depressions described above. In the case of depressions the flow stagnates and forms a "pool". Hildebrand considers this bad form. The flow should nowhere stagnate and smoothly flow over the form[6].

You can easily experience the "depth flow". Simply place a mark on the picture surface. Except for special cases (sometimes visual awarenes attributes the mark to the picture surface) the mark will be carried along with the depth flow as a novel "pictorial object", only to stop when it hits the nearest relief. This is often used by street artists who paint mustaches under politicians noses through putting black marks on the surface of some election billboard. This shows evidently that depth flow exists.

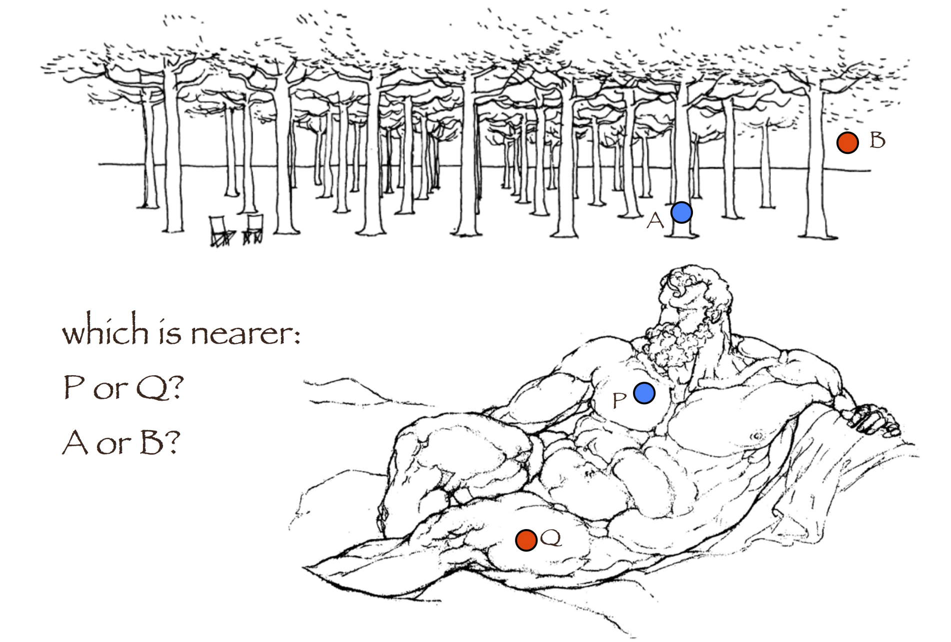
MEASUREMENT IN PICTORIAL SPACE

Measurement in pictorial space is a key issue. Only when you are able to measure its properties are you set to venture empirical research of the phenomenon. Thus one of the major problems in the science of pictorial space is the development of effective measuring techniques.

Techniques of measurement

In physics a measurement determines the ratio of a physical quantity to a conventional unit. The key example is the application of a yardstick to an object in order to determine its length. In order to perform a measurement you thus require a standard unit and a procedure that allows you to determine the ratio. In psychology this is problematic, because the entities of interest are mental, rather than physical (I ignore the numerous physiological measurements performed in the psychological laboratory, these are essentially physical measurements). There are no standard units, and no conventional procedures.

One time-proven method in physics is the *method of comparison*. Suppose I manage to show that "A matches B" in some objectively specified way. Then I may take A as my standard and regard the matching as a measurement of B with the (obvious) result "one standard unit". The beauty of this is that I don't need to have a theory of the property being measured. As long as I can *judge the match* I need no further information.



Which is nearer, the blue point or the red point? The question is easy to answer in case of the lower figure (the red point of course!), much harder (impossible even) in case of the picture at top. There the red point might be a distant star, or a bug buzzing among the trees.

Examples in physics abound. In order to measure a length we hold a standard stick next to it. If the end points match up the object is "one stick long"[7].

[6] This is perhaps inspired by the well know rule that statues should be such that no pools form as water rains down on them as this would evidently detract from their useful lifespan.

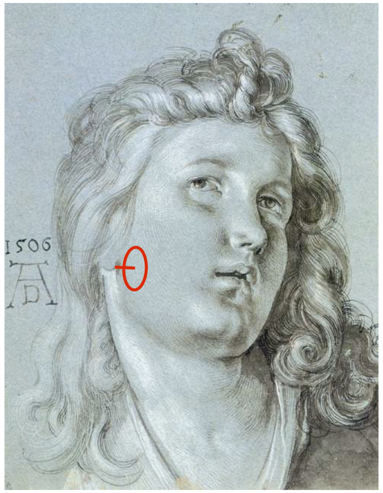
[7] It is an easy matter to extend this to "two sticks", "half a stick", and so forth.

In order to measure weight I use scales, placing a standard stone on one scale, the object of the measurement on the other scale. If the scales don't tip the object "weighs one stone". In order to measure luminous power I compare the brighnesses due to a standard candle and the source to be measured, equating distances, etc. If the brightnesses match the source has a "strrength of one candle". And so forth, and so forth. I don't need much of a theory of space, mass and gravity, or electromagnetic radiation to perform such measurements. In fact, measurements were performed before theories were being framed, and indeed were instrumental in the development of such theories. I propose to use this *Principle of Matching* to develop methods in psychology. A few instances have already been developed, for instance, the field of colorimetry as developed by Maxwell (1831–1879) and Helmholtz (1821–1894) in the nineteenth century is a key example.

In order to perform measurements in pictorial space you have to be able to put some standard in pictorial space.

The black spot at right was put on the picture surface. Yet it appears as a "beauty spot" on the cheek of a lady (actually a pattern of gray tones). Because the spot is evidently a flat figure, you are likely to be able to see it in either way. If you see it as a mark on the picture surface it apparently "hoovers in space" in front of the lady. If you see it as attached to the cheek it travels back into depth until it "gets stuck". THIS IS WHAT MAKES EXPERIMENTAL PHENOMENOLOGY POSSIBLE.

The first problem this introduces is that you have to be able to put something into pictorial space in the first place! That might be supposed hard, since pictorial space is not something physical, but something mental. Here you may take a cue from the painter: as the painter puts a stroke on the canvas, the painter achieves essentially two goals. One is to indeed attach a spot of color to the canvas. The other is to put something in pictorial space. Thus you can "reach" pictorial space by way of the picture surface. This is indeed obvious from the frequent mustaches and black teeth you see in billboards of politicians around town.



The gauge figure in use. Notice how it "fits" the cheek of Albrecht Dürer's Angel. *When attached to other parts of the face you would have to adjust the spatial attitude of the figure. That is the basic principle used to probe the local attitude of pictorial relief.*

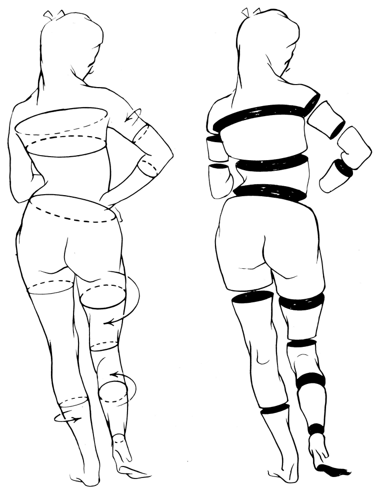
Since we are attempting measurements of a geometrical nature, I will refer to the standard as a "gauge figure". It is the analogon of the yardstick, dividers, template, sieve, and so forth used in geometrical measurements in physical space. Then we need to establish these:

- a suitable gauge figure. This will depend upon the type of quantity to be measured;
- a way to place the gauge figure in pictorial space. Here we may use Hildebrand's depth flow to good advantage;
- a way to manipulate the gauge figure so as to enable a "match". Since the gauge figure is in pictorial space, we can use computer graphics to

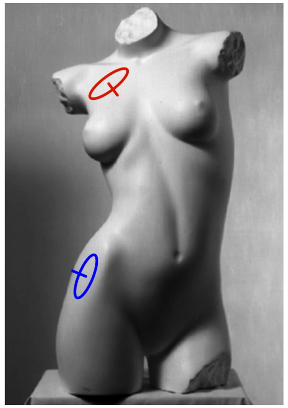
implement this;

- a way to judge the "fit" of the gauge figure. This is the critical issue. For a valid psychological method the fit should be obvious in immediate visual awareness. Since pre-cognitive, it cannot be "explained".

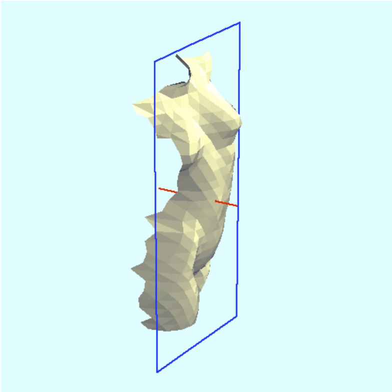
The latter point is absolutely crucial. This is in no way cognitive science. The fit is an immediate visual judgment.



Ellipses are a standard too of the draughtsman. They are used to force the viewer to "see" intended spatial attitudes. The ellipses are sometimes explicitly drawn, but usually they are more implicit, artfully hidden.



This photograph of a torso was used as the "stimulus" using the gauge figure task. Notice the two wireframe figures that have been superimposed. They are computer graphics renderings of a circle with a line segment the length of the radius sticking out of the plane of the circle at right angles. They have been chosen such that the blue one should look like a "reasonable fit" (the circle looks like "painted on the pictorial relief"), whereas the red one is obviously not on the pictorial relief at all.



Here is the "response" to the "stimulus" show in the previous piocture. The blue wireframe rectangle is the boundary of the picture frame, the red line at right angles to it is the depth dimension. Notice that the measurement apparently achieved a 2D → 3D transformation.(Video clip.)

Notice that the fit is supposed to be in pictorial space, not in the visual field. For instance, suppose the gauge figure is a dot and you are supposed to judge whether the dot "fits" ("is on") a line in the image. The result might work out very differently in the visual field and in pictorial space. Two coinciding points (trivially) "fit" in the visual field, whereas they may be far apart in pictorial space. There is no easy way to figure out what observers are doing. Speed is the main criterion.

The foot of the man and the contour of the leaning tower of Pisa coincide in the pictorial plane. They could be far apart in pictorial space, or not, it depends upon your immediate visual awareness. Pictorial space and the (probable) physical scene at the time of exposure are two distinct entities.

Some people are effectively blind, in the sense that they have to reason anything out. My artist friends say that these people apparently "see with their ears". Such people cannot be trusted with the fit in pictorial space. They consider the layout in the picture plane and start calculations of some kind in reflective thought in order to arrive at a "fit". You easily spot people like that in the laboratory because they take ten times as much time as normals on the task. (So having MANY trials in a session helps to suppress the urge to think.) Perhaps unfortunately, people like that seem to prefer careers in experimental psychology. There is probably some reason for that, thought it beats me what. They couldn't survive long outside our protective Western societies, since they are unlikely to last through their first street fight. By the time they have a suitable response ready as a result of reflective thought their adversary is already many moves ahead of them. This is an important point, unfortunately rarely appreciated by "cognitive scientists".

According to needs the gauge figure will be parameterized in some particular way. In a trial the observer is asked to set the parameters such that a fit is obtained. Apart from the need to be able to judge the fit visually, this requires a "natural" interface. Here is another implementation problem. Typically the fit can be obtained via numerous possible implementations that are formally equivalent. It is often not appreciated that "formal equivalence" by no means implies "equally natural" in terms of the interface. For instance, there is a popular kid's toy in which a writing implement can be moved via two knobs, controlling the X and Y coordinates in a Cartesian parameterization. Most adults are quite unable to write their name with such a contraption. Try it when you can. It tends to be a painful experience. The toy is popular exactly *because* it is such an unnatural interface. Unfortunately, many interfaces one encounters in the psychological laboratory are easily as unnatural (I've seen even worse). This fully destroys the efficaciousness of any method of fit. It is absolutely crucial that the interface be natural.

There are various variations on the method of fit. For instance, one might have two gauge figures, one a fiducial, the other adjustable, and require that the observer match the fiducial in some way. The result will depend on the locations and embeddings of the gauge figures in pictorial space.

Pictorial relief

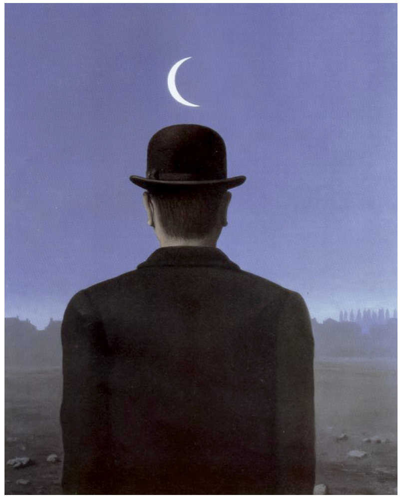
"Pictorial reliefs" are surfaces of solid, opaque pictorial objects. A local patch of pictorial relief exists at a certain location in the visual field and pictorial space. The difference is that it is merely a patch of the pictorial plane in the former case, but has both a depth and a spatial attitude in pictorial space. The depth is already a property of a point, whereas the spatial attitude is a local property of pictorial relief. It is something you might well desire to measure.

An example of "pictorial relief". This is a photograph of a huge chunk of marble known as the Farnese Hercules. You are aware of a "pictorial object" bounded by a "pictorial surface", of which you see only the front side (there simply is no backside in the photograph, although there is in the scene). This is the "pictorial relief".

It is of little interest to attempt to measure "depth" as such, because depth is a volatile entity, in a sense is indeterminate. Only depth relations have some meaning, e.g., point A might look farther away than point B, even if both points individually are at some indeterminate depth. That is why the spatial attitude of a patch of pictorial relief is one of the simplest geometrical entities you may want to measure.

Different from the previous case, this is a drawing. The backside never existed. What you are aware of is pictorial relief.

The spatial attitude of a surface patch is primarily characterized by a slant, or obliqueness. The slant ranges from nothing (frontoparallel) to edge-on. the surface may be slanted into different directions. This is characterized by the tilt. The tilt is a direction in the visual field. It can be represented by a Euclidean angle relative to some fiducial direction. Convenient fiducial directions are the horizontal (towards the right) or the vertical (upwards), for instance. The tilt apparently varies in the range $0°$–$360°$. The numerical representation of the slant is less immediate. A natural way to parameterize the slant is by the positive real numbers, zero standing for frontoparallel, infinity for edge-on. This would be like the tangent of a Euclidean angle in the range $0°$–$90°$. That relation may be used in the generation of suitable gauge figures, for that we need some Euclidean parameterization.

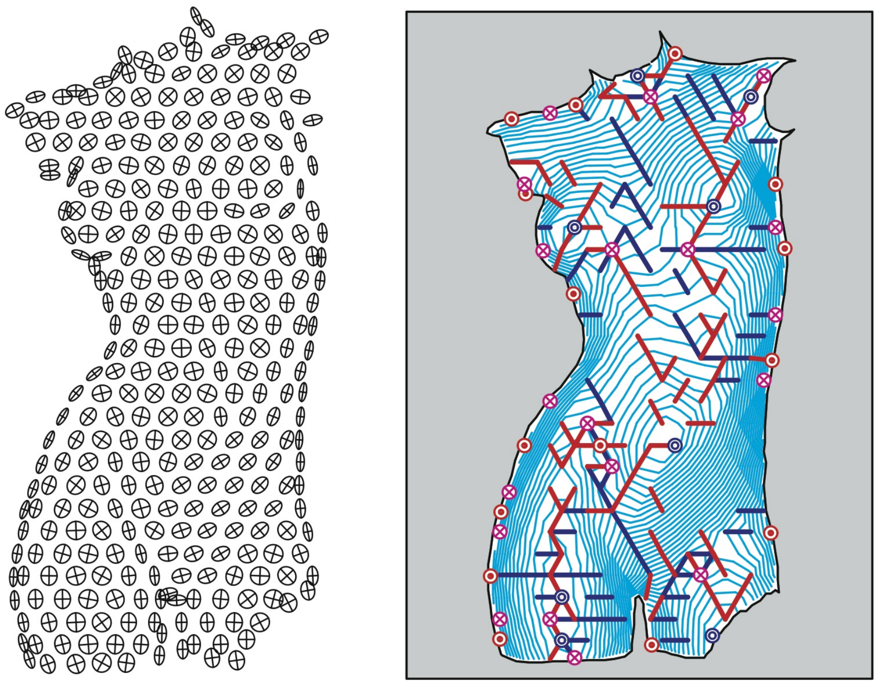


Most people desire to see the face of Magritte's "Schoolmaster". Alas, it was never painted! You can't turn around in pictorial space, you can only turn around in the picture plane. The latter won't reveal the face though.(Video clip.)

This is a rotation about the vertical in physical space, as seen from some fixed point. Now you can see all sides (of the torso in this case). Notice the difference with the previous case. (Video clip.)

One convenient gauge figure is the orthographic projection of a circle. In order to avoid front–back ambiguity, I add a line element (the length of the radius of the circle), sticking out of the plane of the circle at right angles. The "fit" would be to see the circle as "painted upon the pictorial surface". This gauge figure was inspired by a common technique in drawing, where one uses elliptical depictions of circular elements in the scene as a cue to spatial attitude. The gauge figure is this cue "used in reverse" so to speak.

The gauge figure is a wireframe computer graphics rendering of a circular wheel with axle sticking out at right angles. The task is to make it appear as a circle painted upon the pictorial surface. The result of the measurement are the slant and tilt angles used to parameterize the computer graphics rendering. This result is *defined as the spatial attitude of the patch of pictorial relief.*



At left the individual responses that were obtained one at a time, in random order. This result is perhaps half an hour's worth of interaction with the picture. At right a geometrical analysis of the pictorial relief in terms of equal depth loci, water divides and water courses. Pits, peaks and saddles of the relief are also indicated.

This is crucial. There have been attempts to "calibrate" the result of such a measurement. Such attempts show a misunderstanding of the principle of measurement by comparison. There is nothing more to calibrate. The "seen" surface attitude is only known through the measurement. It does not even *exist* apart from the measurement. Thus there is nothing to "calibrate".

This may need serious thinking over. Although simple enough, it has frequently been misinterpreted in the literature.

This method was designed by us in the early nineteen-nineties. It has yielded many interesting results, and it still remains easily the best tool for the task. Many different methods are conceivable though. The interested reader is invited to exercise hers or his creative mind. This is only a young science, and there remains much more to be done!

Deep space

Pictorial reliefs are coherent, extended entities. Thus they are locally specified by derivatives, like the spatial attitude or curvature. But what about locations in pictorial space that are far apart and not on some common surface? One evidently requires different methods to measure relations between these.

It is not hard to design methods that apply to these cases. I describe two. Many others can be conceived of and might work just as well. Again, the interested reader is invited to exercise hers or his creative mind!

POINTING FROM HITHER TO YONDER is a simple and effective method to probe the spatial relation between two points in pictorial space. In *physical space* this is the method of "exocentric pointing". Consider what this involves. There are a pointing device (an arrow say) and a target at locations in the scene at some distance from you. You have remote control over the spatial attitude of the pointing device. The task is to let the pointer point at the target. Notice that this involves more than what you need in egocentric pointing, like in pointing a gun. The fact that the pointer is at some remote location renders the task nontrivial. However, I find that most people have no problem doing it. Of course they commit significant systematic errors, but that only shows that visual space and physical space are different entities. We already knew that.

Here is an introductory scene from Sergio Leone's Once upon a time in the West. I attached a pointer to the head of the man at left, a target to the man in the background. The observer controls the spatial attitude of the pointer.

A deep space (landscape painting Corinth with Akrocorinth by Carl Rottmann).

You should have some intuitive feeling for how good you are at such a task. In a company you routinely notice "who is looking at whom". That is exactly what the exocentric pointing task is about.

You can do exactly the same task in pictorial space. Instead of a pointing device you use a picture of the pointing device, and instead of the target the picture of a target. Otherwise you're all set!

RELATIVE APPARENT SIZE is a good probe for depth. It exploits the observation that two objects in the visual field that differ only in size are often seen as "equally large", but at different distances. This works well in pictorial space, even though the eye is not to be found in pictorial space, thus "distances from the eye" being indeterminate.

"The baths of Caracalla" *(1899, Alma Tadema) with superimposed disks. The observer controls the size ratio of the disks, their (geometrical) mean size being constant. The task is to let the disks look as if they were of equal size in pictorial space.*

The method is rather simple to implement. For instance, one places two circular disks on the the picture plane. Their geometrical mean is kept constant, whereas their size ratio is made adjustable. When these are seen as similar disks or spheres in pictorial space, it makes sense to ask an observer to set the size ratio of the disks so as to make the objects in pictorial space identical.

The logarithm of this size ratio is taken as the depth difference. It is a real number, the sign is evidently relevant.

This definition is reasonable because it makes the depth different independent of absolute size, and it makes that depth differences are additive.

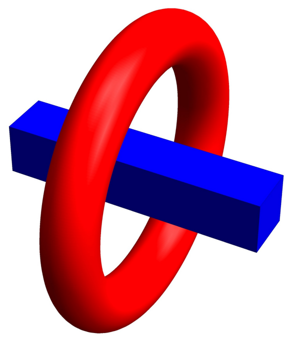
FORMAL STRUCTURE OF PICTORIAL SPACE

The formal structure of pictorial space is a priori unknown. It is easy enough to show that pictorial space cannot have a Euclidean structure, like the space you move in. Otherwise no holds are barred. An initial approach may be based upon an analysis of the well understood invariances under various operations. I freely use findings that derive from various studies that use the methods of measurement discussed above.

Depth

"Depth" is a sense of remoteness. It admits of degrees, some things feel "near", others "remote". It is important not to confuse depth with "range", or to expect a monotonic relation between range and depth. Range is distance from the viewpoint (as your eye, or—in the case of far distances—your body). You can measure range with a measuring tape, or (more conveniently) with optical devices like the range finder on the Leica cameras. You cannot possibly measure depth by some physical means. Depth exists only in visual awareness.

In many cases (but not necessarily so) depth is reasonably well developed. It is apparently "hallucinated[8]" in microgeny so as to account for the optical structure. The optical structure may be understood in terms of "cues", guesstimates based on assumed physical causes. For example, small angular size might be interpreted as great remoteness (the repeated experience "people seen from afar look like ants"). This might give rise to a "size cue" in the form "people that look like ants" (or perhaps "ants that look like people") "are at great depth".



Is the annulus in front or behind the rod? Depth is obviously a point property, not a property of pictorial objects.

A formal analysis of the size cue gives some additional insight. Consider the stimulus patterns $\boxed{\text{XXXX}}$ and $\boxed{\small\text{XXX}}$, clearly the second one looks more remote than the first one. Let's say there is a depth difference D, say. Now if the shape cue is coherent, it is reasonable to expect that the depth difference between $\boxed{\text{XX}}$ and $\boxed{\small\text{XX}}$ is also D. This is simply to say that the parts of objects should be at the same depth as the objects themselves. Let the depth difference between $\boxed{\small\text{XXX}}$ and $\boxed{\scriptstyle\text{XXXX}}$ be $\overline{D}$. Again, if the shape cue is coherent, it is reasonable to expect that the depth difference between $\boxed{\text{XX}}$ and $\boxed{\scriptstyle\text{XX}}$ is $D + \overline{D}$. This is simply to say that the depth difference D_{AC} in a depth sequence A—B—C equals the sum of the partial depth differences D_{AB} and D_{BC}. This defines functional equations that force the depth difference between such things as $\boxed{\text{X}}$ and $\boxed{\scriptstyle\text{X}}$ to equal the logarithm of the ratio of their sizes (you may want to prove this to your satisfaction).

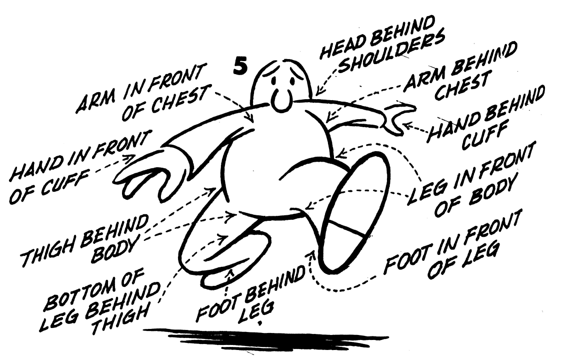
Notice that I have not referred to optics for this derivation, only to the structure of the depth domain as a coherent interval scale. However, it is easy enough to do the optical analysis and show that it suggests the same result. Suppose that two identical objects yield images $\boxed{\text{X}}$ and $\boxed{\scriptstyle\text{X}}$. The angular

[8] If you feel ashamed (as a scientist) to "hallucinate" you might prefer the term "analysis by synthesis".

size of the image of an object is (to an excellent approximation if the objects are sufficiently distant) the physical size of the object divided by its range (distance from the eye). Thus the ratio of the angular sizes of the two objects that correspond to the observed images is the reciprocal of the ratio of their ranges. Thus you obtain an interval scale if you define the depth difference as the logarithm of the angular size ratio.

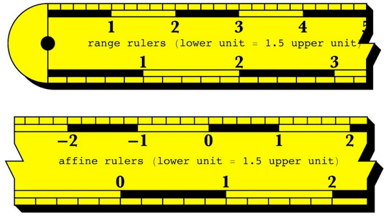
Although this looks like a derivation it really is not. The optical analysis has nothing to do with perception per se, it merely yields a hypothesis that might be put to the test. It suggests a fixed functional relation between range and depth. We already know that such a fixed functional relation breaks down in many cases. Moreover, it would be an odd relation between a physical and a mental magnitude, an ontological paradox.

The phenomenological derivation depends only upon a "view from the inside" so to speak. Thus it avoids crossing the ontological gap. That is why it applies to pictorial depth, whereas the optical derivation does not. Since the eye is not in pictorial space, range is indeterminate, and the optical reasoning cannot even start.



Example of a cue, this is the "occlusion cue". Notice how simple it is to use the cue and how effective it is! This is a figure from a "how to draw" book. It has neither algorithms, nor formulas. It makes its points very clearly though.

The fact that the depth domain is an interval scale is important. It implies that there exists neither a natural origin (that would be the eye, but the eye is not in pictorial space at all), nor a natural unit. Only "affine properties" such as the bisection of a finite stretch are well defined. Formally, one says that the depth domain has the structure of the "affine line" $\mathbb{A}^1$ (where the upper index "1" denotes one-dimensionality), or simply $\mathbb{A}$ (the dimensionality being understood).



Two rulers. The top one is a "ratio scale", you can use it for range. Notice the origin (the hole). The two scales are equivalent, they are in proportion and share the origin. The bottom one is an "interval scale", you can use it for depth. There is no origin. The two scales are equivalent, they are related by an affinity.

Many of such cues can be packaged as simple algorithms, often called "SFX" ("Shape From X" algorithms). For instance, the size cue would be packaged as "range is inversely related to angular size". In most cases such algorithms, as developed in computer vision, hardly pertain to the case of human vision. Sometimes they do, somewhat. Most cues are of a somewhat different, simpler and more qualitative nature. This is not necessarily a bad thing either, for qualitative rules tend to be much more robust against all kinds of relaxations of the prior assumptions than quantitative recipes can be. And—of course—the prior assumptions are just that, that is *guesses*, they can never be checked.

I have ignored the cues based on multiple viewpoints and the so called physiological cues here.

The cues based on multiple viewpoints are of various types. The most obvious (at least as judging from the pop-science press) is the fact that we have two eyes.

Most vertebrates have two functional eyes. It fits the generic bilateral symmetry. However, the eyes are put to rather different uses. Roughly spoken, there exist two types, the predators and the prey animals, let's say the tiger and the lamb.

The tiger and the lamb. Notice the placement of the eyes in the skull and the relative directions of the optical axes.

The lamb has laterally placed eyes. Each looks in a different direction, there is hardly any binocular overlap. As a result the lamb has almost panoramic vision in the horizontal plane. The eyes are specialized for wide-angle vision. Eye movements are hardly necessary. This is very useful if you are a potential prey, nobody can sneak up behind you and remain unseen. Predators get a run for their money because you (or one of your friends, the prey animals often seek company) see them coming. But because there is hardly any binocular overlap, there is no binocular stereopsis[9]. If you take the pop-science accounts literally this would imply that such animals "see no depth". Of course this is nonsense. These animals can generate as many viewpoints as they want by moving.

The tiger has frontally placed eyes with a huge binocular overlap. The eyes are always used in unison and fixate the same target. Moreover, the eyes are specialized for focal vision. This means that eye movements are necessary and very frequent. The doubled viewpoint enables depth discrimination in near space without the need for movement. This is useful if you sneak up on a prey and need to prepare for the final attack. The disadvantage is that such animals have a huge blind spot behind them. They have detailed vision in a narrow cone that can be used to probe through successive fixations. They have low resolution vision in the frontal half-space without the need for eye-movements, thus simultaneously. But they have to watch their back by way of head movements. They develop habits that keep their backs covered.

Humans are like the predators and are able to use binocular stereopsis. Moreover, like virtually all animals, they frequently generate fresh viewpoints via body movements (as locomotion). Humans, like the other primates, are good with their hands and are able to rotate smallish objects so as to view them from all sides.

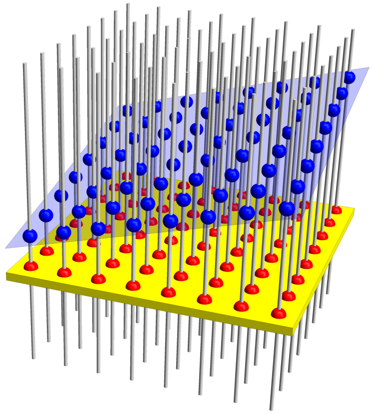
The so called physiological cues are of relatively minor importance, and only work on short distances. They have to do with the optical system, think of the accommodation of the eye and the effect of defocus blur. In this eBook I only consider "pictorial cues". These are often called "monocular cues", but the physiological cues are also monocular, so I prefer "pictorial". A key example of a pictorial cue is the shading cue, which I will consider later on.

Pictorial space as a fiber bundle

"Pictorial space" is a composite of two very different components. One is the visual field, which is a two-dimensional manifold with a geometrical structure (due to local sign and mensuration by way of eye movements) that is perhaps not unlike that of a patch of the Euclidean plane. The other is depth.

Depth is a rather volatile dimension, and it may well vary from place to place in the visual field. An example is the common observation in mountainous region that if there is a light, low hanging fog, with a clearly visible mountain top sticking out above it, the distant mountain may look nearer than the landscape at a few hundred meters away.

[9] "Stereopsis" means literally "spatial vision".



A "fiber bundle". Here the base space is yellow, the fibers are drawn as gray rods. Their mutual relation is purely formal, don't "take the finger for the moon"! I show a planar cross section in blue. The blue spheres are the "beads" of the Glasperlenspiel *played by microgenesis. (Video clip.)*

The formal description should take account of this. A simple formal, initial model of pictorial space may be constructed as follows. The "base space" for pictorial space is the Euclidean plane, symbolized as $\mathbb{E}^2$, where the upper index 2 denotes the two-dimensionality. At each point of the base space we attach a one-dimensional space, the depth dimension. Depth is at best represented on an interval scale[10], I will thus formalize it as the "affine line" $\mathbb{A}^1$, or just $\mathbb{A}$. The construction then is $\mathbb{P} = \mathbb{E}^2 \times \mathbb{A}$, where $\mathbb{P}$ denotes pictorial space[11]. Such a construction is known as a "fiber bundle", the "fibers" being the depth spaces. Notice that there are infinitely many fibers, one for each point of the base space $\mathbb{E}^2$.

A point in pictorial space lies on some fiber. The point that the fiber is attached to is the "canonical projection" of the point, I will refer to it as the "trace" of the point. The "bundle projection" thus simply "forgets depth".

The idea that the fibers are "attached" to the base space is pure formal. You should not think of any particular point of the fiber as being "in" the base space. Moreover, it is not like the fibers stick out of the plane at right angles (although I will often draw them that way). They make no particular angle with the plane. It is only that each fiber is related to a particular point of the base space.

[10]Notice that range would be represented on a ratio scale.

[11]Be aware that it is conventional to use $\mathbb{P}^n$ for n-dimensional projective space. It will be no problem here, because projective spaces do not figure in this eBook.

A "pictorial relief" is a surface in the fiber bundle, a smooth assignment of a point on each fiber. This is technically known as a "cross section" of the bundle.

Microgenesis somehow assigns depth values, that is points on the depth fibers. Intuitively, this might be pictured as a "glass bead game" (*Glasperlenspiel*) in which microgenesis shifts beads along the depth fibers until it is satisfied with the resulting configuration. The glass bead game apparently has some formal rules to it. One rule of the game is the affine structure of the fibers. Another is the special role of planar cross sections as I will argue in the next subsection.

Constraints posed by the "cues"

I consider a single nontrivial cue by way of an example. It is SFS ("Shape From Shading"). Here the prior assumptions are rather involved. Suppose you have some irradiance pattern on your retina. Than you might venture the following assumptions

- the pattern is due to the scattering of radiation from a smoothly curved surface (no forest scene);
- the surface is of a uniform nature (no zebras);
- the surface scattering is Lambertian (this actually rules out *any* actual surface; however, something like blotting paper or plaster comes close);
- the surface is illuminated by a uniform, unidirectional source (in practice sunlight qualifies);
- the surface itself does not occlude parts of the source (this rules out illuminated by the overcast sky in most cases);
- there is no multiple surface scattering (this rules out non-convex surfaces).

This is not a complete list, but it will do here. The upshot is that the probability that these assumptions will ever apply is virtually zero. They may apply approximately for low relief surfaces though. In case they do, there exists a simple relation ("Lambert's Law") between the radiance scattered to your eye, and the spatial attitude of the corresponding surface element with respect to the source. This allows you to start inverse optics calculations!

However, it turns out that the inverse optics calculations lead to a set of infinitely many equally applicable, but mutually very different solutions. So what is the “right” solution? There is no way to find out on the basis of the SFS cue, you have already exhausted its power.

Thus the shading cue is not particularly useful as a prospecting technique, although it works somewhat, in some circumstances, if you are ready to happily accept ambiguity and low precision.

It is possible to extract some useful qualitative facts from this though. One such a fact is that a planar, uniform surface, illuminated by a uniform source, will lead to a uniform patch of retinal illumination. Thus a uniform patch of retinal illumination may well be interpreted as the image of a planar surface. Thus you can “see planarity” under the assumption that you are looking at a uniform surface illuminated by a uniform source. These are conditions that do not necessarily apply (think of the blue sky), but that can usually be checked by independent means.

It turns out that this is a conclusion that generalizes to many other cues. You are “optically sensitive” to planarity. But notice that the spatial attitude of the plane is fully ambiguous. Any attitude will do. There is no way to infer the attitude by optical means.

Basic invariances

Consider a lonely eye, fixed to some “vantage point”. Let the eye be able to monitor all directions. Let there be a scenic layout for it to sample. Now consider two simple transformations, a rigid rotation of the world about the vantage point, and a uniform dilation of contraction of the world about the vantage point.

How do you rotate the world? Very simple, you rotate the eye! Optically this makes no difference whatsoever. If the eye has no fiducial direction, then it will be blind to rotations. If it has, it can undo the rotation with an eye movement. Thus rotations do not yield any exterospecific information.

How do you shrink or dilate the world about the vantage point? Again, no problem, you simply image it to happen! Optically it makes no difference whether you really dilate the world or merely imagine it. Thus such uniform magnifications yield no exterospecific information at all. Lilliput and Brobdignac look the same, it needs a Gulliver to reveal the difference.

Thus you conclude that vision is insensitive to arbitrary rotation-dilations about the vantage point. This basic invariance has important implications for the glass bead game.

Alice doesn’t fit the environment, so to speak. Yet it is impossible to decide whether Alice is too large, or the environment too small. There is no absolute unit of size in geometrical optics.

Pirouette, or whirl, a controlled turn on one leg, starting with one or both legs in plié *and rising onto* pointe. *(Notice that this decreases the moment of inertia, and thus speeds up the rotation.) Turning technique includes spotting, in which a dancer executes a periodic, rapid rotation of the head that fixes the gaze on a single spot, thus nullifying the optical consequence of the rotation. (video clip).*

In order to reflect the most basic structure of optics[12], such rotation-dilations should appear as *congruences*, or *proper movements*, in pictorial space.

This implies that the structure of $\mathbb{P}$ is non-Euclidean. For instance, the curves that are shifted along their own course by movements are not the straight lines of $\mathbb{E}^3$, but (as you may want to prove) planar logarithmic spirals of arbitrary pitch and center at the vantage point.

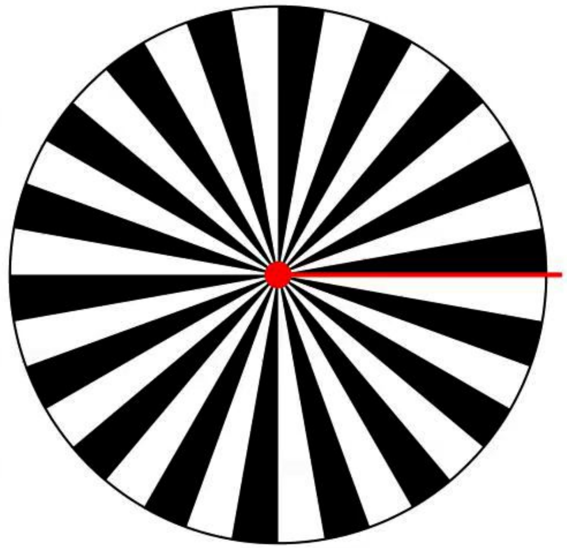
I will not develop this general geometry here. For the present purpose it is more important to consider rather limited fields of view. It is easy enough to show that this implies that the movements and similarities of $\mathbb{P}$ are of two kinds (and combinations of those), namely[13]

- Euclidean movements in the base plane
- "Hildebrand-type transformations" in the depth domains

The former are well understood, so I'll only discuss the latter. Formally, the "Hildebrand-type" transformations are of three types, namely

- shifts along the depth fibers
- dilations of the depth fibers
- "rotations" changing the attitudes of planar sections

The overall shifts (relative shifts are covered by the rotations) are irrelevant since the depth domain has no natural origin anyway. The dilations were clearly described by Hildebrand. They are changes in the "unit depth difference". Since in no way implied by any cue, we may expect these to be idiosyncratic. When comparing results obtained from different observers we should compare *modulo* overall dilations. The "rotations" are of a non-Euclidean kind[14]. They change the apparent frontoparallel plane. Since not implied by any cue, we should expect them to be isiosyncratic. When comparing results obtained from different observers we should compare *modulo* rotations. I will discuss empirical results that bear on this in a later chapter.



A Euclidean planar rotation. It is periodic. (Video clip.)

The fact that dilations and rotations are idiosyncratic has huge consequences for psychophysics. It implies that it may make little sense to compare depth values of different observers, obtained in the same experiment, directly. Comparison should be *modulo* arbitrary proper movements. This sometimes makes the difference between insignificant correlation and almost perfect agreement! It is a fact that is unfortunately rarely appreciated in the current literature.

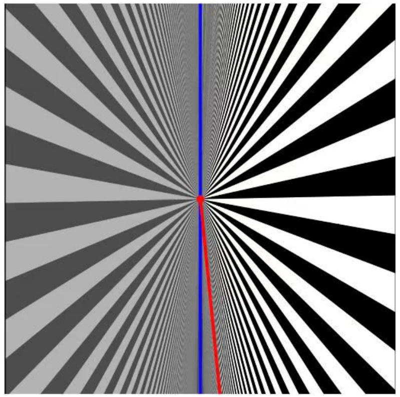
[12] Here "optics" is the optics of Euclid, a kind of information theory rather than a theory of the propagation of radiant power. We need only consider the bundle of "visual rays", that are the half-lines with apex at the vantage point.

[13] Let the Cartesian coordinates of $\mathbb{E}^2$ be denoted $\{x, y\}$, the coordinate of $\mathbb{A}$ be denoted z. Then the transformations can be parameterized as

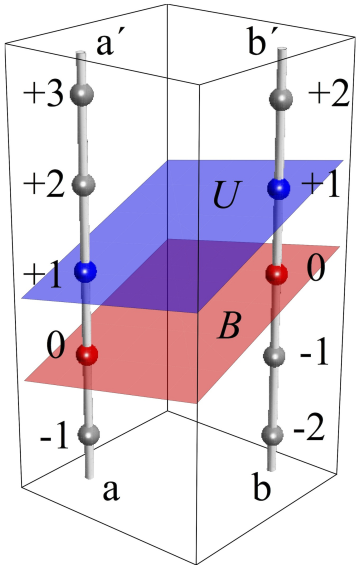
$$\begin{aligned} x' &= x\cos\varphi - y\sin\varphi + t_x \\ y' &= x\sin\varphi + y\cos\varphi + t_y \\ z' &= g_x x + g_y y + sz + r \end{aligned}$$

Here φ parameterizes a Euclidean rotation, $\mathbf{t} = \{t_x, t_y\}$ a Eclidean translation, whereas $\mathbf{g} = \{g_x, g_y\}$ parameterizes a rotation in depth, s a depth scaling, and r a depth shift. In most cases of interest $x' = x$, $y' = y$, and the only interesting components are the Hildebrand transformations given by the third equation. Since the depth orgin is irrelevant anyway, it usually suffices to consider the three-parameter transformations $z' = g_x x + g_y y + sz$.

[14] Technically, these rotations are "parabolic", whereas Euclidean rotations are "elliptic".



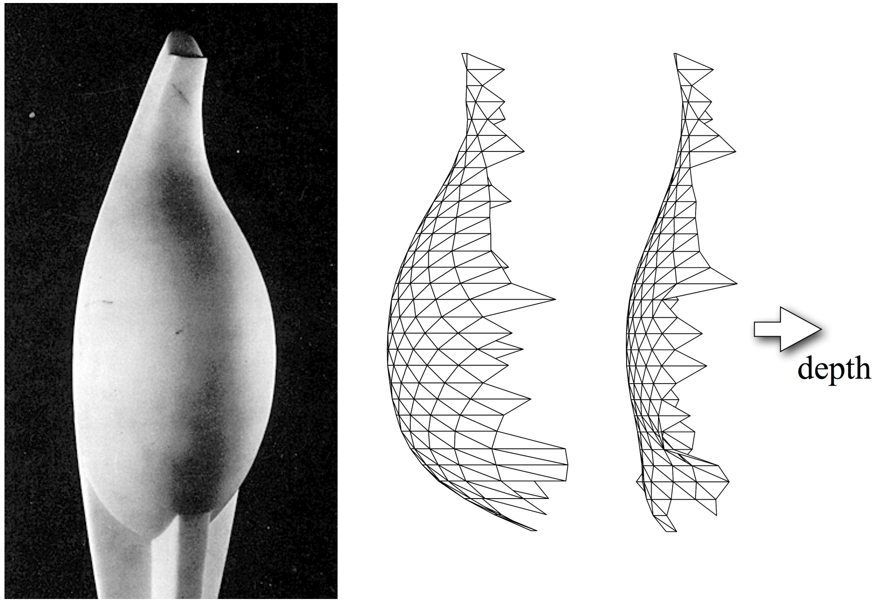
A non-Euclidean planar rotation. It is not periodic (doesn't "turn around"). It takes forever to rotate from minus to plus infinity. The "speeding up" and "slowing down" are illusory though. Formally this is a uniform (constant speed) rotation. (Video clip.)



A gauge. The depth fibers a and b are mutally independent. The cross section B synchronizes them by defining an origin (the "0" of the depth scale) in each. The parallel cross section U likewise synchronizes their units (the "1", etc., of the scales). Thus a pair of mutually paralllel planar cross sections synchronizes all depth fibers globally. Notice that the "gauge" is not specified by any depth cue, it is supplied by the microgenesis of the observer. In experiments one finds that observers make idiosyncratic choices of gauge. Comparison of observers only makes sense modulo *an arbitrary gauge change.*

Gauges and gauge transformations

A "gauge" is a prescription to coordinate the depth fibers. The simplest gauge is geometrically composed of two mutually parallel planar cross sections. The first plane meets the fibers in what we define as the fiducial origin, the second as the fiducial unit point. Given such fiducial origin and unit point any point on the fiber can be assigned a number, its coordinate in the fiducial coordinate system. Since, in an absolute sense, affine lines have neither a natural origin, nor a natural unit, such fiducial coordinate systems are mere conveniences. You may chose them as you see fit. The simplest gauges at least coordinate things between fibers, such that planar sections have simple descriptions in terms of the coordinates. For instance, the planes of the gauge are characterized by coordinate values zero and one *for all fibers* in the bundle[15].



The stimulus is a photograph of one of Constantin Brancusi's "Birds". The two reliefs are by different observers. Notice that the "depth of relief" is quite different, although the reliefs are very similar (roughly they differ by a depth scaling).

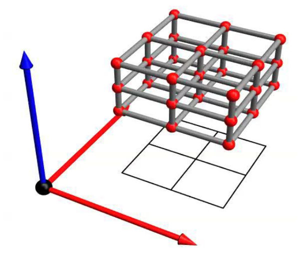
[15]Generic planes are characterized by linear functions of the Cartesian coordinates in their traces.

Some illustrative cases

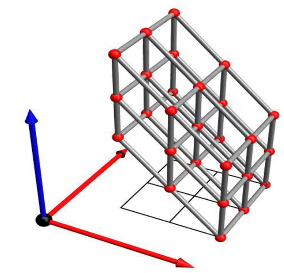
PICTORIAL RELIEF

Pictorial relief can be measured easily and precisely via the gauge figure method. One typically finds the local spatial attitudes of the tangent planes of the pictorial relief in a few hundred points.

Typically observers agree quite well with each other, except for idiosyncratic Hildebrand-type ambiguity transformations. The only reasonable way to compare observers is by factoring out these idiosyncracies. What is left is "pure relief". That people tend to agree on this, is no doubt due to the fact that they recognize the same cues.



A Hildebrand transformation. This is a "similarity of the second kind", which is depth scaling. Here depth is the blue dimension, the picture plane the red dimensions. (Video clip.)

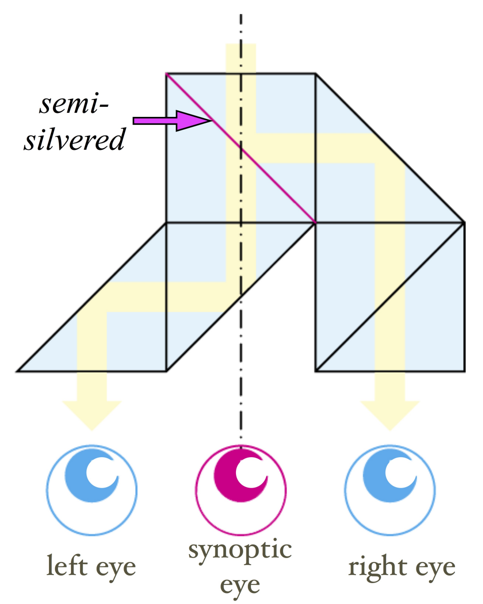


A Hildebrand transformation. This is a "rotation", which looks like depth shear to the Euclidean eye. Here depth is the blue dimension, the picture plane the red dimensions. (Video clip.)

The Hildebrand transformations are either depth dilations or contractions, or non-Euclidean rotations. Such "rotations" affect the apparent frontoparallel planes.

Both transformations can be surprisingly large. The major differences are found when comparing observers, but even a single observer will show this kind of differences when studied at different times.

Large differences are also encountered when you change the viewing conditions for a given observer. For instance, you may let the observer view the picture obliquely, or you may let the observer view the picture with one eye, with two eyes, or "synoptically".

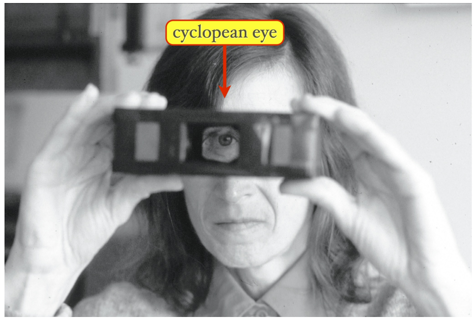


Design for a synopter according to Moritz von Rohr for the Zeiss optical factory at the beginning of the twentieth century. The instrument effectively replaces the two separate eyes by a single (synthetical) "cyclopean eye", thus nullifying binocular disparity.

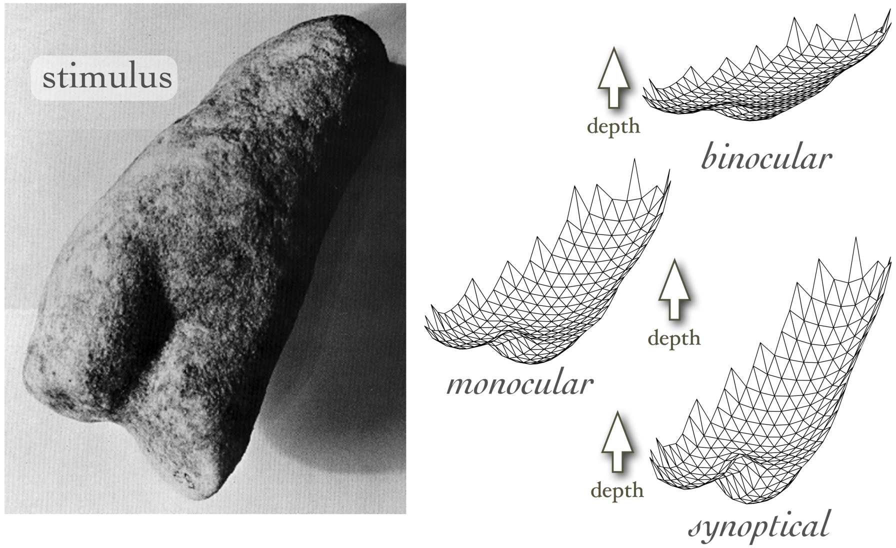
That monocular versus binocular viewing should make a difference has been known for many centuries. Most likely, visual artists were aware of it from the earliest times on.

If you view a scene monocularly, it looks flatter than when you view it binocularly. Artists use this, because the flatter experience makes it much easier to draw the scene.

In contradistinction, when you view a picture monocularly, it looks much more "plastic" than when you view it binocularly. This is apparently due to the fact that binocular vision readily reveals that a picture is actually a flat object.



A synopter constructed from off the shelf optical parts, according to the original von Rohr design. The eye in the photograph is actually a superposition of he left and right eyes. Although the observer uses both eyes, there is no binocular disparity.

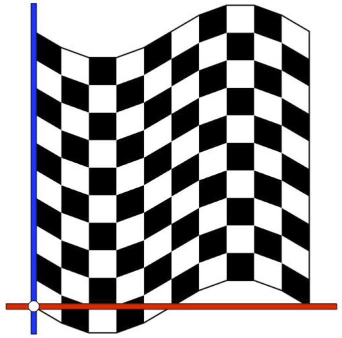


At left the stimulus, a photograph of an eroded Greek statue, found in a stream. At right the reliefs (for a single observer) for monocular viewing, binocular viewing, and synoptical viewing.

Thus you will often see visual artists close one eye. They use this for different reasons when viewing the scene and when viewing their work.

The "synopter" designed by von Rohr supplies the two eyes with identical optical structure. Binocular disparity is simply absent, although the brain expects it. After all, the brain is responsible for closing or opening the eyes! This is (no doubt) the reason for the effect of synoptic viewing.

The empirical results (of course!) agree with what the artists knew all the time. Closing an eye when viewing a picture expands the depth range. Using a synopter expands it even more. For a generic observer the effect is a factor four or even more.



A "lasagna transformation", which is a special transformation that varies smoothly from point to point, but is locally a Hildebrand transformation. This transformation is conformal, it conserves (non-Euclidean) angles in pictorial space. The "blue axis" denotes the depth, the "red axis" the picture plane dimension. (Video clip.)

Isolated locations in deep space

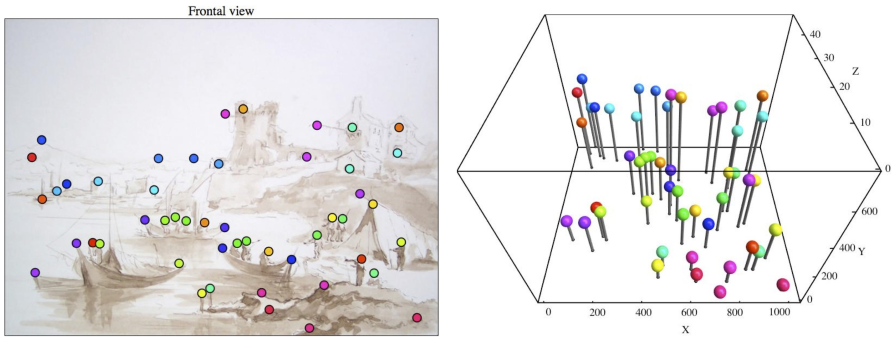
One would suppose that in order to indicate isolated locations one needs a depth value, the location in the picture plane being known already (it identifies the location!). However, a moment's thought reveals that this makes no sense. Depth is only defined *modulo* arbitrary Hildebrand-type transformations. The upshot is that the simples configuration of isolated points that is *unique* modulo Hildebrand-type transformations, consists of four points. Of course, the more points the better!

A configuration of three points is experienced as *planar*. The fourth point can be either *in*, or *in front*, or *behind* the plane defined by the other three. Thus, in comparing observers, we have to compare the invariants with respect to arbitrary Hildebrand-type transformations. This is absolutely *crucial*, yet—perhaps unfortunately—hardly recognized by the contemporary mainstream. They'll need another decade to catch up I guess.

Perhaps the simplest method to find the discriminative power with respect to depth, is to ask observers to estimate which of two points is closer. Notice that this task is actually *indeterminate* when you allow for arbitrary Hildebrand-type transformations! However, it is an empirical "fact" that observers stick to a given gauge (any gauge, there is no way to fix it) during a single session. Granting this "fact", it makes sense to venture the experiment.

We did this task with a number of observers, using about fifty points on a classical landscape drawing. When an observer has compared all pairs you can deduce the overall best depth order (for that observer of course), and find a measure of the discriminative power.

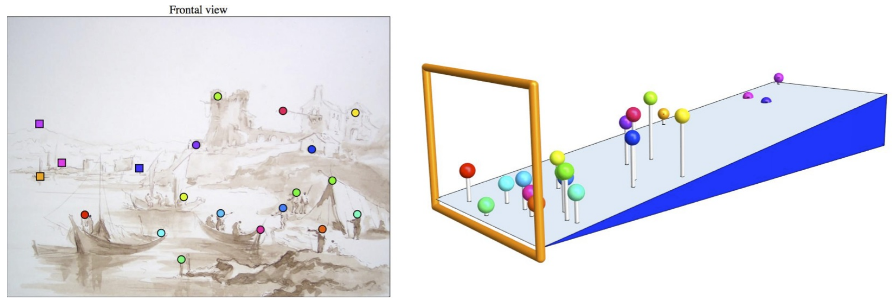
We find that all observers (about a dozen so far) discriminate about fourty depth levels in the drawing. Moreover, they largely agree on their depth orders.



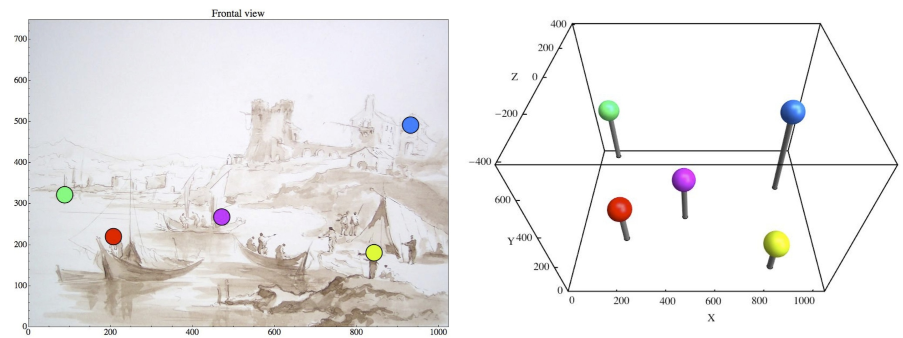
A typical result from pairwise depth comparisons. Single session (of about an hour's worth). Comparing repeated sessions we find that this observer easily discriminates about forty depth layers.

Using a relative size probe it is not possible to use more than about two dozen locations in sessions of about an hour. This is still amply enough to compare with the previous case. We find that the agreement is excellent. The differences between observers are very small when you reckon *modulo* arbitrary Hildebrand transformations. These idiosyncratic transformations are quite appreciable though.

The pointing task is even more time consuming because we always point two-way, say A to B and B to A. This proves to be important, because the to and fro pointings work out differently. It is as if the observers "point by arcs" instead of by straight lines. The results agree very well with those of the other tasks, at least when evaluated *modulo* arbitrary Hildebrand transformations.



A typical result from relative size comparisons. Same observer as for the previous figure. The points are a subset of the points in the previous figure. In the result shown at right I have estimated the "ground plane. The square points indicate the far field. When comparing observers it turns out that observers agree in both near and far fields separately, but that their relative treatment of near and far field is idiosyncratic.

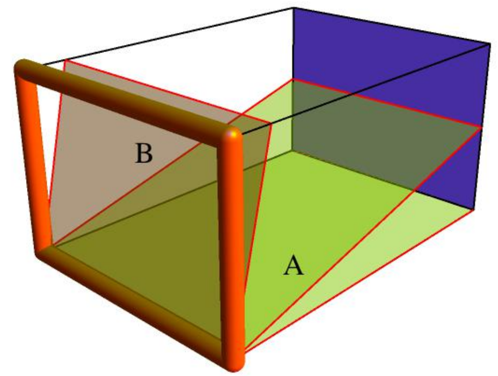


A typical result from pointing. Same observer as in the previous figures. The points are a subset of these in the previous figures.

The depth ranges found with the pointing and relative size tasks can very a lot among observers. This is interesting, because the number of resolved depth levels is almost identical for all observers.

We also find that the ratio between depth ranges from pointing and relative size depend upon the style of the picture. This effect is constant over all observers, irrespective of their relative depth ranges.

So far the laboratory experience with pictorial depth measurements has been very limited. I am convinced that there remains *much* to be learned.



These are the estimated groundplanes for the two extreme observers from a random crowd. Notice the difference: the gounplane B almost coincides with the picture plane, whereas A really runs into depth. However, these observers had identical depth discriminatory power! Depth is a quale.

How different is pictorial vision from generic vision?

The major difference is the effect of multiple viewpoints. Both binocular disparity and motion parallax are of vital importance. The other—also vitally important—aspect is your spatial situational awareness in the environment. As a summary statement one might say that the crucial issue is whether the observer is located in visual space or not. There is a spectrum of possibilities, which ranges from perception in action (say gymnastics) to pure pictorial space.

In daily life the far field is almost invariably much like pictorial space, whereas in near space the enactive element dominates. There is a lot to be done here. Current science is oblivious of many of the key issues.

Other eBooks from The Clootcrans Press:

1. Awareness (2012)
2. MultipleWorlds (2012)
3. ChronoGeometry (2012)
4. Graph Spaces (2012)
5. Pictorial Shape (2012)
6. Shadows of Shape (2012)

(Available for download here.)

About The Clootcrans Press

The Clootcrans Press is a selfpublishing initiative of Jan Koenderink. Notice that the publisher takes no responsibility for the contents, except that he gave it an honest try—as he always does. Since the books are free you should have no reason to complain.

The "Clootcrans" appears on the front page of Simon Stevin's (Brugge, 1548–1620, Den Haag) *De Beghinselen der Weeghconst*, published 1586 at Christoffel Plantijn's Press at Leyden in one volume with *De Weeghdaet, De Beghinselen des Waterwichts*, and a *Anhang*. In 1605 there appeared a suplement *Byvough der Weeghconst* in the *Wisconstige Gedachtenissen*. The text reads "*Wonder en is gheen wonder*". The figure gives an intuitive "eye measure" proof of the parallelogram of forces. The key argument is

> *de cloten sullen uyt haer selven een eeuwich roersel maken, t'welck valsch is.*

Simon Stevin was a Dutch genius, not only a mathematician, but also an engineer with remarkable horse sense. I consider his "clootcrans bewijs" one of the jewels of sixteenth century science. It is "natural philosophy" at its best.