Visually Acquired Information about Rotating Objects

Visueel verkregen informatie over roterende objekten

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de Rector Magnificus, Prof. dr. J. A. van Ginkel
ingevolge het besluit van het College van Decanen
in het openbaar te verdedigen
op maandag 8 januari 1996 des ochtends te 10.30 uur

door

Henricus Antonius Hubertus Cornelis van Veen

geboren op 30 juni 1967, te Reeuwijk

promotor: Prof. Dr. J. J. Koenderink
Faculteit Natuur- en Sterrenkunde
Helmholtz Instituut
Universiteit Utrecht

co-promotor: Dr. A. M. L. Kappers
Faculteit Natuur- en Sterrenkunde
Helmholtz Instituut
Universiteit Utrecht

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Veen, Henricus Antonius Hubertus Cornelis van

Visually acquired information about rotating objects /
Henricus Antonius Hubertus Cornelis van Veen. - Utrecht :
Universiteit Utrecht, Faculteit Natuur- en Sterrenkunde
Thesis Universiteit Utrecht. - With ref. - With summary in Dutch.
ISBN 90-393-1122-6
Subject headings: rotation / visual perception / psychophysics

voor Ilse

CONTENTS

PREFACE

We are usually unaware of the sophistication of our interaction with the surroundings. We navigate our body parts through the physical environment without noticing the unparalleled complexity and extent of the systems that control and co-ordinate our movements. Our sense-organs are continuously collecting vast amounts of information about the perceptible world, including ourselves, but we hardly ever pay attention to the puzzles involved in the interpretation of these endless data streams. Some scientists, however, devote their lives to obtaining an understanding of the workings of the human sensorial and motor processes. As they attempt to unravel the ways humans communicate with the world, they are constantly faced with tantalizing problems. This thesis recounts my personal expedition through this fascinating field of research. I hope that at least some readers will feel stimulated to embark on similar journeys.

Hendrik-Jan van Veen

1

INTRODUCTION AND OVERVIEW

Sensing the visual world is one of our main modes of communication with the environment. At an elementary level of description we humans are equipped with a complicated sense-organ (our eyes) that registers electromagnetic radiation in a certain part of the spectrum. At a much higher level of description our visual sense seems to provide us with an accurate and detailed representation of the objects in the world around us. Human vision scientists strive to relate these extreme notions of the role of vision in human functioning, and they do so by studying all kinds of intermediate levels of description.

Information about the visual world is often classified in cues. A cue is a collection of facts and assumptions pertaining to a certain logically coherent body of evidence. An example of such a cue is the apparent size of something in relation to its distance from the observer; the observed projected size of an object combined with the experience that smaller-looking objects of equal actual size are farther away gives us a clue about the distance between us and this object. Our natural surroundings accommodate many cues like this. Focusing research on one particular cue at a time is a powerful and commonly practised way of tackling the complexity introduced by this richness of our natural surroundings, even though it has some serious drawbacks.

This thesis contains four studies concerned with the processing of the information about the structure and rotation of objects which emerges when those objects move relative to a stationary observer (structure-from-motion cue). The approach will be to show subjects synthetic moving objects (computer graphics), and ask them to respond to certain aspects of the structure and motion of these objects. Use is made of computer generated and displayed objects instead of real physical objects mainly because this makes it possible to explicitly control almost all stimulus parameters. Part of the thesis is devoted to the description of the resulting empirical relationships between stimulus and response. No less exciting, in my opinion, are the sections that attempt to interpret the obtained results in terms of other levels of information processing.

Structure-from-Motion

Not every conceivable relative motion of an object with respect to an observer produces information about the object's structure. It is intuitively clear that we, as observers, need to see different sides of an object to get extra information. Figure 1:1 provides a taxonomy of the different motions, together with a schematic and simplified representation of the corresponding optical flow fields. We will consider the correctness of this picture in a minute, but first we explain the effectiveness of the different motions. Small translations of an object in a direction orthogonal to the direction in which it is seen (panel A) result in uniform flow fields (panel E). It is obvious that this flow field can be described by just two parameters - magnitude and direction of flow - and that these parameters carry information about the translatory movement of the object but not about its structure. Similarly, translations along the direction in which the object is seen (panel B) and rotations about an axis coincident with this direction (panel C) both correspond to simple one-parameter flow fields, which contain information about motion but do not provide information about the object's structure (panel F, pure expansion; panel G, pure rotation). Only when an object rotates about an axis orthogonal to the direction in which it is seen (panel D) does the flow field depend on the structure of the object (panel H). Thus, this type of rotation - loosely called 'rotation in depth' - lies at the root of the structure-from-motion cue. Hence, one of the chief properties common to all stimuli used in the following experiments is that the objects always 'rotate in depth'. Some of the mathematical intricacies involved with the actual recovery of structure from general flow fields are discussed in chapter 3.

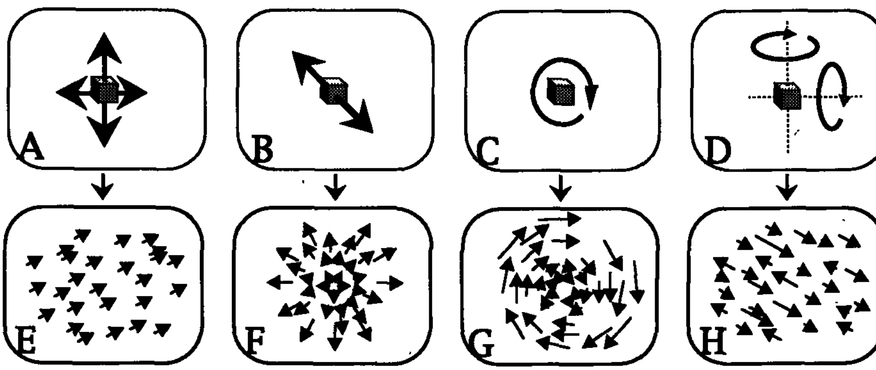


Figure 1.1
Taxonomy of the different types of object motion (top row) and the resulting optic flow fields (bottom row). The flow fields in the bottom panels correspond to arbitrarily chosen examples of the corresponding object motions.

I want to clarify two issues here. Firstly, an object that objectively translates through physical space rotates (and translates) with respect to the direction in which it is seen. Therefore, its flow field is a combination of the fields depicted in panels E, F and H of figure 1.1. This specific incarnation of the structure-from-motion cue is usually called motion parallax. The second issue relates to a simplification that I made in producing the flow fields depicted in figure 1.1. Normally, i.e. in the world outside the laboratory, the structure-from-motion cue is accompanied by a perspective cue. Although perspective deformations do contain information about the structure and motion of a moving object (as such they form a cue), this information is very sensitive to noise and is known to play a role in human vision only for larger fields of view (nearby objects). I have deliberately chosen not to include the perspective cue in the stimuli, the main reason being the substantial simplification of the mathematical model describing the information content of the stimulus. Perspective information would complicate the flow fields depicted in figure 1.1. I refer to chapter 3 for some discussion on this point.

In this thesis I seek to investigate essentially two main issues: to what extent and with what accuracy are humans able to pick up information about the structure and rotation of moving objects, and how does the structure-from-motion cue mediate between these moving objects and the representation of these objects maintained by human observers. Below I briefly indicate how these issues are addressed in the respective chapters.

Structure

Chapter Two comprises a study of the perception of the volume of rotating objects. Volume is one of the most important global structural aspects of an object, but it has not received much attention in the past. The stimulus in this experiment consists of two objects of variable shape and size revolving around a common vertical axis. The subject's task is to indicate which of the two objects has the larger volume. One of the most important findings of this experiment is that the shape of objects influences the accuracy with which subjects are able to compare the volume of objects. Discrimination thresholds for objects without a specific elongation, like cubes, are lower than for objects that do possess a distinct elongation, like rods. Also, the discrimination thresholds for regularly shaped objects like straight boxes are much lower than those for objects that have a more irregular shape (confer the objects depicted in figure 2.1 in chapter 2). The latter result indicates that in this experiment cues other than structure-from-motion play a role as well. Since we ensured that the information provided by the structure-from-motion cue was about equal for the random and regular objects, subjects must have been able to make use of the specific regularities of some of these objects. This knowledge is used later, in chapters 4 and 5, when similar objects are used for other purposes.

Structure and Motion

In the flow field depicted in panel H of figure 1.1 information about the structure and the rotation of the supposititious object is confounded. **Chapter Three** of this thesis is dedicated to the perceptual confusion that arises when these structural and rotational components are only partially disentangled. The stimulus that we show to the subjects consists of two slanted planes oscillating about vertical axes. The subject's task is twofold: match the orientation of one plane to the orientation of the other plane (orientation relative to the axis of rotation) and simultaneously match the magnitudes of rotation. The results show that in certain conditions, depending on the specific orientation and rotation of the planes, subjects inadvertently compensate for errors made in one subtask by making an error in the other subtask. Such a pattern of correlated errors is a sign of a lack of information. A mathematical analysis of the information available in the stimulus reveals that the specific pattern of correlations found can be understood in terms of the low-level motion sensitivities of the human visual system combined with the mathematical nature of the structure-from-motion cue. The less than perfect processing of motion information (the flow field) by the visual system effectively renders the differences within certain families of rotating planes invisible. This conclusion provides a tool for interpreting human performance in many other tasks.

Motion

The last two chapters of this thesis address the processing of visual information about the rotation of objects in space. In general, rotations can be described by the angular speed of rotation and the orientation of the axis of rotation. In **Chapter Four** I consider the perceived angular speed of rotation, while **Chapter Five** is concerned mainly with the orientation of the axis of rotation. Although the rotation of an object about an axis orthogonal to the direction in which the object is seen appears to be crucial for the availability of information about its structure, not much seems to be known about the perception of these rotations. I have therefore conducted a study that investigates how certain spatial aspects of the stimulus influence the perceived angular speed of 3D objects rotating about such axes. The subjects watch two rotating objects (the same random objects that were used in chapter 2) and are required to indicate which of the two objects rotates faster. Among other things, it is found that subjects underestimate the speed of rotation around a vertical axis relative to the speed of rotation around a horizontal or diagonal axis, that they overestimate the speed of rotation of larger objects relative to the speed of rotation of smaller objects, and that they underestimate the speed of rotation of peripherally viewed objects relative to the speed of rotation of foveally viewed objects. Some of these effects can be understood in terms of similar phenomena found in the perception of two-dimensional linear speed and size. Apparently, the perceived angular speed of rotation is subject to many biasing factors. While this result is interesting in itself, given the intimate relationship between structure and motion it also imposes some constraints on the perception of structure (e.g. rigidity).

In chapter 5 I use a somewhat different approach to investigate the extent to which visual information about the rotary movement of objects is processed: subjects are asked to mimic the rotary motion of an object displayed on a computer-monitor by dynamically changing the orientation of a real object that they hold in their hand. This method has several major advantages over the usual methods, one of them being the high dimensionality of the response space; subjects are encouraged and enabled to respond to every change in the observed orientation of the displayed object. The major experimental findings relate to the orientation of the axis of rotation. From the recorded motion of the real object one can compute the average measured axis of rotation. The variability in the orientation of this axis is lowest for elongated objects rotating around their main axis and for rotations about fronto-parallel rotation axes ('rotation in depth'), and is always higher than the variability of the response system itself (the manipulating hand). The pros and cons of this paradigm are discussed at length at several places in chapter 5, but at least I can conclude that because subjects exhibit a high grade of reproducibility and a strong stimulus-response relationship in the experiments the method is an adequate one for investigating visual information processing.

2

VOLUME OF MOTION-DEFINED SOLIDS

Abstract

We investigated the ability of human observers to discriminate an important global 3D structural property, namely volume, of motion-defined objects. We used convex transparent wire-frame objects consisting of about 12 planar triangular facets. Two objects, vertically separated by 7°, were shown simultaneously on a computer display. Both revolved at 67°/sec around a common vertical axis through their centres of mass. Observers watched the objects monocularly for an average of three full rotations before they responded. We measured volume discrimination as a function of absolute volume (3 to 48 cm^3; 1 metre viewing distance) and shape (cubes, rods and slabs of different regularity). We found that: (a) volume discrimination performance can be described by a Weber law, (b) Weber fractions depend strongly on the particular combination of shapes used: regular shapes, especially cubes, are easiest to compare, and similar shapes are easier to compare than different shapes, and (c) humans use a representation of volume which is more veridical and stable in the sense of repeatability than a strategy based on the average visible (2D) area would yield.

Introduction

The volume of a solid object is one of its most fundamental structural features. It determines the space the object occupies and can give some indication of its mass. We use our visual sense continuously to make judgements of volume (and mass): which piece of meat is larger, do I have as much lemonade as she has, can I lift this object, will this pile of clothes fit into that suit-case, etc? How do we go about making these judgements? It is highly probable that we utilise all sorts of experience and prior knowledge. This will not be very helpful, however, when we are dealing with unfamiliar objects. Observing the objects from different viewpoints will help us to gain a better 3D impression. This study reports on the human ability to discriminate the volumes of various objects, when these objects are rotating in front of the observer. The rotation lets the observer see all sides of the objects and make use of structure-from-motion 'computations' to resolve the 3D structure of the objects. We varied the size and shape of the objects and measured the veridicality and reproducibility of volume comparisons.

There has been some previous research in this area, but it has mainly involved the use of stationary, real objects. Lauer (1929) asked subjects to compare spheres, cubes, prisms, spherical sectors and spherical segments and found that shape does effect estimates of volume (e.g. the volumes of cubes were overestimated relative to spheres). He also concluded that disparity of dimensions within an object was critical, because the volumes of elongated prisms were overestimated relative to cube-like prisms, and the volumes of flat spherical segments were underestimated relative to higher segments.

Brunswik (1934, 1956) mentions an unpublished study by Stevenson on a similar theme. In this study prisms of various elongations were compared with cubes of various sizes. Subjects were given a prism and then asked to select from a set of cubes the one with the same volume as the prism. The volumes of long prisms (sticks) were slightly overestimated relative to cubes, whereas the volumes of thin prisms (plates) were underestimated relative to cubes. Brunswik also notes that cube-cube comparisons show much less variability (low threshold) than comparisons of different shapes. The shape dependency of human volume estimation also has some practical applications. In a book called 'Packages That Sell', Franken and Larrabee (1928) discuss some very interesting packaging issues, including the influence of the shape (and colour) of a package on its estimated size. For the specific case that they investigated, they found that customers estimated a flat round can to be larger than a taller one, although the contents of the cans were equal. The manufacturers thereupon decided to sell their 'Cod Fish Cakes' in the flat can. In 1941 Piaget published his famous experiments on the conservation of continuous quantities. He describes how young children seem to base their judgements of the amount of orangeade in a glass on the height of the liquid column, ignoring the width of the column and thus the actual volume of liquid. In the 1960s, the perception of volume was investigated with paradigms comprising ratio and magnitude estimation tasks. The data from these experiments are usually analysed in terms of power functions: judged volume is assumed to be proportional to a certain power of physical volume. Baird (1970) gives a summary of these investigations, concluding that in most cases the exponent of the power function for volume judgements appears to be smaller than 1, typically 0.7, which is smaller than the values found for area and length judgements. He does not give any information about the response variance. Stanek (1968, 1969) published two short papers on surface and volume judgements. He asked subjects to match the volume of prisms and cubes, and found both underestimations and overestimations of prism volume, depending on the specific elongation and size of the prism. Some connections between estimates of volume and mass have also been investigated. The existence of the so-called size-weight illusion has been known for more than a century now (Charpentier, 1891): the perceived weight of an object depends not only on its physical weight, but also on its size. Recently, Ellis and Lederman (1993) published an article about the role of haptic versus visual volume cues in the size-weight illusion.

There are two main differences between our work and previously published studies. Firstly, we investigate volume judgements of *synthetic* (computer graphics) objects which are *rotating* relative to the observer. This approach features structure-from-motion as the main depth cue for estimating volume, while eliminating other (uncontrolled) depth cues that are usually available when real objects are used. Previous research has typically employed fairly natural viewing conditions in which the relative motion of object and observer usually played an inferior role. Secondly, we measure the *variance of responses*, or discrimination thresholds. We have been unable to find any systematic study on volume discrimination thresholds in the existing literature. Our study resembles previous work in that it tries to establish the influence of shape on the judgements of volume, albeit that we vary different aspects of shape.

The paper starts with a description of two psychophysical experiments we did with human subjects. We measured veridicality and variance of relative volume judgements as a function of overall scale in the first experiment, and as a function of shape in the second experiment. Then we consider the possible use of projected area in estimating volume. We present the results of an algorithm that bases its volume estimates on the correlation between projected area and volume, and discuss the implications for the interpretation of the psychophysical data. Finally, in the general discussion, we compare our results with those reported in the literature, and discuss what we have learned about the way in which our visual system represents objects like the ones we use.

Methods

Experimental Set Up

The stimuli were generated on an Apple Macintosh IIfx computer and displayed on a 71 Hz Radius TPD/19 high spatial resolution (82 dpi) grey-scale monitor (1152 by 882 pixels). The subjects were seated in front of the screen at a distance of 1 metre. Head motions were not explicitly restricted, but the subjects were required to keep their right eye just in front of the centre of the screen. The left eye was covered with a black patch. The room was dimly lit such that the subjects were still able to see the monitor and the table on which it stood.

Stimulus

Temporal and spatial parameters

Two objects, vertically separated by 7° (centre-centre), were shown simultaneously on a computer display. Object dimensions were controlled parameters of the experiment, and resulted in projected sizes of between roughly 0.2° and 8° visual angle; the projected sizes depended on shape, volume and instantaneous orientation of the object with respect to the observer. Both objects revolved at a fixed rate of 67°/sec about a common vertical axis through their centres of mass. Every 75 msec a new view (parallel projection) of the objects was shown on the screen. The objects rotated back and forth in depth with an amplitude of 180°. Thus, information about the period of rotation was available to the subjects. Subjects had to watch each display for at least one full rotation (360° in 5.4 seconds). There was no upper limit to the presentation time. We did not prescribe fixation.

It should be noted here that, because we used orthographic projections of wire-frame objects, the direction of rotation was mathematically ambiguous at every instant (see. Ullman, 1979). This effect was confirmed by the subjects; they experienced spontaneous reversals as well as opposite rotation directions for the two objects.

Furthermore, the orthographic projection itself is not sufficient to define the position in depth of the objects. However, the subjects reported that all objects appeared to be located at the computer display, i.e. at 1 metre distance (note that the subjects were able to see the monitor boundary), which is of course in accordance with accommodation information. Apparently, since all objects were seen at the same position in depth, apparent size differences were not interpreted as being due to different positions in depth but were probably interpreted as real size differences.

General object description

We used convex transparent wire-frame objects consisting of about 12 planar triangular facets. They were presented in high contrast as light line segments on a dark background; the lines had a diameter of 1 pixel, and anti-aliasing was not used. The objects were rendered as wire-frames without hidden line removal. This method of rendering can facilitate the perception of the spatial structure, because it allows subjects to see all parts of the objects at the same time. During pilot experiments, subjects sometimes reported they had difficulty in memorising the structure of the hidden part of opaque objects. This effect was strongly coupled to the angular velocity, apparently because slower rotations require comparisons of parts more separated in time. Although the definition of volume seems to be more intuitive in case of opaque objects than in the case of transparent ones, we did not want to lose the advantage of seeing the complete structure at any instant in time. (Todd, Akerstrom, Reichel and Hayes (1988) investigated the difference in the perceived rigidity of opaque and of transparent rotating cylinders. They found minor differences for the specific frame rate that we used. Braunstein and Andersen (1984) also reported only small differences in the veridicality of perception of shape of either opaque or transparent rotating spheres, veridicality being slightly better in the case of transparent objects.) Similar considerations regarding the specific choice of object motion led us to the conclusion that a rotation around a fronto-parallel rotation axis was the most suitable one. Any other rotation, or a translation combined with perspective instead of orthographic projections, would always hide certain parts of the object from direct view. We used a vertical rotation axis, but any other fronto-parallel axis would do.

Object generation, classification and manipulation

For every trial two new objects were created as follows: a set of vertices was generated by randomly picking 10 positions within a unit cube. Next, the convex hull[1] of this set of points was computed. Subsequently, points inside the hull were removed: a set of 10 points usually reduced to about 8 vertices on the hull. This enclosing surface defines our objects, which at this stage have many uncontrolled properties.

[1] We generate and define the convex hull of a set of points in space in two steps. First, we take the planes defined by all possible sets of three different points ($n\cdot(n-1)\cdot(n-2)/3!$ planes for n points). Second, we select those planes that have all points on one side of the plane (points on the plane are ignored). When there are at least four non-coplanar points, these selected planes enclose a volume (intersection of half-spaces), and the enclosing surface is called the convex hull. In our case, this surface always consists of a collection of connected triangular facets. If there are only a few points almost all of them are on the hull, while only a few are inside. The number of facets is always equal to twice the number of vertices minus four.

In our experiments we controlled a few structural parameters of these objects: their volume, centre of mass and elongation and flatness. Elongation and flatness are defined using the global length-width-height ratio (LWH ratio) of an object. Our definition of the LWH ratio for a general object can be found in the APPENDIX. Using this ratio we define three classes of objects: cube-like, *rod*-like and *slab*-like objects, see figure 2.1. A cube has an LWH ratio of about 1:1:1. A rod (elongated or prolate object) has an LWH ratio of about 4:1:1. Finally, a slab (flattened or oblate object) has an LWH ratio of about 4:4:1. The factor 4 was used in all experiments. The preposition 'about' is used to indicate that for an object to be in a certain class the LWH ratio need not be exactly equal to a certain ratio, but may differ by as much as a factor of 2 (which equals the square root of the aforementioned factor 4). Thus, an object with LWH ratio 2:1:1 is considered to be halfway between a cube and a rod, whereas 2:2:1 is at the boundary of cube and slab and 4:2:1 is at the boundary of rod and slab. Clearly, there are many LWH ratios that do not fit into any of the three classes that we defined. For our purposes that does not present a problem. We need only a few well-defined and reasonably different shape categories. (A similar classification scheme, based on the notion of 'extendedness', can be found in Willats, 1985 (also see Willats, 1992). Roughly speaking, cubes, rods and slabs correspond to his 3_{111}, 3_{100} and 3_{110} regions respectively. We prefer the mathematically well-defined moment-based description, as given in the APPENDIX.)

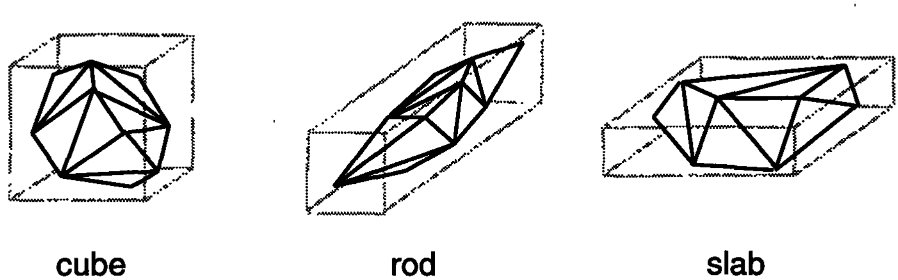


Figure 2.1
The three drawings depict objects with length-width-height ratios of 1:1:1 (cube), 4:1:1 (rod) and 4:4:1 (slab), from left to right respectively. Both the regular variety (outside box) as well as the random variety (inside shape) are shown. The objects have different volumes. (To make the drawing clearer, some line elements are removed. In particular, note that the random shapes are presented as opaque instead of transparent, and that the diagonals of the regular shapes are not drawn.)

To prevent subjects from making use of knowledge about fixed LWH ratios, we scattered all three dimensions by multiplying them by factors uniformly distributed between 1 and 2. In this way all ratios become distorted but the classification cube-, rod- or slab-like is still valid. (A 1:1:1 object (cube-like) can, in an extreme case, become

2:1:1 (cube-rod boundary) or 2:2:1 (cube-slab boundary), but it never becomes truly rod- or slab-like.)

There is one object parameter that we have not yet discussed; we call this the regularity of the objects. In addition to the randomly shaped objects discussed above, in the second experiment we also used a set of rectangular (regular) boxes. These boxes are again described by their volume, centre of mass and cube-rod-slab classification. They consist of exactly 8 corners (vertices) and are triangulated just like the random objects, i.e. each of the six sides is divided into two triangles. This limits the differences between random and regular objects to the regularity alone.

After the object parameters have been defined, it is easy to control them. Elongation and flatness can be manipulated by anisotropic scaling along the principal axes of the object; see APPENDIX. The position of the centre of mass with respect to the observer is easily changed by translating the object in space. The volume can be adjusted by applying an overall scaling of the object. Under these transformations the property of convexity is invariant. Special care was taken to ensure that the object orientation was totally random; in other words each possible orientation had an equal probability of occurrence (the orientation of the objects can be defined using their principal axes). This randomisation of orientation is very important, for two reasons. Firstly, we want to avoid conditions in which both objects have the same orientation, especially when they are both rods or both slabs. Having roughly the same shape and orientation, the volume difference translates into a difference in *projected* size, which makes the task much easier. Secondly, in general the randomisation ensures that the faces of the regular objects (the boxes) are not aligned with the screen or with the (vertical) rotation axis. A situation in which they are aligned is clearly not preferable.

Procedure

The subjects were instructed to discriminate the volume of two simultaneously rotating objects. They watched the objects for an average of about three full rotations, before indicating which of the two had the larger volume by moving a trackball either up or down. No feedback was given.

Each stimulus consisted of a test object and a reference object. Their order on the screen (upper or lower position) was randomly chosen for each trial. Each experimental condition was defined by the volume of the reference object, the regularity of both objects (random or regular), and the cube-rod-slab classification of the reference object and test object separately.

Adaptive Stimulus Control and the Psychometric Curve

For each condition we measured, as a function of the ratio of volumes $\ln(V_{test}/V_{ref})$, the percentage of trials in which the subject indicated that the test object had the larger volume. If we make the assumption that subjects will behave similarly for different absolute volumes (scale-invariance), then we have to use a logarithmic scale. (The use of this log is based on a result from decision theory (known as Jeffrey's rule); see Jaynes, 1968. In fact, the assumption was confirmed by the experimental data, as we shall point out later.) An adaptive psychometric procedure (Werkhoven and Snippe, in press) was used to estimate the two parameters that characterise our psychometric curve (cumulative normal distribution): the point of subjective equality (PSE) μ, which is the physical log of the volume ratio at which the two objects appear equal in size (the PSE represents a certain bias in the subjective volume of two different objects); and the threshold σ, which is the difference in the log of the physical volume ratio relative to the PSE needed to successfully discriminate the two volumes (84% correct). After each trial, new maximum likelihood estimates of the PSE and threshold were calculated. The next trial was placed either at $\mu+\sigma$ or at $\mu-\sigma$, with equal probability. We used 100 trials per condition per subject. The adaptive procedure usually converged within 50 to 70 trials. Only one adaptive procedure out of 69 did not converge sufficiently and was discarded. For each psychometric curve, we computed statistical minimum estimates of the measurement errors in both μ and σ. To allow for an easier interpretation of the results, we do not present the logarithmic μ and σ but we translate their values into *subjectively equal volume ratios* and *volume increment thresholds* respectively.

Subjects

Four subjects participated in our experiments: HV, SP, MH and IH. The subjects SP, MH and IH were emmetropes. The author HV was myopic but spectacle corrections were used throughout. All subjects had had previous experience with structure-from-motion tasks, but except for the author HV, they were only slightly acquainted with the topic of research.

Experiment 1

Volume discrimination as a function of overall scale

Experiment 1 was designed to get a general impression of volume discrimination performance. We varied the absolute size of the objects and used a number of different shapes.

Design

We measured the volume discrimination performance of three subjects (SP, MH and IH) as a function of reference volume. We defined 15 experimental conditions. Five different values for V_{ref} were used: 3, 6, 12, 24 and 48 cm^3. (To prevent subjects from using simple cues based on absolute size, in each trial we chose the actual value of V_{ref}

from a uniform distribution of plus and minus 30% around one of these values.) Three different shape conditions were used: the reference object was either cube-like, rod-like or slab-like, whereas the test object was always cube-like. We used random shapes only. All 15 conditions were mixed together, resulting in a total set of 1500 trials per subject. Subjects completed the experiment in about three two-hour sessions.

Results

The upper panels of figure 2.2 show the increment threshold values for volume discrimination for three subjects. The first observation is that the thresholds are independent of the value of the reference volume, except possibly for subject IH in the cube vs. slab condition. Thus a Weber law for volume discrimination holds within this range of volumes. The average Weber fractions vary from 15% (SP; cube vs. cube) to 80% (IH; cube vs. slab), depending both on subject and shape-condition. A result common to all subjects is that the cube vs. cube condition results in significantly lower Weber fractions (21±2%; averaged over subjects and reference volumes) than the other two shape conditions (cube vs. rod 45±3%; cube vs. slab 55±6%).

The subjectively equal volume ratios are presented in the lower panels. It is clear that there are many inter-subject variations. For subject SP the ratios are almost one, whereas for IH and MH the ratio of the cube vs. slab condition is about two thirds, indicating a relative overestimation of 50% of cube volume relative to slab volume! For the other shape conditions there is no consistent deviation from veridicality, although subject MH does estimate rods to be slightly larger than cubes (ratio 1.13±0.03; averaged over reference volumes). We can summarise the data on subjectively equal volume ratios by stating that a significant deviation from unity (the veridical value) can be found for two out of three subjects, but only in the slab vs. cube condition: in that case the volume of the slab seems to be underestimated relative to the volume of the cube.

Discussion

Firstly, volume increment thresholds of up to 80% are found, and even for the cube vs. cube condition the average increment threshold is about 20%. In view of our everyday experience with estimating weights (of ingredients for our meals for instance) one may find these values surprisingly high. On the other hand it is a well-known psychophysical result that discrimination of 3D lengths also reveals high thresholds, typical values are about 25% (Norman, Todd, Perotti & Title, in press, used stimuli similar to ours to estimate 3D length discrimination under varying experimental conditions. Werkhoven & van Veen, 1995, found similar thresholds for relief discrimination, which is equivalent to an affine 3D length task.). Since we can combine three length estimates to get a volume estimate, high thresholds for 3D length discrimination could have some predictive value for volume discrimination, although the numerous possibilities of combining lengths into volume impede actual quantitative predictions[2]. Nevertheless, the high thresholds may also result from subjects having conceptual difficulties in dealing with the random shapes. It might be a good idea to use objects that are easier to recognise. We will try this in Experiment 2.

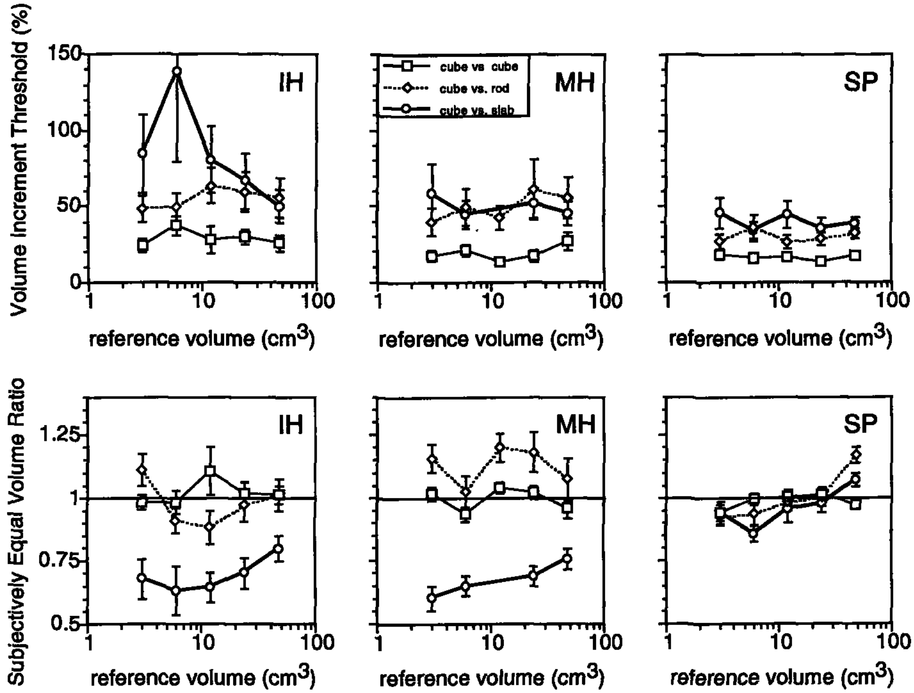


Figure 2.2
This figure shows the results of experiment 1. The upper row of panels shows, from left to right, the volume increment threshold percentages for subjects IH, MH and SP respectively. The lower row shows the corresponding subjectively equal volume ratios. The horizontal axis indicates the volume of the reference object on a log scale. The test object was always a cube-like random object, whereas the reference object was either a cube-like (square symbol), rod-like (diamonds) or slab-like (circles) random object. Error bars represent minimum statistical estimates of the measurement error.

---

[2] A full metric representation of the objects is not necessary for comparing their volumes. A collective affine representation would suffice, since the ratio of volumes is a relative affine invariant of weight 1 (See Gurevich, 1964. Also see Koenderink and van Doorn, 1991). Todd and Norman (1991) tried to find out whether the representation of visual space that humans use in structure-from-motion experiments is closer to an affine representation or to a metric one. Their conclusion is in favour of an affine representation, partly because they *found the lowest thresholds with tasks requiring only an affine representation instead of a* full metric representation. The high thresholds that we find seem to contradict this conclusion. To the very least, it shows the difficulty of a quantitative comparison of thresholds obtained with affine and metric tasks. Recently, van Veen and Werkhoven (in press) showed that within one experiment, thus comprising one task and one type of stimulus, responses reminiscent of an affine as well as of a metric representation could be found, depending on specific stimulus parameters.

Secondly, by inspecting the values of the subjectively equal volume ratios we see that very large deviations from veridicality do occur. These deviations are subject dependent. If we extend the set of shape combinations, we might be able to show consistency of the ratios within each subject. This issue will also be addressed in Experiment 2.

Our assumption that performance would be independent of reference volume is clearly proven by the data: the Weber fractions are independent of the absolute volume. Supplementary evidence for this was supplied by a pilot experiment in which we only used symmetrical conditions, i.e. test and reference object were both cube-like, both rod-like or both slab-like. With such a set-up, asymmetries of the underlying psychometric curve show up as differences between objectively and subjectively equal volume ratios (bias). Such an asymmetry indicates that Weber's law has been violated. When we analysed the data (2065 trials) in terms of volume ratios instead of the logarithm of those ratios, a small but significant (three sigma) deviation appeared; this was absent when we used the logarithmic scale. Apparently the psychometric curve is best described (it is symmetric) by a cumulative normal distribution of the logarithm of the volume ratio, not of the ratio itself.

Experiment 2

Shape dependence

Experiment 2 was designed so that we could investigate more closely the differences in subjects' performance with differently shaped objects.

Design

The reference volume was fixed at 24 cm^3 (and scattered as before). We measured the volume discrimination performance of two subjects (HV and SP) as a function of two shape parameters: the amount of regularity of the objects and the cube-rod-slab classification of each object. We defined 12 experimental conditions. Two different levels of regularity were used: either random shapes (like in Experiment 1) or regular shapes were shown. Furthermore, each object was either cube-like, rod-like or slab-like, resulting in cube-cube, rod-rod, slab-slab, cube-rod, cube-slab and rod-slab combinations for test and reference shape. All 12 conditions were mixed together, resulting in a total set of 1200 trials per subject. It took subjects about three sessions to complete the experiment.

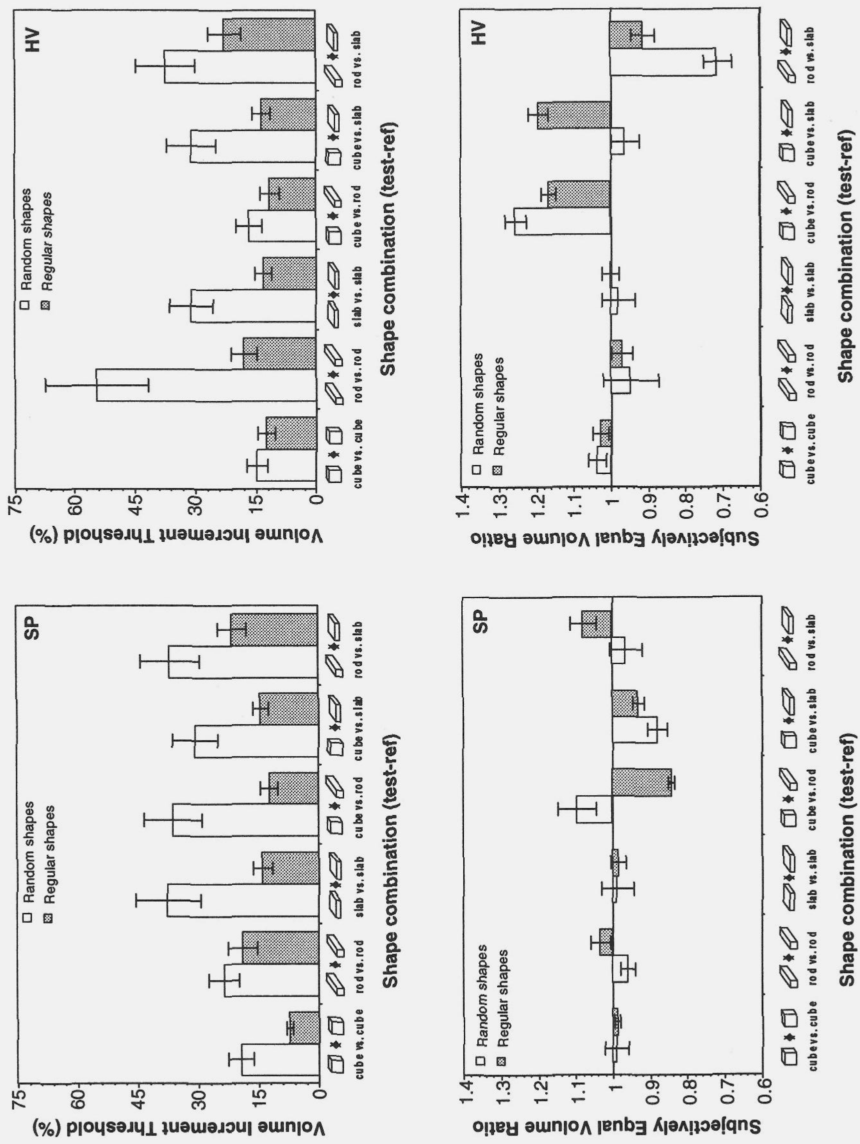


Figure 2.3
The figure presents the results of experiment 2. The upper row of panels shows the volume increment threshold percentages for subjects SP (left) and HV (right). The lower row shows the corresponding subjectively equal volume ratios. The data are presented as a function of the shape condition. The various cube-rod-slab combinations are specified along the horizontal axis. The two bars for each shape combination correspond to random and regular shapes (see legend). Error bars represent minimum statistical estimates of the measurement error.

Results

As we can see in figure 2.3, the most important result of this experiment is that the performance of the subjects improves dramatically when they were tested with regular shapes instead of random shapes. In almost all conditions and for both subjects the Weber fractions are roughly halved: the average ratio of increment thresholds (regular over random) is 0.49±0.07 for subject SP, and 0.56±0.08 for HV. With the regular shapes the lowest thresholds are about 10%! Because we measured each condition only once for both subjects, we do not have enough data to test the significance of the small differences between the thresholds for different cube-rod-slab combinations.

The subjectively equal volume ratios show much variation. Within a subject the ratios for random and regular shapes are not the same. These within-subject differences seem to be almost as large as those between subjects. No systematic difference between the ratios of regular and random shapes could be found. Note that the subjectively equal volume ratios of the symmetric conditions ought to be unity on average; therefore, inspection of the measured ratios for these conditions gives us an impression of the accuracy of the data and the validity of the estimated measurement errors: of the 12 ratios (2 subjects*3 symmetric conditions*2 levels of regularity), 6 deviate from unity by less than one measurement error and only 1 differs from unity by more than two measurement errors; in all this indicates a very slight underestimation of the measurement error.

In the discussion of Experiment 1 we mentioned the possibility of performing a consistency check for the ratios within a subject. To measure consistency we define a closure-relation for the subjectively equal volume ratios R:

$$C = R_{cube-rod} \bullet R_{rod-slab} \bullet R_{slab-cube}$$

This consistency definition is based on the assumption that the ratio of volumes and not of volume itself is the basic entity describing human volume discrimination. This view is supported by the results of Experiment 1: Weber fractions and subjectively equal volume ratios are almost independent of absolute volume over at least one decade. When the ratios of the three mixed shape conditions are consistent within a subject, the value of C should be unity. We found 0.93±0.07 (random), 0.89±0.04 (regular) for subject HV, and 1.20±0.09 (random), 0.97±0.04 (regular) for subject SP. These values are not close enough to unity to demonstrate consistency of shape-bounded overestimations and underestimations of volume within each subject.

Discussion

What could be the reason for the differences in subjects' performance in the regular and the random shape conditions? The complexity of the projected scenes is equal, because we use equal numbers of line elements for both types of objects. The main differences are in 3D: the regular objects are more symmetrical than the random ones. It cannot be deduced from the current set of experiments whether the improved performance is due to the subjects making good use of the parallelism of facets, the equal length of parallel edges or the right angles while 'computing' volume, or to the easier recognition and remembrance of more symmetric objects in general. One would need to test objects with characteristics in the range between those of the regular and the random ones. An alternative explanation based on area estimates (2D-cue) and the variation of shapes within a shape category will be explored in the next section.

Simulations

Area-based volume judgements

One aspect of our experiments that we have not yet discussed is whether area cues may perhaps be used in estimating volume. The argument is that projected area is related to surface area, and the latter is related to volume. An observer might find it easier to estimate projected area than estimating volume directly. In this section we make an attempt to establish the usefulness of this cue. We test an algorithm that uses the average projected area during a single rotation about a fixed axis as the direct correlate of the volume of the corresponding object.

The relation between area and volume

The objects that we used in the previous experiments are all convex and therefore their 3D surface area is finite. For these types of objects there is a theorem (see van de Hulst (1964), p.110-111) which states that the average projected area (2D) of such an object (averaged over all possible orientations of the object in 3D space; *not* an average during one rotation!) is one quarter of the 3D surface area. If, in addition, we appreciate the fact that surface area is related to volume, a link between projected area and volume emerges. The relation between surface area and volume depends on the shape of the object. To summarise so far, we can write

$$V = \gamma(shape) * S^{3/2}$$

$$S = 4 < A >$$

in which V is the volume, S is the surface area, <A> is the average projected area, and γ is a *shape* dependent factor relating volume to surface area. This factor γ is 0.094 for spheres, 0.068 for cubes, 0.052 for rods and 0.048 for slabs[3]. (The exponent $3/2$ is

[3] This statement is only valid for the *regular* cubes, rods and slabs with LWH-ratios 1:1:1, 4:1:1 and 4:4:1 respectively. The *random* objects as well as objects with other LWH-ratios have slightly different γ values.

easily understood: when we scale each dimension of an object with a factor λ, its surface area scales with λ^2 and its volume with λ^3. So, in order to keep γ scale-independent, we have to take the surface area to the power of $3/2$.)

Of course a few difficulties arise when a subject tries to use the projected area cue. Firstly, the subject needs to know the factors γ for the shapes that he is judging. This would mean judging shape before judging volume. Apparently, this is quite likely, because subjects can recognise the shape of the objects at least to some extent. Secondly, an accurate judgement of volume requires area estimation using projections from as many different directions as possible. This is clearly a big problem, because we presented the subjects with a very limited set of projections, namely those corresponding to a rotation around a fixed axis. Consider the rods for an example: depending on the orientation of the rod relative to the rotation axis, its projected area can be constant and at its maximum during the rotation (axis of rotation aligned with the length-axis of the rod), or it can vary from maximum to minimum (axis of rotation orthogonal to the length-axis of the rod), or anything in-between (random axis of rotation). Clearly this will result in very large differences in the average projected area of one and the same object! From this we must conclude that the estimation of volume based solely on the average projected area during one single rotation is a very unreliable method. However, it is much easier to find out how the method works for a large number of trials. The average projected area in such a large group of trials, and therefore also in a large group of relative orientations, eventually becomes equal to one quarter of the 3D surface area. Therefore a reasonable average performance might develop. To check this 'average usefulness' of the area cue, we ran several simulations of the previous psychophysical experiment, replacing the human subject by an algorithm that used the area cue. Although it is hard to anticipate the outcome of the simulations with regard to the thresholds, the values of the subjectively equal volume ratios for the different shape combinations can be predicted from the definition of the algorithm. For the regular shapes (with the standard LWH-ratios) we can compute these ratios: 1.19 (cube-rod), 1.26 (cube-slab) and 1.06 (rod-slab). Even before actually running the simulations we can see that these values are quite different from the values found with human subjects. Predictions for the random shapes are much harder to make because their shapes are only defined statistically.

Methods

Stimulus, procedure and design were the same as in Experiment 2. The subject was replaced by an algorithm that computed the projected area of each object, averaged over one complete rotation. Thus, the judgement of volume was replaced by a judgement of average projected area. Our decision to average over exactly one full rotation is somewhat arbitrary; subjects always watched at least one full rotation, but typically somewhere between two and three rotations. The use of either more or less than one full rotation will obviously influence the results of the simulations. However, in comparing the results of the algorithm with those obtained with human subjects, we decided that the most important point was to keep the number of different projections equal.

Results

The results are presented in figure 2.4. The different stimulus conditions have two marked effects on the threshold levels. First, the thresholds for the cube-cube combination are 2 to 4 times lower than the thresholds for the other shape combinations. Second, thresholds are significantly lower for regular objects than for random ones; the average ratio of regular threshold over random threshold is 0.80±0.10.

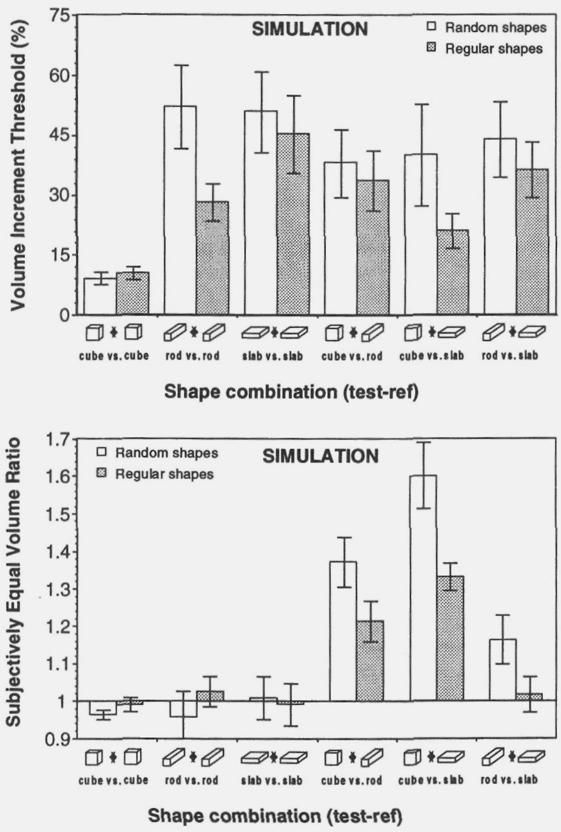


Figure 2.4
The figure depicts the results of the simulations. The upper panel shows the volume increment threshold precentages and the lower panel shows the corresponding subjectively equal volume ratios. The data are presented as a function of the shape condition. The various cube-rod-slab combinations are specified along the horizontal axis. The two bars for each shape combination correspond to random and regular shapes (see legend). Error bars represent minimum statistical estimates of the measurement error.

We computed the factor γ for each individual trial during the simulations and found a range of almost 1.5 decades. This means that for each single trial volume judgements based on the average projected area can be wrong by more than an order of magnitude! However, after averaging over 100 trials per condition, we conclude that the algorithm is reasonably accurate.

The subjectively equal volume ratios of the symmetric conditions are equal to unity, as we anticipated. We can compute the closure of the ratios of the non-symmetric conditions to check for internal consistency. Of course we do not expect the algorithm to yield any significant deviation. We found 1.00±0.09 (random) and 0.93±0.07 (regular). These values are indeed close to unity, closer than the results with human subjects in Experiment 2. The subjectively equal volume ratios that we measured for the regular mixed shape conditions are 1.21±0.05 (cube-rod), 1.33±0.04 (cube-slab) and 1.02±0.05 (rod-slab), which is in agreement with the values that we predicted above using the γ factors.

Discussion

The main goal of running the simulations was to find out something about the amount of volume information carried by projected area, which then indicates the usefulness of this area-cue for human observers. We already concluded that it is mainly average information which is conveyed; the area cue is not very effective for a single volume judgement. When average performance over 100 trials is compared, the algorithm shows somewhat higher thresholds than do the human subjects, except in the cube vs. cube conditions. The improvement in performance from random to regular shapes is less drastic in the case of the algorithm than in the case of human subjects: the average Weber fraction drops to 80% of the random object fractions, whereas it drops to 53% for the human subjects. It is still surprising, however, that the absolute thresholds produced by the simulations, which are based on a rather primitive and imprecise strategy, are only moderately larger than the human thresholds.

A direct comparison of the subjectively equal volume ratios that we measured with human subjects with those produced by the algorithm will not be very fruitful, because the differences between subjects are quite large. Furthermore, since the subjects clearly recognise the objects in some sense (cube-like etc.), their judgements might be affected by feedback and training. A deviation from veridicality then has more to do with subjects' past experiences and less to do with a specific strategy in judging volume. We did not verify whether feedback can indeed influence a subject's behaviour. Nevertheless, the fact that the subjectively equal volume ratios differ between the human and algorithmic discrimination suggests that humans do use a representation of volume which is not based on projected area alone. Moreover, this representation is on average more veridical than if based on projected area alone (the ratios are closer to 1 than found with the algorithm; compare figures 2.3 and 2.4, lower panels).

We must conclude that in machine vision projected area can indeed be used as an estimator for volume, but that it is not sufficient to describe human performance in volume discrimination tasks.

General Discussion

Summary

We presented results of psychophysical experiments testing the abilities of human observers to discriminate the volumes of rotating objects, as affected by the shape of the objects. We found that the volume discrimination thresholds could be described by a Weber law. The Weber fractions depended on the particular combination of shapes such that the task was easiest for regular objects, especially cubes (Weber fraction ~10%), and most difficult for irregular objects with different length-width-height ratios (Weber fractions sometimes up to 80%). Furthermore, we designed an algorithm using the average projected area of a rotating object to compute its volume and tried to address the relevance of this relationship to human observers. We found that the repeatability of human volume judgements varies with shape in a way which is rather well predicted by the algorithm. We also found that human observers are generally more veridical than the algorithm predicts.

Discussion

It is almost impossible to compare our results with previous literature on volume judgements, because we use a completely different experimental set-up and analysis. However, as has been reported in the literature, we do indeed find that shape influences volume judgements. As we have argued before, the effect of shape on relative underestimations and overestimations of volume is too variable over subjects and conditions for us to draw detailed conclusions. Past research has almost always used large groups of subjects and has reported (stable) average biases instead of individual ones. For instance, Stanek (1969) found that stationary real objects, much like our regular rods of Experiment 2, were judged to have the same volume as cubes that were 10% smaller. This is an average over sixty subjects. Comparing this result with our results for two subjects, summarised in the lower panels of figure 2.3, will not teach us much. However, the result found by Lauer (1929) and Stevenson (in: Brunswik, 1934, 1956) that the volume of flat objects tends to be underestimated with respect to that of more cube-like objects might be reflected in the data of Experiment 1, in which two out of three observers were found to significantly underestimate the volume of slabs relative to cubes. We found one study which actually reported on the response variability; this response variability is closely related to our discrimination thresholds. Stevenson (mentioned in Brunswik, 1934, 1956) noted that in his experiments cube-cube comparisons had much less variability (lower threshold) than comparisons of different shapes, which include rod- and slab-like objects. This is indeed in agreement with our results.

Area-based Volume Judgements

Using projected area measures as a replacement of volume can be interpreted as an ignorance of the third dimension. We could equally well say, however, that we did not ignore the third dimension but instead assumed the unknown depth of an object to be directly related to its width, or to the square root of its projected area. Whether we compare projected area or projected area to the power of $3/2$ is clearly irrelevant. Such relations between depth and width have been proposed in the literature before (Caudek & Proffitt, 1993 and Rubin, Solomon & Hochstein, 1995). Whatever the interpretation, the experimental results prove that human subjects do better than a simple area-based analysis of volume. To obtain more reliable judgements with area-based algorithms, one could for instance incorporate some extra information about the γ factor (which requires a shape judgement) preferably depending on the specific orientation of the object (which requires an orientation judgement as well). This requirement of shape and orientation judgements in addition to an area estimate means making quite a complex extension of the basic method and we will not discuss this any further. There might be some practically relevant applications of area-based volume computations in general. A machine vision system could combine these with knowledge about the chemical constituents of an object in order to compute its mass. The advantage of using vision to judge mass instead of measuring weight would be that vision can be used at any distance from the object, whereas the measurement of weight requires some force-exerting device near the object. It is obvious that the area method fails with non-convex objects, but in general at least an upper bound of volume (mass) can be computed.

Appendix

How to define length-width-height ratios.

It is easy to assign a length, width and height to regular objects like boxes or ellipsoids. But what about irregularly shaped objects? Below we will use a decomposition into 3D solid moments to find a generalisation of the length-width-height ratio which is useful for a large set of objects.

A well-known class of methods used for object description and recognition is based on moment invariants (Hu, 1962; Prokop & Reeves, 1992; Dudani, Breeding & McGhee, 1977). These moments describe object (or image) properties that are invariant under certain types of transformations, typically rotations, translations, scaling, etc. This invariance is indeed very useful in dealing with certain recognition problems. Moments are also commonly used in mechanics to characterise bodies by their spatial distributions of mass (see e.g., Banach, 1951, or Routh, 1960). We will use moments of three different orders to describe the objects that we generate in our experiments. An analogy with the use of moments in mechanics will serve as an aid to understanding the mathematics.

The space occupied by a three-dimensional geometric solid can be described by a distribution function $\rho(x,y,z)$ that is 0 outside the object and 1 inside. We can define a set of Cartesian[4] solid[5] moments $\{M_{pqr} | p,q,r = 0,1,2..\infty\}$ of this distribution function; this set in general forms a complete description of $\rho(x,y,z)$ (see Hu, 1962 and Sadjadi & Hall, 1980):

$$M_{pqr} = \iiint_{space} x^p y^q z^r \rho(x,y,z) dx dy dz, \qquad p,q,r=0,1,2,...$$

These moments can be grouped on the basis of the sum of their indices $s = p+q+r$, also called the order of the moments. In mechanics, usually only orders up to and including 2 are used; this is also sufficient for our purposes.

The zero-th order moment M_{000} (s=0) equals the *volume* of the object (or its mass, depending on the way $\rho(x,y,z)$ is defined). The first order term (s=1) is a vector in space: $\begin{pmatrix} M_{100} \\ M_{010} \\ M_{001} \end{pmatrix}$, which equals M_{000} times the *centre-of-mass* vector. Clearly we can use these two terms to describe the size and position of an object. To adjust the values of

[4] Some people use rotational, orthogonal or complex moments instead of Cartesian moments, depending on their specific needs. For instance, the use of orthogonal moments allows for an easier inverse transformation, i.e. computing the distribution function from a given (limited) set of orthogonal moments. See Prokop & Reeves, 1992.

[5] *Solid* moments as opposed to *surface* moments; we integrate over 3D space, not over the surface of the objects.

these parameters to fit our needs, we can apply isotropic scaling and translation to the object.

The second order term ($s=2$) is somewhat more complex; it is represented by a tensor of rank 2, which can be denoted as a 3×3 real valued symmetric matrix. This matrix has three orthogonal eigenvectors, known in mechanics as the *principal axes*, corresponding to three eigenvalues known as the *principal moments of inertia* (with reference to planes perpendicular to these axes; these values are scaled by a factor M_{000}). The direction of the principal axes can be used to define the orientation of an object (see Prokop & Reeves, 1992). Since the moments of inertia are proportional to the square of the dimensions of the object measured along the corresponding principal axes[6], these moments can be used to define the length-width-height ratio of an object. As an example consider an ellipsoid: the principal axes coincide with the three axes of symmetry of the ellipsoid, and the ratio of the moments of inertia is equal to the ratio of the three radii squared. Thus, we will use the (double) ratio of the square roots of the principal moments of inertia as our definition of a length-width-height ratio. This ratio can be manipulated by anisotropic scaling along the principal axes.

[6] This is only true when we compute *central* moments, i.e. moments relative to planes passing through the centre-of-mass of the object. However, we can easily correct for the difference.

3

METAMERISMS IN STRUCTURE-FROM-MOTION PERCEPTION

Abstract

As a 3D object is moving through our world, we generally obtain a vivid impression of both its structure and its motion through space. The time-course of 2D projections of the scene (optic flow) is important in conveying this 3D information to us. The extent to which we can solve this specific inverse problem, i.e. infer a 3D scene from 2D flow, depends on the accuracy with which the required flow characteristics are processed by our visual system. Inadequate 2D processing can lead to incomplete representations of the 3D world (3D metric information is lost). Then the motion and structure of objects can no longer be recovered uniquely. Consequently, metameric classes of 3D representations arise (e.g. only affine properties are conserved). This study investigates under what conditions we find metameric combinations of the perceived attitude and perceived rotation of a plane.

Our subjects are presented with stimuli consisting of two horizontally separated planar patches rotating back and forth in depth about vertical axes. Subjects are required to match both the attitude and the rotation magnitude of these two patches. We vary the attitude from 15° to 60° vertical slant, and the rotation magnitude from 28° to 98°. We find that the matched slant and rotation settings vary widely. For high slant values and for small rotations, attitude and rotation settings become highly correlated, suggesting metamery. For low slant values and for large rotations, the correlation almost disappears, suggesting that both quantities are estimated independently and uniquely.

Our paradigm reveals that with one task and one type of stimulus a gradual transition occurs from unique settings (metric representations) to metameric classes of settings (e.g. affine representations).

Introduction

Motion parallax is considered to be an important cue for human perception of the three-dimensional structure of objects and their motion through a scene. It has been known for a long time that humans can get a fairly accurate three-dimensional impression of a rotating object, just from the shadow it casts on a screen. In 1953, Wallach and O'Connell were the first to investigate this effect systematically. They termed it the 'kinetic depth effect'. This paper deals with the human perception of motion and structure of rotating planar objects, defined solely by motion parallax information. All other possible depth cues are absent. The objective of the study is to reveal the interdependence of perceived attitude and perceived rotation.

Modelling human performance with SFM tasks

There is an ongoing search for a model that will describe human structure-from-motion perception. On the one hand there is a vast literature concerning the theoretical possibilities of reconstructing the spatial layout of a scene from the changing projections of such a scene. Different approaches require different numbers of views and markers (or different spatio-temporal derivatives of the optic flow field), and are based on different assumptions about the underlying structure of the scene, the transformations between the views and the type of projection. For an extensive review of a large number of these computational methods, see Aggarwal & Nandhakumar (1988). On the other hand there is a large collection of experimental (psychophysical) data that proposes certain reconstruction schemes and disproves others. For instance, read Simpson (1993).

We now give a short introduction to two important issues that motivated the research described in this paper.

Affine-in-depth and metric representations

It is theoretically impossible to compute a full metric representation of a scene from only two orthographic views (ignoring all cues but structure-from-motion)[7]. One needs at least three views of four points to do that (Ullman, 1979). This raises a question about how humans possibly can perceive a 3D structure with only two orthographic views, a fact for which abundant evidence has been collected (e.g. Ullman, 1984; Hildreth, Grzywacz, Adelson & Inada, 1990; Todd & Bressan, 1990; Todd and Norman, 1991; Liter, Braunstein & Hoffman, 1993; Werkhoven & van Veen, 1995). We must ask ourselves what information *can* be extracted from two views only. Several researchers have addressed this question (Ullman, 1983; Bennett, Hoffman, Nicola & Prakash, 1989; Todd & Bressan, 1990; Koenderink & van Doorn, 1991); it turns out that we can obtain a one-parameter family of solutions. This we call the affine-in-depth[8] representation of

[7] In principle, the combination of structure-from-motion with perspective deformations allows for the full reconstruction of a scene from only two views of five points (except for an overall scale parameter; for example see Faugeras & Maybank, 1990; also see Ullman, 1979 and Longuet-Higgins, 1988). However, this reconstruction from two perspective views is known to be very sensitive to noise and is only accurate for large visual fields (high perspective). Read Ullman (1983) and Kanatani (1993) for some discussion on this. We will further ignore perspective information and instead discuss orthographic projections.

[8] Several authors have given names to different representations of the structure and motion of three-dimensional objects. We give a short explanation here. An *affine* representation means effectively that structure and motion are not disentangled at all. No metric aspects of the structure are represented; only affine properties (like parallelism of line segments) and affine quantities (like the ratio of distances along a certain direction (relief), or the ratio of volumes) are represented. Motions are undefined. This is stratum I of Koenderink and van Doorn's analysis (1991). An *affine-in-depth* representation encompasses a nearly perfect separation of structure and motion. Both are almost completely known except for one parameter which couples possible interpretations of structure to those of motion. A full metric is obtained in the fronto-parallel directions but only a special affine representation is obtained along the viewing direction. Note that sometimes this type of representation is also called affine, which is rather confusing. This is stratum II in Koenderink and van Doorn's analysis which is also described by Bennett et al. (1989) and Todd and Bressan (1990). In a *metric* representation, motion and structure are no longer confounded. We know all distances between points in 3D and the object motion is also fully recovered. This is the stage which was historically believed to be the one the visual system aims to reach for. However, it has become clear recently that many tasks examined in past experiments do not require a metric representation at all (Todd & Bressan, 1990).

a scene, as opposed to the metric representation. In principle, an affine-in-depth representation can be upgraded to a metric representation with the use of other sources of information, such as some other visual cue like binocular stereopsis (e.g. Richards, 1985) or prior information and assumptions about object shape and motion.

It is not yet clear whether an affine-in-depth or a metric description (or something else) most closely resembles the representation that we use. That we already seem to have some representation of visual space with only two views can be used as an argument in favour of the hypothesis that such a representation must resemble an affine-in-depth representation. However, instead of measuring the representation that a subject has as a function of the number of views, we should measure it as a function of the information available to the subject to perform a specific task. The problem then is to specify what exactly that information is. In the general discussion we will develop some ideas about the sorts of information that might play a role in controlling the type of representation available to the subject, based on the results of the experiments that we did.

Motion & Structure

A second important topic is the relation between perception of motion and perception of structure. In an affine-in-depth representation, these entities are, theoretically at least, strongly coupled. That is, misinterpretations of the motion will lead to corresponding misinterpretations of the structural aspects.

Whether human performance does indeed reflect this relation is an interesting question which will be reviewed in the general discussion at the end of this paper. We can hypothesise two distinct mechanisms for instance, one for the extraction of motion information and one for the extraction of structure information. In such a configuration, errors in one system will not be correlated with errors made in the other system. Because the affine-in-depth theory lets us make strong predictions about this issue, it makes sense to investigate this question psychophysically.

This paper

The paper starts with a short analysis of a method for recovering structure and motion from two parallel views. We reduce the problem to its basics, and pay special attention to the connection between recovered motion and recovered structure.

We then report on new psychophysical experiments, in which we investigate the relation between perception of surface attitude and perception of rotation magnitude. We probe human performance using a double-matching procedure. First we show that the errors in the perception of surface attitude and rotation magnitude can become quite large. Then we present clear results which indicate that, under certain conditions, there is only one combined perceptual correlate of these two separate entities: attitude and rotation become fully interchangeable. This supports the existence of mechanisms that compute motion and structure simultaneously. In practice, the results imply that an overestimation of depth is accompanied by an underestimation of motion. This perceptual mixing (metamerism) is to be understood theoretically in terms of a so-called affine-in-depth representation.

Finally, in the general discussion, we adopt an approach that differs from the usual geometric approach to structure-from-motion. The connection between the subject's use of two or more views in structure-from-motion tasks and the subject's performance with affine and metric tasks respectively is usually based on theoretically derived minimum configurations for the recovery of spatial layout and motion (e.g. three orthographic views are required for a metric representation, which is needed for good performance with metric tasks). We argue that our experiments show that the difference between affine-in-depth and metric representations is not as strict as generally believed. Whether an affine-in-depth or a metric reconstruction can be made from the stimulus and whether human performance is affine-in-depth or metric is basically a matter of noise tolerance, not of geometry.

Reconstruction of 3D structure and motion from two parallel projections

This section explains some essential aspects of the theories concerning the reconstruction of 3D object motion and layout from parallel projections. We consider the discrete case of moving identifiable object points. We treat the two-views case, which results in an affine-in-depth representation.

The stimuli that we use in the present experiments always consist of planar objects rotating around fronto-parallel axes, which is a generic example of the structure-from-motion problem. We clarify this by making a decomposition of the 3D motion, following Koenderink and van Doorn (1991). We will show that with an affine-in-depth representation, underestimation of the amount of rotation in depth corresponds to overestimation of the depth of the rotating structure.

The general case

Consider a moving object in space. The projections of its identifiable points[9] carry information about the spatio-temporal relation between these points and the observer. Several authors have already described in detail some techniques for making the 3D reconstruction. We stress here that the outcomes of their different methods should be the same in all respects, provided the information and assumptions used are identical. In our analysis we use the set of projected positions in different parallel projected views (defining the structure-from-motion cue and ignoring any perspective cues) and we make the commonly used rigidity assumption.

A convenient approach then is to consider the relative motion to be the combination of a translation of the object centre and a rotation of the object around an axis through the same centre, both with respect to the viewer-fixed reference frame[10]. Following Koenderink & van Doorn (1991), who gave an elegant stratification of the structure-from-motion problem, we first treat the translation part of the motion. (Koenderink and van Doorn took the semi-parallel approach whereas we discuss the parallel case. This leads to small differences in some points.) A translation in depth, i.e. along the viewing direction, is *not* visible in the projection; that is, global translations in depth are mathematically indistinguishable from no translation at all. Translations in a fronto-parallel direction, however, (orthogonal to the viewing direction) have a strong and simple impact on the image: everything moves with a fixed amount in the same direction as if there were a translation of the image plane. We can define the amount and direction of translation by selecting some point of the object and interpret its displacement from view 1 to view 2 as translation alone. The axis of rotation now passes through that point. This is what we earlier on called 'the object centre'; which point we actually choose is not important. To summarise, we can say that the translations are either not noticeable (along the viewing direction) or easily extracted (along a fronto-parallel direction) and do not reveal any information about the structure of the object.

Let us assume the translation part has been dealt with[11] and that we have thus defined an origin through which the rotation axis passes. We can now decompose the remaining object rotation into two successive rotations: one rotation around an axis orthogonal to the viewing direction followed by a rotation around the viewing direction (see APPENDIX A). The latter rotation is equivalent to a rotation of the image plane and therefore contains no information about the spatial layout of the object under consideration. In view of the fact that the other rotation component causes projected displacements in one direction only (orthogonal to the axis direction), all we

[9] When discussing reconstruction schemes, most authors start at the point where the correspondence between the projections of each marker in different views has been established. Solving that problem is interesting in itself, and has been investigated by Scott & Longuet-Higgins (1991), Aggarwal, Davis & Martin (1981) and Ullman (1979), amongst others. We will not discuss this any further and instead assume that the correspondence has been established.

[10] We are using a specific degree of freedom here (Chasles' theorem): in general, a motion in space can be described by a rotation followed by a translation. The *position* of the rotation axis is not defined and can be chosen anywhere. For each choice there is one corresponding translation. This is just a matter of selecting a convenient co-ordinate system.

[11] When we watch a moving object, we typically follow it with our eyes as it translates through space. This effectively removes the translation component orthogonal to the viewing direction.

need to do is to find the specific amount of image plane rotation that makes the remaining flow (displacement field) parallel[12]. To do this we need at least two points visible in both views, in addition to the point that we defined as origin. The solution can easily be found[13]. After we have corrected the views for this image plane orientation difference, we are left with a parallel flow. This corresponds to a rotation in depth around an axis parallel to the fronto-parallel plane and perpendicular to the flow. Note that as soon as we have reduced the problem to this parallel flow situation, such an interpretation in terms of a rigid object rotating in depth is always possible.

Although the previous steps are all quite interesting in themselves, the information about the structure is only revealed through the rotation in depth. We will therefore focus our attention on this component alone.

The generic problem: a planar patch rotating in depth

At this point we have a parallel displacement field, specified by at least three moving points: an origin and two others to establish the parallelism of flow. How should we proceed from here? In APPENDIX B it is shown that any reconstruction of this generic structure-from-motion problem will be ambiguous both in depth and motion. Adding more object points does not make any difference. We conclude that the basic gist of the structure-from-motion problem is contained in a planar patch rotating in depth (which is specified by three non-collinear points in two views), at least for parallel projections. Furthermore, we conclude that from two parallel projected views neither motion or depth can be fully determined; only a relation between them can be established. In other words, we have a metameric class of solutions of motion and structure.

In the rest of this paper we will discuss experiments with such planar patches rotating in depth. The angle between a planar patch and its rotation axis is an invariant during the rotation. We call this angle the *elevation* θ of the plane with respect to the rotation axis[14]. See figure 3.1.

---

[12] A perspective projection of an object that translates in depth changes in size. This scale problem is totally ignored in the parallel projection case. The scaling causes the flow field to be non-parallel after removal of translation and image-plane rotation, which makes the decomposition somewhat more complicated, but it is nevertheless possible to find both the image plane rotation and the scaling factor if we have at least four points in both views (Koenderink et al., 1991).

[13] Suppose we have plotted these two views on a pair of transparencies. After choosing the origin (pick any one) we can put the transparencies on top of each other with their origins coincident. Now just rotate one transparency (keep on matching the origins) until all displacements are parallel. You will find two solutions which differ by 180 degrees but that is just a matter of signs. Moreover, it is readily seen that this is a method which is quite tolerant to noise: just find the amount of rotation which maximises parallelism of flow (e.g. maximise the sum of the dot-products between all displacement vectors). If we cannot find such an amount of rotation, we must conclude that the object underwent a non-rigid transformation between the two views. This would falsify our assumption of a single object *moving* in space (rigidity assumption!).

[14] $\theta = ArcTan\left(\frac{Gx}{\sqrt{1+Gy^2}}\right)$. Gx and Gy are the x and y components of the depth gradient. They are larger when the plane is more slanted in the corresponding direction. Note that the (vertical) rotation axis points in the x direction!

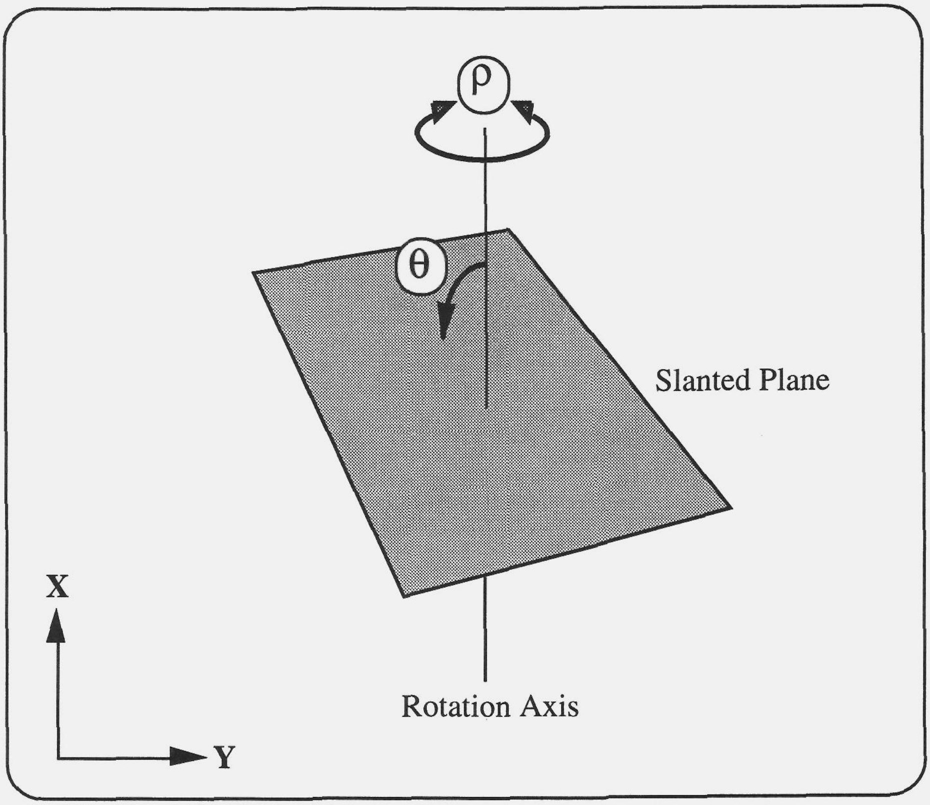


Figure 3.1
A diagram depicting a backwards slanted plane rotating around a vertical axis. The subject's direction of view coincides with the positive z-direction. (Note that this sketch by no means resembles the actual stimulus; for the sake of a clearer drawing the plane has been drawn in perspective, which was not the case in the experiment. Furthermore, in the real experiment the contours of the plane were masked and the plane was textured with light dots on a dark background.) The elevation of the plane is indicated with θ and the magnitude of rotation is denoted as ρ. For the particular plane depicted, the depth gradient in the x direction Gx is larger than the depth gradient in the y direction Gy, because the plane is clearly more slanted backwards than sidewards.

Using the results from APPENDIX B, a (rather complex) relation between reconstructed elevation θ^{rec} and rotation magnitude ρ^{rec} can be derived. To make this relation easier to understand, we consider a range of reconstructed rotation magnitudes around the veridical magnitude. In that case a much simpler dependence emerges. We distinguish two extreme situations:

A: Gy » 1 (the slant in the horizontal direction, orthogonal to the rotation axis, is very large); in this situation the reconstructed elevation coincides with the veridical value, independent of the reconstructed rotation magnitude (still assuming rotation magnitudes close to the veridical value!). This result is easily understood when one realises that Gy->∞ corresponds to a side-view of the plane. The projection of the plane reduces to a line, in which case the elevation simplifies to the 2D angle between the rotation axis and that line.

B: Gy « 1 (there is almost no slant in the horizontal direction); in this situation the relation between reconstructed and veridical elevation and rotation magnitude is given by

$$Tan(\theta^{rec}) = Tan(\theta^{ver}) \cdot \frac{Sin(\rho^{ver})}{Sin(\rho^{rec})},$$

in which the superscripts *rec* and *ver* denote *rec*onstructed and *ver*idical quantities respectively. It is clear that the relation between reconstructed elevation and reconstructed rotation is approximately hyperbolic (inversely proportional). This means that an *underestimation of the amount of rotation* will result in an overestimation of the elevation, which *corresponds to more depth.*

For intermediate values of Gy there is a general transition between the two extremes. In general, an inverse relationship is a good description of the relation between reconstructed rotation magnitude and reconstructed elevation. It means that when reconstructed elevation is plotted against reconstructed rotation magnitude, the metameric class is depicted as a hyperbolic curve passing through the veridical point.

The above formulas are exact only when the reconstructed values are close enough to the veridical ones. When the deviations get too large, the formulas become more complex and the inverse relationship becomes distorted. However, in interpreting the experiments described in this paper it suffices to keep the inverse relationship in mind. Whenever a quantitative result is needed to interpret the data we revert to the exact formulas from the APPENDIX.

Remarks

The reduction of the structure-from-motion problem using motion decomposition seems to be an interesting approach. We are not suggesting however, that the human visual system does indeed perform this decomposition or reduction of the problem.

The extension to the three-or-more-views case was not discussed here, but it suffices to know that a metric representation can be obtained from three views by the application of intersection of constraints (Koenderink and van Doorn, 1991).

Methods

Subjects had to simultaneously match the magnitudes of rotation and the elevations of two rotating planar patches.

Apparatus

The stimuli were shown on a Radius TPD/19 high resolution (82 dpi) grey-scale monitor at 71 Hz, connected to an Apple Macintosh IIfx computer.

Stimulus

Two rotating planar objects with certain attitudes were simultaneously displayed on a computer screen. The rotation axes were always vertical and through the centres of the objects. The planes were textured with light random dots (dot size 43 arcsec) on a dark background, and were presented with high contrast. Both planes were shown behind circular masks with diameters of 3.4°. The masks were not directly visible because they had the same luminance as the background. The horizontal separation of the masks (and the object centres) was 6.9°. Because the planes were much larger than the mask size, no contour information was available in the displays, except when the amount of rotation approached 180°, causing a side-view of the plane. This large amount of rotation was far outside the range used to generate the stimuli, but could in principle occur when a subject overestimates the rotation magnitude dramatically. However, it turned out that the subjects very rarely encountered this effect while adjusting the magnitude of rotation. The number of visible dots varied during the rotation, but at the most frontal orientation[15] of the planes about 60 dots were visible (~1‰ dot density). Thus, if the plane is rotated 60° relative to the most frontal condition, then the number of visible dots will double. The planes were presented in oscillatory motion, with initially different magnitudes of rotation. We generated 20 new frames per second, thus creating a rather smooth motion sequence. New sets of random dots were generated for each trial and for each plane.

During a trial both planes were shown simultaneously, except in the first second when only the reference plane was shown. This procedure was used to avoid possible confusion about which plane could be adjusted: the test plane could be either on the left or the right of the screen.

Dot density can be a cue for the instantaneous attitude of a plane. We made sure that the dot density was always chosen independent of the elevation: in each trial, a random number of dots (between 45 and 75) was visible in the most frontal orientation

[15] The orientation at which the normal of the planar patch is within the plane defined by the viewing direction and the rotation axis is called the *most frontal orientation* during the rotation. Gy=0 in that case.

of a plane. Moreover, this dot density was kept independent of elevation even during subjects' adjustments of elevation. This sometimes caused a subject to gain a somewhat unnatural impression of the behaviour of the dots in the plane while he or she made these adjustments: when the subject increased the elevation of the plane, the dots stayed at the same height on the screen instead of being compressed towards the centre. During such an adjustment, the subject sometimes had the impression that the dots were moving inside the plane which is of course exactly what we put in. Because subjects reported this phenomenon only occasionally (and it was only visible while adjusting) we decided to maintain the precaution. Pilot studies did not show any significant change in performance that could be attributed to this phenomenon.

It should be noted again that the dot density is of course not constant during the rotation, which is a potential cue to the rotation magnitude. As said above, we have randomised the number of dots available in the most frontal view, but there is no way to neutralise the dot density cue completely. When subjects could make use of this cue, they would probably end up avoiding extremely high rotation magnitudes.

We randomised the period of oscillation of the test object, to prevent the subjects from matching either the elevation or the rotation magnitude by simply matching a single image velocity at a certain location inside the masks. Whereas the period of oscillation for the reference object was always 28 frames (1.4 sec), for the test object it was (with equal probability) either 20, 24, 28, 32, 36 or 40 frames (between 1 and 2 sec). Moreover, the phase of the two oscillations was randomised, which is only a relevant factor for equal periods. Pilot studies did not show any significant change in performance which could be attributed to randomisation of the period of oscillation.

Procedure

The subjects were seated in front of the computer screen at a distance of 150 cm. Head movements were not restricted, but subjects were required to keep their right eye just in front of the centre of the screen. The left eye was covered with a black eye patch. The room was almost completely dark, but subjects were still able to see the monitor and the table on which it was placed.

Each stimulus consisted of a test object and a reference object. A stimulus is specified by six parameters: the angle that each plane makes with the (vertical) rotation axis, also called the *elevation* of the plane (zero means an upright plane); the *magnitude* of oscillation of each plane (twice the amplitude); and the average orientation of each plane with respect to the most frontal condition, also called the *twist* of that plane. The values of the reference object parameters depended on the specific condition measured. The values of the test object parameters were chosen randomly, except for the twist, which was always the same for reference and test.

The task of the subjects was twofold: they had to simultaneously adjust both the elevation and the magnitude of rotation of the test plane until these parameters matched those of the reference plane. They used a trackerball to control the parameters; up-down movement of the hand corresponded to an increase-decrease in elevation of the test plane, whereas left-right movement corresponded to a decrease-increase in the magnitude of rotation of the test plane. It was possible for subjects to match both parameters within one single movement. One full rotation of the ball corresponded to a 30° change in elevation (0.05° angular resolution). The resolution of the magnitude of rotation depended on the frequency of oscillation, but was always between 0.11° (high frequency) and 0.21° (low frequency). When subjects were satisfied with the adjustments they had made, they indicated this by pressing a button, whereupon a new stimulus appeared. Presentation time was not limited and on average subjects took about one minute per trial.

Subjects

Three subjects participated in our experiments: HJ(author), MH and IV. All had normal or corrected-to-normal vision. HJ and MH had had extensive previous experience with structure-from-motion tasks and were fully aware of the objectives of the research, whereas IV was only slightly acquainted with this type of task and the topic of research.

Design

We defined 32 experimental conditions. We varied the elevation of the reference plane (15°, 30°, 45° and 60°), the magnitude of its rotation (28°, 42°, 70° and 98°) and its twist (0° or 30°). The elevation of the test plane was (initially) always randomly chosen between -15° and 70°, and its magnitude of rotation was between 14° and 126°. As mentioned before, the twist of the test plane was always the same as that of the reference plane.

These 32 conditions were measured in two separate experiments, one in which all planes had twist 0° (subjects HJ, MH and IV), and a second one in which all planes had twist 30° (subject MH). Each of the 16 conditions was measured about 20 times per subject. These 320 trials were presented to the subject in random order, although pilot studies have shown no significant effect of randomisation. In fact, measuring one condition in isolation (thus with a fixed and known reference object) gave the same results as mixing it with many others. Subjects completed an experiment in about four sessions.

Results

Method of analysis

The raw data that we acquired consisted of paired settings of elevation and rotation magnitude. First, we took the absolute value of the recorded rotation magnitudes; subjects could 'go' to negative values during adjustment, which happened occasionally with the smaller rotation magnitudes. Taking the absolute value is allowed because it is effectively the same as a 180° phase change, which has no effect on a stimulus in which phase is randomised. Secondly, we mapped all recorded elevation angles onto angles between 0° and 90°. A sign inversion of the elevation is equivalent to a mirror inversion in the fronto-parallel plane, which is a well-known mathematical ambiguity for displays generated with parallel projection.

To facilitate the interpretation of these corrected data, we performed several statistical analyses on the settings. The data were grouped per condition and subject (~20 trials) and plotted in a two-dimensional graph as elevation (y) versus rotation magnitude (x). Variances and covariances were computed; this enabled us to draw the covariance ellipse in the same graph. The covariance ellipse is a contour line of the two-dimensional Gaussian probability distribution fitted to the data. If any settings were outside the 2.5 sigma contour (4% probability), these settings were discarded and a new covariance ellipse was computed. Each covariance ellipse was specified by the two-dimensional median of the settings (the ellipse centre), its major and minor axis lengths (the width of the Gaussian in that direction) and its orientation. We also computed the median, mean and standard deviations of both the elevation and rotation magnitude settings, as well as the linear correlation between these settings.

Figure 3.2 shows the complete dataset obtained from subject IV at twist 0°, illustrating the methods of the above analysis. Because of limited space, only this example is given. The results for all subjects are summarised below.

Introspective results

All subjects reported seeing slanted random-dot textured planes rotating in depth in almost every trial. Sometimes subjects reported having problems with too high image speeds in conditions in which both the elevation and the rotation magnitude were large. Subjects remarked that while they were adjusting the elevation, the rotation magnitude sometimes seemed to change as well. In some cases when the subjects moved the trackerball along a diagonal, thus objectively adjusting both parameters at the same time, they had the impression they were changing only one parameter at a time. One subject reported sometimes having the impression that while the plane was moving the dots were also moving over the planar surface.

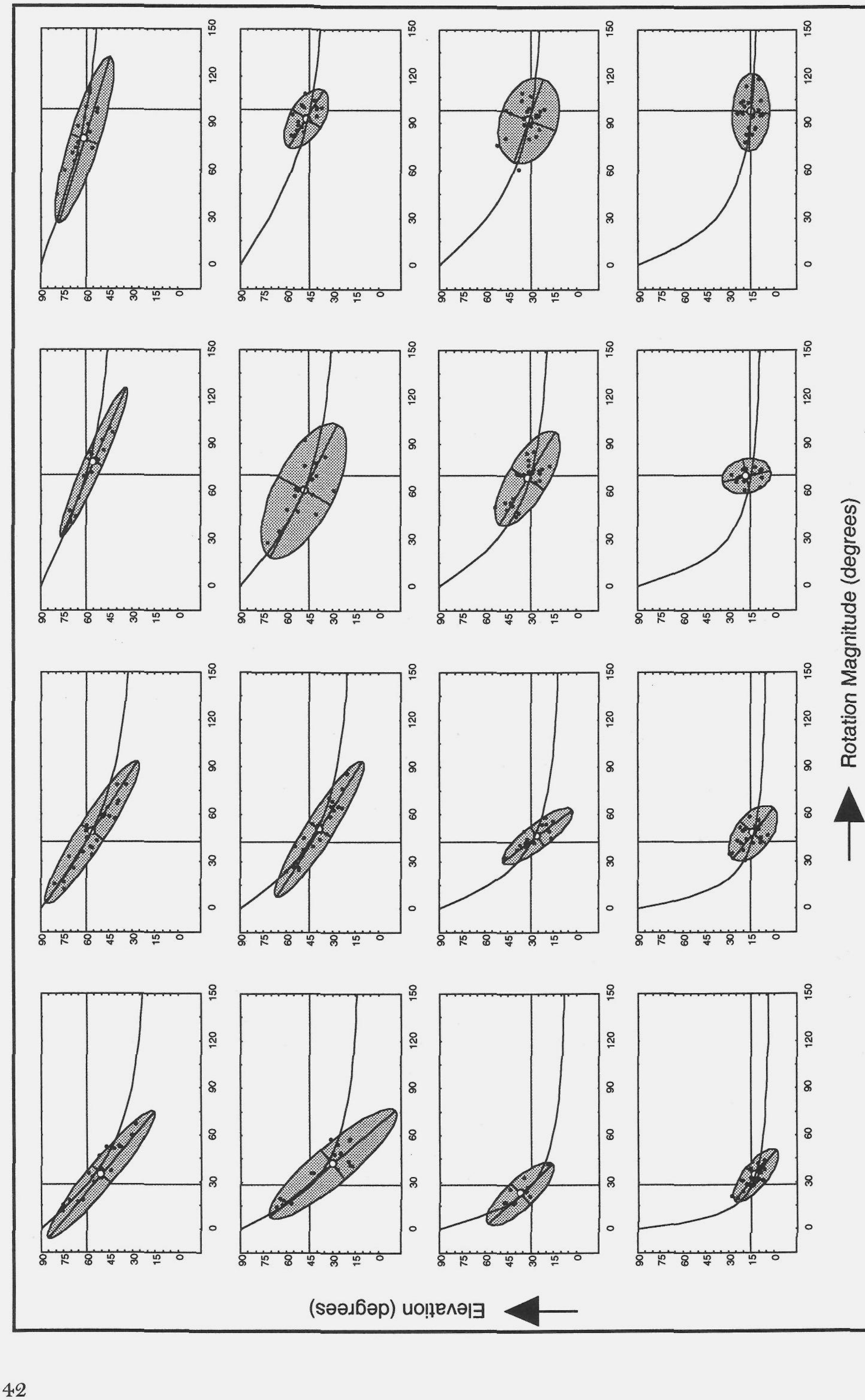

Rotation Magnitude (degrees)
Elevation (degrees)

Figure 3.2 (left page)
These graphs present raw data for one subject (IV, twist 0°). Matched elevation is plotted versus matched rotation magnitude. All panels are on the same scale. *Explanation of symbols:* The small black dots in each graph represent the data points. The crosshairs represent the veridical values of elevation and rotation magnitude. Different panels correspond to different veridical values. For each row the elevation is constant: from lower to upper 15°, 30°, 45° and 60°. For each column the rotation magnitude is constant: from left to right 28°, 42°, 70° and 98°. The grey region is the computed covariance ellipse of the data at 2.5 sigma. Its centre is located at the median of the settings, denoted by a small white circle. The curve that is shown is the affine-in-depth solution computed from two views: the most frontal one and one of the two extreme views. *Interpretation:* The lower right panel is a clear example of what we call a metric response. The upper left one is a typical example of what we call an affine-in-depth response.

Medians of the settings of elevation and rotation magnitude

Introduction

The medians of the settings cannot be given a proper meaning in this experiment, because the reference stimulus and the test stimulus were almost completely identical. Measures of overestimation or underestimation can only be obtained if the test stimulus is defined differently. For instance, we could use a stereo-defined test stimulus, or ask subjects to estimate numerical values of elevation and rotation. Therefore, in this experiment, medians deviating from the real values that were put into the displays can be due only to particular asymmetries between test and reference stimulus, or to response biases of the subjects. We will discuss these possibilities briefly.

First, there is one major difference between test and reference stimulus, namely the frequency of oscillation. The reference plane rotates at 1 cycle per 1.4 seconds, whereas the test plane rotates at 1 cycle per period of 1 to 2 seconds. This asymmetry could lead to perceptual biasing of the elevation or of the rotation magnitude. For instance, higher frequencies of oscillation, and thus higher image velocities, might give the subject an impression of a larger rotation magnitude.

Second, we give an example of possible response biases. Suppose that subjects avoid the range of adjustments that result in extreme image velocities, either very low or very high ones. If during a certain adjustment the image velocities become too exceptional (due to the specific rotation magnitude, elevation or frequency of oscillation), the subjects will effectively adjust the rotation magnitude and/or the elevation to moderate the image velocities. The resulting median of the settings will then become biased.

Results

For each subject, the 16 conditions per experiment resulted in 16 median values for both matched elevation and matched rotation magnitude. No significant effects of the reference rotation magnitude on the median of the elevation settings could be found. The same holds for effects of reference elevation on the median of the rotation magnitude settings. Effects of the frequency of oscillation on elevation or rotation magnitude settings could not be found. We computed the linear regression of the median of matched elevation as a function of the reference stimulus elevation, and of the median of matched rotation magnitude as a function of the reference rotation magnitude. The results of these analyses are shown in table 3.1. A compression of the range of both the rotation magnitude and elevation responses is clearly visible, at least for subjects MH and IV. Further inspection shows that the compression of rotation magnitude range is caused by overestimation of small magnitudes and underestimation of large magnitudes. Inspection of the elevation results reveals a compression towards 0° over the whole range. The results for subject MH in the twist 30° conditions are similar to the results for the twist 0° conditions.

	Rotation magnitude (°)			Elevation (°)		
subject	r^2 (%)	constant (s.e.)	regression coefficient (s.e.)	r^2 (%)	constant (s.e.)	regression coefficient (s.e.)
HJ - twist 0°	98	3 (3)	0.95 (0.04)	97	4.9 (2.0)	0.94 (0.05)
MH - twist 0°	92	13 (4)	0.73 (0.07)	96	4.7 (2.1)	0.87 (0.05)
IV - twist 0°	93	13 (4)	0.80 (0.06)	91	3 (3)	0.88 (0.08)
MH - twist 30°	99	5.4 (2.0)	0.87 (0.03)	97	1.1 (1.8)	0.89 (0.04)

Table 3.1
The regression results per subject for the median setting of rotation magnitude as a function of reference rotation magnitude, and for the median setting of elevation as a function of reference elevation. Standard errors are given in brackets.

Conclusions

The fact that the recorded elevation and rotation magnitude (i.e. the median of settings) of the test stimulus are highly correlated with the reference stimulus parameters proves that subjects are quite capable of performing the task. Moreover, the subjects demonstrate only a slight tendency to compress the range of responses with respect to the stimulus range; this suggests that both the differences between test and reference object and the response biases are negligible. The small range compression is compatible with a response bias caused by subjects' regression away from extreme image velocities.

Variance of elevation and rotation magnitude settings

The variances in the settings of the two parameters give us some indication about the discriminability of elevation and rotation magnitude.

Results for the rotation magnitude settings

As can be seen in figure 3.3, the standard deviation of the rotation magnitude settings (s.d.) does not increase with increasing absolute values of rotation magnitude, except possibly for the 60° elevation conditions. The influence of the different elevations on the level of the s.d. depends on the subject: MH shows virtually no effects (both twist conditions), whereas subjects HJ and IV exhibit an increase in the s.d. with increasing elevation.

The average s.d. (average of all conditions) was quite constant over subjects: 11° (HJ, MH twist 0°) and 12° (IV). Subject MH at twist 30° showed a lower value, namely 8°. Note that these averages correspond to a relative s.d. of 40% at 28° rotation magnitude, going down to 10% and below at 98° rotation magnitude.

Results for the elevation settings

As we can see in figure 3.4, the standard deviation of the elevation settings (s.d.) does not increase with increasing absolute values of elevation beyond 30° (subject IV beyond 45°). No significant effects of the rotation magnitude on the s.d. could be found.

The average s.d. (average of all conditions) was higher for subjects HJ and IV than for subject MH: 9° (HJ, IV) versus 7° (MH twist 0°) and 6° (MH twist 30°).

Conclusions

The variances in the settings of the two parameters are fairly consistent over subjects and are more or less independent of the specific condition measured. Subject MH (who tested both twist conditions) showed slightly better performance with the twist 30° condition than with the twist 0° condition. This might however be due to a learning effect, since the twist 30° experiment was done after the twist 0° experiment. If we interpret the results in terms of a discrimination experiment, differential thresholds of 11° for the rotation magnitude and 8° for the elevation seem to be reasonable averages. The elevation threshold at 15° elevation is somewhat lower, namely about 6°. Since these values are so constant, relative thresholds are quite high for small rotation magnitude or elevation; both thresholds can become as large as 40%.

Correlation between the settings

Introduction

In our introduction to this paper we brought up the subject of the relation between perception of motion and perception of structure. In the theory section we then discussed the affine-in-depth representation and showed that it is equivalent to a one-parameter metameric class of solutions of motion and depth. If, on the one hand, the visual system maintains an *affine-in-depth* representation of a stimulus, misjudgements of motion will be accompanied by corresponding misjudgements of depth. The relation between the misjudgements can then be predicted using the equations from APPENDIX B. If, on the other hand, the visual system maintains a *metric* representation of a stimulus, misjudgements of motion and depth will not be correlated; they will then be treated by the visual system as if they were independent entities.

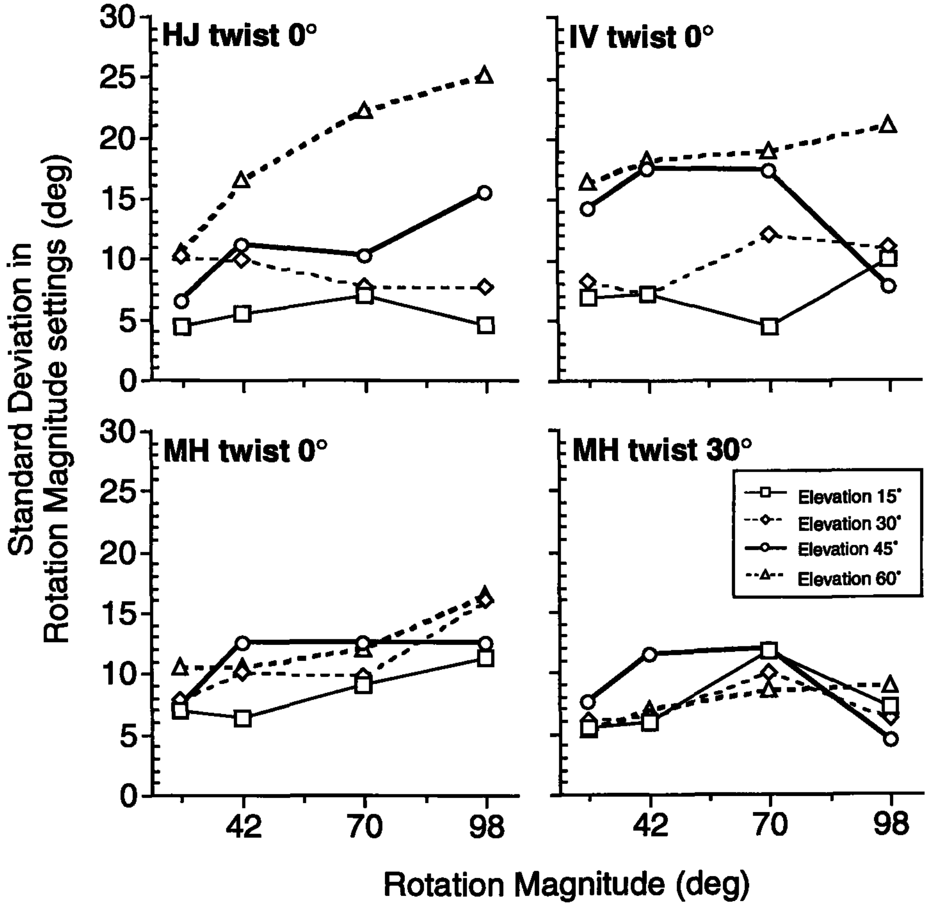


Figure 3.3
The standard deviations in the settings of rotation magnitude as a function of the real rotation magnitude. Various symbols represent different elevations (see legend inside lower right panel). Different panels correspond to different subjects.

In the experiment we varied the stimulus by varying the twist, the rotation magnitude and the elevation. When the settings of elevation and rotation magnitude are correlated (per condition), we are in fact probing a metameric representation. When the settings are not correlated, a metric representation is probed. We will present our results for this correlation demonstrating that there is a gradual transition between metric and affine-in-depth related behaviour, depending on the specific stimulus condition used. We also show that the correlation, when significantly different from zero, is indeed in accordance with a relation expected on the basis of an affine-in-depth representation.

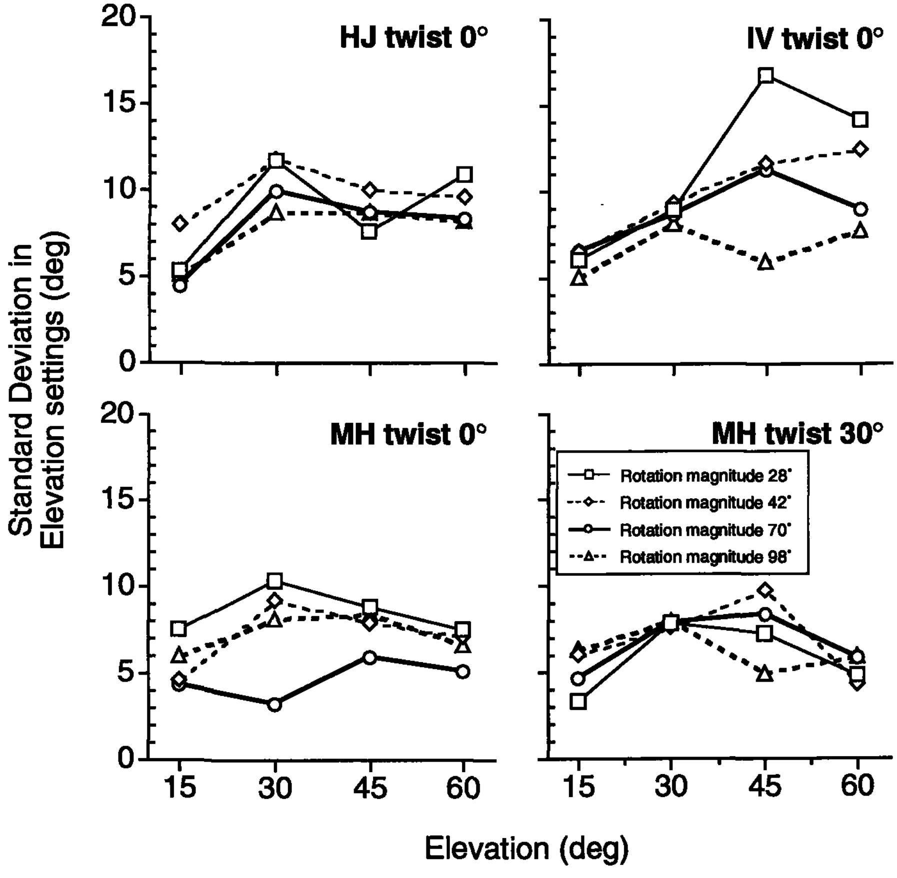


Figure 3.4
The standard deviations in the settings of elevation as a function of the real elevation. Various symbols represent different rotation magnitudes (see legend inside lower right panel). Different panels correspond to different subjects.

Correlation results

We computed the linear correlation between matched elevation and matched rotation magnitude for all conditions and for each subject. For each correlation we computed the probability that that specific value could originate from an uncorrelated distribution. This allows us to estimate the significance of a certain correlation differing from zero. In figure 3.5 we present all the results. The first and major result is that *all correlations are negative.* Moreover, 46 out of 64 are significantly different from zero at the 5% level. Linear correlations are plotted versus the elevation for all subjects. Note that correlations of up to 95% are reported. It is clearly visible that elevation has a large effect on the correlation: larger elevations result in correlations closer to -1. The main difference between the subjects concerning this point is the behaviour at low elevations: subject HJ shows virtually no correlation at 15° elevation, whereas the other two subjects do show a correlation. The same results are found for 30° elevation. The behaviour at 45° and 60° elevation is very similar amongst subjects. The effects of rotation magnitude are smaller than those of elevation, but it is evident that larger rotation magnitudes prohibit high correlations, at least for small elevations. Finally, for 30° twist the correlations are all shifted towards zero relative to the 0° twist condition. In this case the increase in correlation with elevation already levels off at 30° elevation.

Orientation results

To test whether the reported correlations are in agreement with the predictions of an affine-in-depth representation, we compared the orientation of the covariance ellipse with the orientation of the affine-in-depth curve at the veridical point. We compared only those orientations for which the correlation was significantly different from zero, because for the uncorrelated conditions the orientation becomes undefined (anyway, we then have a metric representation). As mentioned before, the theoretical curves that we use (see figure 3.2) are based on the combination of the most frontal view with the outermost one. This is not the only possible choice, since every pair of views yields a corresponding curve. We choose this particular pair because it represents the complete motion sequence (the other extreme view is in fact a mirror image of the first one, at least for twist 0°). The orientations of the different possible curves can vary a lot, but the rank correlation with the data is always about the same. Because of the limited statistics and the lack of an adequate description of how these curves can be combined, we will not try to find a best fitting curve.

In figure 3.2 we already saw that for subject IV the agreement is reasonable: the orientation of the curve at the veridical point coincides with the orientation of the covariance ellipse in most conditions. The Spearman rank correlation of the measured orientations with the theoretical ones is indeed reasonably high for twist 0°: 0.83 (HJ), 0.79 (IV), 0.70 (MH), but close to zero for twist 30°: 0.18 (MH). This is probably due to the fact that the range of theoretical orientations (slopes) is much smaller at 30° than at 0° twist.

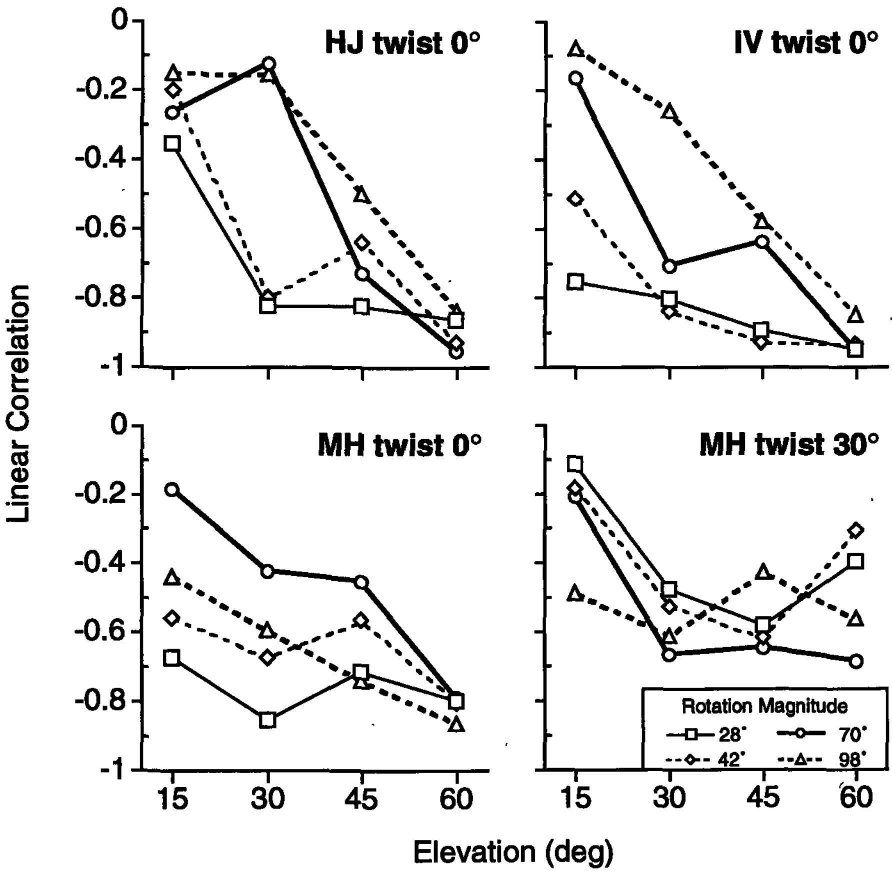


Figure 3.5
Linear correlation between matched elevation and matched rotation magnitude, plotted as a function of elevation. Various symbols represent different rotation magnitudes (see legend inside lower right panel). Different panels correspond to different subjects.

Conclusions

The fact that all reported correlations are negative supports the theoretically predicted inverse relationship between estimated elevation (depth) and rotation magnitude. This is the main finding. Since a broad range of correlations, from less than 10% up to 95%, has been found, this means that in our interpretation we have probed the full spectrum from metric to affine-in-depth interpretations. This spectrum is clearly depicted in figure 3.2. We can see the correlation grow from upper left to lower right just by looking at the shape and size of the ellipse.

We have to make an additional remark here. The absence of a correlation does not exclude the possibility of a metameric class in which the degree of freedom has shifted completely to one of the variables. In fact, the lower right panels in figure 3.2 show an affine-in-depth curve with almost zero slope at the veridical point, which is exactly such a special metameric class. This means that both a metric and an affine-in-depth representation are compatible with a zero correlation result in those conditions. As a consequence, to distinguish between an affine-in-depth and a metric representation one should take a look at the variances itself. We consider the fact that the areas of the ellipses in figure 3.2 clearly co-vary with the correlation as additional support (but not proof) for our classification of the data in affine-in-depth and metric representations.

General Discussion

Summary of experimental results

We investigated the relation between the perception of motion and the perception of structure, using stimuli consisting of planar patches rotating in depth. Subjects had to match both depth and motion of two of these patches. What we found is, that although the attitude and rotation of a plane are sometimes poorly perceived, the relation between these entities is usually very well perceived. For large elevations and small rotations, the information available in the stimulus seems to be too noisy to allow for a full disentangling of motion and structure from the image flow. Thus we have an affine-in-depth-like representation. We confirmed this by showing that the correlation between matched elevation and matched rotation magnitude becomes very large under certain conditions. For larger rotation magnitudes and smaller elevations, subjects performed as if they were using a metric representation.

Geometry versus Tolerances

In the introduction to this paper we mentioned the 'affine versus metric' discussion. We indicated that the usual reasoning about representations is based on minimum configurations and geometric arguments: two parallel views are sufficient for an affine-in-depth representation, whereas three views are needed for a metric representation. In our experiments, however, we used a much larger number of views and still were able to find metric as well as affine-in-depth-like behaviour. We will try to explain what we think is the basic theory behind these results.

Suppose we are presented with two different but similar views of a spatial structure, one after the other. What if we add another view *between* these two views? We might get a more fluent impression of motion (assuming the total presentation time stays the same), but not too much new information is added. We are basically provided with a more accurate sampling of the same information. All of this of course depends on the definition of 'similar views'. So, what is a reasonable definition?

Every pair of views defines a velocity field. Every triplet of views also defines an acceleration field. Whereas the human visual system exhibits reasonable accuracy in velocity extraction, at least in optimised conditions (McKee, 1981), the detection of acceleration has been shown to be quite poor (Snowden and Braddick, 1991; Werkhoven, Snippe and Toet, 1992; Snippe and Werkhoven, 1993). We know that acceleration information is needed to derive a full (metric) reconstruction of a moving object, if motion parallax is the only cue available. Therefore, a poor coding of accelerations in the visual system would severely impair this reconstruction, effectively leaving a time-sequence of two-view reconstructions. We think that this continuum of affine-in-depth representations is the main descriptor of human structure-from-motion performance. The extent to which the (poorly coded) acceleration information can be utilised to distinguish individual affine-in-depth representations in such a sequence determines the amount of metric information that is available. Eagle & Blake (1995) succeeded in modelling the performance with certain metric and affine-in-depth (they call this relief) tasks in terms of such low-level motion sensitivities. They compared thresholds for discriminating dihedral angles (a metric task) and detection of non-planarity (an affine-in-depth task) and concluded that no significant differences in threshold were found between the metric- and relief-structures tasks *when plotted in terms of image-motion sensitivity*. Although their analysis is not easily applied to our data, their approach is a good example of how one might incorporate knowledge about image-motion sensitivities into an analysis of 3D structure-from-motion perception. A similar analysis was published by Werkhoven and van Veen (1995). They showed that the characteristics of 'low-level' velocity extraction determines those of 'high-level' relief extraction.

From the above arguments we conclude that those stimulus conditions that display an affine-in-depth-like behaviour effectively consist of two different views with a large number of 'similar' views in-between. It is currently not clear which parameters actually control this similarity. An analysis similar to Eagle & Blake or Werkhoven & van Veen might be useful. We have some indications that the differential invariants of the flow fields (see appendix B) play a role as well. This is certainly an issue that should be pursued in further research. The important message is that the performance with these 3D tasks is controlled by 2D motion sensitivities (noise tolerance).

Simulations

Let us take a more detailed look at the continuum of affine-in-depth representations proposed above. For each pair of consecutive views, there is one structure-motion relation between the views. If we plot the reconstructed elevation of the plane versus the reconstructed amount of rotation between the views, the metameric structure-motion class is represented by a single curve. The next pair of frames is also represented by a curve. The combined graph of all pairs will look like figure 3.6. We can see that the curves intersect at one point, which is the veridical (or metric) solution that we used to generate the graph. The curves differ of course, but the differences might be too small to be distinguishable from noise effects in the visual system. The distance between two curves in the graph is in general not a good measure of their difference in terms of a tolerance analysis.

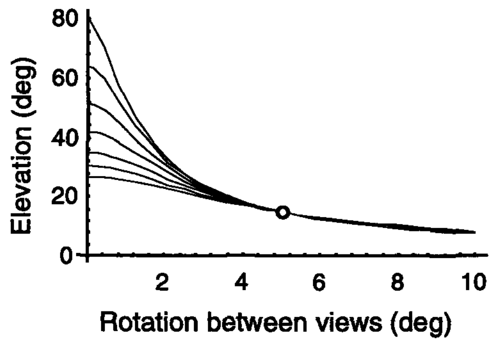


Figure 3.6
Set of reconstructed metameric structure-motion relations for a rotating planar patch. The recovered elevation is plotted against the recovered amount of rotation between two views. Elevation of the plane was 15°, twist was 0° and the rotation magnitude equalled 70°. Fifteen consecutive views differing by 5° were used. Only 7 curves are visible because the first part of the motion mirrors the second part.

To get an impression of the noise sensitivities involved in the reconstruction process, we ran several simulations. We had to make some arbitrary choices here (noise model, reconstruction algorithm), but the results were the same, at least qualitatively, for a large range of choices. The quantitative differences between the results obtained with different models make it very hard to compare the results of simulations with those of human observers. To avoid this problem, we merely use these simulations as an illustration of the questions at hand. Our algorithm was supplied with three views of a rotating planar patch to which we had added relative velocity noise (better: displacement noise). The algorithm computed two affine-in-depth solutions from the two pairs of views and then computed the metric solution from the intersection of constraints (see Koenderink and van Doorn, 1991). This method makes no assumptions about the constancy of angular velocity or about fixed-axis motion. We also tried the Hoffman and Bennett algorithm (1986) and obtained very similar results.

Figure 3.7 shows some representative results of these simulations (30% velocity noise). We see a cloud of solutions stretched along the direction of the affine-in-depth curves. This is what we would expect on the basis of the above arguments. The difference between the two metameric structure-motion relations (two relations, because we have three views) is of course reflected in the shape and size of the covariance ellipse. The size of the ellipse is a measure of how well the metric solution is represented (note that the median of the solutions generated by the algorithm is close to the metric solution). The amount of elongation along the affine-in-depth curves is a measure of how similar the pairs of views are, at least in the sense of a tolerance analysis. The linear correlation of the reconstructed elevation and rotation magnitude is -0.41, -0.82 and -0.50, for examples A, B and C respectively, all significantly different from zero at the 1% level (in fact much better) and all negative! From this we conclude that example B is closest to an affine-in-depth representation.

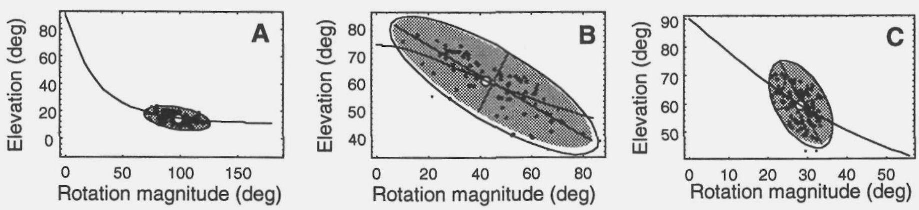


Figure 3.7
A few representative results of the simulations are shown. All three examples show 100 trials, using 30% velocity noise. For an explanation of symbols refer to figure 3.2. **A:** 15° elevation, 98° rotation magnitude, 0° twist. **B:** 60° elevation, 42° rotation magnitude, 30° twist. [16 out of 100 trials did not result in a solution (no intersection of constraints). These points were discarded.] **C:** 60° elevation, 28° rotation magnitude, 0° twist.

Results from the literature

The relation between perceived motion and perceived depth has been reported on before. In 1991, Cortese and Andersen published a study on the recovery of 3D shape from deforming contours. Using the silhouettes of rotating ellipsoids, they presented results confirming that the '...recovery of 3D shape from smooth, deforming contours is dependent on the perceived extent of rotation'. Petersik, 1980, concluded that 'The two judgements [of rotation direction and mean depth rating] are not completely independent'. In his 1991 paper, Petersik investigated whether information regarding rotation magnitude develops prior to, in parallel with, or after the recovery of structure. No definitive results on this were presented, but he argues in favour of the latter two alternatives. Liter and Braunstein (1994) found that in structure-from-motion displays in which the most frontal view of an object was not shown, the rotation magnitude was overestimated and at the same time the perceived depth was underestimated (dihedral angle was judged to be flatter). Our results are compatible with all these findings. Furthermore, for the specific conditions that we measured, we have shown to be able to understand some of the dependencies between motion and structure perception in terms of a mathematical model.

In 1957 Gibson & Gibson, using horizontally slanted planes mounted on a turntable, reported that 'Misidentification of the shape [=depth gradient in direction of motion] ... was accompanied by anomalies in the perception of [change in] slant [=depth gradient]'. In other words, they found a relation between errors made in judging slant and in judging the temporal derivative of slant. This raises a related and interesting question about the interpretation of our experimental data: did subjects match rotation magnitude (motion) or did they match extreme views (orientation)? There is no way in which the current data can distinguish between these two possibilities, although the subjects were formally instructed to match rotation magnitude. Matching the horizontal component of the depth gradient of the planes (Gy) in their most extreme views is equivalent to a match of the rotation magnitudes, at least in this experiment[16]. Note, however, that because the two planes rotate out of phase, the extreme views occur at different points in time, which prevents a direct comparison of these views. The ultimate consequence would be that we can not distinguish between matching of motion (rotation magnitude) and structure (elevation), and matching of structure alone (complete attitude of the plane in its extreme view). Although this might be injurious to an interpretation of the data in terms of the visual system's processing of structure information and motion information and the connections between those assumed processing pathways, it does not impair our conclusions about the metamerism as such. Furthermore, based on the introspective reports from the subjects, we have no reason to believe that they did anything else but matching of rotation magnitude.

We strongly believe that we have demonstrated a marked inverse relationship between perceived rotation magnitude ('motion') and perceived elevation ('depth'). The fact that all 64 experimental conditions (4 subjects times 16 conditions) showed a negative correlation between the two settings is a very powerful illustration of this point. This exercise shows that the distinction between 'metric' and 'affine-in-depth' performance is not as strict as we used to believe. When the tolerance of the visual system to noise-effects is taken into account, or the tolerance of any algorithm for that matter, a gradual transition occurs between the two representations.

[16] In this experiment, the possibility of matching the orientation of the extreme views can not be distinguished from matching rotation magnitude and elevation, because the twist of both planes is equal. Thus, one has to manipulate the twist in order to discriminate between the two possibilities. We chose not to do so because of the additional complexity of the mathematical description; when the twists differ, the affine-in-depth curves corresponding to the test and reference stimulus differ as well. This severely hampers a correct interpretation of the matching data.

Appendices

A: decomposition of a general rotation

Any rotation can be written as the product of a rotation around a fronto-parallel axis followed by a rotation around the viewing direction. To understand this decomposition one should realise that any rotation R_ρ can be written as a reflection M_U in a plane U containing the rotation axis ρ, followed by a second reflection M_W in a plane W, also containing the rotation axis. The dihedral angle between U and W has to equal half the magnitude of the rotation. Second, we can always choose the plane U such that it also contains the viewing direction (z axis). This fixes the plane W and thus the intersection ω of W with the fronto-parallel plane (xy plane). Third, let an extra plane V be the plane containing both the z axis and the ω axis. We can then write the following:

$$\begin{aligned} R_\rho &= R_{U\cap W} = M_U \circ M_W \\ &= M_U \circ I \circ M_W = (M_U \circ (M_V \circ M_V) \circ M_W) \\ &= (M_U \circ M_V) \circ (M_V \circ M_W) = R_{U\cap V} \circ R_{V\cap W} \\ &= R_z \circ R_\omega \end{aligned}$$

q.e.d.

B: planar patch rotating in depth

Suppose we have a parallel displacement field, specified by three moving object points (shown in two parallel projected views): an origin and two others, P and Q, needed to establish the parallelism of flow. Let us define the flow direction to be the y direction. We can assume that from view 1 to view 2 a point P underwent a rotation around the x-axis with magnitude ρ. We can then write down an equation relating the two projected y co-ordinates of P with each other (the x co-ordinates do not change of course):

$$Py_2 = Py_1 * Cos(\rho) - Sin(\rho) * Pz_1.$$

This can be rewritten in terms of the depth co-ordinate:

$$Pz_1 = \frac{Py_1 * Cos(\rho) - Py_2}{Sin(\rho)}.$$

So, if we know (or guessed) the amount of rotation ρ, then we would know the depth of point P. So far so good, but how do we know about ρ? If we try to combine these equations for several points, we see that we add one equation and one new unknown (the z-value) for each point. It is clear now that unless we have more information, we cannot solve the puzzle completely. We might know ρ from other cues[17], and it would also be very helpful to have specific information about the structure, such as the angle between two given line segments. We must conclude that any solution that can be found will be ambiguous both in depth ('z') and motion ('ρ'). The basic problem that is left now consists of only two views of three points, which is exactly defined by a planar patch rotating in depth. If there are more points in the projection, we can simply treat each triplet of points as such a planar patch.

Let us therefore assume that we have two parallel projected views (1 and 2) of a planar patch which differ only by a rotation with magnitude ρ around an axis in the image plane; again we call this axis the x-axis. The patch is described by its position (it contains the origin) and its depth gradient in view 1:

$$\vec{G}_1 = \begin{pmatrix} G_{x_1} \\ G_{y_1} \end{pmatrix}.$$

This means that the depth 'z' of each point P in the patch is given by

$$Pz_1 = \begin{pmatrix} Px_1 \\ Py_1 \end{pmatrix} \bullet \vec{G}_1 .$$

In the previous paragraph we wrote down an equation relating the two y co-ordinates of a point in two different views. If we do this for two points P and Q (which, together with a third point defining the origin in both views, constitute a planar patch) we get:

$$\begin{aligned} Py_2 &= Py_1 * Cos(\rho) - Sin(\rho) * \left(\begin{pmatrix} Px_1 \\ Py_1 \end{pmatrix} \bullet \vec{G}_1 \right) \\ Qy_2 &= Qy_1 * Cos(\rho) - Sin(\rho) * \left(\begin{pmatrix} Qx_1 \\ Qy_1 \end{pmatrix} \bullet \vec{G}_1 \right) \end{aligned}.$$

With some algebraic manipulation, this pair of equations transforms into another set, which is equivalent but easier to use and understand:

$$\vec{G}_1 = -\frac{1}{Sin(\rho)} \begin{pmatrix} a \\ b - Cos(\rho) \end{pmatrix}$$

$$a = \frac{Py_2 \cdot Qy_1 - Py_1 \cdot Qy_2}{Px_1 \cdot Qy_1 - Py_1 \cdot Qx_1} \quad , \quad b = \frac{Px_1 \cdot Qy_2 - Py_2 \cdot Qx_1}{Px_1 \cdot Qy_1 - Py_1 \cdot Qx_1}$$

[17] In an ego-motion condition the observer moves around the object instead of the object rotating in front of the observer. For an analysis of optic flow this makes no difference at all, since only relative motion can be considered. However, a moving observer might know his own motion (and thus ρ) from other sources of information, like proprioceptive information from the muscles. In binocular stereo, knowing ρ boils down to knowing the vergence angle.

The expression for the depth gradient of the patch is now in a very simple form, depending only on two functions of known projected co-ordinates and on the unknown rotation magnitude ρ. If we realise that the relation between depth gradient and rotation magnitude also holds for the veridical values of these parameters (the values that were used to generate the displays in the first place), we can write

$$G_x^{rec} \cdot Sin(\rho^{rec}) = G_x^{ver} \cdot Sin(\rho^{ver}) = -a$$
$$G_y^{rec} \cdot Sin(\rho^{rec}) - Cos(\rho^{rec}) = G_y^{ver} \cdot Sin(\rho^{ver}) - Cos(\rho^{ver}) = -b$$

in which the superscripts 'rec' and 'ver' denote reconstructed and veridical respectively. These equations clearly show the inverse relation between reconstructed depth gradient and recovered amount of rotation, at least for small rotations. The parameters **a** and **b** can be identified as first order spatial derivatives (differential invariants) of the finite (because of finite rotations) flow field: (-CURL) and (DIV+1) respectively (see Koenderink, 1986). The other two first order differential invariants, DEF and the orientation of the axis of deformation ∂, are coupled to DIV and CURL. This coupling between the invariants is due to the fact that we use parallel projection instead of perspective projection. DEF equals the square root of the sum of squares of DIV and CURL, whereas ∂ equals half of the arc tangent of CURL/DIV.

If we write down the displacement along the y co-ordinate of any point R of the planar patch, we get

$$Ry_2 - Ry_1 = -CURL \cdot Rx_1 + DIV \cdot Ry_1 .$$

Thus, if we use the origin and two other points, we can predict what the projected displacement of a fourth point R has to be, to be identified as part of the same planar patch. If we reverse the argument here, we can make use of these extra points to get better estimates of the DIV and CURL, since these entities are the same for all triplets of points that are part of one and the same planar patch. Of course we need to know that all these points do indeed belong to the same planar patch in the first place!

If we extend this argument a little further, we see that when we have a smooth flow field instead of a set of moving feature points, local estimates of DIV and CURL would be sufficient information to be able to compute the above relation between local depth gradient (attitude) and rotation magnitude. In order to make that local analysis, we have to make the reduction of the problem (removing translation and image-plane rotation) on a local basis as well. Translations seem to be no problem because they can easily be corrected for, but image plane rotations do cause CURL changes. However, since DEF is coupled to DIV and CURL in the reduced case of such a planar patch rotating in depth, we are able to detect this and correct for it. We can even invert the problem and use this relation to estimate the (local) amount of image plane rotation itself. If we combine the local analyses at two different parts of a rigidly moving object, we can compare the local attitudes without knowing them.

4

SPATIAL CONFIGURATION AFFECTS JUDGEMENTS OF 3D ANGULAR SPEED

Abstract

In a series of experiments subjects view two three-dimensional objects presented simultaneously on a computer screen, each object rotating around an axis in the fronto-parallel plane. Subjects are required to judge which of the two objects has the higher angular speed. The results enable us to investigate how the spatial configuration of these objects, relative to each other and to the observer, affects the observer's speed judgements. The objects used are random polyhedrons with approximately equal extension in all directions. We find that the angular speed of rotation around a vertical axis is underestimated relative to the angular speed of rotation around a horizontal or a diagonal axis. This anisotropy is slightly smaller when the objects are presented above each other than when they are presented side-by-side. In further experiments subjects view one object foveally and one object peripherally. We find that subjects underestimate the angular speed of peripheral objects compared to foveal ones, and that the underestimation increases with increasing eccentricity. However, this effect can be cancelled by appropriate spatial scaling of the objects. The discrimination threshold is found to be largely independent of the specific spatial configuration employed; it is typically 20%. An attempt is made to interpret the results in terms of known two-dimensional image processing properties of the visual system. Some implications for the perception of three-dimensional structure are discussed.

Introduction

The structure of a moving object can only be recovered from optic flow if the object's orientation changes with respect to the viewing direction[18]. Consequently, such relative rotations between object and observer are of substantial importance for an observer (either an artificial or a biological visual system) interested in the three-dimensional spatio-temporal layout of a scene. It is therefore quite remarkable that the perception of 3D rotation has received so little attention in the past (we believe that the work on mental rotation is an independent area of research).

Rotations are commonly described by two separate quantities: the orientation of the axis of rotation and the angular speed. We are aware of two studies that have analysed the perception of the axis of rotation. Shiffrar and Shepard (1991) investigated the role of reference frames in the perception of cube rotations. Employing a same-different paradigm they found that the speed and accuracy of observers' comparisons of cube rotations decreased as the axes of successive rotations departed from the canonical axes of the environment, or even more sharply, from the symmetry axes of the cube. Pollick,

[18] Viewing direction should be read as the visual direction in which the object is seen; the viewing direction is not necessarily equal to the fixation direction of a human eye. An object that objectively *translates* through space always *rotates* relative to the viewing direction, except when it is moving along the viewing direction itself.

Nishida, Koike & Kawato (1994) used a pointing task to investigate the perceived rotation axis of moving constellations of dots as a function of axis orientation and type of display. They found that subjects were sensitive to both the slant and tilt of the rotation axis, and that the bias in judged slant and the variability in judged tilt depended on the type of structure-from-motion display. Perceived angular speed was studied by several other authors. We summarise the clearest findings. Kaiser (1990) found that observers are able to discriminate angular velocities with a competence approaching that of linear velocities. She also reported that larger objects appear to rotate faster than smaller ones. Kaiser and Calderone (1991) found that edge-transition rate and projected velocities of surface elements positively bias perceived angular velocity, although judgements are mainly determined by the true angular velocity. Petersik (1991) basically confirmed the findings of Kaiser et al. (1990, 1991). Liter, Braunstein and Hoffman (1993) found that judged rotation magnitude increased both with increasing objective rotation magnitude and with increasing depth of the rotating object (both corresponding to increased relative image motion), in accordance with the other studies.

Our study concentrates on the influence of spatial configuration on the perceived angular speed of 3D objects. We believe that a few basic spatial aspects deserve to be studied more systematically, as we will point out in the introductions to the various experiments. We present results about the effects of the orientation of the axis of rotation, spatial ordering, retinal eccentricity and size on perceived angular speed. Where possible, we attempt to interpret our findings in terms of known 2D processing characteristics, and we speculate how these effects might influence the perception of moving objects in general.

Experiment 1

Axis of rotation

This experiment was designed to investigate the possible existence of an anisotropy in the perception of angular speed as a function of the orientation of the (fronto-parallel) axis of rotation. Although this problem has been studied before, the results are not very convincing, as we will now explain. Kaiser (1990) found no evidence for an anisotropy. In her study she used computer-generated rotating cubes, whose faces were distinguishable only by their different luminance levels. The orientation of these cubes was highly constrained: the axis of rotation was always perpendicular to one of the faces, and at the onset of rotation one of the other faces was either aligned with the projection plane or at a 45° angle with it. It is not clear whether or not the fact that the shape and orientation of these objects are easily recognised (equal sized edges and facets, 90° angles, etc.) calibrates the information available from the structure-from-motion cue in such a way that potential anisotropies are masked. In one condition she used a bisphenoid (double tetrahedron) for which similar arguments hold. By using random polyhedral objects with roughly the same extension in all directions we try to prevent these confounding effects of shape recognition and the occurrence of specific orientations. Van Veen, Kappers, Koenderink and Werkhoven (in press) have shown recently that the distinction between such random objects and (easily recognised) cubes can indeed lead to large perceptual differences. The only other study that has investigated the effect of orientation of the axis of rotation on judged angular speed is that conducted by Petersik (1991). He used rotating spheres, which were either textured with random dots or filled with random dots. In his first experiment (discrimination task) he found that the axis of rotation (either horizontal or vertical) had weak effects on the point of subjective equality (PSE). However, he also found significant differences between PSE and POE (point of objective equality) in symmetric conditions (both objects rotating around the same axis), which testifies to the existence of a response bias. This severely weakens his conclusions regarding possible effects of axis orientation on perceived angular speed. In our experiment we deliberately include symmetric conditions, which allows us to establish the reliability of the data.

Methods

Experimental Set-Up

The stimuli were generated on an Apple Macintosh IIfx computer and displayed on a 71 Hz Radius TPD/19 high spatial resolution (82 dpi) grey-scale monitor (1152 by 882 pixels), subtending approximately 20° by 16° visual angle at a viewing distance of 1 metre. There was no mechanical restriction on head movements, but the subjects were required to keep their preferred eye just in front of the centre of the screen at the prescribed distance of 1 metre. The non-preferred eye was covered with a black eye patch. The room was dimly lit such that the subjects were still able to see the monitor and the table on which it stood.

Stimulus

Two convex polyhedral objects, horizontally separated by 4° (centre-centre), were shown simultaneously on a computer display (see figure 4.1). They had approximately the same extension in all directions; hence their size can be described with one parameter: the volume was fixed at 8.2 cm^3, yielding a maximum projected diameter of 2° visual angle. Each object had 12 planar triangular facets, the edges of which were presented in high contrast as light line segments on a dark background. We used parallel projection (70 msec per frame), hidden lines were removed (resulting in opaque, wire-frame-like objects) and anti-aliasing was not applied. New objects were generated for each trial to ensure that the configuration of facets was never the same. Each stimulus consisted of a test object and a reference object. Their order (i.e. left or right position) on the screen was chosen randomly for each trial. The precise description and method of generation of these randomised wire-frame objects can be found in Van Veen et al. (in press).

The objects are made to rotate in depth for a short period (5 sec) and at moderate rates; the reference object rotated with a speed chosen randomly between 43 and 71°/sec, and the speed of the test object was varied in relation to this. Pilot experiments had shown that a five second presentation time was about the optimal value: performance in general did not improve with longer presentation times. Both objects rotated about their own rotation axis, which passed through their respective centres-of-mass. All rotation axes were fronto-parallel and could have one of three different orientations: horizontal (0°), vertical (90°) or diagonal (45°). See figure 4.1. The orientations could be different for the two axes. The direction of rotation was not varied; the movement of the frontal part of the objects was always rightwards (0° axis), downwards (90° axis), or rightwards and downwards (45° axis).

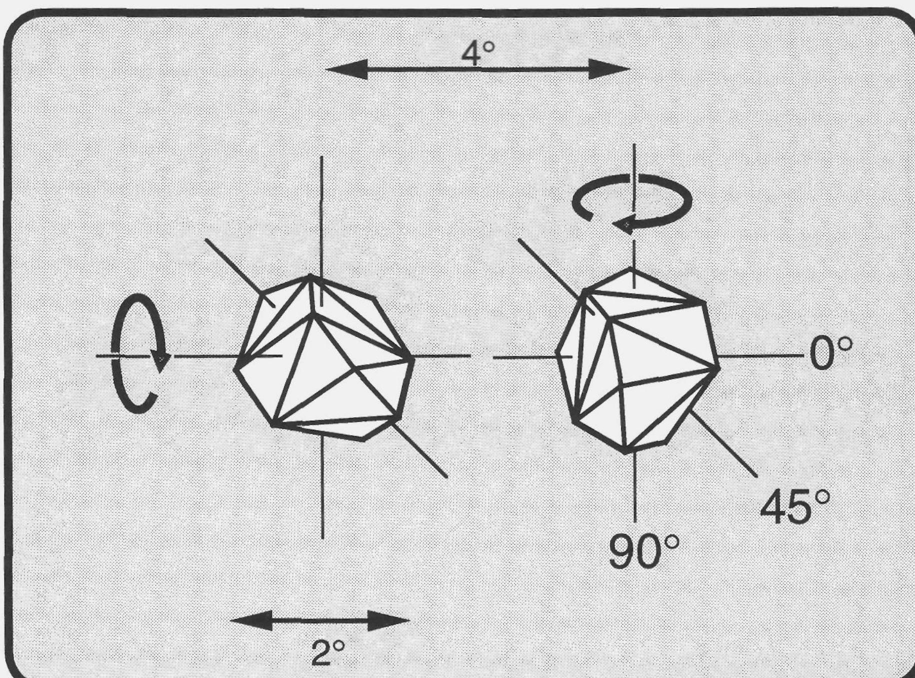


Figure 4.1
This drawing shows the stimulus geometry utilised in Experiment 1. Two horizontally separated polyhedrons are simultaneously projected on the screen. The object surfaces consist of randomly shaped triangular facets. These random polyhedrons have no specific orientation because they have approximately equal extension in all directions. Their maximum projected size is about 2° and their centre-centre separation is 4°. Each object rotates around its own fronto-parallel rotation axis, which is either horizontal (0°), vertical (90°) or diagonal (45°).

Procedure

The subjects were asked to indicate, via the arrow keys on the keyboard, which of two simultaneously rotating objects was rotating faster. Subjects were free to move their eye from one object to the other. No feedback was given.

For each condition we measured the percentage of trials in which the subject indicated that the test object had the greater angular speed, as a function of the (logarithmic) ratio of the angular speeds of test and reference object. An adaptive psychometric procedure (Werkhoven and Snippe, in press) was used to estimate the two parameters that characterise the psychometric curve (cumulative normal distribution): the Point of Subjective Equality (PSE) μ and the threshold σ. The PSE equals the log of the objective speed ratio at which the two objects appear to rotate equally fast; a value different from unity represents a certain bias in the subjective angular speed of objects rotating around different rotation axes. The threshold is the difference between the PSE and the log of the objective speed ratio needed to successfully discriminate the two angular speeds (84% correct criterion). After each trial, new maximum likelihood estimates of the PSE and threshold were calculated. The next trial was placed either at $\mu+\sigma$ or at $\mu-\sigma$, with equal probability (except for the first 15 trials for which the log of the objective speed ratio was picked randomly from the range -ln(2)..ln(2)). We used 100 trials per condition per subject. We checked the convergence of each sequence of trials and (sporadically) rejected the result when necessary. To allow for an easier interpretation of the results, we do not present the logarithmic μ and σ, but we translate their values into angular speed ratios (also called PSE) and increment threshold percentages respectively.

Subjects

Three subjects participated in this experiment: HV, SP and NW. All subjects were myopic. Their corrected vision was normal. All subjects had had previous experience with similar displays and tasks. The three subjects viewed the stimuli with their right eye.

Design

The combination of 3 possible orientations for both rotation axes defined 9 different conditions: three symmetric ones (0-0, 45-45 & 90-90) and six asymmetric ones (0-45, 45-0, 45-90, 90-45, 90-0 & 0-90). The asymmetric conditions can be grouped into three pairs (e.g. 45-90 & 90-45). As we expected, we never found any systematic difference between the members of a pair, so we decided to use the average values (taking sign into account). The resulting 6 conditions were measured sequentially in random order in 15 minute blocks. Each condition was measured three times by each subject.

Results

The left panel of figure 4.2 depicts the increment thresholds. The values are roughly between 15% and 25%, with an overall average of 18.3±0.7%. The results fail to reveal significant effects of rotation axis orientation on the thresholds. The right panel of figure 4.2 shows the PSEs for the different conditions and subjects. As we expected, the symmetric conditions (0-0, 45-45 & 90-90) demonstrate PSEs compatible with one, which is the veridical value or POE. This is evidence in support of the symmetry of the psychometric curves and the absence of response biases. The PSE is significantly different from the POE in the asymmetric conditions (0-45, 45-90 & 90-0). All three subjects show a remarkable underestimation of the 3D angular speed of rotation around a vertical axis when compared with rotation around a horizontal (90-0) or diagonal (45-90) axis. They also underestimate the angular speed of rotation around a diagonal axis relative to rotation around a horizontal axis (0-45), but this effect is somewhat weaker. Apparently, the vertical rotation axis is 'the odd one out'.

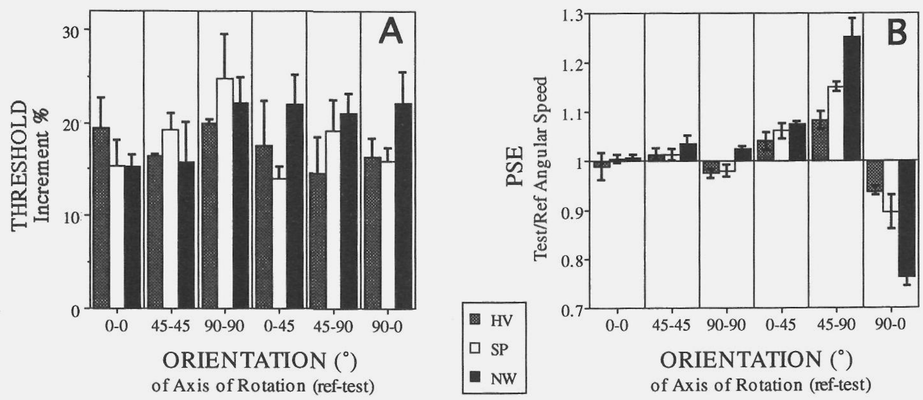


Figure 4.2
Results of Experiment 1. Different subjects (3) are indicated by differently textured bars. The horizontal axis lists the six different conditions, denoted by the orientations of the reference objects' rotation axis and test objects' rotation axis respectively. Error bars represent one standard error of the mean. Panel A: Increment threshold percentages, i.e. the percentage increase in test object angular speed above the PSE needed to judge the test object's angular speed to be faster than the reference object's angular speed in 84% of the trials. Panel B: Point of subjective equality, i.e. the objective ratio of test object's angular speed divided by reference object's angular speed at which the objects subjectively appear to rotate equally fast. Thus, a value smaller than unity refers to a relative underestimation of the reference object's angular speed.

Discussion

Kaiser (1990) and Petersik (1991) both report average differential thresholds of 12%. Note, however, that their criterion is different from ours in the sense that their differential threshold corresponds to two-thirds of the sigma of the (supposedly) underlying cumulative normal distribution, whereas our increment threshold corresponds to one sigma. The factor two-thirds precisely explains the difference between their threshold values and ours, and we therefore conclude that the results of these three different experiments confirm each other, as far as the discriminability of 3D angular speed is concerned.

Although the strength of the effect of axis orientation on the PSEs is different for different observers, the results undeniably prove the existence of an anisotropy in the judged angular speed as a function of the orientation of the rotation axis. As we discussed in the introduction to this experiment, previous studies by Kaiser (1990) and Petersik (1991) did not reveal a significant anisotropy of this kind. Although we pointed out that certain aspects of stimulus and methodology could have masked potential biases in those studies, it is currently not clear which parameters control the presence or absence of such an anisotropy.

Interestingly, a similar anisotropy has been found for linear velocities. Brown (1928, 1931) investigated the influence of the direction of movement on the matching of velocities in the fronto-parallel plane. On the basis of a rather limited set of data he concluded that movement in the vertical direction was perceptually overestimated with respect to horizontal and diagonal (45°) movement. The bias was quite large, approximately 30-40% and 10% respectively. This corresponds remarkably well with what we found, namely that the angular speed of rotation around a horizontal axis (vertical image motion!) is overestimated with respect to the angular speed of rotation about vertical and diagonal axes. Even Brown's observation that the difference between vertical and horizontal movements is larger than the difference between vertical and diagonal movements is confirmed by our data. Thus, the anisotropy of perceived *2D linear* velocity as a function of movement direction appears to be the underlying factor that controls the anisotropy of perceived *3D angular* velocity as a function of the orientation of the rotation axis. However, this might not be the complete explanation. To compute angular velocity one should scale linear velocity with the distance to the rotation axis. One way of doing this, is to divide the maximum projected linear velocity of a feature point by the total length of its projected trajectory. However, vertical dimensions are often overestimated relative to horizontal ones. The first observation of this so-called Horizontal-Vertical Illusion (HVI) is usually credited to Fick (1852). Two interesting and more recent studies that summarise the older literature and present quantitative results are reported by Künnapas (1955) and Cormack & Cormack (1974). Depending on various known and unknown factors the magnitude of the effect ranges from a few percent up to 20%. Enormous differences between subjects are also very common (e.g. see Pollock and Chapanis, 1952). If the HVI is applicable to our stimuli, it partially cancels the linear velocity anisotropy. To test whether the above-mentioned relationships between these entities are still valid in a complex stimulus situation like the one we used one needs to perform an experiment that simultaneously measures an object's perceived linear velocity, perceived angular velocity and perceived size.

Experiment 2

Spatial Ordering

At first sight the spatial ordering of objects on the screen is irrelevant for judgements of relative 3D angular velocity. However, in line with the discussion about the HVI started in the previous section, we should bear in mind that spatial ordering has been shown to influence the magnitude of the horizontal-vertical illusion. Pollock and Chapanis (1952) measured the apparent length of black lines on a white background, and found that the overestimations of vertical length were smaller when the two lines are above one other instead of side-by-side[19]. Thus, we decided to verify whether the relative underestimation of 3D angular speed of rotation around *vertical* axes is connected with the *horizontal* separation of the objects. We repeated Experiment 1 with the objects separated vertically instead of horizontally, and we also repeated the original Experiment 1 with the horizontally separated objects to check consistency over time.

Methods

The methods were completely identical to those employed in Experiment 1, except for the positioning of objects on the screen. In Experiment 1 the objects were horizontally separated (run 1). We completed two additional runs; run 2, in which the objects were separated vertically; and run 3, in which the objects were separated horizontally again. Runs 2 and 3 were not started before Experiment 1 was finished, and were completed by the same three subjects who performed Experiment 1.

[19] There is a hidden subtlety here. Contrary to most experimental investigations of the HVI, the lines used by Pollock and Chapanis were *not connected*. This is an important factor, because in most studies T- or L-shaped figures are used, and it is known that in such figures (consisting of *connected* line-elements) an additional illusion makes the situation more complicated. The so-called 'overestimation of the dividing line' (Künnapas, 1955) causes the ovestimation of the length of the vertical line in ⊢ to be much less than in ⊥. Although this effect is in the same direction as the effect found by Pollock and Chapanis for unconnected lines, we think it has a different cause.

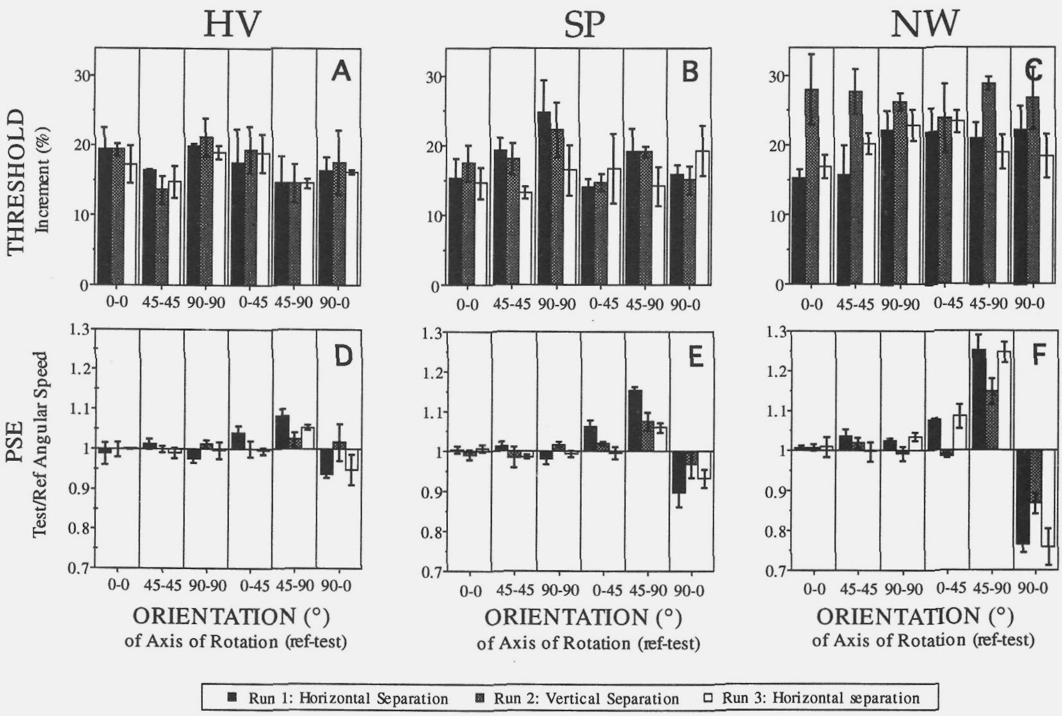


Figure 4.3
Results of Experiment 2. The different bars indicate different runs: in run 1 the objects were separated horizontally (these are actually data from experiment 1!); in run 2 the objects were separated vertically; and run 3 is a repetition of run 1. The horizontal axis lists the six different conditions, denoted by the orientations of the reference objects' rotation axis and test objects' rotation axis respectively. Error bars represent one standard error of the mean. The upper row of panels (A, B & C) gives the increment threshold percentages for the different conditions for subjects HV, SP and NW. The lower row of panels (D, E & F) gives the corresponding PSEs. A value smaller than 1 refers to a relative underestimation of the reference object.

Results

The threshold data presented in the upper panels of figure 4.3 show almost no effect of run number. Only for subject NW does the second run appear to be slightly more difficult than the other runs (highest thresholds in all 6 conditions). Clear effects of spatial ordering or time (learning effects) cannot be identified. This is in contrast to the results found for the PSEs (lower panels). All three subjects show a smaller anisotropy with vertically separated objects (run 2) than with the original horizontal separation (run 1). A small subject-dependent learning effect can also be identified. The difference between the first and last run varies between subjects from considerable (SP) to small (HV) or none at all (NW). We must emphasize that the authors HV and NW participated in many pilot-experiments before they were used as subjects in Experiments 1 and 2, so we did not expect them to display considerable learning effects.

Discussion

Although the effect of spatial ordering is not equally strong for all three subjects, the anisotropy with vertical separation is clearly less than it is with horizontal separation. We can only speculate about the reason for this. However, it is clear that the reduction of the HVI with a vertical separation instead of a horizontal one, as reported by Pollock & Chapanis, certainly does not explain the result we find here. On the contrary, such an effect would enhance the anisotropy rather than reduce it.

Note that a within-objects rigidity constraint, which occurs when both objects rotate about the same axis (0-0 in runs 1 and 3, and 90-90 in run 2), does not lower the thresholds, except possibly in the case of subject NW whose 0-0 thresholds are much lower in runs 1 and 3 than in run 2.

Experiment 3

Eccentricity and Size

When looking at the world around us we continuously make eye-movements, trying ceaselessly to direct our gaze towards interesting events in the visual field. For instance, while analysing the structure and motion of the rotating objects used in the previous experiments, subjects made pursuit eye-movements while tracking feature points, and they made saccadic eye-movements to other feature points to collect information from various places. Tracking feature points with our eyes reduces common motion components and enhances our ability to perceive the small differences in velocity that can reveal to us the three-dimensional structure of the environment. If we make the reasonable assumption that the basic pieces of information in structure-from-motion tasks are collected in small portions by the fovea and not in the whole visual field at once, we are confronted with the following problem: how is this spatially and temporally scattered collection of bits of information integrated into a judgement of, let's say, 3D angular speed? Although this is of course a very intriguing question, we will not try to answer it here. We prefer to go back to the assumption that information is collected mainly with the fovea. Hence, it might be interesting to see what happens to the judgements of angular velocity when the eye-movements are restricted. In the next experiment subjects are not allowed to make eye-movements from one object to the other. However, eye-movements in the neighbourhood of one of the objects are still allowed. Thus, features of this foveal object can still be tracked. This means that one object can be investigated with the same accuracy and strategy as previously (Experiments 1 and 2), whereas the other, peripheral object must be analysed with the eccentric parts of the retina. We vary the eccentricity and size of these objects and analyse how these factors affect the variability of the judgements of 3D angular speed and the perceived angular speed itself.

Methods

Unless mentioned otherwise, the methods were identical to those employed in Experiment 1.

Stimulus

Instead of varying the orientation of the rotation axes, we varied the size of the objects and their horizontal separation. We chose to make the rotation axes vertical. Furthermore, since we prescribed fixation in this experiment, the reference object was always in the foveal field of view whereas the test object was in the nasal field of view at different horizontal eccentricities. We made the objects transparent (i.e. hidden lines were not removed), because we wanted to avoid a situation in which the retinal image motion was predominantly in one fixed direction for longer periods of time. Of course, making the objects transparent meant that the direction of rotation became undefined (remember that we used parallel projections), but this did not appear to be a problem. The frame duration had to be increased slightly, namely to 80 msec per frame, so that larger objects could be dealt with. The angular speed of the foveal object was chosen randomly between 38 and 63°/sec.

Procedure

Subjects were not allowed to move their eye from one object to the other, in contrast to the first two experiments. More precisely, eye movements were permitted, but only within the vicinity of the foveal target. We call this type of fixation 'object fixation'. During the 400 msec preceding the appearance of the objects a small light square was shown to help the subjects direct their view to the foveal object (again, this is *not* a fixation mark in the usual sense). Although this foveal object always appeared at the same location, this location is not easily remembered on an empty screen in a dark environment. Because of the light square, subjects did not make unnecessary eye-movements at the beginning of the stimulus period. From observing the subjects' eye-movements during the experiment we learned that subjects were indeed able to restrict their view as prescribed.

Subjects

Two subjects participated in this experiment: GV and NW. Both subjects were myopic. Their corrected vision was normal. NW had been a subject in the previous experiments and was selected because out of the three subjects he showed the largest variability in the PSE. GV was a paid subject and ignorant about the goals of the experiment. Subject NW viewed the stimuli with his right eye whereas subject GV used his left eye.

Design

We varied 4 parameters defining 24 different conditions. The size of the foveal object could be either 1 or 2° (size is defined as the maximum projected diameter, see methods Exp. 1). The peripheral object was 1, 2 or 4° in combination with a 1° foveal object, and 2 or 4° in combination with a 2° foveal object. This gives us the combinations F1-P1, F1-P2, F1-P4, F2-P2 and F2-P4. The third parameter that we varied was the eccentricity of the peripheral object. For each size combination we picked four values ranging from as small as possible (the objects nearly overlap) to an eccentricity that was about four times the size of the peripheral object. This upper limit was chosen because the subjects complained that at still higher eccentricities they could no longer see a 3D object rotating in depth, but could see only motion in the image plane. Note that the transparency of the objects involves a redoubling of the number of visible line segments relative to the opaque objects case, which potentially hampers the correpondence of features between different frames. Nevertheless, the perceptual transition from a 3D rotating object to 2D image motion with increasing eccentricity appeared to the subjects to be similar for transparent and opaque objects. Finally, the fourth parameter that we varied was the viewing distance. At the standard distance of 2 metres the monitor subtended only 10°, which is not sufficient to support the 10° and 16° eccentricity conditions. Therefore, the conditions F1-P4 and F2-P4 were measured at 2 metres with eccentricities 3° and 6°, and at 1 metre with eccentricities 3, 6, 10 and 16°. The overlap allows us to check for effects of viewing distance. We could have measured all conditions at a viewing distance of 1 metre, but the angular resolution of the screen was of course much better when viewed from a distance of 2 metres. The 24 different conditions were measured sequentially in random order in 15 minute blocks. Each condition was measured three times by both subjects.

Subject NW measured 8 extra control conditions in a separate experiment (two repetitions). In this control experiment he was allowed to move his eyes freely. The conditions were: viewing distance 2 metres, both objects 1° and separations 1, 2, 4 and 8° (D2-F1-P1); and viewing distance 1 metre, both objects 2° and separations 2, 4, 8 and 16° (D1-F2-P2).

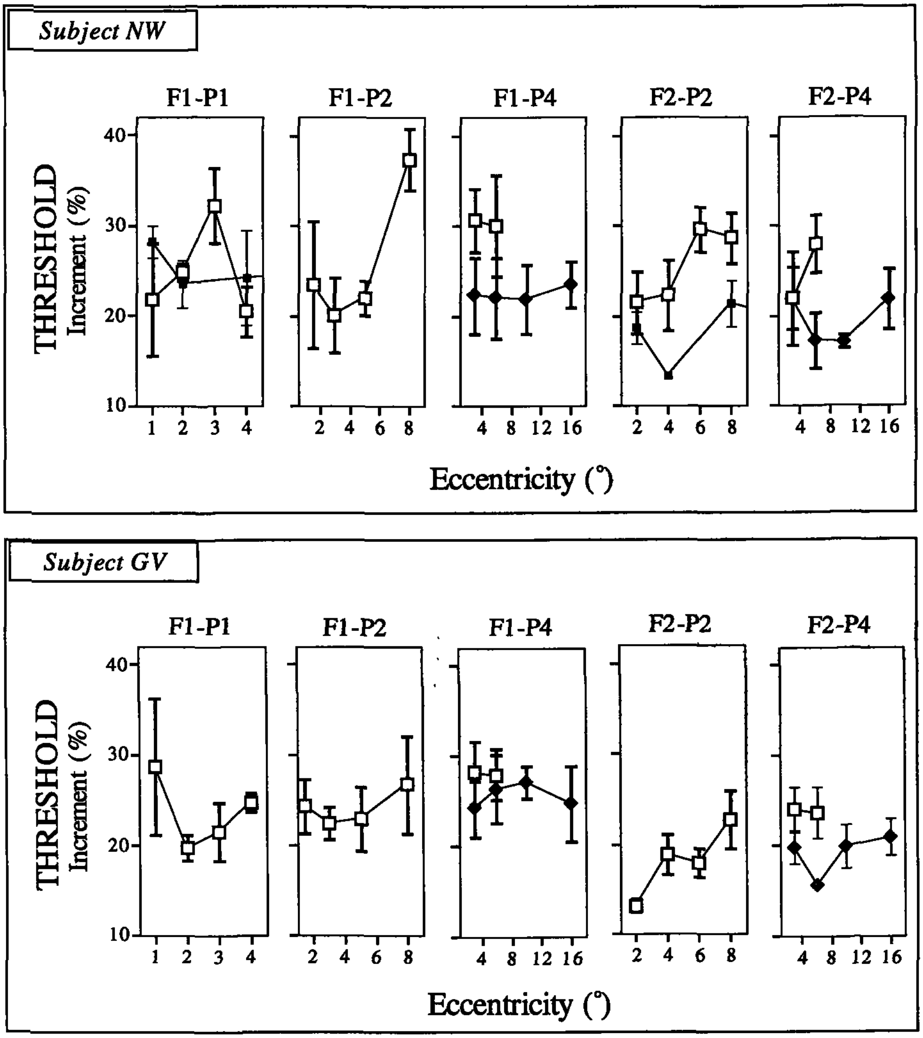


Figure 4.4
Increment thresholds found in Experiment 3. The upper and lower panels present the results for two different subjects. Each subpanel summarises the results for one combination of object sizes as a function of eccentricity. For example: the upper right subpanel, denoted F2-P4, plots the increment threshold percentages versus the eccentricity of the peripheral object for the conditions in which the foveal object's diameter is 2° visual angle (F2) and the peripheral object's diameter is 4° (P4). Error bars represent one standard error of the mean. The open symbols (squares) correspond to a viewing distance of 1 metre, whereas the filled symbols (diamonds; only in F1-P4 & F2-P4) correspond to a viewing distance of 2 metres. The small grey squares connected with dotted lines in the subpanels F1-P1 and F2-P2 of subject NW depict the control conditions without object fixation (see Methods section).

Results

As can be seen in figure 4.4, the thresholds are remarkably independent of eccentricity, with a few minor exceptions. One should bear in mind that the eccentricity values used are restricted to the range in which subjects reported that they indeed see a 3D object rotating in depth. At still higher eccentricities this percept changes to that of image plane motion. Most thresholds are somewhere between 20% and 30%, slightly higher than in the previous experiments. The thresholds found in the control conditions without the eye-movement restriction measured by subject NW (small symbols in subpanels F1-P1 and F2-P2) are in agreement with this conclusion. We also see an effect of viewing distance; for all overlap conditions (see subpanels F1-P4 and F2-P4) thresholds are slightly lower at 2 metres than at 1 metre. This effect is rather small and can probably be attributed to the increased angular resolution of the display when seen from 2 metres.

Figure 4.5 shows the PSEs as a function of the eccentricity of the peripheral object. All PSEs increase with increasing eccentricity, reflecting the reduction in the judged 3D angular speed of the peripheral object relative to the foveal object at higher eccentricities. This is in strong contrast to the PSEs found for the two control conditions without eye-movement restrictions (subject NW only), which show no significant deviation from unity at any eccentricity. There are several other details in figure 4.5 that should be mentioned. The conditions in which both objects have the same size result in PSEs larger than unity, which means that the 3D angular speed of a peripheral object is underestimated relative to the 3D angular speed of a foveal object of equal size. When one shifts one's gaze from the nasal side of the foveal object to the temporal side one can actually experience dynamically the slowing down of the perceived speed of the peripheral object! A similar effect can be noticed when a voluntary (but forbidden) eye movement is made to the peripheral object; its perceived speed suddenly increases dramatically. Note that subject GV seems to have PSE's lower than unity at eccentricities approaching zero. This might be caused by a small response bias, but we have no way of confirming this hypothesis.

Another interesting aspect of the data is the downward shift of the curves when the foveal object's size is halved (everything else being kept the same). Thus, a smaller foveal object introduces an accompanying shift in such a way that the foveal object appears to rotate more slowly. We are allowed to speak about a 'downward shift' only because the shape of the curves appears to be relatively unaffected by the change in size of the foveal object. Note that halving the size of the foveal object has roughly - but not exactly - the same effect as doubling the size of the peripheral object (compare D2-F2-P2 with D2-F1-P2 and D2-F2-P4 respectively). The strength of the size-effect varies over conditions and subjects; we estimate the decrease in the PSE to be roughly 0.07-0.15 for GV and 0.2-0.4 for NW, the lower values corresponding to the 1 metre viewing distance conditions. The size-effect actually compensates for the slowing down of the peripheral object caused by its eccentricity. For instance, the curve D2-F1-P2, describing a condition in which the peripheral object is twice the size of the foveal one, starts below unity at small eccentricities and rises above unity at high eccentricities, demonstrating that approximately 5-8° eccentricity (depending on subject) is sufficient to cancel a factor of two in object size. A similar result is found for the conditions with foveal size 2° and peripheral size 4°.

The effect of viewing distance is somewhat unclear. For NW the curves measured at a distance of 1 metre lie slightly below the corresponding curves measured at 2 metres, whereas for subject GV the order of the curves is the other way around. Moreover, for subject NW the differences between F2-P4 measured at 1 metre and at 2 metres are remarkable. To investigate the effects of viewing distance properly one clearly needs more data.

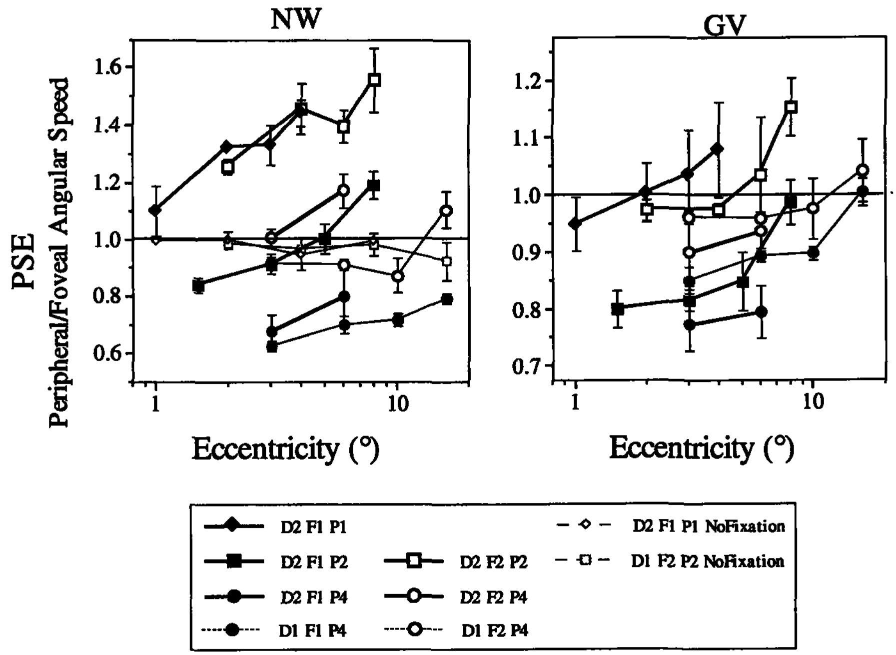


Figure 4.5
Points of subjective equality measured in Experiment 3. The left and right panels present the results for subject NW and GV respectively. A PSE of 1 corresponds to the veridical value or POE. Values larger than unity indicate a relative underestimation of the 3D angular speed of peripheral objects compared to that of foveal objects. Error bars represent one standard error of the mean. Note that the horizontal eccentricity scale is logarithmic and that the vertical scales are different for the two subjects. The black symbols represent conditions with a foveal object of 1° visual angle, whereas the open symbols correspond to conditions with a 2° foveal object. The size of the peripheral object is coded by the shape of the symbol: a diamond for 1°, squares for 2° and circles for 4°. Solid connecting lines indicate a viewing distance of 2 metres, thin dotted lines correspond to a distance of 1 metre. Finally, the control conditions without eye-movement restrictions (NW only) are indicated by smaller open symbols and thin dashed lines. Note that the viewing distance in metres, and the sizes of the foveal and peripheral objects in degrees are indicated in the legend by the capitals D, F and P respectively.

Discussion

The degree of eccentricity of the peripheral object appears to be unimportant for the variability of judgements; thresholds do not change as long as the peripheral object can be recognised as a 3D object rotating in depth. In this experiment the increment thresholds seem to be slightly higher than in the previous two experiments, but this may be due to the eye-movement restriction (no tracking of feature points of the peripheral object), although other explanations are conceivable.

The effect of object size on the PSEs has been reported before by Kaiser (1990) and Petersik (1991). Kaiser found that the angular velocity of a small cube needed to be 18.5% higher than the angular velocity of a cube twice as large for the subject to perceive the velocities as equal. This value is in reasonable agreement with the magnitude of the size effect that we find. Petersik observed similar effects of object size on angular velocity estimates. Note that the size-effect cannot be explained by the increased *objective* values of the projected velocities alone; a doubling of size results in a 100% increase in the objective projected speeds, which is a great deal more than the 18.5% quoted above for perceived angular speed. This is an argument against the possibility that subjects were unwittingly judging linear projected velocity instead of the prescribed 3D angular velocity. See Kaiser (1991) for more discussion on this.

The eccentricity dependence of the PSE is not too surprising. Many characteristics of the human visual system vary with eccentricity, and the perception of linear velocity is no exception. In order to judge 3D angular velocity we need to divide linear velocitiy by some size-measure, and on the basis of this intimate relationship between linear velocity and 3D angular velocity, one might expect similar effects of eccentricity on both linear velocity and 3D angular velocity. There are a few highly relevant studies that report on the eccentricity dependence of 2D velocity perception. Tynan and Sekuler (1982) found that peripheral moving dots appeared to move more slowly than foveal ones, and that this slowing effect increased with eccentricity and decreased with increasing pattern velocity. The projected velocities in our displays are in the range where they found strong effects of eccentricity. McKee and Nakayama (1984) reported that velocity discrimination is just as precise in the periphery as it is in the fovea, but the velocity at which the optimum performance is reached shifts to faster velocities with increasing eccentricity. These optimum velocities are much higher than the ones that occur in our stimuli, especially in the case of the smallest objects that we used. The Weber fractions for velocity discrimination that McKee and Nakayama found increase considerably with velocities below the optimum value, and this might explain why our thresholds are rather high compared to their optimum increment thresholds of 6% (80% correct criterion). A study by Orban, Van Calenbergh, De Bruyn and Maes (1985) confirmed the McKee and Nakayama results. Johnston and Wright (1986) published a study about the matching of foveal speeds with peripheral ones, which is an approach that is actually closer to what we did than the studies by McKee & Nakayama and Orban et al., because the latter authors studied speed discrimination *at specific eccentricities.* In agreement with Tynan and Sekuler, Johnston and Wright found once more that sinusoidal gratings appeared to move more slowly in the periphery than in the fovea (according to Brown, 1931, the peripheral slowing down was first reported by Czermak as early as 1857). In addition, they concluded that the apparent velocity of a peripheral grating can be matched to that of a foveal grating by a purely *spatial* transformation of the retinal stimulus parameters of the peripheral stimulus, all *temporal* parameters being kept constant. This is a very interesting result which fits our data quite well. We find that a factor two in size can cancel an eccentricity of 5-8° , in the sense that the slowing down of a peripheral object is compensated by its larger size. Johnston and Wright found a characteristic eccentricity of approximately 6°, surprisingly close to the value we found. Note that the characteristic eccentricity has been shown to vary over an enormous range (over 100-fold) depending on the specific task (Whitaker, Mäkelä, Rovamo and Latham, 1992).

General Discussion

Summary of Experimental Data

We investigated effects of spatial configuration on the perceived 3D angular speed of objects rotating around fronto-parallel rotation axes. It was shown that 3D angular speed of rotation around a vertical axis is underestimated relative to the speed of rotation around a diagonal and a horizontal axis. The underestimation was smaller when the objects were presented above each other than when presented side-by-side. Objects presented in the peripheral part of the visual field appear to rotate more slowly than a similar object seen foveally. At 5-8° eccentricity the reduction of perceived angular speed could be cancelled by making the size of the peripheral object twice that of the foveal object. The increment threshold for discrimination of angular speed was almost independent of spatial configuration and was approximately equal to 20%. Apparently, spatial aspects of the stimulus greatly influence the subjective 3D angular speed.

2D Processing Characteristics

Our attempts to interpret the results of the 3D angular speed task in terms of known 2D processing characteristics proved to be rather successful. To some extent this should not be too surprising, since our visual system samples the outside world with essentially two spatial dimensions at a time. However, note that all the papers refered to that studied the 2D characteristics utilised stimuli that do not even remotely resemble the rotating objects that we present to our subjects. We like to mention two relevant studies that describe attempts to link human performance in 3D tasks comparable to ours with the 2D processing capabilities of the visual system. Eagle and Blake (1995) investigated the perception of planarity of 3D objects and of dihedral angles between planes. They successfully and quantitatively modelled the measured performance in terms of known image motion sensitivity. Werkhoven and Van Veen (1995) studied the human perception of relief, and used a simple relief-from-motion computation based on local velocity estimates to model the performance with the 3D relief task. The model assumed a Weber's law for velocity estimation and an eccentricity-dependent underestimation of velocity. They were able to make a quantitative prediction of the 3D performance based on these known 2D processing properties of the visual system. Both studies as well as the current one involve relatively simple tasks, which might have made it easier to determine the relationships with 2D processing. It might be worth doing similar studies involving more complicated tasks.

Consequences of our Results for Structure Recovery from Optic Flow

Algorithmic recovery of the structure and motion of moving objects from optic flow information is a well-studied branch of science (e.g. see reviews by Aggarwal & Nandhakumar, 1988, Simpson, 1993, and many others). Optic flow confounds structure and motion information and it is for this reason that these three-dimensional quantities are recovered simultaneously by most algorithms. In some specific cases optic flow is insufficient for a complete reconstruction and in other cases a full reconstruction is feasible although very sensitive to noise. In such cases other information sources are needed to calibrate the recovered spatial-temporal layout of the scene. There is growing evidence that human perception of 3D motion and structure exhibits similar characteristics, in the sense that the disentanglement of motion and structure is sometimes incomplete in human vision too (Petersik, 1980; Van Veen & Werkhoven, in press). Therefore, errors in perceived motion might influence perceived structure, and vice versa. The impressive effect of eye-movement restrictions on the perceived angular speed clearly proves that a comparison of foveal 3D angular speed with peripheral 3D angular speed is really different from a comparison of two foveally presented 3D angular speeds at different locations. We will briefly discuss what this could mean for the perception of the rigidity of moving objects.

An extreme consequence of the results of Experiment 3 could be that a rotating object can only appear to be rigid to us when we make eye-movements, inspecting different parts of the object with our fovea. Otherwise, different parts of the object would seem to rotate at different angular speeds (which is proof of non-rigidity) because they are at different eccentricities. We mentioned above that a peripheral object appears to slow down when the gaze is shifted from the nasal side to the temporal side of the foveal object. This gaze-shift effectively increases the eccentricity of the peripheral object. But although the eccentricity surely varies within an object (even without eye-movements), none of the objects ever appeared to be rotating non-rigidly. The fact that effects of eccentricity already emerge at very small eccentricities (Figure 4.5 subject NW, conditions D2-F1-P1 & D2-F2-P2, 25% effect at 2° eccentricity) strongly supports this possibility. However, it is possible that the connections between the different parts of an object ensures a non-biased perception of angular speed within one object. A report by Van Doorn and Koenderink (1982) sheds some light on this. They studied the detection of differences in the magnitude of velocity of moving random dot patterns and found very large Weber-fractions, close to 100% (without eye-movements). Being aware of the reasonably good accuracy of human depth perception from optic flow, they concluded that tracking movements of the eyes are necessary to be able to perceive the small velocity differences needed to extract (relative) depth with high accuracy. Furthermore, from their study it appears that, in order to be perceivable, velocity differences between adjacent patterns, as occurs *within* an object, must be about as large as the common velocity component. Effects of eccentricity within an object might simply be masked by this effect.

Future Research

We have studied only part of the perception of the rotation of moving objects. The angular speed of objects rotating around non-fronto-parallel axes, the axis of rotation itself, and situations with more complicated, time-dependent rotations all deserve more attention. We will pursue this in future investigations.

5

MIMICKING THE ROTARY MOVEMENT OF OBJECTS IN SPACE

Abstract

In a series of experiments we asked subjects to mimic the rotary movement of objects displayed on a computer-monitor by dynamically changing the orientation of a real object that they hold in their hand. We found that subjects can indeed accomplish this task in a fairly accurate way, as far as the orientation of the rotation axis is concerned. The variability in the measured orientation of the axis of rotation was lowest for elongated objects rotating about their main axis and for objects rotating about fronto-parallel rotation axes, and was always higher than the baseline response accuracy of the motor system involved in manipulating the real object. The standard deviation of signed angular error in arbitrary directions was as low as 5° in these optimum conditions. For two out of three subjects the variability of the orientation of the rotation axis increased drastically when the axis of rotation made a 45° angle with the object's symmetry axis. Significant systematic deviations from veridicality were found (veridicality in terms of the geometric model used to generate the stimuli). These biases can be described by an overall rotation of response space relative to stimulus space, combined with a reduction in the range of measured slants of the rotation axes relative to the stimulus range of slants of the rotation axes. The experimental data prove the feasibility of using a motor action as a basis for investigating visual information processing.

Introduction

Rotations can be very interesting and mystifying; think about the weird motion of spinning tops or coins, or watch how a thrown boomerang loops back to you. Observe the tumbling of a bottle thrown in the air, the spin of a curve-ball, the bouncing of billiard-balls and the curved paths they sometimes follow. Under certain circumstances we can even experience rotations as enjoyable or as frightening, as anyone who has been for a ride with a roller-coaster might argue. And as children, didn't we always beg our parents to let us have a go on those beautiful merry-go-rounds? Rotations play an important role in many situations encountered in daily life. When one is manipulating an object, accurate perception and control of the orientation of the object is often quite important for a successful manoeuvre. Perception and control of orientation allow us to put a teapot on a table without spilling the tea, to park our car without hitting other cars, to put a postcard in the letter-box without bending it, etc.

Although we are constantly confronted with rotary movement, little is known about human perception of rotations in space. Some work has been done on the perception of angular speed of rotation of objects rotating in three dimensions (Petersik, 1980, 1991; Kaiser, 1990; Kaiser & Calderone, 1991; Liter, Braunstein & Hoffman, 1993; Van Veen, Kappers, Koenderink & Van Woerkom, submitted), but only for rotations about fronto-parallel axes and for rotations about the viewing direction. These studies have shown that the perceived angular speed of an object is influenced by factors as diverse as the size of the object, the 'transition rate of edges' of the object, the orientation of the axis of rotation in the fronto-parallel plane, the eccentricity of the object in the visual field, etc. The studies have also demonstrated that under optimised conditions the discrimination threshold for 3D angular speed approaches the discrimination thresholds found for linear velocity. Shiffrar and Shepard (1991) have investigated the role of the rotation axis in the perception of rotation as a whole. They found that subjects are faster and better at comparing two subsequent rotations when the axis of rotation is aligned with the environment or, even better, when it is aligned with the object itself. Pollick, Nishida, Koike & Kawato (1994) have published a study on the perception of the axis of rotation of rotating configurations of dots and ellipsoids of revolution. They asked subjects to indicate with their index finger the direction of the rotation axis. They found that subjects were sensitive to both the slant and the tilt of the rotation axis, and that the bias in reported slant and the variability in reported tilt depended on the type of structure-from-motion display and on the slant of the simulated rotation axis.

We adopt a slightly different approach to the perception of rotations. Firstly, we use a new paradigm: we ask subjects to mimic the entire rotary movement of an object that they see by dynamically changing the orientation of a real object they have to hold in their hand. Every subject is willing to do this, and to most subjects the task appears to be reasonably pleasant and natural. This shouldn't surprise us; from our earliest childhood we have been manipulating the objects around us, often watching other people showing us how something is done. As we grow up, we practise sports or learn ballroom dancing; our teachers show us how to do the high-jump or how to move our feet in certain dances. Learning by example is clearly an important aspect of our behaviour. When watching a rotating object, mimicking its movement seems to be a natural way of expressing what we see. A major advantage of our paradigm is that it enables the subjects to respond to virtually all aspects of the stimulus motion. For instance, the subject might have the impression that the rotation axis varies over time. If we were to ask the subject to estimate the orientation of the (invisible) axis of rotation, the subject would be faced with a conceptual problem: how to respond to the orientation when the orientation varies?[20] This problem is elegantly avoided with our paradigm, because it does indeed allow the subject to respond with a varying rotation axis. It is up to the scientists to decide whether they want to consider this aspect of the data. We point out here that in the current study we do not investigate the translational component of motion (the subject is required to hold his hand at a fixed position in space), although the set-up certainly would allow us to do so. A second major difference between our approach and other approaches is that we ask subjects to respond to the *integral motion*, not to its *components*. The difference is crucial; whereas we all are familiar with the changes in orientation of objects, whether we turn them with our hands or tools or whether they spin on their own, the concept of an abstract rotation axis and its orientation in space is not generally encountered in daily life (except in

[20] None of the experimental conditions in our experiments features a varying rotation axis, nor does the task require the subjects to judge the orientation of this axis directly. However, when asked to do so, subjects report that in certain conditions the axis of rotation seems to vary considerably (notably in cases where the axis of rotation is not aligned with the symmetry axis of the revolving object; see the last experiment in this chapter).

situations in which the rotation axis is actually visible, i.e. where there are screws and wheels). Of course, we will still have to analyse the results in terms of the orientation of the rotation axis and the magnitude of rotation, simply because they constitute the rotational motion. These components, however, have now become quantities that are implicitly defined by the motion of the real object. A third advantage of our method is that we can address the veridicality of the judgements because we probe the perceived rotation movement with a real world action. The term 'veridical' refers to the geometric model that is used to generate the stimulus; this notion of veridicality, however, differs slightly from the veridicality that can be defined when actual physical scenes are used.

After a comprehensive description of the experimental methods - we explain and illustrate which choices we have made in setting up the experiment and how we analyse the results - four different experiments are discussed. We start with a pilot-study in which the objects rotate about axes in the fronto-parallel plane. This study is used to test the feasibility and reliability of the method in general and to optimise certain experimental parameters. The results prove that subjects can respond quite accurately to the stimulus and that we can measure these responses with high precision. In the second experiment we ask subjects to rotate the object in their hand about one of four pre-defined axes but do not provide the subjects with a visual stimulus. This experiment is designed to establish a baseline response accuracy of the subjects' motor system. It turns out that the variability in movement is remarkably low. The third experiment (with visual stimulus again) extends the range of rotations with rotations about slanted axes. We analyse how the reproducibility and veridicality of the subjects' responses depend on the orientation of the axis of rotation. Finally, in the fourth experiment we display objects rotating about axes that make a certain angle with the object's symmetry axis. We show how such a misalignment between object axis and rotation axis can have a considerable influence on the mimicked motion.

General Method

Basic Task and Procedure

We ask subjects to mimic the complete rotation movement of an object displayed on a computer-monitor by dynamically changing the orientation of a dummy-object they hold in their hand. From here on we use the term 'dummy-object' instead of 'real object' to point out that the physical object that subjects hold in their hand is not at all the same as the one that is displayed on the screen; the dummy is just a substitute that provides the means to make rotary movements, not a replica of the actual stimulus. Due to the mechanical constraints of the limbs the rotations have to be restrained in magnitude. We decided to use periodic rotations (i.e. the stimulus object is rotating back and forth and so is the hand with the dummy-object) because they allow continuous measurements for longer periods. At the start of each trial the subject takes a few seconds to match the motion of the dummy-object in his hand to that of the stimulus object on the screen. As soon as the subject is ready, he indicates this to the experimenter, who then starts recording the movement. During the recording, which lasts for 18 seconds (10 seconds in Experiment 1), the subject is not allowed to change the way he is rotating the dummy. The end of a trial is indicated by a short sound. After each trial subjects lower their arm to relax the muscles. Whenever subjects feel the need to take a longer break they are allowed to do so.

Experimental Set-Up

The basic experimental setting is depicted in figure 5.1. The observer stands in front of a computer monitor on which rotating objects can be displayed. The centre of the monitor is about 175 cm away and slightly below eye-level. In one of his hands (always the right hand, except for the first experiment in which the observers were free to use the hand they preferred, which could differ from trial to trial) the observer holds a rubber-like ball (weight 245 gr.; diameter 6.5 cm; mass density about that of a brick). The ball contains an electromagnetic sensor which is connected with an electrical wire to the electronics unit of the tracking system. The subject holds the ball between his fingers at a more or less constant position in the median plane, 10-30 cm below eye-level and 30-50 cm in front of the body (we allow subjects to select a convenient position within these margins). With this set-up, the observer can easily alternate his gaze between the object on the screen and the hand manipulating the ball. Except in the first experiment, in which binocular viewing was permitted, vision is restricted to one eye (the right one). The other eye is covered in such a way that both stimulus and hand are only seen monocularly. About 40 cm away from the ball, slightly to the left of the observer and upwards, the electromagnetic source (field generator) of the tracking system is suspended from a rail. In this way the source is as close to the sensor as possible (which is needed for good accuracy), yet does not block important parts of the observer's visual field or interfere with his arm and hand movements. Subject's posture is not restricted mechanically. However, the position of the feet is prescribed and indicated by a line on the floor, and the subject is required to stand up straight without torsion of head and shoulders relative to the feet. The experimenter monitors the subject's posture during the experiment and gives further instructions whenever necessary.

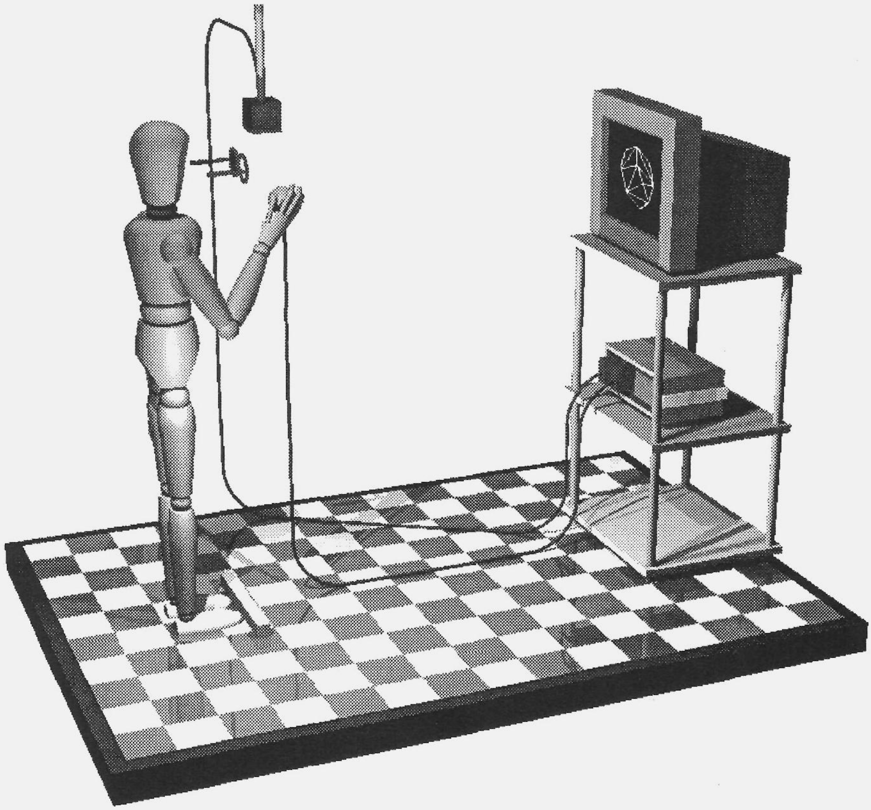


Figure 5.1
This picture illustrates the general experimental set-up. An observer stands in front of a computer-monitor on which rotating objects are displayed (the rotation axis is never actually visible, contrary to what is depicted here). In his hand the subject holds a spherical dummy-object, the orientation of which is measured by an electromagnetic tracking system. The *subject's task is to mimic* the rotary motion of the object on the screen by turning the ball in his hand in the same way.

Equipment

The computer display is a 71 Hz Radius TPD/19 high spatial resolution (82 dpi) grey-scale monitor (1152 by 882 pixels) driven by an Apple Macintosh IIfx computer. The electromagnetic tracking system (3SPACE, made by Polhemus) is connected to the same computer by means of a fast parallel interface. The tracker measures all 6 degrees of freedom of the ball (position and orientation of the sensor relative to the source) and transmits them to the computer at an update rate of 60 Hz. Spatial and angular resolution are 0.005 cm and 0.007° respectively. Spatial and angular accuracy are specified as 0.04 cm and 0.1° r.m.s. These specifications are valid up to a separation of approximately 70 cm between source and sensor. Because of the electromagnetic nature of the device special care has to be taken to avoid disturbances from conductive materials and electromagnetic sources in the vicinity of the system. Therefore, the fixture of the source on the rail and the rail itself are made of pvc and other plastics, both source and sensor are far away from the monitor, and no conductive materials are near or in-between the source and the sensor. We have checked in several ways that these precautions are sufficient for the system to meet its specifications. The optional adaptive filter provided by the tracker was turned off, which allowed us to control the filtering ourselves (off-line).

Reference Systems

The subject is always required to reproduce the rotary motion of the stimulus object in absolute space (i.e. in the earth-fixed reference system, as opposed to observer-related reference systems). To avoid any confusion on this point we have aligned all potential reference systems as much as possible: the computer display is parallel to the walls of the (rectangular) room, and the subject-display direction (the viewing direction) is orthogonal to the front of the display (although the lights in the room are dimmed, the monitor and the walls of the room are clearly visible). Even more important, the subject's fronto-parallel and median planes are aligned with the room and with the viewing direction.

Calibration

In order to be able to compare the motion of the ball in the observer's hand with the motion of the object on the screen both need to be specified in the same co-ordinate system. Because of the limited range of the tracking system it is not possible to position the electromagnetic source such that the position and orientation of the screen and the ball can be measured at the same time. For that reason we use a calibration procedure that requires a few intermediate positions of the source and a movable reference system (a cube mounted on a tripod). It takes about 45 minutes to completely calibrate the set-up. We practised the procedure several times and have checked the results in various ways. The systematic orientation error introduced by the calibration procedure itself is not more than 1°, and probably considerably less. Position is ignored as we will only analyse the changes in the orientation of the ball.

Data Processing

The basic quantity that we are interested in is the *change in the orientation* of the ball. The absolute orientation of the ball is irrelevant since the way in which the observer holds the ball in his hand varies from trial to trial, and, apart from the projecting wire, the ball is isotropic. Likewise, the position of the ball is disregarded. The tracker specifies the orientation of the ball in terms of unit quaternions. To get a more stable signal we pass these quaternions off-line through a high-pass third order (software) filter with a cut-off frequency of 12 Hz. It is important to use quaternions in the filtering stage and not the more commonly used Euler-angles, because the latter co-ordinate system suffers from so-called *gimbal-lock*: at specific orientations of the ball a slight change in orientation can induce a huge and discontinuous change in the Euler-angles (see Shoemake, 1985, and many others). Meaningful filtering requires that the signal represents the range uniformly, which is obviously not the case with Euler-angles. Each trial consists of a sequence of rotations with alternating direction. We split this sequence into the separate rotations (typically 16 per trial) using the minima of rate-of-change of the orientation of the ball. For each single rotation we compute the total change of orientation from initial to final orientation, which can be expressed by a vector quantity, the rotation vector. This vector points in the direction of the axis of rotation, and its length equals the magnitude of rotation. The sign ambiguity of the rotation vector is resolved by selecting the orientation closest to the axis of rotation of the stimulus. This latter axis is always in the upper half of the frontal hemisphere of visual directions (note that we are not bothered by the problem of orthographic inversion, because the objects are opaque). We average these vectors over all rotations in the sequence, leaving out the first two and last two of each sequence. In short, for each trial this procedure yields an average axis of rotation and magnitude of rotation. If we compare the axis of rotation of the individual rotations in the sequence to their sequence-average, the standard deviation of the angular difference (in any direction) turns out to be 1-2° on average, depending on subject and specific experiment; this means that subjects are quite able to rotate the ball in a stable manner during a trial. The value also sets the lower limit for the corresponding standard deviation over trials.

Stimulus

The stimulus shown on the computer display always consists of one convex polyhedral object rotating about a fixed axis through its centre-of-mass. The objects consist of 16 to 20 triangular facets the edges of which are visible as light lines on the dark background of the monitor. The edges of back-facing facets are not drawn and the luminance of the facets is equal to that of the background. The overall impression is that of an opaque solid revolving in space. The facets are generated randomly for each trial, volume and global shape being kept the same. Shape is controlled by the length-width-height ratio of an object. We speak of cube-like objects when length, width and height are equal. These are the objects that we will use in the first experiments described in this paper. In the last experiment we will use elongated, rod-like objects, with a length-width-height ratio of 3:1:1. All objects have a volume of 100 cm^3 (which is slightly smaller than the size of the dummy-object that subjects hold in their hand), which yields an apparent size of approximately 2° visual angle (cube-like objects). More details about the properties and generation of these objects can be found in Van Veen, Kappers, Koenderink and Werkhoven (in press).

Every 50 msec the object is rotated 4° (7.5° in the first experiment) with respect to its previous orientation, and a new image (parallel projection) is displayed on the monitor. If we disregard the noise introduced by the display device (pixel rounding and unknown anisotropies of the pixel array), then the set of projections of corners of the object in three or more different frames is theoretically sufficient to allow for a complete recovery of the motion and structure of the object (see Ullman, 1979). The axis of rotation is a controlled parameter of the experiment. We specify its orientation in terms of slant and tilt, quantities that are traditionally used to describe the orientation of a plane with respect to the viewing direction. We define the slant of an axis of rotation as its angular deviation from the fronto-parallel plane. Thus, a fronto-parallel axis has a slant of 0°, whereas a rotation axis aligned with the straight ahead direction has a slant of 90°. The tilt of an axis is defined as the angle that its projection on the fronto-parallel plane makes with the rightward horizontal. Thus, a rotation axis in the horizontal plane has a tilt of either 0° or 180°, and an upward vertical axis has a tilt of 90°. All rotation axes used in the experiments are in the upper half of the frontal hemisphere of visual directions. Thus, slant ranges from 0° to 90°, and tilt ranges from 0° to 180°.

In two of the four sessions of the first experiment, when we were still experimenting with the paradigm, the object rotated continuously instead of rotating in an oscillatory manner. In the other two sessions the magnitude of rotation is fixed at 90° and the frequency of oscillation is 0.9 Hz. In the rest of the experiments the magnitude of rotation is scattered between 90 and 130° from trial to trial (360° would be a full rotation) resulting in a frequency of oscillation of 0.31-0.46 Hz (one cycle consists of two equal rotations in opposite directions). These particular values for the magnitude and speed of rotation are chosen because they appear to be maximally convenient to the observers. Faster rotations prohibit accurate control of subjects' movements, while slower ones become annoying after a while (note that subjects are required to mimic the complete rotation, including amplitude and frequency of oscillation). Likewise, larger magnitudes of rotation will interfere with mechanical constraints of the forearm; subjects easily twist their arm 180°, which means that we are on the safe side with the 90 to 130° range. Much smaller magnitudes of rotation are not desirable either, for two reasons: they reduce the clear impression of a solid revolving in space, and the frequency of oscillation becomes too high for accurate control of movement.

It should be clear by now that we have optimised the movement of the stimulus object to allow for an optimally accurate and comfortable handling of the dummy-object by the observer. This is a basic choice, which imposes strict constraints on the admissible motions of the stimulus. We are convinced, however, that a high reproducibility of movement is required for a sensible investigation of the mimicked rotational motion. The actual accuracy of movement will be discussed further on in this paper.

Subjects

Three male observers participated in the experiments reported in this paper. The authors HV and NW were of course fully aware of the specific stimulus conditions employed, whereas subject FO was not. HV and NW are myopic, and their corrected vision is normal. All subjects are right-handed and none of them has a history of any motoric disorder. Several other subjects were asked to participate in the experiment for a few trials. They had no problems in understanding the task and were immediately prepared to act as required. It usually takes a few trials to get used to the periodicity of movement and to find a convenient way to hold and manipulate the dummy-object.

Method of Analysis

The slanted rotation axes pose some difficulties for the analysis of the data. Tilt of a rotation axis becomes gradually more poorly defined with increasing slant and measures of variation of tilt are therefore not very useful for describing rotation axes outside the fronto-parallel plane. Another problem is that of visualisation of the data; since the orientation of the axes varies both in tilt and slant, a concise way of presenting the results is required. The solutions that we have chosen are explained below. This particular method of analysis and data visualisation will be used in all but the first experiment described in this chapter.

Averaging Rotations over Trials

To be able to compare the 'responded' rotations with the veridical ones it is most helpful to compute an average responded rotation. This is rather a straightforward matter for a one-dimensional quantity like rotation magnitude, but averaging a vector quantity like the orientation of the axis of rotation requires some extra attention. An axis of rotation can be denoted as a unit vector in three-dimensional space. A set of rotation axes is then most easily visualised as points on a unit sphere of visual directions; the points indicate the end-points of the unit vectors. Given this representation a natural definition of the average rotation axis is the two-dimensional median of the cloud of points on the unit sphere of directions. Such a median can be computed in almost the same way as in planar cases, as long as one takes the metric of the sphere into account. In short, the median point is defined as that point on the sphere that has equal numbers of data points on both sides of each great-circle passing through that point. Obviously, for a limited set of data points not all great-circles passing through the median point will have equal numbers of data points on each side. In practice we use only three great-circles, which are chosen to be maximally different with respect to their orientation. An example of the outcome of this procedure is depicted in figure 5.2. The method appears to be very robust, and produces excellent results in all conditions that we have measured in this experiment and in the other ones described in this chapter. Moreover, it is quite an objective method, not least because it does not rely on the choice of a particular co-ordinate system, which would be the case if one, for example, were averaging slant and tilt.

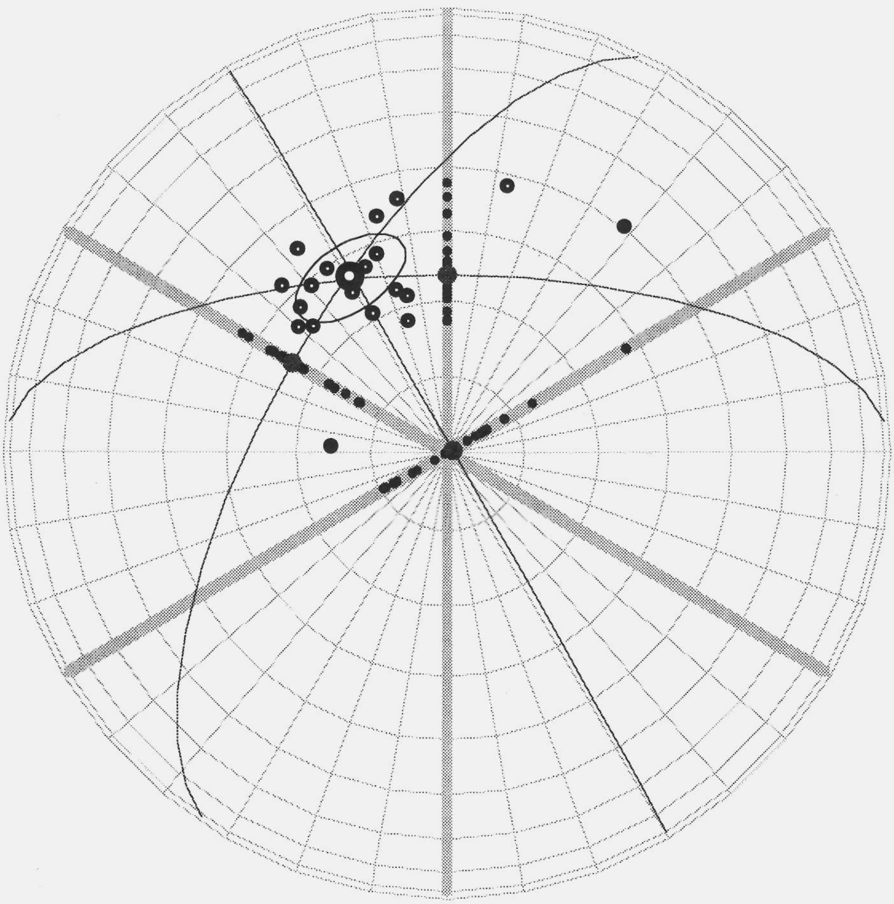


Figure 5.2
Orthographic view of the frontal part of the unit sphere of visual directions. The straight ahead direction is in the centre of the picture. The drawing illustrates the geometric construction of the median axis of rotation (large black disk with white centre) from a set of measurements (endpoints of unit rotation vectors). Step 1: the data points (smaller black disks with and without central white spot) are projected onto three primary meridians (thick straight grey lines). Step 2: along each meridian, the median location (large grey disks) of the projected points (smallest black disks on the meridians) is computed. Step 3: three new secondary meridians (thin curved black lines) are constructed orthogonal to the primary meridians and passing through the median points. Step 4: the median axis of rotation is constructed by taking the mean of the mutual points of intersection of the secondary meridians.

After finding the median axis of rotation, we compute for each data point the signed angular deviation from this median in two orthogonal directions on the sphere. We compute the variances and covariance and construct the corresponding variance ellipse. We throw away outliers, which we define as data points outside the 2.5 sigma contour (in a genuine two-dimensional Gaussian distribution approximately 4% of the points are outside the 2.5 sigma range). This means that usually one but sometimes zero or two data points are removed. Finally, we recompute the median axis and the variance ellipse based on the reduced dataset, and consider these as the most reliable estimates that we can get[21]. This last ellipse is shown in figure 5.2. The variance ellipse is in fact a contour line of the two-dimensional Gaussian probability distribution fitted to the data. The ellipse is centred on the median rotation axis. We always show the one-sigma contour. Note that for a two-dimensional Gaussian approximately 39% of the data points should be within the one-sigma range, as opposed to 68% for a one-dimensional Gaussian.

Solid Angle and Equivalent Standard Deviation

The variance ellipses introduced in the previous paragraph are in principle described by their orientation on the sphere and the length of the major and minor axes. It turns out, however, that in most (but not all) conditions in the following experiments, the amount of data points is not sufficient to significantly distinguish the ellipses from circles. This implies that the area of the ellipses can be used as a good description of the variance of responses. Areas on the unit sphere are conveniently expressed in steradians (sr); the total area of a sphere is 4π sr. If we visualise the ellipses as the intersections of elliptical cones with the unit sphere, we can see that another interpretation of the variance ellipse is that of a sheaf of rotation axes within a certain solid angle. Solid angle is numerically equal to the corresponding area on the unit sphere and these quantities can be used interchangeably.

Although solid angle is a most adequate and intuitive way of indicating the variance in responded rotation axis, it is not easy to compare directly variance in solid angle with variance in planar angle. However, we can make use of the approximate circular-symmetry of the ellipses to introduce a more easily accessible quantity. The *equivalent standard deviation* of responses is defined as the geometric mean of the length of the major and minor axes. Equivalent standard deviation also has a very intuitive meaning: it represents the standard deviation in degrees of the signed angular error with respect to the median axis in any direction on the sphere. In the sequel, we will show all variance ellipses graphically, which allows one to appreciate the differences in size, shape and orientation. Whenever quantitative results are needed, we will use both solid angle and equivalent standard deviation.

---

[21] The method described here is not necessarily the best method for averaging rotations in other situations. Since we are averaging the magnitudes of rotation separately from the direction of the rotation axes, in situations in which the magnitude of rotation varies considerably from trial to trial our method can result in peculiar behaviour of the average direction of the rotation axis. We have chosen to average direction and magnitude separately because this seems to be the most intuitive way to present the results, and we are allowed to do this because the variation in the magnitude is rather small. Note that the average rotation *per trial* is computed in a different (but faster) way (average rotation vector); variation within a trial is so small that different methods give nearly identical results.

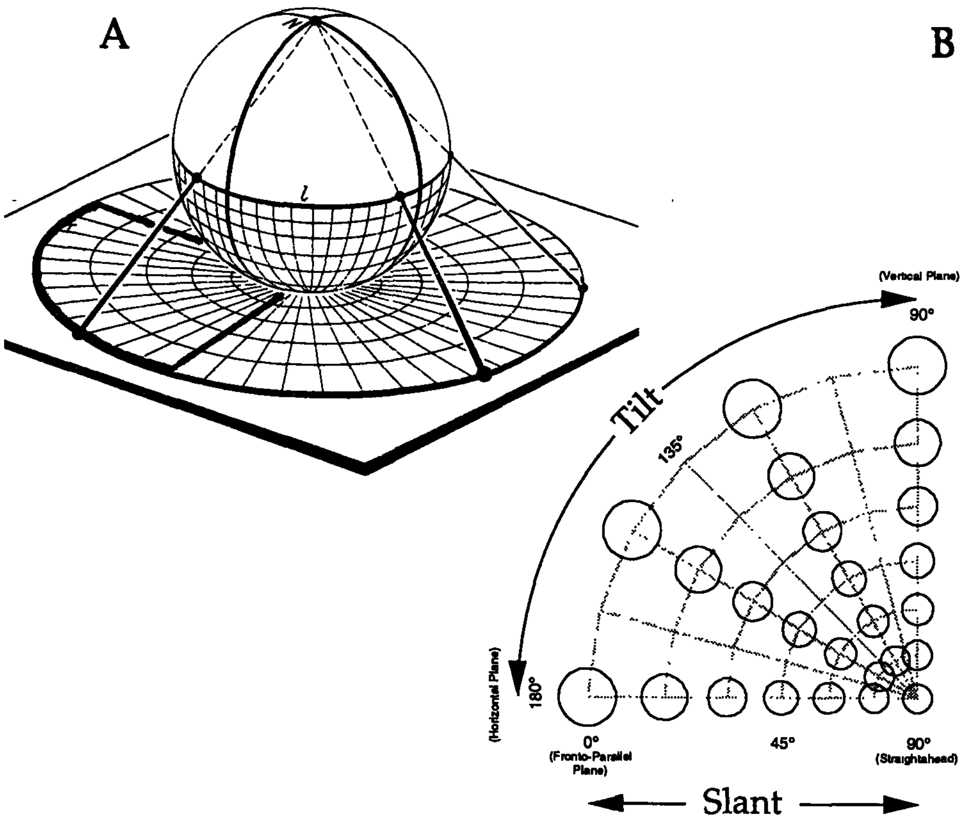


Figure 5.3
We will use stereographic projections of the sphere of directions to present the results of the experiments. Part A explains the general principle (reproduced from D. Hilbert and S. Cohn-Vossen, Anschauliche Geometrie, Dover Publications, New York, 1944). The frontal part of the sphere of visual directions is drawn on the sphere as well as on the projection plane. The top of the sphere corresponds to the direction opposite to the subject's straight ahead direction. Part B gives an idealised example of the actual diagrams that we use. Only the upper left quadrant of the frontal sphere of visual directions is shown. Slant of the axis of rotation varies in the radial direction, whereas tilt varies in the angular direction. The 25 circles correspond to circles of 5° radius on the unit sphere, and are provided to give an idea of the scale errors caused by the projection. In order to get a veridical impression of the sphere, one should view this diagram should from a distance equal to its width.

Stereographic Projection

Visualising the experimental data in a comprehensible way is essential for a thorough understanding of the results. As we indicated in the two preceding paragraphs, the average rotation axis can be represented by a point on the unit sphere, and the variance of responses can be depicted by an ellipse on the sphere drawn around this point. To investigate this picture one would preferably hold such a unit sphere in one's hand and manipulate it to look at it from different angles. The restrictions of flat paper imply, however, that we have to resort to a projection of the unit sphere on a plane. We have decided to use the so-called stereographic projection (Hilbert and Cohn-Vossen, 1944). This projection is essentially a polar projection from one point on the sphere (the viewing point) onto the tangent plane on the other side of the sphere (see figure 5.3a). The advantage of the stereographic projection is that it is a conformal mapping: it leaves the shape of figures on the sphere invariant (but not their size!). We have chosen the viewing point (the pole) to be that point on the unit sphere which corresponds to the direction of the observer as seen from the stimulus. As a result the straight ahead direction (as seen from the observer) is in the centre of the projection and the fronto-parallel rotation axes project onto a large circle around this central point. Figure 5.3b shows an example of the stereographic projection. Only the upper left quadrant of the frontal hemisphere of directions is drawn, because in all experiments in this paper (except for the first one) all stimulus rotation axes are in this range. The small circles in the projection are actually circles on the sphere with a radius of 5°, corresponding to a solid angle of 0.024 sr. Note that the projected size of the ellipses increases slightly for axes approaching the fronto-parallel plane. Figure 5.3 is provided to acquaint the reader with the specific type of projection.

Manipulating the Dummy-Object

We have experimented with the size and weight of the ball. The current choice is a compromise between comfort and accuracy. If we make the object lighter its position and orientation become more sensitive to muscle tremor and other disturbances; the electrical wire emerging from the ball will carry a major part of the moment of inertia of ball+wire, which is clearly undesirable. In this sense a heavier object is always better. The current object appears to have enough weight to mask the existence of the wire, because manipulation is not noticeably obstructed by it. Some of the subjects use their free hand to guide the electrical wire, which further reduces its influence. Making the object heavier than it is at the moment severely reduces comfort and in the end accuracy of movement as well. The size of the ball permits it to be easily grasped by most adults.

In principle, the ball can be rotated by turning it between the fingers, by supination and pronation of the forearm or by rotation of the complete arm around the shoulder. All three situations occur, but virtually all movements are twists of the forearm, sometimes combined with a slight flexion and extension of the wrist and slight turning of the ball between the fingers. Consequently, in most trials the orientation of the forearm parallels the axis of rotation to some extent. Rotations of the complete arm around the shoulder occur sporadically when the axis of rotation appears to the subject to be close to the straight ahead direction. The constraint of keeping the ball at a fixed position in the median plane impedes the rotation of the forearm in such conditions, a problem that is typically solved by twisting the shoulders a little, or by bending the torso to the side opposite to the manipulating arm. In general, however, the positional restriction of the hand causes the subject no problems. This might relate to an observation made by Soechting (1987), who concluded that object location and orientation seem to represent two separate characteristics of the organisation of reaching movements. Shoulder and elbow motion, resulting in a trajectory of the wrist through space, seem to be co-ordinated separately from the wrist angular motion needed to match the orientation of the hand to the orientation of an object that is being reached for (Lacquaniti and Soechting, 1982).

In some of the earliest pilot sessions, the object on the screen rotated continuously and the subjects were asked to rotate the dummy-object about the same axis as the stimulus object, but with a frequency and magnitude of their own choice (two such sessions are discussed in the first experiment). This procedure resulted in a situation in which the subjects were rotating the dummy object with a high frequency of rotation, typically 1 Hz (back and forth within one second), which turns out to be quite tiring after a while. Furthermore, subjects appeared to be unable to control the magnitude of rotation accurately with such a high movement frequency. Later we decided that the subjects should also respond to the magnitude and frequency of rotation. Therefore, in the rest of the experiments described in this paper, the stimulus object is set in oscillatory motion and its frequency of rotation is lowered to enable subjects to respond to the rotation magnitude. The phase of rotation is not prescribed, but subjects always try to move in phase with the stimulus, probably because this gives them the highest confidence that they are making an accurate reproduction. We do not record the phase relationship between stimulus and response.

We have observed subjects' eye-movements while they were doing the task. Subjects seem to pay very little (visual) attention to their moving hand. At the beginning of a trial, when they try to match the motion of the dummy-object with the motion of the object on the screen, subjects tend to look back and forth between hand and display. After a short period all attention is directed towards the stimulus and the movement of the hand is seen only peripherally. Then, during the recording period, the subjects continue to watch the stimulus object, probably because it seemed to be almost impossible for them to keep moving in phase with the stimulus object while looking at their hand instead of at the display. Subjects usually only glimpse at their hand, possibly trying to make sure that its position in space is still what it should be. Note that the recording period lasts for 18 seconds; after a while the arm tends to lower somewhat due to gravity, and vision can be used to correct for this.

Exp. I: Fronto-parallel Rotation Axes

Introduction

A change in the orientation of an object with respect to an observer can create on the observer's retina a flow pattern that contains information about the structure of the object. In general, such rotations can be decomposed into a rotation about the line-of-sight and a rotation about an axis in the fronto-parallel plane. Only the latter produces optical flow containing information about the structure of the object (Koenderink, 1985, 1986). In this sense, rotations about fronto-parallel axes are of special interest and they are commonly used in experimental studies concerned with the structure-from-motion phenomenon. In the first experiment of the series reported in this paper we concentrate on this special class of rotations about fronto-parallel axes. The following experiments will deal with more general rotations.

Stimulus, Procedure and Experimental Design

The cube-like object on the screen rotates about one of 6 different fronto-parallel axes, having tilts of 0, 30, 60, 90, 120 or 150° (slant is always zero). Each of these 6 conditions is presented 10 times in each session, in random order and without feedback. Subject HV performed three sessions on different days, and prior to each of these sessions the set-up was slightly rearranged and recalibrated. Subject NW participated in one session. The rotation magnitude differs from session to session. In NW's session and in HV's first session the rotation magnitude of the stimulus object was fixed at 90°. In HV's second and third session the object on the screen rotated continuously instead of rotating in an oscillatory manner.

In this experiment subjects are free to look with both eyes and they are allowed to use the hand they prefer, which can differ from trial to trial. This situation is inevitable, since rotations about rightward pointing axes (tilt 0-90°) are hard to make with the right hand, and likewise, rotations about leftward pointing axes (tilt 90-180°) are hard to make with the left hand. It turns out that subjects always use their left hand for rotations about axes with a tilt of 30° or 60°, whereas they use their right hand for rotations about axes with a tilt of 120° or 150°. HV uses both hands (one at a time) for the directions 0° (which is the same as 180°) and 90°. NW always uses his right hand for these directions. For the rest of the experiments in this paper the range of tilts will be restricted to one quadrant (tilt 90-180°) and only the right hand may be used.

Results and Discussion

For every trial we compute the slant and tilt of the measured rotation axis and subtract the slant and tilt of the rotation axis of the stimulus object. For each of the 6 different conditions we then compute the median and standard deviation of these slant and tilt errors (this method is rather useful for fronto-parallel rotation axes, but it should not be used in the general case of slanted axes). The results are summarised in figure 5.4. The upper panels indicate the systematic errors in slant and tilt. It is clear that substantial deviations from veridicality do occur. The large positive error in the slant means that subjects respond with a rotation about an axis that is slanted away from the fronto-parallel plane towards the straight ahead. The cause of this effect is not really clear. It has been found in several studies that the apparent vertical is rather close to the objective vertical (see Howard, 1982). However, it has been known since the 19th century that the vertical horopter is strongly tilted in the median plane, in the same direction as we find for the rotation axes. More recently, Cogan (1979) found that the vertical horopter tilted some 27° away from the observer (under normal binocular viewing conditions). Also, the horizontal-vertical-illusion is sometimes explained by stating that the length of the vertical line is relatively overestimated because we perceive that line to be slanted backwards (top of the line away from the observer), which means that its real length is longer than its visible length. However, given the fact that we are studying the orientation of abstract rotation axes instead of visible lines, the only conclusion that we can draw from these results is that the deviations that we find might be related to existing evidence that vertical directions tend to be perceived inclined towards the straight ahead direction. That the bias in slant is less for tilts of 120° and 150° than for tilts of 30° and 60° is conceivably due to the use of different hands in those conditions. In addition to a small positive offset, the bias in the tilt shows a periodic behaviour that can be summarised as a tendency towards the horizontal direction: tilt between 0° and 90° is biased towards 0°, whereas tilt between 90° and 180° is biased towards 180°. This can be understood as an anchoring phenomenon; apparently, the horizontal direction, and possibly the vertical direction as well, serve as a reference (anchor) to the subject. Lederman and Taylor (1969) found a similar but slightly smaller anchoring effect for the visual perception of line orientation (they also studied an active touch condition, for which a different and stronger anchoring effect was found). In our case, the references or anchors might be provided by the room, by the direction of gravity (see Howard, 1982) or by our own body.

A rather remarkable result is the similarity of biases found in different sessions. HV's three sessions resulted in nearly identical biases, even though the task was slightly different in these sessions. Apparently, subjects can respond to the stimuli in a highly reproducible manner. This result proves the reliability of the experimental procedure, and particularly the calibration of the set-up. Note that the set-up was rearranged and recalibrated prior to each of HV's sessions, which involved repositioning of the equipment. The similarity between the two subjects forces us to believe that there is some underlying mechanism that dictates the bias. From the current data it cannot be decided whether kinematic restrictions of the limb movements or erroneous visual or kinaesthetic perception cause the non-veridicalities of the response. Subjects are convinced that the object that they see on the screen and the one that they hold in their hand are rotating similarly. Both can be inspected visually, which means that either visual perception is different for distant objects on a screen and nearby objects in one's hand, or visual information about the hand-held object is overruled by kinaesthetic information. Our observation that subjects pay only minor (visual) attention to their hand and mainly look at the screen favours the interpretation that we are basically probing differences between a visual and a kinaesthetic reference system. We will keep this in mind when interpreting the results of the following experiments.

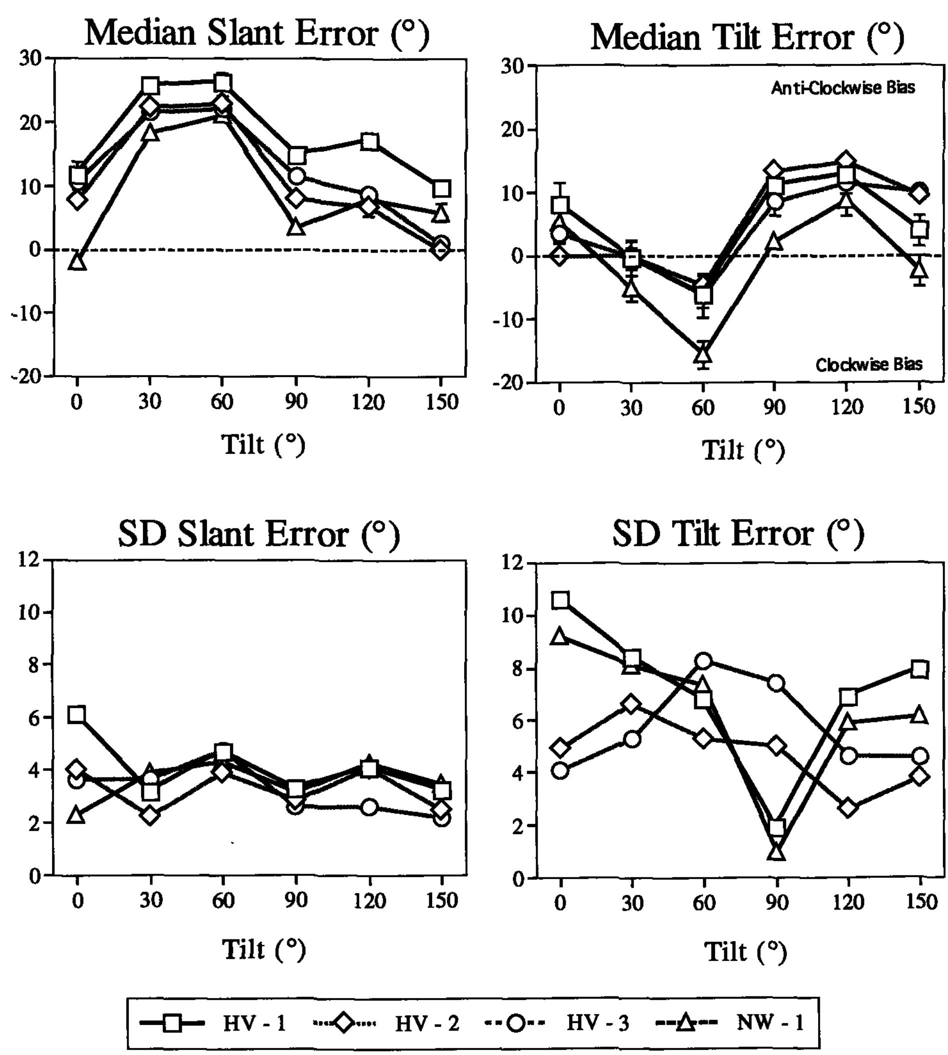


Figure 5.4
Results of Experiment I (fronto-parallel axes of rotation). The top row shows the systematic errors (median error) in slant and tilt, whereas the bottom row shows the variable errors (standard deviations) in slant and tilt. Different symbols correspond to different sessions of subject HV (square, diamond and circle) and NW (triangle). Error bars represent standard error of the mean.

The standard deviation of slant error is quite low (typically 4°) and independent of the tilt of the stimulus rotation axis (lower left panel of figure 5.4). Slant is not varied in this experiment and this might explain the high reproducibility that we find, but the large biases that we have found for slant suggest otherwise. The standard deviation of tilt error (lower right panel of figure 5.4) is also fairly low (between 4–8° in most cases) and roughly independent of the tilt of the stimulus axis of rotation. There is one major exception: the standard deviation for tilt is down to 2° and below for a tilt of 90° in NW's session and in HV's first session (remember that HV's other sessions featured a slightly different task). Apparently, subjects are able to rotate the dummy about a vertical axis with a very high reproducibility. For subject HV this result becomes even more remarkable when one realises that he uses his left and right hand alternately in this condition.

In this first experiment the projected trajectories of the vertices of the rotating objects are always straight lines orthogonal to the (fronto-parallel) axes of rotation. Consequently, the tilt of the rotation axis can be obtained from the orientation of these trajectories. Heeley and Buchanan-Smith (1990), summarising the more recent literature, report that the orientation acuity thresholds for the tilt of visual lines and gratings are generally found to be below 3°. The standard deviations for errors in the tilt of the rotation axis that we find are somewhat higher than this value, which is not too surprising when the experimental differences are taken into account. We conclude that subjects are quite able to perform the task in a reliable and reproducible manner, which stimulates us to continue the current approach.

Exp. II: Rotation about Remembered Axes

Introduction

Our primary goal is to investigate visual perception of rotary motion in space. However, we make use of an action of the motor system to probe visual perception. It is not known how accurately this system can respond to the specific task in the first place, and although the previous experiment showed promising results, a dedicated and more precise estimation of the baseline response accuracy is needed. In the current experiment we ask the subjects to rotate the dummy-object about one of four verbally specified and remembered axes, without providing a visual stimulus. The rationale behind this experiment is that by excluding the visual stimulus the accuracy of the motor system alone is probed. Naturally, we need to be sure that the subjects are confident about the axis of rotation that they are supposed to respond to. Due to the absence of a visual reference deviations from veridicality are probably inevitable. The variance of responses could be pretty low though, in particular because, as we already pointed out in the general method section, the within-trial variability of the orientation of the responded rotation axis is quite low.

Procedure and Experimental Design

The subject is asked to rotate the dummy-object about one of four remembered axes. These axes are the three canonical ones: the horizontal (slant 0°, tilt 180°), the vertical (slant 0°, tilt 90°) and the straight ahead (slant 90°, tilt undefined), and the maximum oblique axis, which is the vector average of the three canonical axes (slant 35.3°, tilt 135°). Subjects are carefully instructed about the orientation of these axes and they seem to have no particular problems memorising them. Because there is no visual stimulus prescribing the magnitude or frequency of rotation, we use a metronome adjusted to 0.8 Hz to indicate the movement frequency (one tick every 1.25 sec to indicate a change of rotation direction; this matches the frequency of oscillation of the rotating stimulus objects in the following experiments). The magnitude of rotation was not controlled but it was similar to the magnitudes used in the other experiments. Instead of showing a visual stimulus, at the start of each trial the experimenter informs the subject about which axis he should rotate the ball. The rest of the procedure is identical to what is described in the general method section. The four axes are run in fixed order, 10 repetitions each. All three subjects participated in one session.

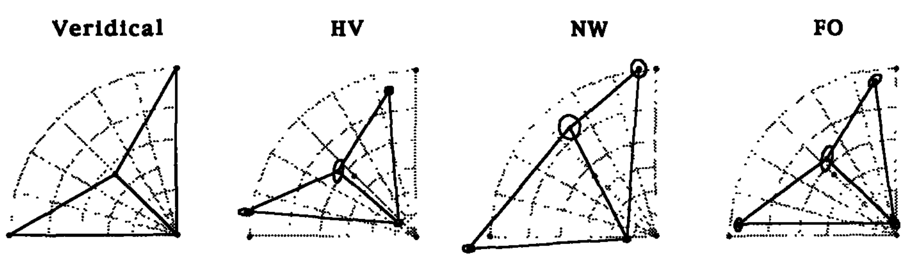


Figure 5.5
Graphical presentation of the results of Experiment II (rotation about remembered axes). The left diagram shows a grid connecting the axes of rotation that were presented to the subject. The other diagrams show the results for the three subjects. The ellipses around the median rotation axes depict the one-sigma contour of the (supposedly) underlying Gaussian probability distribution. The light grey grid in each diagram consists of radial lines of equal tilt and circular lines of equal slant, both with a 15° spacing. The small black dots on this grid correspond to the veridical axes of rotation.

Results

The diagrams in figure 5.5 present the complete results for our three subjects. Bias can be observed as the deviation of the median axis of rotation from the corresponding veridical axis. Also, compare the grids connecting the axes, which gives an overall impression of the distortion of space, or bias. Variance is indicated by the one-sigma variance ellipses. It is immediately clear from these diagrams that significant deviations from veridicality exist (the veridical axis is far outside the variance ellipse around the median axis), and that these deviations differ from subject to subject. It is also clear that the response is highly correlated with the stimulus: the variance ellipses are extremely small and never overlap. The size of the ellipses is tabulated in table 5.1. There are minor differences between the variances related to the three canonical axes, but one should take into account that each of these values is based on only 10 data points. Subjects reported having the greatest difficulty with rotations about the straight ahead direction, primarily because these rotations require a slightly different way of movement (as we explained before). However, the data show that the variance is highest for the maximum oblique axis: the equivalent standard deviation is roughly twice that found for the canonical axes (note that solid angle quadruples when equivalent standard deviation doubles).

	HV		NW		FO	
	Solid Angle (sr)	Equiv. SD (°)	Solid Angle (sr)	Equiv. SD (°)	Solid Angle (sr)	Equiv. SD (°)
horizontal	0.0023	1.6	0.0016	1.3	0.0026	1.7
vertical	0.0021	1.5	0.0069	2.7	0.0036	2.0
straight ahead	0.0046	2.2	0.0036	2.0	0.0101	3.3
maximum oblique	0.0120	3.6	0.0199	4.6	0.0151	4.0

Table 5.1
Quantitative results of Experiment II, tabulated per subject and per axis of rotation.

Discussion

The most important outcome of this experiment is that the variances are rather low. Apparently, the motor system is quite capable of reproducing the same rotational motion every time it is asked to do so. As we noted before, due to the way the movements are executed, the orientation of the forearm resembles to some extent the orientation of the momentary axis of rotation. Several studies have investigated the perception of arm and hand orientation. Although none of those studies investigated dynamic limb movements, it is worth comparing the results of these studies with ours. Soechting (1982) tested subjects' abilities to match elbow joint angles and limb orientation of their two arms (no vision permitted). The standard deviation of errors was found to be lower for matching of limb orientation (6.7° on average) than for matching of elbow joint angle (9.6° on average). Darling and Miller (1995) asked subjects to match arm orientation to a previously memorised orientation (no vision permitted). Variable errors for yaw (≈15°) and elevation (≈11°) were almost independent of whether the match was made in an intrinsic (trunk-fixed) or extrinsic (earth-fixed) co-ordinate system, as was verified by manipulating the orientation of the trunk. In a slightly different experiment, Darling and Miller (1993) tested subjects' ability to reach to remembered object locations and orientations. Variable errors for arm yaw (≈9°) and elevation (≈4.5°) were somewhat smaller than for hand yaw (≈14°), elevation (≈9°), and roll (≈15°), and were 1-2° lower for reaches without visual feedback than with visual feedback. These values are all much higher than those we find in the current experiment. Pollick et al. (1994) tested subjects' ability to match finger orientation to an oriented straight metal rod attached to a robot arm in an unrestricted visual environment. Over a large range of orientations the standard deviation of angular difference between rod and finger orientation was approximately 3.8°, while the mean error angle was about 9°. Pointing in a canonical direction was more accurate; mean angular error was 2-5° with a standard deviation of 0.8-2.5°. Finger orientation can be matched to rod orientation by making a direct comparison of two visual 'lines' in the subject's fronto-parallel plane; this seems to be much easier than rotating an object about an invisible axis. Although matching finger orientation is really quite different from the motor action that we exploit in the present experiment, the findings of Pollick et al. provide at least some comparison values for the reproducibility of spatial orientation over trials. Note that in both experiments arm and hand are repositioned at the beginning of each trial, and small body movements occur in both cases.

We can think of two possible reasons why the variance is higher for the maximum oblique axis than for the canonical axes: either the motor system is less accurate for rotations about non-canonical axes, or information about the axis of rotation itself is less accurately conveyed to the motor system due to the lack of any other reference than a verbal instruction. Support for the latter hypothesis might be found in Heeley and Buchanan-Smith (1990). Studying the perception of tilt of visual lines, they found some evidence for the hypothesis that the lower orientation thresholds that are generally found for vertical and horizontal directions compared to oblique directions are partly a result of the recognition of these canonical directions. This explanation would also clarify why in a certain condition in the previous experiment (vertical rotation axis) the variance in tilt response was extremely low: subjects merely recognised this condition as being 'the one with the vertical rotation axis', and instead started to rotate the dummy-object about this recognised axis, no longer paying attention to the stimulus. In retrospect, this is indeed what the subjects reported. If we accept the response to the maximum oblique axis as representative for most other rotation axes (except for the canonical axes of course), the equivalent standard deviation of 4° is low enough to make further investigations of perceived rotary movement of objects in space feasible.

It is not clear whether we can expect the biases found in this experiment to have any predictive value for biases that occur in experiments with a visual stimulus. In the experiment with the fronto-parallel axes we found very similar behaviour of the two subjects with respect to bias, whereas the same two subjects respond rather differently in the current experiment. The difference between a visual stimulus and a mental image of a rotation axis might be important here. We will pursue this issue a little further in the following experiments.

Exp. III: Slanted Rotation Axes

Introduction

In general, an object that rotates in space can revolve around any conceivable axis of rotation. So far, we have only rotated the stimulus objects about fronto-parallel axes. In this third experiment we show the observers objects rotating about slanted axes; this allows us to investigate the effects of slant of the rotation axis on the variability and veridicality of responses. It is not really clear what we should expect. Slanting the axis of rotation while keeping the amount of rotation equal effectively increases the component of rotation about the line-of-sight, and at the same time reduces the component of rotation about a fronto-parallel axis. The latter component is the one that conveys the information about the structure of the rotating object. Therefore, we might expect the perception of a rigid object rotating in space to deteriorate gradually with increasing slant of the rotation axis, and finally to break down when the axis approaches the viewing direction. It is not unlikely, however, that subjects obtain information about the axis of rotation in a different way. Any point in space (like a corner of an object) that revolves around a slanted axis traces out an elliptical path on the computer display. The eccentricity of this ellipse directly relates to the slant of the rotation axis: the ellipse is a circle when the axis of rotation parallels the viewing-direction, and it becomes a straight line when the axis of rotation becomes fronto-parallel. Irrespective of whether an observer perceives a moving object in space, he could in principle use the shape of the trajectories to compute the orientation of the axis of rotation. We have no way of predicting how subjects' ability to derive the orientation from these trajectories might vary with slant and how this information might influence the mimicked motions, but the possibility should be kept in mind, even though we believe that subjects' responses are certainly not based on such a mechanistic deduction (also, see Todd, 1982).

Procedure and Experimental Design

We define 9 different axes of rotation by combining three different tilts (90°, 135° & 180°) with three different slants (15°, 45° & 75°). These axes are all in the upper left quadrant of the frontal hemisphere of visual directions, which allows us to ask the subjects to use their right hand only. The rest of the procedure is described in the general method section. The 9 axes are presented to the subject in random order, 20 repetitions each. All three subjects took part in one session.

Results

The diagrams in figure 5.6, depicting the results using stereographic projections of the sphere of directions, are drawn on the same scale as the ones in figure 5.5. This means that they can be compared directly. From the size of the ellipses it is immediately clear that the variances found in this experiment are much larger than the baseline response accuracy of the motoric system found in the previous experiment. To allow for a quantitative comparison, we have plotted the size of the variance ellipses as a function of the slant of the axis of rotation in figure 5.7. The numerical values in this graph can be compared with those in table 5.1. Note that subject NW's variance is very low for the two conditions in which the axis has a slant of 15° and a tilt of 90° or 180°: in these conditions the variances are similar to the baseline values for the maximum oblique axis found in the previous experiment. For all other conditions the variance is much higher than the baseline variance. The variance is lowest for an axis with a slant of 15° and it is somewhat higher at 45° slant. The amount of variance at 75° slant relative to the variances at 15° and 45° slant varies strongly from subject to subject, possibly indicating variation of strategy between or even within subjects. The effect of tilt on the variance is somewhat unclear but it is not particularly strong. From the size of the variance ellipses and their overlap we can deduce that the subjects were able to distinguish the different conditions quite reliably, except for some of the slant 75° conditions (notably for subjects NW and FO).

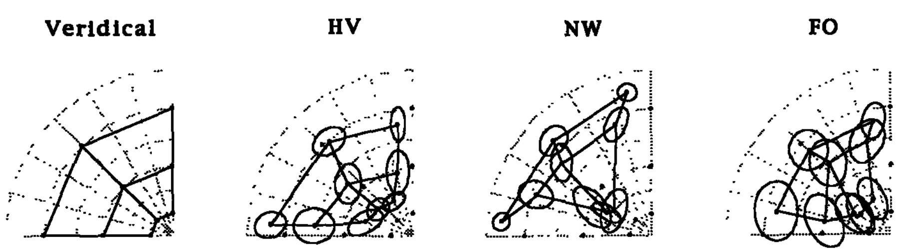


Figure 5.6
Graphical presentation of the results of Experiment III (slanted rotation axes). See caption of figure 5.5 for more details.

Figure 5.6 also reveals some interesting aspects of the veridicality of the mimicked rotation. The range of responses is clearly reduced: if we compute the area (solid angle) enclosed by the grids in the different panels of figure 5.6, we find 1.1 sr for the veridical grid, and only 0.74, 0.53 and 0.57 sr for the measured grids of subjects HV, NW and FO respectively. This reduction seems to be apparent in the tilt range as well as in the slant range.

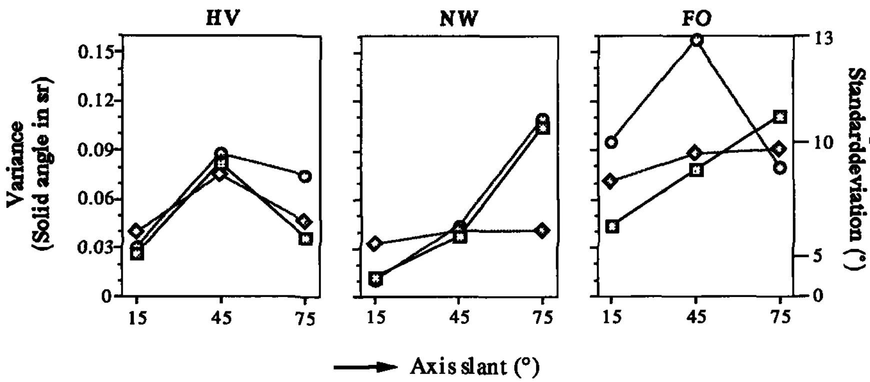


Figure 5.7
Variation in the size of the variance ellipse with the slant of the axis of rotation found in Experiment III. Different symbols correspond to different values of the tilt: 90° (squares), 135° (diamonds) and 180° (circles). The left side shows a graduation in terms of solid angle. On the right side some corresponding values of the equivalent standard deviation are given.

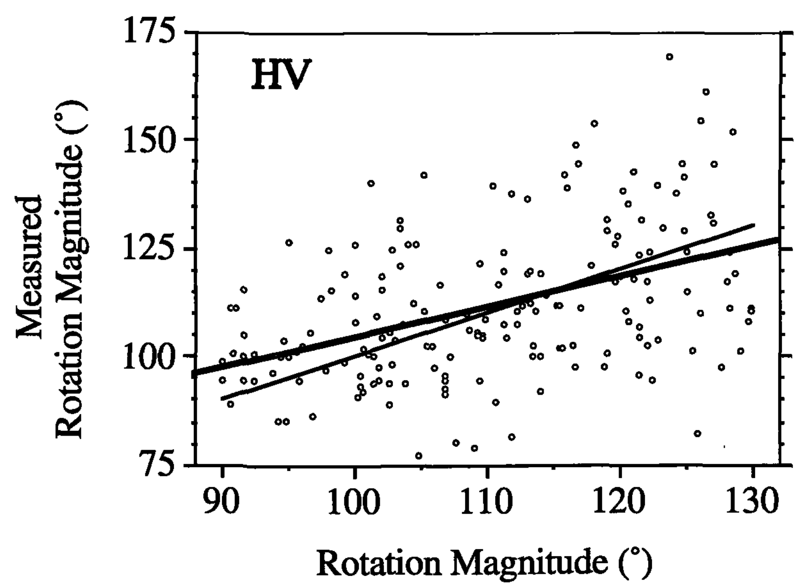


Figure 5.8
Scatter plot of the measured magnitude of rotation versus the magnitude of rotation of the stimulus object for subject HV in Experiment III. Other subjects show similar behaviour. The thick line indicates the linear regression through the data points (slope 0.69), whereas the thin line corresponds to a perfect match.

Contrary to the previous experiments in this experiment subjects were also asked to pay attention to the magnitude of rotation of the stimulus object. Figure 5.8 gives the relation between the magnitude of rotation of the stimulus object and the magnitude of rotation of the dummy-object (as manipulated by subject HV; other subjects show similar behaviour). We have combined the data from all nine conditions, because due to the high amount of variability in the data, differences between conditions turn out to be hardly significant. When averaged over conditions, the standard deviation of the magnitude error is approximately 16° for subject HV, 21° for NW (17° for slant 15° and 45°, but 28° for slant 75°), and 15° for FO. Subjects NW and FO overestimate the magnitude on average (mean magnitude errors of 18° and 7° respectively), whereas subject HV's mean magnitude error is negligible (-1°). The slope of the regression line drawn in figure 5.8 is 0.69. For subjects NW and FO the slope is 0.61 and 0.90 respectively.

Discussion

The angular speed of rotation of the stimulus object was always the same. Therefore, it is not unlikely that subjects were simply rotating the dummy with a constant (average) speed too, with the variation in responded magnitude originating primarily from the confounded variation of the period of oscillation. This might explain why mimicking the magnitude of rotation is not really successful: the range of responses is much larger than the stimulus range, although there is still a weak relation between stimulus and response. A larger range of rotation magnitudes and independent manipulation of angular speed and magnitude of rotation are needed for a more detailed investigation. Such a set of parameters, however, would interfere with the optimisation of comfort and accuracy of subjects' movements. Therefore, we conclude that magnitude of rotation is not adequately addressed in the current paradigm. From now on we will ignore rotation magnitude and concentrate on the orientation of the rotation axis.

The variability of response to the orientation of the rotation axis is typically in the range of 5-10° equivalent standard deviation. The variability in slant found by Pollick et al. for matching finger orientation to the axis of rotation was typically 7°. However, their definition of variability is slightly different from ours: they define variability as the average unsigned angular distance between the centroid of responses and the individual responses. Assuming a Gaussian distribution of errors, variability in their terms is therefore expected to be 79% of the value found with our definition (Yule and Kendall, 1953). In all, this means that the variances that we find are about equal to the variances found by Pollick et al. Note that the two studies employ quite different paradigms and stimuli (their rotating object consisted of 21 dots revolving around a common axis).

It is hard to say what mechanism might underlie the fact that the variance in responded rotation axis depends on the slant of the stimulus rotation axis. Variance does increase with increasing slant, as can be seen from a comparison of the results found with slant 15° and 45° in this experiment and with 0° slant in the first experiment, but for slant 75° the results become unclear. When mimicking the motion of the stimulus object, different subjects might have used different pieces of (stimulus) information in different conditions: complete spatio-temporal reconstruction of structure and motion; shape of the trajectories of object-points; recognition of a specific condition, etc. Altogether we can conclude that subjects are indeed able to reproduce the orientation of the motion of an object rotating about a slanted axis. Large biases do occur, notably a compression of both slant and tilt to the mean stimulus value, but subjects are quite efficient in distinguishing the different rotation axes.

Exp. IV: The Role of Object Orientation

Introduction

When we rotate an object that has a marked extension in a certain direction, the orientation of the object with respect to the axis of rotation has a distinct effect on the appearance of the motion. It makes quite a difference whether a bottle thrown in the air spins about its symmetry axis or about an axis orthogonal to that axis, even if the axis of rotation is objectively the same in both situations. If we force a bottle to rotate about an axis that makes a certain angle with the object's symmetry axis, its motion looks even more unusual. This observation raises the question of whether or not the apparent difference between these motions can be quantified psychophysically. In this fourth experiment we show the observers elongated objects rotating about the same axes as in the previous experiment, but with a variable orientation with respect to these rotation axes.

Procedure and Experimental Design

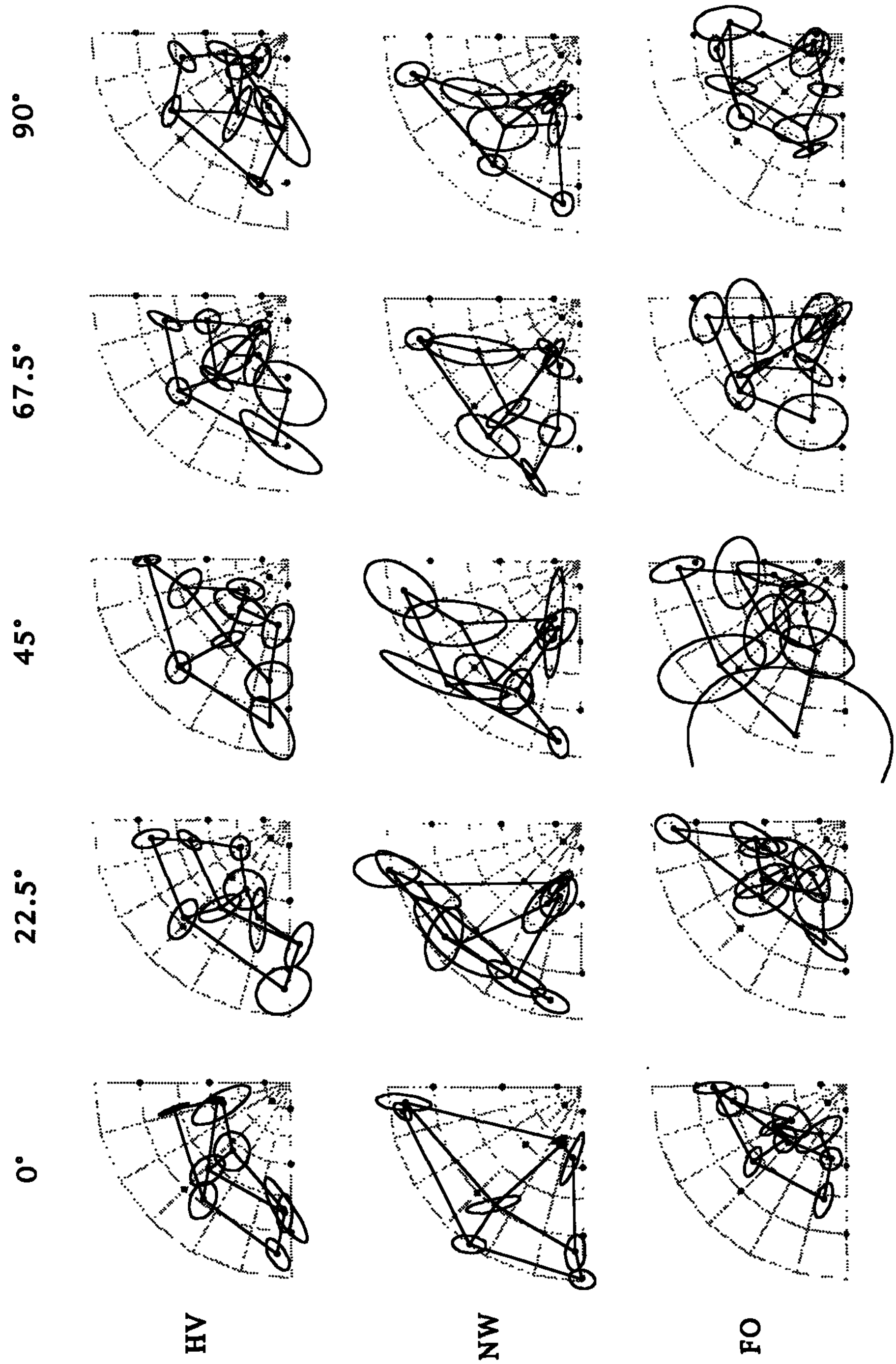
We use basically the same objects as in the previous experiments, but we stretch them by a factor three in one direction (after which they are re-sized to keep the volume the same), which gives them a distinct elongation and orientation. Altogether 45 different conditions are defined. We use the same nine axes of rotation as in the previous experiment: tilt 90°, 135° or 180°, and slant 15°, 45° or 75°. In addition, we vary the angle between the objects' main axis and the axis of rotation: this 'misalignment angle' is either 0°, 22.5°, 45°, 67.5° or 90°, which covers the complete range of possible misalignment angles. We ensure that the object's main axis is slanted away from the rotation axis in a random direction. Each condition was presented 8 times to each of the three subjects.

Results

For each subject we have made five diagrams with the median axes and variance ellipses, corresponding to the five different misalignment angles. These diagrams are shown in figure 5.9. From left to right the misalignment increases from 0° to 90° in this graph. It is immediately clear that several aspects vary with the misalignment angle. First of all, the size of the ellipses appears to be lowest for 0° and 90° misalignment (4–6° equivalent SD), whereas much higher values are found for the angles in-between. Figure 5.10 quantifies this dependency. It shows the average size of the variance ellipse (the median value of the size of the nine ellipses per diagram) as a function of the misalignment angle. There is indeed a very strong effect on the variance of response, at least for two out of three subjects; subject HV shows no significant change of variance with increasing misalignment angle. The average size of the variance ellipse found in Experiment III is plotted in the left part of each of the sub-graphs of figure 5.10. This value is higher than the value that we find in the current experiment, except when the misalignment angle is 45° (and 22.5° for subject NW). For subject HV the differences are not significant. Note that, for the sake of clarity, we have averaged the size of the ellipses over all axis orientations, thereby ignoring the effects of slant and tilt which we proved to exist in the previous experiment.

The deviations from veridicality constitute another interesting part of the results shown in figure 5.9. From a comparison of the grids (which connect the median rotation axes in each diagram) in figure 5.9 with those in figure 5.6, a characteristic distortion of space for each of the subjects seems to emerge. The similarity is most apparent between figure 5.6 and the 45° misalignment condition in figure 5.9. For example, the fact that subject NW makes almost no distinction between a slant of 15° and a slant of 45° is apparent in several of his diagrams, whereas such a distinction is not to be seen in the other subjects' diagrams. In accordance with the previous experiment, we again find smaller measured ranges of slant and tilt than those presented to the subjects. We investigated this a little further after making the following observation: in most diagrams, the line connecting the measured rotation axes corresponding to stimulus rotation axes with a tilt of 90° and the one corresponding to stimulus axes with a tilt of 180° are approximately at right angles. Knowing the properties of the stereographic projection, one realises that these lines are orthogonal on the sphere of directions as well. In other words, for each diagram there is a rotation of the sphere of directions which approximately maps the measured values of the tilt on the veridical values. It turns out that this mapping is roughly the same for each of a subject's diagrams (i.e. for each misalignment), and that the mapping is slightly different for the three subjects. Thus, we hypothesise that subjects respond to a stimulus with a movement in terms of a reference system that is rotated with respect to the stimulus reference system. Figure 5.11 shows how these rotated reference systems are oriented in space for each of the three subjects. If we replot the measured rotation axes in terms of these rotated reference systems we obtain figure 5.12.

Figure 5.9 (next page)
Graphical presentation of the results of Experiment IV. Different rows correspond to different subjects, and different columns correspond to the five different misalignment angles. The grid connecting the axes of rotation that were presented to the subject is identical to the one in Experiment III, and is shown in the left part of figure 5.6.


90°
67.5°
45°
22.5°
0°
HV
NW
FO

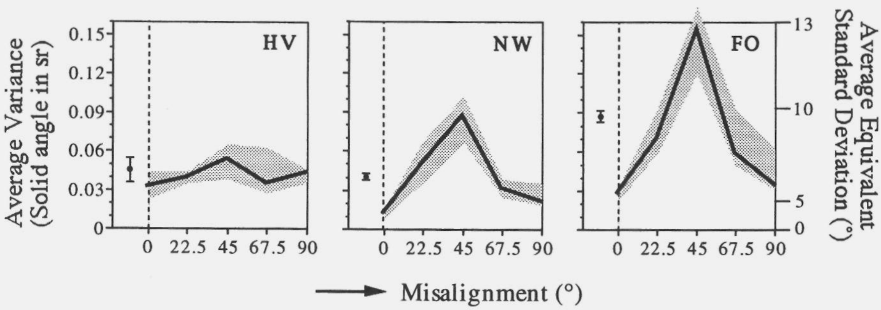


Figure 5.10
Average size of the variance ellipse as a function of the misalignment angle (Experiment IV). The solid black line indicates the median size of the nine ellipses, whereas the grey region marks the semi-interquartile range, which is used as a robust estimator for the standard error of the mean. For the convenience of comparison we have reproduced the corresponding values found in Experiment III in the left part of each sub graph.

Each of the rotated reference systems (solid cubes) depicted in figure 5.11 is obtained by rotating the corresponding stimulus reference system (wire-frame cubes) about an axis in the fronto-parallel plane of the observer. This axis of rotation was tilted away from the horizontal by 63° (HV), 68° (NW), or 41° (FO). The amount of rotation, which can serve as a measure of the mismatch between the two reference systems, was 17° (HV), 36° (NW), and 19° (FO). It is not clear why it appears to be sufficient to rotate the reference system about a fronto-parallel axis and not about a slanted axis. Nevertheless, the effect of comparing the measured rotation axes with the rotated veridical axes instead of the original ones is quite impressive: see the difference between figure 5.9 and figure 5.12. To quantify the remaining deviation we cannot simply compute the errors in tilt because tilt becomes increasingly poorly defined for directions approaching the straight ahead. Instead, we compute the remaining 'tilt' error as the angle that a rotation axis makes with either the vertical, the horizontal or the diagonal plane, depending on which plane the axis should be in. Averaging over all 45 conditions per subject, we find the mean 'tilt' error (±1°) to be -1.0 (HV), -1.2 (NW) and -4.0 (FO). The rather small negative residual indicates that for an optimum match a minor anti-clockwise rotation about the line-of-sight might be needed after all. The mere conclusion that this rotation of space is an appropriate description proves that trying to map the tilts onto the veridical values is not some arbitrary procedure that could have been substituted by, for instance, a mapping of the slants on their veridical values. If we compute the corresponding 'tilt' errors using the uncorrected data in figure 5.9, we find values of +17° for the 90° tilt axes and -14° for the 180° tilt axes, which provides further support for the reliability of the mapping.

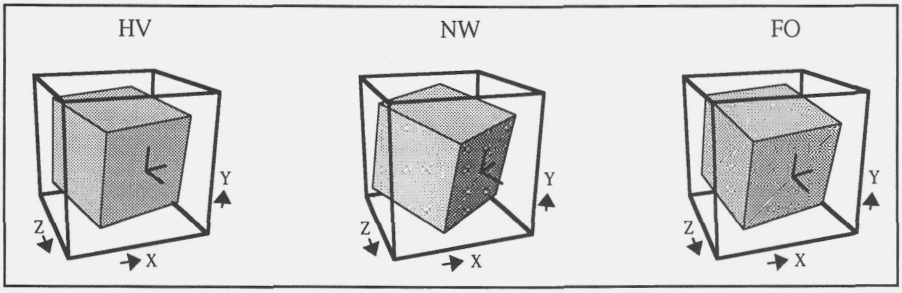


Figure 5.11
These pictures show for each subject the difference in orientation between the stimulus reference system (outside cube) and the reference system in which the movements seem to be made (inside cube). Subject's viewing direction (the straight ahead direction) is the *negative* z-direction.

After fixing the orientation difference between the stimulus reference system and the reference system in which the movements seem to be represented, the remaining deviation from veridicality is almost exclusively in the slant direction. It seems useful to summarise in a separate graph the deviations in slant manifest in figure 5.12. Figure 5.13 shows for each subject the mean slant (collapsed over tilt) as a function of the misalignment angle. The compression of the slant range is visible as the height of the grey area at a certain misalignment angle divided by the distance between the two horizontal lines at 15° and 75°. The effects of misalignment on slant and compression differ from subject to subject.

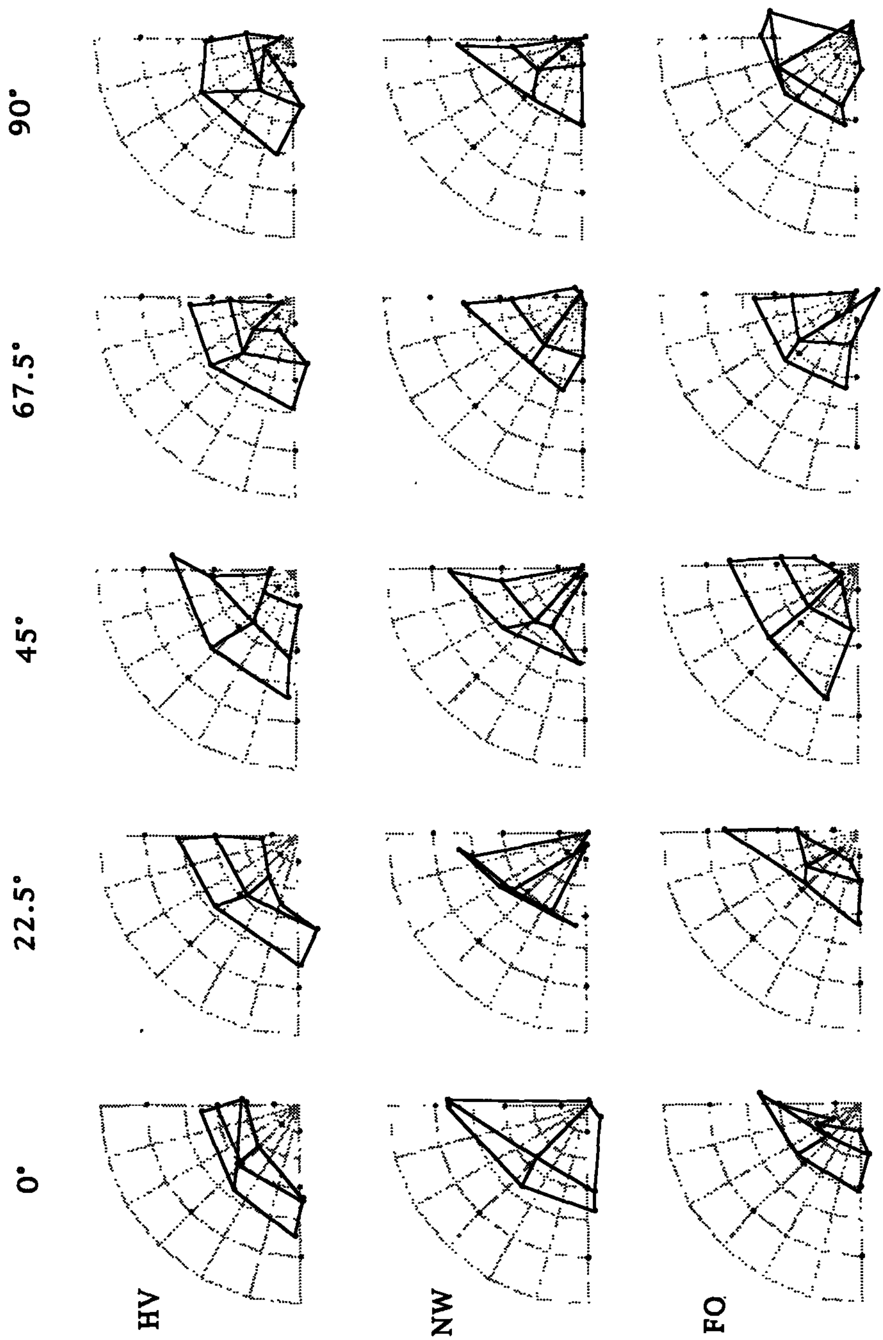


Figure 5.12
Graphical presentation of the results of Experiment IV after correction for the difference in orientation between the stimulus reference system and the reference system in which the movements seem to be made (compare this figure with figure 5.9). To improve clarity we have not drawn the variance ellipses.

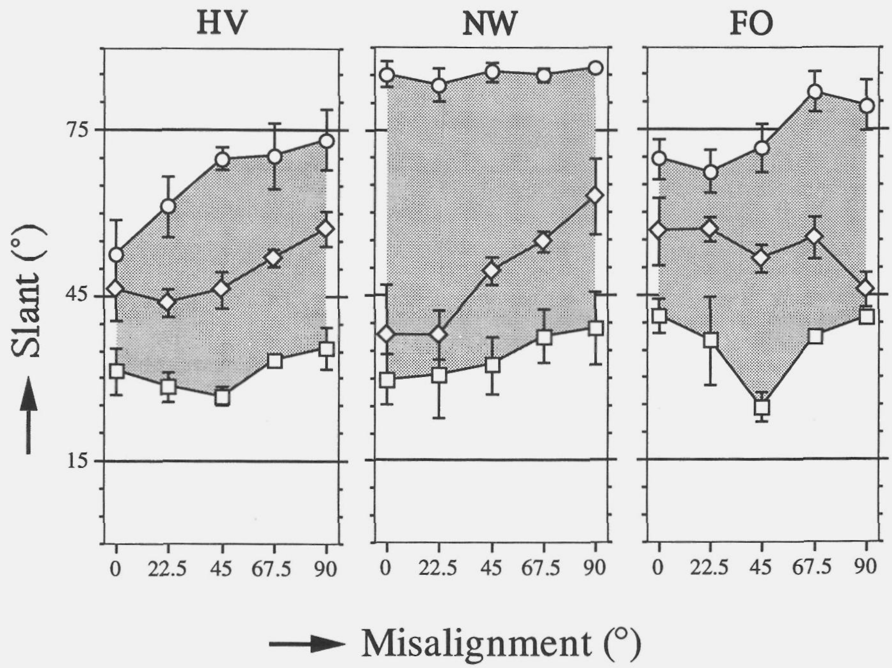


Figure 5.13
Average slant (collapsed over tilt) as a function of the misalignment angle, based on the corrected data presented in figure 5.12. Different symbols correspond to different slants of the stimulus rotation axis: 15° (squares), 45° (diamonds) and 75° (circles). The grey area indicates the range of slant responses for each value of the misalignment angle.

Discussion

The difference between an object rotating about its symmetry axis and the same object rotating about an axis that makes an angle with this symmetry axis is most evident in the increased variability of the motion as mimicked by the observers. What exactly causes this effect is not clear. Subjects might become less confident about what they see and how they should respond when the axis of rotation is not aligned with the object. Another possible explanation is that the elongation of the object pulls or pushes the orientation of the rotation axis in a certain direction. For example, the motion of an object rotating about a 'misaligned' axis can be decomposed into two rotation components: a change in the orientation of the objects' main axis and a rotation of the object around this axis. A change in the orientation of the object's main axis is readily visible as a change of the projected contour of the (approximately symmetric) object, whereas the component of rotation around the object's main axis has to be deduced from changes in the optical structure inside this contour. It is possible that subjects under-or over-estimated the component of rotation around an object's main axis due to the limited surface structure of the objects that we showed. This effect, if it exists, depends strongly on the orientations of symmetry axis and rotation axis in relation to the fronto-parallel plane. The way we control the object orientation combined with the limited number of only 8 trials per condition prevents us from analysing this effect, but it is definitely an interesting possibility. Finally, a third way to account for the effects of misalignment could be connected with the way subjects handle the objects. An intuitive way to hold the dummy-object is to grasp it as if it has the same instantaneous orientation as the elongated object on the screen. Different orientations of the objects then give rise to different grasps, possibly resulting in a bias that varies from trial to trial, which in turn causes an increase in the variance of responses over trials. We have not recorded the way subjects hold the dummy-object, but from observing the subjects in action we conclude that small differences in grasp do indeed seem to be present.

A similar mis-orientation of the reference system in which the movements seem to be made relative to the stimulus reference system was also found in Experiment III. In fact, when we apply the rotations of the reference systems found in this experiment to the data of Experiment III, the tilt errors nearly disappear. Taking into account that there was a 1-3 month gap between experiments III and IV, during which, among other things, all the equipment was repositioned, the rotation of space seems to be an intrinsic property of the movement characteristics of a subject. Why the rotation of space is in this particular direction is of course an interesting question. The orientation of the rotated straight ahead direction is roughly parallel to the orientation of the right forearm during the movement, as well as to the imaginable line through the hand holding the dummy-object and the shoulder. Also, the rotated straight ahead is roughly in-between the veridical straight ahead and the mean orientation of the rotation axis during the experiment. We have thought of several experiments that could be done to distinguish between these different possible explanations, one of them being a repetition of the same experiments with the left hand and rightward pointing axes of rotation, but we have not yet attempted to perform such experiments.

Concluding Remarks

In principle, the subjects can utilise the motion of the object on the display in several ways to obtain the rotation parameters. There is abundant information for a complete (metric) reconstruction of the motion and structure of the object, although human subjects generally seem to ignore part of this information, making errors in the absolute depth-scale of the stimulus (Todd & Bressan, 1990; Liter, Braunstein and Hoffman, 1993; Werkhoven & Van Veen, 1995; Van Veen & Werkhoven, in press; Norman, Todd, Perrotti & Tittle, in press; Hogervorst, Kappers & Koenderink, submitted). Such errors are expected to give rise to an increasing variability of response when the axes of rotation slant away from the fronto-parallel plane, possibly accompanied by large biases in the slant (but not in the tilt) of the perceived rotation axis. The results that we found do indeed display these characteristics. Other sources of information about the orientation of the axis of rotation include the shape of the projected trajectories of the moving points of an object, and the deformation of the object's occluding contour (Giblin, Pollick & Rycroft, 1994).

One of the difficulties of the current approach is that it remains somewhat unclear whether we are probing a sensorimotor or a cognitive mode of spatial representation (or a combination of both; also see Paillard, 1987). Phenomena like 'blindsight' (Weiskrantz, Warrington, Sanders and Marshall, 1974) prove that in some cases visual information can appear to be absent at the cognitive level and yet still be present in the sensorimotor system, which shows that differences between the two representations exist. As far as our experiments are concerned, however, the fact that the mimicking approach seems to work pretty well (people show a high ability) is sufficient for our objectives: any method that reveals a high grade of reproducibility and a firm relationship between stimulus and response is adequate to investigate the extent to which humans are capable of processing the visual information. Nevertheless, it is possible to interpret the roles of the visual and kinaesthetic systems in this experiment at least to some extent. In short, both our results and the results of Pollick et al. (1994) show that actively responding to the orientation of the axis of rotation results in an increase in the variability of responses relative to the baseline response accuracy of the corresponding motor system, even though pointing and rotations of the forearm are quite different actions. Thus, we seem to be justified in concluding that both studies are primarily probing aspects of visual information processing, rather than probing the properties of the motor system. The role of kinaesthetic information is most apparent in the pattern of deviations from veridicality that we find. Subjects are continuously able to see their hand and the object in it, which means that they can visually compare what they are doing with what they are supposed to mimic. In theory, such a comparison would calibrate most systematic differences between stimulus and response. The fact that we nonetheless find severe deviations from veridicality signals a dissociation of the (visual) stimulus reference system and the (primarily kinaesthetic) movement reference system. The rotation of space depicted in figure 5.11 provides further support for this view.

In the domain of mental rotation imagining rotary motion has been studied extensively (e.g. see the interesting paper by Pani and Dupree (1994) on the role of reference systems in mental rotation tasks). It is not clear whether the results of these studies can be transferred to the quite different incarnation of rotary motion that we have investigated. For instance, the effort and accuracy with which rotations can be imagined depend strongly on the orientation of the axis of rotation with respect to local and global reference systems: imagining the rotation of an object about a slanted axis is much less accurate than imagining the same object rotating about a vertical axis aligned with the visual environment. In contrast, in our experiments we do not find such a strong effect of orientation of the rotation axis on the variability of the responses, although minor individual effects of slant do exist. On the other hand, alignment of the principal directions of an object with the axis and planes of rotation is known to be very helpful in the imagining of rotations (Pani and Dupree, 1994, and references therein), a finding which is clearly confirmed by the results of our fourth experiment. The characteristics of imagined rotations seem to differ somewhat from the characteristics of actual rotations; this might be because mental rotations are often more concerned with one or two given or imagined orientations than with the actual change of orientation over time.

Mimicking rotary movements has certainly proved to be a worthwhile method for investigating the processing of visually acquired information. We have only just begun to explore this method of researching visual processing, and many other experiments can be conjectured. Realistic possibilities include the study of the perception of more complicated rotary movements such as those involving time-dependent rotation parameters or rotation axes that do not pass through an object's centre-of-mass. It should also be feasible to incorporate into such experiments translational components of motion.

Similar approaches in other domains of motor action are promising too. For example, Ringach, Hawken and Shapley (1995) have used vergence eye-movements to show that subjects do indeed perceive a 3D structure when monocularly observing random dot kinematograms. Both in their study and in ours a motor action is used to probe a subject's spatial perception. The relevance of this approach is well illustrated by a quotation taken from Von Hofsten (1986): '..spatial perception has been designed to serve action and only in the context of action can spatial perception be optimally tested'.

REFERENCES

Aggarwal, J. K., Davis, L. S. & Martin, W. N. (**1981**). Correspondence processes in dynamic scene analysis. *Proceedings of the IEEE, 69, 6,* 562-572.
Aggarwal, J. K. & Nandhakumar, N. (**1988**). On the computation of motion from sequences of images: a review. *Proceedings of the IEEE, 76, 8,* 917-935.

Baird, J. C. (**1970**). *Psychophysical Analysis of Visual Space.* Pergamon Press, Oxford.
Banach, S. (**1951**). *Mechanics.* Warszawa, Poland.
Bennett, B. M., Hoffman, D.D., Nicola, J. E. & Prakash, C. (**1989**). Structure from two orthographic views of rigid motion. *Journal of the Optical Society of America, A, 6, 7,* 1052-1069.
Braunstein, M. L. & Andersen, G. J. (**1984**). Shape and depth perception from parallel projections of three-dimensional motion. *Journal of Experimental Psychology: Human Perception & Performance, 10, 6,* 749-760.
Brown, J. F. (**1928**). Über gesehene Geschwindigkeiten (transl.: About observed velocities). *Psychologische Forschung, 10,* 84-101.
Brown, J. F. (**1931**). The visual perception of velocity. *Psychologische Forschung, 14,* 199-232.
Brunswik, E. (**1934**). *Wahrnehmung und Gegenstandswelt; Grundlegung einer Psychologie vom Gegenstand her (transl.: Perception and the world of objects; foundation of a psychology based on objects).* Franz Deuticke, Wien.
Brunswik, E. (**1956**). *Perception and the Representative Design of Psychological Experiments.* University of California Press, Berkeley.

Caudek, C. & Proffitt, D.R. (**1993**). Depth perception in motion parallax and stereokinesis. *Journal of Experimental Psychology: Human Perception & Performance, 19, 1*, 32-47.
Charpentier, A. (**1891**). Analyse expérimentale de quelques éléments de la sensation de poids (transl.: Experimental analysis of several aspects of the perception of weight). *Archives de Physiologie Normales et Pathologiques, 1.3, 23*, 122-135.
Cogan, A. I. (**1979**). The relationship between the apparent vertical and the vertical horopter. *Vision Research, 19*, 655-665.
Cormack, E. O. & Cormack, R. H. (**1974**). Stimulus configuration and line orientation in the horizontal-vertical illusion. *Perception & Psychophysics, 16 (2)*, 208-212.
Cortese, J. M. & Andersen, G. J. (**1991**). Recovery of 3-D shape from deforming contours. *Perception & Psychophysics, 49(4)*, 315-327.

Darling, W. G. & Miller, G. F. (**1993**). Transformations between visual and kinesthetic coordinate systems in reaches to remembered object locations and orientations. *Experimental Brain Research, 93*, 534-547.
Darling, W. G. & Miller, G. F. (**1995**). Perception of arm orientation in three-dimensional space. *Experimental Brain Research, 102*, 495-502.
van **Doorn**, A. J. & Koenderink, J. J. (**1982**). Visibility of Movement Gradients. *Biological Cybernetics, 44*, 167-175.
Dudani, S. A., Breeding, K. J., & McGhee, R. B. (**1977**). Aircraft identification by moment invariants. *IEEE Transactions on Computers, C-26, 1*, 39-45.

Eagle, R. A. & Blake, A. (**1995**). Two-dimensional constraints on three-dimensional structure from motion tasks. *Vision Research, 35, 20*, 2927-2941.
Ellis, R. R. & Lederman, S. J. (**1993**). The role of haptic versus visual volume cues in the size-weight illusion. *Perception & Psychophysics, 53 (3)*, 315-324.

Faugeras, O. D. & Maybank, S. (**1990**). Motion from point matches: multiplicity of solutions. *International Journal of Computer Vision, 4*, 225-246.
Fick, A. (**1852**). Erörterung eines physiologisch-optischen Phänomens (transl.: Discussion of a physiological-optical phenomenon). *Zeitschrift für rationelle Medicin, 2*, 83-94.
Franken, R. B. & Larrabee, C. B. (**1928**). *Packages that sell.* Harper & Brothers Publishers, New York.

Giblin, P. J., Pollick, F. E. & Rycroft, J. E. (**1994**). Recovery of an unknown axis of rotation from the profiles of a rotating surface. *Preprint Journal of the Optical Society of America, A.*
Gibson, J. J. & Gibson, E. J. (**1957**). Continuous perspective transformations and the perception of rigid motion. *Journal of Experimental Psychology, 54, 2*, 129-138.
Gurevich, G. B. (**1964**). *Foundations of the theory of algebraic invariants.* P. Noordhoff Ltd, Groningen, The Netherlands.

Heeley, D. W. & Buchanan-Smith, H. M. (**1990**). Recognition of stimulus orientation. *Vision Research, 30, 10*, 1429-1437.

Hilbert, D. and Cohn-Vossen, S. (**1944**). *Anschauliche Geometrie*, Dover Publications, New York.
Hildreth, E. C., Grzywacz, N. M., Adelson, E. H. & Inada, V. K. (**1990**). The perceptual buildup of three-dimensional structure from motion. *Perception & Psychophysics, 48(1)*, 19-36.
Hoffman, D. D. & Bennett, B. M. (**1986**). The computation of structure from fixed-axis motion: rigid structures. *Biological Cybernetics, 54*, 71-83.
von Hofsten, C. (**1986**). Early spatial perception taken in reference to manual action. *Acta Psychologica 63*, 323-335.
Hogervorst, M. A., Kappers, A. M. L. & Koenderink, J. J. (**submitted**). Discrimination of hinged planes with different depth-to-width ratios from motion parallax. *Technical Report, Helmholtz Instituut, Utrecht, The Netherlands. (Submitted for publication).*
Howard, I. P. (**1982**). *Human Visual Orientation*, John Wiley & Sons, Toronto.
Hu, M. K. (**1962**). Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory, 8*, 179-187.
van de Hulst, H. C. (**1964**). *Light scattering by small particles.* John Wiley & Sons, New York, third printing.

Jaynes, E. T. (**1968**). Prior Probabilities. In: Rosenkrantz, R. D. (Ed.), *E.T. Jaynes: Papers on Probability, Statistics and Statistical Physics*, p. 114-130. Dordrecht 1983, Holland; D. Reichel publishing company. (Reprinted from IEEE transactions on systems science and cybernetics, SSC-4, 1968, 227-241.)
Johnston, A. & Wright, M. J. (**1986**). Matching velocity in central and peripheral vision. *Vision Research, 26*, *7*, 1099-1109.

Kaiser, M. K. (**1990**). Angular velocity discrimination. *Perception & Psychophysics, 47*, 149-156.
Kaiser, M. K. & Calderone, J. B. (**1991**). Factors influencing perceived angular velocity. *Perception & Psychophysics, 50 (5)*, 428-434.
Kanatani, K. (**1993**). Unbiased estimation and statistical analysis of 3-D rigid motion from two views. *IEEE Transactions on Pattern Analysis & Machine Intelligence, 15*, 37-50.
Koenderink, J. J. (**1985**). Space, form and optical deformations. In: Ingle, D. J., Jeannerod, M. & Lee, D. N. (eds). *Brain Mechanisms and Spatial Vision*, Martinus Nijhoff Publishers, Dordrecht, The Netherlands.
Koenderink, J. J. (**1986**). Optic flow. *Vision Research, 26, 1*, 161-180.
Koenderink, J. J. & van Doorn, A. J. (**1991**). Affine structure from motion. *Journal of the Optical Society of America, A, 8, 2*, 377-385.
Künnapas, T. M. (**1955**). An analysis of the "vertical-horizontal illusion". *Journal of Experimental Psychology, 49, 2*, 134-140.

Lacquaniti, F. & Soechting, J. F. (**1982**). Coordination of arm and wrist motion during a reaching task. *Journal of Neuroscience, 2, 4*, 399-408.
Lauer, L. (**1929**). Untersuchungen über die scheinbare Größe von Körpern (transl.: Investigations concerning the apparent size of objects). *Archiv für die Gesamte Psychologie, 68*, 295-324.
Lederman, S. J. & Taylor, M. M. (**1969**). Perception of interpolated position and orientation by vision and active touch. *Perception & Psychophysics, 6 (3)*, 153-159.

Liter, J. C., Braunstein, M. L. & Hoffman, D. D. (**1993**). Inferring structure from motion in two-view and multiview displays. *Perception, 22*, 1441-1465.
Liter, J. C. & Braunstein, M. (**1994**). Judged rotation magnitude in structure from motion with and without frontal views. *Investigative Ophthalmology & Visual Science, 35*, *4*, p. 1742.
Longuet-Higgins, H. C. (**1988**). Multiple interpretations of a pair of images of a surface. *Proceedings of the Royal Society of London, A, 418*, 1-15.

McKee, S. P. (**1981**). A local mechanism for differential velocity detection. *Vision Research, 21, 4*, 491-500.
McKee, S. P. & Nakayama, K. (**1984**). The detection of motion in the peripheral visual field. *Vision Research, 24, 1*, 25-32.

Norman, J. F., Todd, J. T., Perotti, V. J. & Tittle, J. S. (**in press**). The visual perception of 3D length. (*Accepted for publication in Journal of Experimental Psychology: Human Perception & Performance*)

Orban, G. A., van Calenbergh, F., de Bruyn, B. & Maes, H. (**1985**). Velocity discrimination in central and peripheral visual field. *Journal of the Optical Society of America A, 2, 1*, 1836-1847.

Paillard, J. (**1987**). Cognitive versus sensorimotor encoding of spatial information. In: Ellen, P. & Thinus-Blanc, C. (eds). *Cognitive Processes and Spatial Orientation in Animal and Man, Volume II: Neurophysiology and Developmental Aspects*, Martinus Nijhoff Publishers, Dordrecht, The Netherlands, p. 43-77.
Pani, J. R. and Dupree, D. (**1994**). Spatial reference systems in the comprehension of rotational motion. *Perception, 23*, 929-946.
Petersik, J. T. (**1980**). Rotation judgements and depth judgements: Separate or dependent processes? *Perception & Psychophysics, 27(6)*, 588-590.
Petersik, J. T. (**1991**). Perception of three-dimensional angular rotation. *Perception & Psychophysics, 50(5)*, 465-474.
Piaget, J. (**1941**). *The child's conception of number*. London; Routledge & Kegan Paul Ltd.
Pollick, F. E., Nishida, S., Koike, Y. & Kawato, M. (**1994**). Perceived motion in structure from motion: pointing responses to the axis of rotation. *Perception & Psychophysics, 56 (1)*, 91-109.
Pollock, W. T. & Chapanis, A. (**1952**). The apparent length of a line as a function of its inclination. *Quarterly Journal of Experimental Psychology, 4*, 170-178.
Prokop, R. J. & Reeves, A. P. (**1992**). A survey of moment-based techniques for unoccluded object representation and recognition. *Comput. Vision Graphics Image Process: Graphical Models and Image Processing, 54, 5*, 438-460.

Richards, W. (**1985**). Structure from stereo and motion. *Journal of the Optical Society of America, A, 2, 2*, 343-349.
Ringach, D. L., Hawken, M. J. and Shapley, R. (**1995**). Perception of three-dimensional structure from motion elicits vergence eye movements in humans. *Investigative Ophthalmology & Visual Science, 36, 4*, p. 418.

Routh, E. J. (**1960**). *Dynamics of a System of Rigid Bodies.* Dover Publications Inc., New York.
Rubin, N., Solomon, S., & Hochstein, S. (**1995**). Restricted ability to recover three-dimensional global motion from one-dimensional local signals: theoretical observations. *Vision Research, 35, 4,* 569-578.

Sadjadi, F. A. & Hall, E. L. (**1980**). Three-dimensional moment invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-2, 2,* 127-136.
Scott, G. L. & Longuet-Higgins, H. C. (**1991**). An algorithm for associating the features of two images. *Proceedings of the Royal Society of London, B, 244,* 21-26.
Shiffrar, M. M. & Shepard, R. N. (**1991**). Comparison of cube rotations around axes inclined relative to the environment or to the cube. *Journal of Experimental Psychology: Human Perception and Performance, 17, 1,* 44-54.
Shoemake, K. (**1985**). Animating rotation with quaternion curves. *Computer Graphics, 19, 3,* 245-254.
Simpson, W. A. (**1993**). Optic flow and depth perception. *Spatial Vision, 7, 1,* 35-75.
Snippe, H. P. & Werkhoven, P. (**1993**). Pulse modulation detection in human motion vision. *Vision Research, 33, 5/6,* 647-656.
Snowden, R. J. & Braddick, O.J. (**1991**). The temporal integration and resolution of velocity signals. *Vision Research, 31, 5,* 907-914.
Soechting, J. F. (**1982**). Does position sense at the elbow reflect a sense of elbow joint angle or one of limb orientation? *Brain Research, 248,* 392-395.
Soechting, J. F. (**1987**). Coordinate representations underlying arm movements in three-dimensional space. In: Ellen, P. & Thinus-Blanc, C. (eds). *Cognitive Processes and Spatial Orientation in Animal and Man, Volume II: Neurophysiology and Developmental Aspects,* Martinus Nijhoff Publishers, Dordrecht, The Netherlands, p. 35-42.
Stanek, R. J. (**1968**). Surface and volume judgements of three-dimensional shapes. *Psychonomic Science, 11(4),* 121-122.
Stanek, R. J. (**1969**). A parametric study of volume and surface judgements. *Perception & Psychophysics, 6(1),* 16-18.

Todd, J. T. (**1982**). Visual information about rigid and nonrigid motion: a geometric analysis. *Journal of Experimental Psychology: Human Perception and Performance, 8, 2,* 238-252.
Todd, J. T., Akerstrom, R. A., Reichel, F. D. & Hayes, W. (**1988**). Apparent rotation in three-dimensional space: Effects of temporal, spatial and structural factors. *Perception & Psychophysics, 43,* 179-188.
Todd, J. T. & Bressan, P. (**1990**). The perception of three-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics, 48(5),* 419-430.
Todd, J. T. & Norman, J. F. (**1991**). The visual perception of smoothly curved surfaces from minimal apparent motion sequences, *Perception & Psychophysics, 50 (6),* 509-523.
Tynan, P. D. & Sekuler, R. (**1982**). Motion processing in peripheral vision: reaction time and perceived velocity. *Vision Research, 22,* 61-68.

Ullman, S. (**1979**). *The Interpretation of Visual Motion.* Cambridge, MA, MIT Press.

Ullman, S. (**1983**). Recent computational studies in the interpretation of structure from motion. In Beck, J., Hope, B. & Rosenfeld, A. (Eds), *Human and Machine Vision*, p. 459-480. New York, Academic Press.

Ullman, S. (**1984**). Maximizing rigidity: the incremental recovery of 3-D structure from rigid and non-rigid motion. *Perception, 13*, 255-274.

van Veen, H. A. H. C., Kappers, A. M. L., Koenderink, J. J. & Werkhoven, P. (**in press**). Discriminating the volume of motion-defined solids. *(Accepted for publication in Perception & Psychophysics; also see Chapter II of this thesis).*

van Veen, H. A. H. C. & Werkhoven, P. (**in press**). Metamerisms in structure-from-motion perception. *(Accepted for publication in Vision Research; also see Chapter III of this thesis).*

van Veen, H. A. H. C., Kappers, A. M. L., Koenderink, J. J. & van Woerkom, N. (**submitted**). Spatial configuration affects judgements of 3D angular speed. *Technical Report, Helmholtz Instituut, Utrecht, The Netherlands. (Submitted for publication; also see Chapter IV of this thesis).*

Wallach, H. & O'Connell, D. (**1953**). The kinetic depth effect. *Journal of Experimental Psychology, 45*, 4, 205-217.

Weiskrantz, L., Warrington, E. K., Sanders, M. D. & Marshall, J. (**1974**). Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain, 97*, 709-728.

Werkhoven, P., Snippe H. P. & Toet, A. (**1992**). Visual processing of optic acceleration. *Vision Research, 32, 23*, 2313-2329.

Werkhoven, P. & van Veen, H. A. H. C. (**1995**). Extraction of relief from visual motion. *Perception & Psychophysics, 57 (5)*, 645-656.

Werkhoven, P. & Snippe, H. P. (**in press**). An efficient adaptive procedure for psychophysical discrimination experiments. *(Accepted for publication in Behavior Research Methods, Instruments & Computers).*

Whitaker, D., Mäkelä, P., Rovamo, J. & Latham, K. (**1992**). The influence of eccentricity on position and movement acuities as revealed by spatial scaling. *Vision Research, 32, 10*, 1913-1930.

Willats, J. (**1985**). Drawing systems revisited: The role of denotation systems in children's figure drawings. In Freeman, N. H. & Cox, M. V. (Eds), *Visual order: The nature and development of pictorial representation*, pp 78-100. Cambridge; Cambridge University Press.

Willats, J. (**1992**). Seeing lumps, sticks and slabs in silhouettes. *Perception, 21*, 481-496.

Yule, G. U. & Kendall, M. G. (**1953**). *An Introduction to the Theory of Statistics.* Charles Griffin & Company Ltd., London, p. 140.

PUBLICATIONS

Most of the work described in this thesis has been accepted or submitted for publication in refereed journals, or has been presented at international conferences.

CHAPTER 2:

van Veen, H. A. H. C., Kappers, A. M. L., Koenderink, J.J. & Werkhoven, P. (**in press**). Discriminating the Volume of Motion-Defined Solids. *(Accepted for publication in Perception & Psychophysics).*
van Veen, H. A. H. C., Werkhoven, P., Kappers, A. M. L. & Koenderink, J.J. (**1993**). Global Structure-from-Motion: Psychophysics of Volume Discrimination. *Perception, vol. 22 supplement, #100C.* (Poster presented at the 16th European Conference on Visual Perception, Edinburgh, Scotland, 25-29 August 1993).

CHAPTER 3:

van Veen, H. A. H. C. & Werkhoven, P. (**in press**). Metamerisms in Structure-from-Motion Perception. *(Accepted for publication in Vision Research).*

van Veen, H. A. H. C., Werkhoven, P., Kappers, A. M. L. & Koenderink, J.J. (**1994**). Metamerisms in Structure-from-Motion Perception. *Perception, vol. 23 supplement, #38A.* (Poster presented at the 17th European Conference on Visual Perception, Eindhoven, The Netherlands, 4-8 September 1994).

CHAPTER 4:

van Veen, H. A. H. C., Kappers, A. M. L., Koenderink, J.J. & van Woerkom, N. (**submitted**). Spatial Configuration Affects Judgements of 3D Angular Speed. *(Submitted for Publication).*

van Veen, H. A. H. C., Kappers, A. M. L., Koenderink, J.J. & van Woerkom, N. (**1995**). Discrimination of Angular Velocities in 3D-Space. *Investigative Ophthalmology & Visual Science, Vol. 36, No. 4, #1702.* (Poster presented at the 1995 Annual Meeting of The Association for Research in Vision and Ophthalmology, Fort Lauderdale, Florida, 14-19 May 1995).

van Veen, Hendrik-Jan (**1995**). Waarneming van drie-dimensionale beweging (transl.: Perception of three-dimensional motion). *Nieuwsbrief Nederlandse Vereniging voor Psychonomie, Augustus 1995, nr. 2,* 12-13.

CHAPTER 5:

van Veen, H. A. H. C., Kappers, A. M. L., Koenderink, J.J. & van Woerkom, N. (**in preparation**). Mimicking the Rotary Movement of Objects in Space. *(Being prepared for submission).*

Closely related publications of which the author of this thesis is co-author:

Werkhoven, P. & **van Veen,** H. A. H. C. (**1995**). Extraction of Relief from Visual Motion. *Perception & Psychophysics, 57 (5),* 645-656.

Werkhoven, P. & **van Veen,** H. A. H. C. (**1993**). Relief-Recovery from Motion. *Perception, vol. 22, supplement, #10B.* (Paper presented at 16th European Conference on Visual Perception, Edinburgh, United Kingdom, 25-29 August 1993).

NEDERLANDSE SAMENVATTING

Het waarnemen van de visuele wereld is een van onze belangrijkste middelen om te communiceren met de omgeving. Op een elementair fysisch beschrijvingsniveau zijn mensen uitgerust met een complex zintuig (onze ogen) dat gevoelig is voor straling in een bepaald gedeelte van het electromagnetisch spectrum. Op een meer abstract niveau lijkt het visuele zintuig ons te voorzien van een preciese en gedetailleerde beschrijving van de voorwerpen in de wereld om ons heen. Wetenschappers die zich bezighouden met de werking van de menselijke visuele waarneming streven er naar om deze ver uit elkaar liggende benaderingen van de rol van het visuele zintuig in het functioneren van mensen aan elkaar te relateren. Zij doen dat onder andere door allerlei tussenliggende beschrijvingsniveaus te bestuderen.

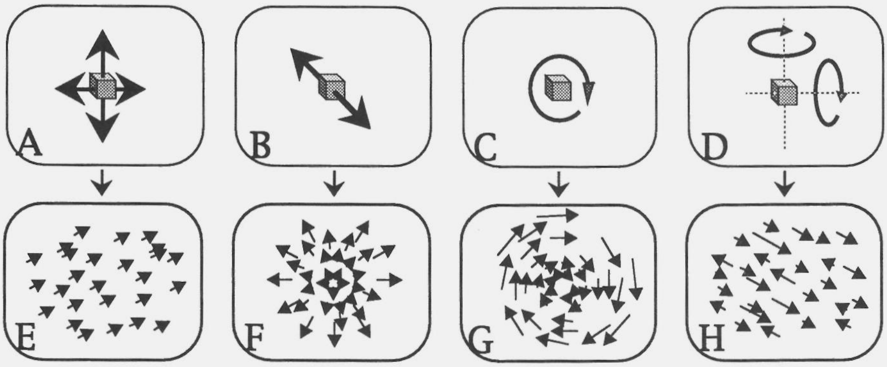
Informatie over de visueel waarneembare wereld wordt vaak geordend in 'cues'. Een cue is een verzameling feiten en aannames met een zekere logische samenhang. Een voorbeeld van zo'n cue is de schijnbare grootte van een objekt in relatie tot zijn afstand tot de waarnemer; de waargenomen geprojekteerde grootte van een objekt gecombineerd met de ervaring die we hebben dat kleiner lijkende objekten van gelijke grootte verder weg zijn geeft ons informatie over de afstand tot het objekt. Onze natuurlijke omgeving herbergt veel van dit soort cues. Het bestuderen van één bepaalde cue in isolatie is een krachtige en veel gebruikte manier om de complexiteit die door deze rijkdom aan cues wordt veroorzaakt te lijf te gaan, ook al heeft zo'n benadering een aantal nadelen.

Dit proefschrift bevat vier studies die betrekking hebben op de verwerking van de informatie over struktuur en beweging van objekten die beschikbaar komt wanneer deze objekten bewegen ten opzichte van een stilstaande waarnemer (structure-from-motion cue). De aanpak is steeds om proefpersonen bewegende synthetische objekten te laten zien (de stimulus), om ze vervolgens te laten reageren op bepaalde aspecten van de struktuur en/of de beweging van deze objekten (de respons). In plaats van echte objekten te gebruiken, benutten we een computer om objekten te gegeneren en te tonen, voornamelijk omdat dat de mogelijkheid schept vrijwel alle stimulus parameters expliciet te controleren. Dit proefschrift beschrijft zowel de gevonden empirische stimulus-respons relaties, als de pogingen om deze te begrijpen vanuit diverse niveaus van informatieverwerking.

Struktuur uit Beweging (structure-from-motion)

Niet iedere denkbare relatieve beweging tussen objekt en waarnemer levert de waarnemer informatie op over de ruimtelijke struktuur van zo'n objekt. Het is intuïtief duidelijk dat het daarvoor nodig is om het objekt van meerdere kanten te zien. De onderstaande figuur bevat een overzicht van de verschillende bewegingen tezamen met een schematische en versimpelde weergave van de corresponderende optische stroomvelden. Verderop komen we terug op de juistheid van dit beeld, maar eerst bespreken we de bruikbaarheid van de verschillende bewegingen voor het afleiden van ruimtelijke struktuur.

Een kleine translatie van een objekt in een richting loodrecht op de richting waarin het objekt wordt gezien (paneel A) resulteert in een uniform optisch stroomveld (paneel E). Het is duidelijk dat zo'n stroomveld volledig kan worden beschreven met slechts twee parameters - richting en grootte van de stroom - en dat deze parameters informatie bevatten over de beweging van het objekt maar niet over zijn struktuur. Op vergelijkbare wijze corresponderen translaties in de richting waarin het objekt wordt gezien (paneel B) en rotaties om een as die samenvalt met deze richting (paneel C) met eenvoudige één-parameter stroomvelden die geen informatie bevatten over de struktuur van het objekt doch slechts over zijn beweging (paneel F, pure expansie; paneel G, pure rotatie). Slechts wanneer een objekt roteert om een as die loodrecht staat op de richting waarin het objekt wordt gezien (paneel D) bevat het stroomveld informatie over de ruimtelijke struktuur van het objekt (paneel H). We concluderen dus dat dit type rotatie - rotatie 'in de diepte' - de kern vormt van de structure-from-motion cue. Een van de belangrijkste gemeenschappelijke aspecten van alle stimuli die worden gebruikt in het onderzoek beschreven in dit proefschrift is dan ook dat de objekten altijd roteren in de diepte. Enkele mathematische aspecten van de daadwerkelijke reconstructie van de struktuur van een objekt uit een algemeen optisch stroomveld worden besproken in hoofdstuk 3.



Overzicht van de verschillende typen objekt-beweging (bovenste rij) en de resulterende optische stroomvelden (onderste rij). De stroomvelden corresponderen steeds met een willekeurig gekozen voorbeeld van de bijbehorende objekt bewegingen.

Ik wil twee punten verder toelichten. Ten eerste, een objekt dat in werkelijkheid transleert door de fysische ruimte, roteert (en transleert) ten opzichte van de richting waarin het wordt gezien. Het bijbehorende stroomveld is dus een combinatie van de velden weergegeven in de panelen E, F en H. Deze specifieke verschijningsvorm van de structure-from-motion cue wordt ook wel met de term bewegingsparallax aangeduid. Het tweede punt betreft een simplificatie die ik heb toegepast bij het tekenen van de stroomvelden in bovenstaande figuur. Normaal gesproken, dat wil zeggen in de wereld buiten het laboratorium, wordt de structure-from-motion cue altijd vergezeld van een perspectivische cue. Perspectivische vervormingen bevatten informatie over de beweging en struktuur van een bewegend objekt en als zodanig definiëren ze dus een cue. Deze informatie is echter erg gevoelig voor ruis en speelt verder alleen een rol in de menselijke waarneming wanneer er sprake is van zeer grote visuele hoeken (zeer nabije objekten). Ik heb er met opzet voor gekozen om de perspectivische cue niet mee te nemen in de stimuli, met name vanwege de grote vereenvoudiging van de mathematische beschrijving van de informatie inhoud van de stimulus. Perspectivische vervormingen maken de stroomvelden een stuk ingewikkelder. Ik verwijs naar hoofdstuk 3 voor meer discussie hierover.

In dit proefschrift probeer ik hoofdzakelijk twee belangrijke kwesties te onderzoeken: in welke mate en met welke nauwkeurigheid zijn menselijke waarnemers in staat informatie op te pikken over de struktuur en rotatie van bewegende objekten en wat is de rol van de structure-from-motion cue bij het opbouwen van een ruimtelijke representatie van deze objekten door de waarnemers. Ik zal hierna kort aangeven hoe deze kwesties worden benaderd in de verschillende hoofdstukken.

Struktuur

Hoofdstuk 2 bevat een studie van de perceptie van het volume van draaiende objekten. Volume is een van de belangrijkste globale struktuur eigenschappen van een objekt, maar is in het verleden nauwelijks bestudeerd. De stimulus die we gebruiken in dit experiment bestaat uit twee objekten van variabele vorm en grootte die roteren om een gemeenschappelijke vertikale as. De taak van de proefpersoon is om aan te geven welke van de twee objekten het grootste volume heeft. Een van de belangrijkste bevindingen is dat de vorm van een objekt in sterke mate de nauwkeurigheid waarmee proefpersonen antwoorden beinvloedt. Discriminatiedrempels voor het volume van objekten zonder een specifieke uitgestrektheid in een bepaalde richting, zoals kubussen, zijn lager dan die voor langwerpige objekten. Evenzo zijn de drempels voor regelmatig gevormde objekten, zoals rechthoekige dozen, aanmerkelijk lager dan die voor objekten die een meer onregelmatige struktuur hebben (vergelijk de objekten weergegeven in figuur 2.1 in hoofdstuk 2). Dit laatste resultaat duidt er op dat in dit experiment naast de structure-from-motion cue ook nog andere cues een rol spelen. Omdat we er voor hebben gezorgd dat de informatie over volume die beschikbaar is via de structure-from-motion cue grofweg hetzelfde is voor de regelmatig en onregelmatig gevormde objekten, kunnen we concluderen dat proefpersonen in staat zijn gebleken gebruik te maken van de specifieke eigenschappen van sommige van deze objekten. Deze kennis wordt gebruikt in de hoofdstukken 4 en 5, wanneer we vergelijkbare objekten gebruiken voor andere doeleinden.

Struktuur en Beweging

In het stroomveld in paneel H van de figuur op de vorige pagina is informatie over de struktuur van het achterliggende objekt vermengd met informatie over zijn roterende beweging. Hoofdstuk 3 van dit proefschrift is geheel gewijd aan de perceptuele verwarring die kan onstaan wanneer deze componenten slechts ten dele worden gescheiden door het visuele systeem. De stimulus die we aan de proefpersonen laten zien bestaat uit twee gekantelde platte vlakken die oscilleren rond vertikale assen. De taak voor de proefpersoon is tweeledig: maak de orientatie van het ene vlak gelijk aan die van het andere vlak (orientatie ten opzichte van de rotatie-as) en maak tevens de amplitudes van de bewegingen gelijk. De resultaten laten zien dat proefpersonen in bepaalde condities, afhankelijk van de specifieke stand en rotatiemagnitude van de vlakken, fouten die ze maken in de ene deeltaak compenseren met fouten in de andere deeltaak. Zo'n patroon van gecorreleerde fouten duidt op een tekort aan informatie. Een mathematische analyse van de in de stimulus beschikbare informatie brengt aan het licht dat de correlaties kunnen worden begrepen in termen van reeds bekende gevoeligheden voor snelheidswaarneming door het menselijk visueel systeem, gecombineerd met de mathematische struktuur van de structure-from-motion cue. De minder dan perfekte verwerking van bewegingsinformatie (het optische stroomveld) door het visueel systeem zorgt er effectief voor dat verschillen tussen bepaalde klassen van roterende vlakken niet waarneembaar zijn. Deze conclusie biedt ons aanknopingspunten om bestaande experimentele resultaten met andere taken beter te begrijpen.

Beweging

De laatste twee hoofdstukken van dit proefschrift behandelen de verwerking van visuele informatie over de rotatie-beweging van objekten in de drie-dimensionale ruimte. Rotaties kunnen in het algemeen worden beschreven met een hoeksnelheid en met de stand van de rotatie-as. In hoofdstuk 4 bestudeer ik de waarneming van de hoeksnelheid, terwijl hoofdstuk 5 voornamelijk handelt over de stand van de rotatie-as. Alhoewel de rotatie van een objekt rond een as die loodrecht staat op de richting waarin het objekt wordt gezien cruciaal blijkt te zijn voor de beschikbaarheid van informatie over zijn ruimtelijke struktuur, is er nog niet veel bekend over de menselijke perceptie van dergelijke rotaties. Ik heb daarom onderzocht hoe bepaalde ruimtelijke aspecten van de stimulus de waargenomen hoeksnelheid van zulke rotaties beinvloeden. De proefpersonen bekijken steeds twee roterende objekten (dezelfde onregelmatig gevormde objekten als welke werden gebruikt in hoofdstuk 2) en worden gevraagd aan te geven welke van deze objekten het snelst draait. Ik vind ondermeer dat proefpersonen de draaisnelheid van objekten die roteren rond vertikale assen onderschatten ten opzichte van de draaisnelheid bij rotatie om horizontale of diagonale assen, dat ze de hoeksnelheid van grotere draaiende objekten groter vinden dan die van kleinere draaiende objekten, en dat de draaisnelheid van perifeer waargenomen objekten wordt onderschat ten opzichte van de draaisnelheid van foveaal waargenomen objekten. Een aantal van deze effecten kan worden begrepen in termen van vergelijkbare fenomenen die zijn gevonden voor de perceptie van twee-dimensionale lineaire snelheid en van grootte. Blijkbaar wordt de subjectieve hoeksnelheid niet alleen bepaald door de objektieve hoeksnelheid maar ook door een aantal andere factoren. Dit is een interessant resultaat dat gegeven de sterke relatie tussen struktuur- en bewegingsinformatie ook betekenis heeft voor de perceptie van struktuur (zoals bijvoorbeeld rigiditeit).

In hoofdstuk 5 maak ik gebruik van een andere experimentele methode om de verwerking van visuele informatie over de rotatie-beweging van objekten te onderzoeken: proefpersonen worden gevraagd de beweging van een op het computerscherm getoond draaiend objekt na te bootsen door met hun hand een bolvormig echt objekt op overeenkomstige wijze te manipuleren. Deze benadering heeft een aantal belangrijke voordelen boven de gangbare methoden, bijvoorbeeld de hoge dimensionaliteit van de respons- of antwoordruimte: proefpersonen worden aangemoedigd en in staat gesteld te reageren op elke verandering van de waargenomen (instantane) stand van het getoonde objekt. Uit de gemeten beweging van het echte objekt (in de hand van de proefpersoon) kan de gemiddelde stand van de 'geantwoorde' rotatie-as worden berekend. De variabiliteit in de orientatie van deze as blijkt het laagst te zijn voor langgerekte objekten die draaien om hun lengte-as en voor rotaties om fronto-parallelle assen ('rotatie in de diepte') en is altijd hoger dan de variabiliteit inherent aan het motorische systeem dat wordt gebruikt om te antwoorden (de manipulerende hand). De voor- en nadelen van dit paradigma worden uitvoerig besproken in hoofdstuk 5. Ik concludeer dat de experimentele resultaten minimaal aantonen dat proefpersonen een grote mate van reproduceerbaarheid vertonen bij het uitvoerer van deze taak en dat de sterke stimulus-respons relatie deze methode geschikt maakt voor het onderzoeken van de verwerking van visuele informatie door menselijke waarnemers.

NAWOORD

De promotietijd is bij uitstek een periode waarin iemands individuele kwaliteiten op de voorgrond treden. Als promovendus wordt je geacht zelfstandig wetenschappelijk onderzoek te verrichten en daarover op internationaal geaccepteerde wijze te publiceren. Je bent geheel verantwoordelijk voor je eigen onderzoek en de keuzes die je daar in maakt horen principieel de jouwe te zijn. Ook het op schrift stellen van de onderzoeksresultaten is een behoorlijk eenzame bezigheid. Daar staat tegenover dat de beloning, in de vorm van plezier hebben in het werk, het ontvangen van een titel en het tot stand komen van een proefschrift (en de hopelijk glansrijke verdediging daarvan), ook geheel en al aan de promovendus zelf ten deel valt.

Iedere promovendus zal echter beamen dat de steun van mensen in de directe omgeving en het weerwoord van wetenschappelijke collega's essentieel zijn voor het tot een goed einde brengen van de promotietijd. Daar moet men niet te licht over denken. Ik ben erg dankbaar voor het feit dat ik het in dit opzicht goed heb getroffen. Op deze plaats wil ik daarom graag mijn dank uitspreken aan iedereen die een bijdrage heeft geleverd aan de totstandkoming van dit proefschrift.

Ik wil niet nalaten diverse mensen met naam en toenaam te noemen. Allereerst bedank ik mijn promotor. Jan, ik heb altijd geprobeerd gebruik te maken van je inzichten. Wanneer ik mijn verhaal aardig rond dacht te hebben bleek je immer in staat een probleem vanuit een andere invalshoek te benaderen. Je hebt me daardoor vaak met relatief korte opmerkingen zeer lange tijd bezig gehouden en dat beschouw ik als een grote verdienste. In het bijzonder waardeer ik de vrijheid die je me gegeven hebt om dàt onderzoek te doen wat ík leuk en interessant vond, zolang ik mijn keuze maar kon verdedigen. De bij tijd en wijle zeer stevige en soms moeilijke discussies die zo onstonden hebben zeker bijgedragen aan de kwaliteit van dit proefschift. Vervolgens bedank ik ook mijn co-promotor. Astrid, de ijver en nauwgezetheid waarmee je hebt meegewerkt aan de totstandkoming van de diverse manuscripten worden erg gewaardeerd. Vele malen wist je vage redeneringen en zwakke plekken in de opbouw van een verhaal aan te duiden. Ook je aandacht voor concrete en meer aardse zaken als tijdschema's en deadlines hebben mij, zeker in de laatste fase, geholpen alles op tijd rond te krijgen.

Ook Maarten Hogervorst ('Vanmiddag? Ben ik aan de beurt dan...?') wil ik hier bedanken. Maarten, de vele 'het kan wel - het kan niet' sessies die wij hebben gehad zullen mij eeuwig bijblijven, met name ook vanwege het feit dat jij achteraf vrijwel altijd gelijk bleek te hebben. Je wist vaak mijn intuïtie te corrigeren en dat was uitermate nuttig. We hebben samen ook veel meegemaakt. Natuurlijk op conferenties, maar met name juist in de gewone dagelijkse werkzaamheden. Ook al zijn we er nooit toe gekomen voor langere tijd samen aan een project te werken, veel ideeën die we nu hebben zijn voortgekomen uit onze gezamelijke discussies. Vrijwel tegelijkertijd zijn we begonnen aan grotendeels identieke projecten en ik ben daarom blij dat we nu ook op dezelfde dag promoveren. Samen uit, samen thuis. Ik wil ook graag Peter Werkhoven bedanken. Peter, ik heb altijd erg veel plezier beleefd aan onze samenwerking. Je enthousiasme en daadkracht hebben mij sterk gestimuleerd en ik ben er van overtuigd dat mijn keuze om in dit vak te blijven, om dus onderzoek te blijven doen, voor een belangrijk deel aan jou te danken is. Onze samenwerking heeft tot nog toe geleid tot een aantal leuke publicaties en ik hoop oprecht dat we daar in de toekomst nog wat meer aan zullen toevoegen.

Vele andere namen schieten mij te binnen. Zo zijn daar natuurlijk Bert Mooi, Erik van Wijk, Hans Kolijn, Peter Faber, Pieter Schiphorst, Ronald Kassies, Wim Aarnink en de diverse mensen van de werkplaats. Ieder op zijn eigen wijze hebben jullie bijgedragen aan de realisatie van mijn plannen. Vele uren zijn er opgeofferd aan het draaiend krijgen van de 'Polhemus opstelling' (die uiteindelijk fantastisch werkte) en de projectieschermen (waar iedereen mee lijkt te werken behalve ik). Ik hoop dat jullie het met me eens zijn dat het het allemaal waard was. Ik kom graag nog eens langs, zomaar, als er toevallig weer eens taart is ofzo.. Ik bedank ook Mieke: jij wist mij altijd weer langs allerlei administratieve barrières te loodsen en had immer tijd voor een kletspraatje.

Ans, ik wil jou graag bedanken voor je behulpzaamheid bij het verbeteren van de manuscripten en voor je heerlijk directe kritieken. Dat werkte vaak heel verhelderend. Raymond, bedankt voor je ontwapende eerlijkheid en onze fantastische trip door de States (met Maarten). Hopelijk wordt het floridabolussen-project gauw vervolgd! Lon, bedankt voor je eeuwige toewijding aan de Mac. Het heeft mij en vele anderen erg geholpen. En dat je zelfs op zaterdag komt om je collega uit de brand te helpen mag ook wel eens worden gezegd. Susan, als laatst overgebleven kamergenote wil ik je bedanken voor je hulp als praatpaal en meedenktank. Het was heel gezellig en ik zal dat gaan missen straks. Sylvia, bedankt voor je fratsen en je schaterlach. Onze gesprekken waren niet altijd even gestruktureerd, maar dat mocht de pret niet drukken. Alle anderen uit de visuele groep: Ingrid, Chris II, Ruth, Yong, Maria, Bram en alle studenten, bedankt voor jullie gezellige en stimulerende aanwezigheid. Dat geldt ook voor de 'andere helft van de etage': Casper (vooral over stereo hadden wij een aantal felle doch leerzame en openhartige discussies), Evert-Jan, Ignace (altijd tijd, nooit chagrijnig en immer bezig zijn ideeën te verkondigen), Chris I (onze koekjesman), Jasper en Martin.

En dan al die vakgroepsleden uit een verder verleden: Aartjan van Pelt, André Noest (zelden zo'n berg wetenschappelijke creativiteit bij elkaar gezien, bedankt dat ik er van mocht proeven), Bert van Velzen, Erik de Haan (onze Edinburgh-trip was wonderschoon), Erik Kouwenhoven, Jane Boulton, Harm Jonker, Herman Snippe (part-time kamergenoot), Jeroen Smeets (die de liefde voor het vak al tijdens mijn studie bij mij wist boven te halen; dus eigenlijk is dit allemaal jouw schuld), Phil Koken, Roderik Erens, Sjoerd de Vries (die mij nu nog af en toe helpt met Mac-problemen), allemaal bedankt! Alle Helmholtzers (teveel om op te noemen) die hun best deden om de maandagmiddag tot een succes te maken, bedankt!

Ik wil ook graag mijn studenten bedanken. Michel Lecluse, die een taaie klus had aan de Polhemus. Edwin Roetman, die als talentvol schaatser natuurlijk weinig fout kon doen bij mij. En Norbert van Woerkom natuurlijk, met wie ik een zeer lange tijd heb gewerkt en die aanzienlijk heeft bijgedragen aan de laatste twee hoofdstukken van dit proefschrift (en aan de kwaliteit van vele overhead sheets). Vele gesprekken hebben wij gevoerd, over van alles en nog wat, en ik denk dat het jouw verdienste is dat ik nog weet hoe het studentenleven er uit ziet. Bedankt voor de leuke tijd.

De hoogleraren Han Collewijn, Casper Erkelens, Stan Gielen en Wim van de Grind wil ik bedanken voor hun bereidheid plaats te nemen in de beoordelingscommissie.

Karst Meijer wil ik bedanken voor het geheel belangeloos vervaardigen van figuur 5.1. Sheila McNab bedank ik voor haar grote inzet bij het corrigeren en redigeren van iedere regel Engels in dit proefschrift. Het was prettig met je samen te werken en ik zal je kennis gaan missen.

Dank ook aan mijn familie en aan al mijn vrienden en kennissen die mij, vooral het laatste jaar, nogal eens hebben moeten missen. Vaak wat meer op de achtergrond maar daarom niet minder belangrijk hebben jullie gezorgd voor mijn andere ik. Speciaal wil ik mijn ouders bedanken voor alles wat jullie me hebben meegegeven. Zonder die bagage was dit proefschrift niet verschenen. Ik wil Alexis (bokkebokkebokke) bedanken voor zijn fantastische vriendschap. De donderdag lunch was altijd weer een fijn rustpunt en de vele go-partijen die wij tussen de bedrijven door hebben gespeeld konden vaak de problemen van de dag verzachten. Ik zal je gaan missen, zo straks.

Rest mij nog om Ilse te bedanken. Samen zijn wij door diepe dalen en over grote bergen gegaan en ik hoop dat we dat nog heel lang zo blijven doen. Bedankt voor al je liefde en steun, voor al je begrip en toewijding. Op naar het volgende avontuur!

Hendrik-Jan

CURRICULUM VITAE

Hendrik-Jan van Veen werd geboren op 30 juni 1967 in het dorpje Reeuwijk. De jaren 1979 tot en met 1985 bracht hij door op het Sint Antonius College te Gouda, alwaar hij op 30 mei 1985 het diploma Ongedeeld VWO behaalde. Vervolgens wendde hij zich tot de Rijksuniversiteit Utrecht om zich te verdiepen in de Experimentele Fysica. De keuzevrijheid binnen de studie werd ondermeer gebruikt om zich te bekwamen in de computationele fysica en om meer kennis op te doen over computerarchitectuur. Het afstudeerwerk werd verricht bij de vakgroep kernfysica, alwaar hij deelnam aan een internationaal onderzoeksproject gericht op het bepalen van de interne magnetische structuur van het neutron. Zowel voor de eigenlijke metingen als voor het uitvoeren van een verscheidenheid aan analyses was het noodzakelijk om een groot deel van de afstudeerperiode door te brengen op het kernfysisch versneller instituut NIKHEF-K te Amsterdam. Uiteindelijk slaagde hij op 2 juli 1991 voor het doctoraal examen experimentele fysica (met genoegen). Aansluitend begon hij als Assistent-in-Opleiding bij de vakgroep Medische en Fysiologische Fysica te werken aan een project genaamd 'Optische detectie van 3D structuur'. Bij deze groep, thans omgedoopt tot vakgroep Fysica van de Mens en binnen het kader van het Helmholtz Instituut kwam het voor u liggende proefschrift tot stand. Per 1 januari 1996 zet schrijver dezes zijn wetenschappelijke loopbaan voort als postdoc bij het Max-Planck-Institut für biologische Kybernetik in Tübingen.