

Quantitative analysis of saccadic search strategy

Kwantitatieve analyse van saccadische zoekstrategie
(with a summary in English)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit
Utrecht op gezag van de rector magnificus, prof.dr. W.H.
Gispen, ingevolge het besluit van het college voor
promoties in het openbaar te verdedigen op maandag 18
juni 2007 des middags te 12.45 uur

door

Eelco Abram Boudewijn Over

geboren op 7 mei 1976
te Utrecht

Promotor: Prof. Dr. C.J. Erkelens

Co-promotor: Dr. I.Th.C. Hooge

Dit proefschrift werd mogelijk gemaakt met financiële steun van TNO Defence, Security and Safety, Soesterberg

Quantitative analysis of saccadic search strategy

Eelco Over

ISBN: 978-90-8570-190-3

Table of contents

	Page
Chapter 1. Introduction	7
Chapter 2. A quantitative measure for the uniformity of fixation density: the Voronoi method	19
Chapter 3. Saccade direction in visual search: top-down or bottom-up?	49
Chapter 4. Search area edges affect visual span size, fixation density and saccade direction	73
Chapter 5. Coarse-to-fine eye movement strategy in visual search	115
Chapter 6. Summary	137
Hoofdstuk 7. Samenvatting	143
Publications	149
Curriculum Vitae	151

Chapter 1. Introduction

This thesis investigates strategies for eye movements during visual scanning in general, and visual search in particular. Better understanding of eye movement behavior may provide insight in how to train people to search more effectively and efficiently. Effective and efficient search is especially relevant for medical doctors examining X-rays, soldiers guarding an area, flight controllers, and so on. Before going into the aim of this thesis in more detail, I first want to explain some common terms that are used in the field of eye movement research.

People make many eye movements, on average about three per second. They do that for a reason, and this reason has to do with the structure of the retina. The human eye is a remarkable sensor. Its structure allows for high visual acuity *and* a large field of view. Some animals have a high-resolution retina, but not so large a field of view (e.g. a hawk). Other animals have a larger field of view, but not as high visual acuity (e.g. a rabbit). Humans have an anisotropic retina with high spatial resolution and color sensitivity in the center of the visual field (called the fovea, having a diameter of about 2 degrees of visual angle), surrounded by a large region (the peripheral retina, having a diameter of about 150 degrees of visual angle) that has a resolution that gradually decreases with distance to the fovea (eccentricity). This variable-resolution retina on its own does not guarantee optimal acquisition of visual information. However, the eye can be rotated so that interesting places in the environment are projected on the fovea rapidly, one after another. Visual information at the new interesting location can be sampled in more detail with the high-resolution fovea.

Eye movements are accomplished by a set of six muscles attached to the eyeball. They are arranged in three pairs of agonist-antagonist muscles; roughly speaking, one pair rotates the eye horizontally (left/right), the second rotates the eye vertically (up/down), and the third allows

cyclotorsion, or rotation about the line of sight. The type of eye movement that is most relevant to visual search behavior is the *saccade*. It is the eye movement that quickly moves the point of gaze from one location in the visual environment to another. Vision is suppressed during a saccade. However, saccades are among the fastest movements human beings can make, so only little time is lost making saccades. This leaves more time for acquisition of visual information. The eyes can rotate at a speed of up to 600 deg/s (Collewyn, Erkelens, & Steinman, 1988). Therefore, saccades are usually completed in tens of milliseconds. Despite (or perhaps, because of) their speed, saccade trajectories tend to be remarkably stereotyped, not only within but also between individuals. Humans make well over one hundred thousand saccades daily, which is about three per second on average. Apparently, saccadic eye movements are cheap.

The size of the saccade, defined as its *amplitude*, is usually measured in degrees of rotation of the eyeball. The duration and peak velocity of a saccade both increase monotonically with amplitude of the movement. The relation between duration, peak velocity and amplitude has been called the *main sequence* (Bahill, Clark, & Stark, 1975). The relation between duration and amplitude can be described by a linear function over a wide range of saccade amplitudes: $t = 21 + 2.2a$, t : duration in ms, a : amplitude in degrees. Saccades are not always accurate; actual landing position can deviate from the intended landing position. If the amplitude of a saccade is too small or too large, a corrective saccade has to be made in order to fixate the intended target location. Amplitudes are more often too small than too large (Becker, 1972; Robinson, 1973; Findlay & Brown 2006). There can also be an error in the *direction* of a saccade. However, saccades have usually larger on-axis variability than in the perpendicular direction (Findlay & Brown, 2006). In other words, the direction of a saccade is usually more accurate than its amplitude.

Acquisition of visual information occurs during *fixations*. Fixations are the inter-saccadic intervals in which the eye is relatively stable. Fixation *location* mainly

determines what information is available for acquisition, and fixation *duration* mainly determines how much of that available information can be acquired.

Strategy in eye movement behavior

Eye movement behavior has been subject of quantitative investigation since the previous century. Buswell (1935) established that there is a difference in eye movement behavior between trained artists and untrained subjects while viewing paintings. He concluded that ‘Eye movements are unconscious adjustments to the demands of attention during a visual experience.’ (p. 9). This means that top-down influences play a potentially significant role in eye movement behavior. People move their fixation location to interesting parts of the environment, fixate that point for 300 ms on average and move on to the next interesting point, in this way building up a percept of the visual environment. Yarbus (1967) showed that eye movement behavior is task dependent, which is also a top-down influence. Starting from Buswell’s and Yarbus’ classical work, many studies have suggested that gaze is directed according to the ongoing cognitive demands of the task at hand. For example, Antes (1974) suggested that while viewing pictures, the first fixated regions are highly informative, and later fixated regions are less informative.

These early studies involved eye movement behavior in viewing pictures, either with or without a specific task. In viewing without any specific task, i.e. “free” viewing, the goal of the strategy may be described as to acquire as much information in a time period as short as possible. However, identifying the strategy in free viewing, or other tasks, was difficult. The task demands in these early studies were not under strict control of the experimenter and therefore poorly defined. Eye movement research after these early studies concentrated on a considerably better-controlled task: visual search. Research topics that were investigated with free viewing can also be studied with visual search; it is still assumed today that eye movements are at least to some extent

the reflection of attention and ongoing cognitive processes, as Buswell and Yarbus suggested. There are several reasons why visual search is a better-suited task for eye movement research than free viewing. Firstly, task demands are better defined in a visual search task than in free viewing because the experimenter has control over the composition of stimuli (targets and distracters) and instruction. This control allows for simple stimuli. The task is to find the target as quickly as possible. Secondly, the response in a search task often is either a “target absent/present” response, or a reaction time. These responses are easily analyzed: signal detection theory can be applied, and reaction times can be analyzed quantitatively. The change from free viewing to visual search in eye movement research yielded increased possibilities to analyze eye movement behavior quantitatively, and to relate the output (saccade parameters) in eye movement research to the input (stimulus characteristics). These possibilities have been exploited to analyze eye movement strategies. In this thesis, eye movement strategy is defined as a plan of reproducible action for saccades and fixations. It is assumed that eye movement parameters are the reflection of the search strategy. Eye movement parameters can be measured and analyzed quantitatively.

In the eye movement literature, search strategy has not been a major subject of investigation. There are some reports of results considering search strategy. In several studies it was shown that subjects often made more saccades than strictly necessary (Hooge & Erkelens, 1996; Zelinsky, 1996; Hooge & Erkelens, 1998). Subjects were uncomfortable with the task to minimize the number of fixations. Hooge and Erkelens (1998) showed that subjects rather make a new saccade than complete the analysis of visual information. In their study, completed analysis of available information would have yielded the direction in which to make a saccade in order to find the target. Apparently, strategies that lead to a minimum number of fixations were not used.

The employment of eye movement strategies is restricted by certain boundary conditions. For example, it is

not realistic to expect mean fixation durations in the order of 10 ms. Fixation duration is partly determined by saccadic latency, and latency typically is at least 200 ms (see e.g. Zingale & Kowler, 1987). The employment of some eye movement strategies has been suggested to depend solely on sensory-level factors. Koch and Ullman (1985) suggested a strategy of making a saccade to the most salient location in a stimulus. A saliency map is a combination of several maps for sensory-level features such as color, orientation, luminance, etc.. The most salient location is the result of a winner-take-all process operating on the saliency map. The saliency map is updated after each saccade, suppressing the saliency of the current fixation location and reducing suppressions at previously fixated locations. If the saliency map is used in a model to predict eye movement behavior, simulated eye movements depend on how the saliency map is constituted, and on the parameters quantifying suppression and reduction of suppression. The design of the visual system itself imposes boundary conditions to the employment of strategies too. Rajashekar, Cormack, and Bovik (2004) and Najemnik and Geisler (2005) showed that observers make use of information about the target and the visual system to move the eyes to locations that have high probability to contain the target. Moreover, they do this nearly as well as an ideal observer (Najemnik & Geisler 2005). Experience may change the boundary conditions for the employment of strategies. Experience may cause faster visual processing, or acquisition of specific information only. Thus, shorter fixation durations are possible (Chapman & Underwood, 1998; Williams & Davids, 1998). Experience also may shift the locus of fixation. Experienced soccer players fixate the hip of their opponents more often than inexperienced players do (Williams & Davids, 1998), and experienced drivers fixate lower down (closer to their own vehicle) than novice drivers (Chapman & Underwood, 1998). The fact that experienced observers may fixate other locations with shorter average fixation duration points to a possible difference in strategy due to cognitive-level factors. Itti and Koch (2000), extending the saliency map

model by Koch and Ullman (1985), concluded that the selection process in the saliency map model must reflect properties of the entire visual system, thereby implying that the model cannot be exclusively based on sensory-level factors, as in the model by Koch and Ullman (1985). Henderson, Weeks, and Hollingworth (1999) showed that semantic consistency (which is a characteristic at cognitive level) influences both fixation duration and reaction times. Objects that are inconsistent with their surroundings evoke longer fixation duration and take longer to be found. Scinto, Pillalamari and Karsh (1986) showed that subjects do not easily adopt a particular strategy based on cognitive-level factors, although such a strategy is supposed to be cognitively penetrable. This result was also found by Hooge and Erkelens (1998). Systematic strategies are also based on cognitive-level factors. In regular grids of elements for example, it is possible to employ a “smart”, systematic strategy that reduces memory load. If the grid is scanned in a reading-like manner, it is unnecessary to remember of each element separately whether it is checked already. Instead, the strategy (“reading”) and the currently fixated grid element suffice to deduce which elements are still to be checked. Also, the scene’s gist and common knowledge may guide eye movements during a search task; it is clear that an automobile is not to be found in mid air.

In the literature, mean reaction time (or search time) is one of the few *quantitative* measures for characterizing search strategies. Quantitative measures, such as reaction time, can be used to make a model, and the model allows for prediction of results in new experimental conditions. In some simplified cases of visual search, it is even possible to predict reaction times analytically with the help of models. For example, two models for these simplified cases of visual search are parallel and serial search (Treisman & Gelade 1980). When one makes some assumptions and simplifications, it becomes easier to build these models. A few examples of such assumptions and simplifications are:

- The stimulus usually contains one or zero targets. The required response is a binary answer (target absent/present), or the reaction time is measured.
- It is assumed that one element is analyzed per fixation (or another constant number).
- It is assumed that processing of visual information takes a constant period.
- It is assumed that no errors are made in the visual analysis, so no targets are missed and no false alarms occur. Therefore no refixations are needed.
- Search in target-present trials is self-terminating, not exhaustive. Search is exhaustive in target-absent trials, and the response is “target absent”.

The problem

The above-mentioned simplifications and assumptions apply to many of the studies described in the previous section. However, they are not always justified. For example, if search is carried out on a target-absent trial and the searcher does not remember what locations he has visited already, a response may not come within the allowed viewing time. Moreover, 100% accurate analysis is not realistic. And in the case of searching an unknown number of targets, there is no such response as “target present” or “target absent”. In daily-life search tasks, the search area itself may have no limits since the person or animal conducting the search can move around in his environment.

One of the goals of this thesis is to provide quantitative measures for characterizing search strategies that can be measured with all kinds of tasks and stimuli, and not only when the above-mentioned simplifications and assumptions apply. These measures should be supplements to the quantitative measure of reaction time that is generally used in the literature. Once the input (stimulus characteristics) and output (eye movement parameters) of the visual system can be described quantitatively, one may start to analyze the characteristics of the visual search system itself. In electronics,

there are ways to figure out the properties of an electronic circuit (a “black box”) by means of measuring the impulse response function. For the visual search system the impulse response function cannot be measured, because it is not (yet) known what the impulse input for the visual search system is (if it exists at all). Describing the input for the visual system quantitatively is difficult anyway, especially for “natural” stimuli: they are usually too complex.

The aim of this thesis is twofold:

1. To quantify eye movement parameters that usually are not reported quantitatively: fixation location and saccade direction.
2. To use the quantitative description of eye movement parameters (saccade amplitude and direction, and fixation location and duration) to describe eye movement strategies.

The tools that allow for a more general, quantitative analysis of visual search strategy will be introduced in the following two chapters of this thesis. The later chapters will focus on visual search strategy itself.

Overview of this thesis

In chapter 2 we design a method for quantitative analysis of fixation locations, in particular: the uniformity of fixation density. The method can be applied to any database of fixation locations. It is also possible to compare databases of different sizes. In chapter 3 we extend the method of chapter 2 so that it is possible to create vector plots that are representative of general saccade direction. A major advantage of the method is that not just the starting location of a saccade is coded by the direction of the saccade, but instead the entire area is coded according to the directions of all saccades in the scan path. Calculating the divergence of these vector plots allows for qualitative analysis of saccade directions. Attractors and repellers can be identified in direction-coded maps for stimuli and observers. In chapter 4 we investigate whether

search area edges affect eye movement parameters and search strategy. To explain the experimental results a model is introduced and validated. The model is also used to predict experimental results in new experimental conditions. In chapter 5 we investigate whether eye movement strategy depends on knowledge about target conspicuity. To this means, we analyze the time courses of saccade amplitude and fixation duration.

References

- Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, 103(1), 62-70.
- Bahill, A. T., Clark, M., & Stark, L. (1975). The Main Sequence, a tool for studying human eye movements. *Mathematical Biosciences*, 24, 191-204.
- Becker, W. (1972). The control of eye movements in the saccadic system. *Bibliotheca Ophthalmologica*, 82, 233-243.
- Buswell, G. (1935). *How people look at pictures*. Chicago: University of Chicago Press.
- Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: danger and experience. *Perception*, 27(8), 951-964.
- Collewijn, H., Erkelens, C. J., & Steinman, R. M. (1988). Binocular co-ordination of human horizontal saccadic eye movements. *The Journal of Physiology*, 404, 157-182.
- Findlay, J. M., & Brown, V. (2006). Eye scanning of multi-element displays: II. Saccade planning. *Vision Research*, 46(1-2), 216-227.
- Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 25(1), 210-228.
- Hooge, I. T. C., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Perception & Psychophysics*, 58(7), 969-976.
- Hooge, I. T. C., & Erkelens, C. J. (1998). Adjustment of fixation duration in visual search. *Vision Research*, 38(9), 1295-1302.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40(10-12), 1489-1506.

- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4(4), 219-227.
- Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, 434(7031), 387-391.
- Rajashekar, U., Cormack, L. K., & Bovik, A. C. (2004). Point-of-gaze analysis reveals visual search strategies. *Proceedings of SPIE: Human Vision and Electronic Imaging IX*, 5292, 296-306.
- Robinson, D. A. (1973). Models of the saccadic eye movement control system. *Kybernetik*, 14(2), 71-83.
- Scinto, L. F., Pillalamarri, R., & Karsh, R. (1986). Cognitive strategies for visual search. *Acta Psychologica*, 62(3), 263-292.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97-136.
- Williams, A. M., & Davids, K. (1998). Visual search strategy, selective attention, and expertise in soccer. *Research Quarterly for Exercise and Sport*, 69(2), 111-128.
- Yarbus, A. (1967). *Eye movements and vision*. New York: Plenum Press.
- Zelinsky, G. J. (1996). Using eye saccades to assess the selectivity of search movements. *Vision Research*, 36(14), 2177-2187.
- Zingale, C. M., & Kowler, E. (1987). Planning sequences of saccades. *Vision Research*, 27(8), 1327-1341.

Chapter 2. A quantitative measure for the uniformity of fixation density: the Voronoi method

Introduction

Scanning the visual environment is an ongoing activity of human beings. Visual scanning consists of saccadic eye movements interleaved with periods of fixation. The spatial distribution of fixation locations and its relation to scene content and task are interesting because they may illustrate how the viewer interprets the scene. During various search tasks, for example, people were found to fixate objects rather than the space between objects (Hooge & Erkelens, 1999; Williams, 1966). However, Zelinsky, Rao, Hayhoe, and Ballard (1997) found fixations between objects as well and challenged this view. Since the pioneering work by Yarbus (1967), sequences of visual fixation, later called scan paths (Noton & Stark, 1971a, 1971b), have been generally accepted representations of the viewing behavior of observers. Using scan paths, researchers have demonstrated that schizophrenics (Phillips & David, 1997) and people with brain damage (Zihl & Hebel, 1997) exhibit viewing behavior different from that of normal, healthy people. De Lucio, Mohamadian, Ruiz, Banayas, and Bernaldez (1996) have shown differences in scanning behavior between men and women. Zangemeister, Sherman, and Stark (1995) have described differences both between processes of viewing realistic and abstract art and between the viewing behavior of professional and nonprofessional art viewers. All of these reports referred to temporal differences in viewing behavior (e.g., in fixation duration or order of fixation). Reports referring to spatial differences in viewing behavior, on the other hand, are scarce. This gap in the literature may be caused by the absence of quantitative and objective methods that can be used for characterizing spatial aspects of viewing behavior. Here we

introduce the Voronoi method, which is based on Voronoi diagrams (Voronoi, 1908). The Voronoi method is a quantitative and objective method for characterizing fixation density, and we show that it may be a useful tool in the analysis of spatial viewing behavior. Scan paths capture most features of viewing behavior, because they provide information about both the temporal and the spatial domains. However, analysis of scan paths is usually restricted to subjective, visual inspection. As an exception, Choi, Mosley, and Stark (1995) developed a quantitative method to compare individual scan paths: the string analysis. This method was also used to compare scan paths made during first-time visual inspection with eye movement trajectories made later during imagery of the same scene (Brandt & Stark, 1997; Gbadamosi & Zangemeister, 2001). The string analysis converts scan paths to strings of characters: Each character represents a particular part of the stimulus that has been fixated. After removal of all refixations of the same area, strings are compared to a reference string, and a distance from the scan path to the reference string is calculated. Finally, normalization of this distance by the length of the string yields a number between 0 (identical scan paths) and 1 (completely different scan paths). This analysis is suitable for comparing individual scan paths, and it can also be used to compare scan paths that are generated by computer models to a human reference scan path. However, a disadvantage of the string analysis is that regions of the stimulus have to be designated to represent characters. Each of these regions is often called a region of interest (ROI) or area of interest (AOI). The subdivision into ROIs (see, e.g., Gbadamosi & Zangemeister, 2001; Laeng & Teodorescu, 2002; Pelz, Hayhoe, & Loeber, 2001; Shinoda, Hayhoe, & Shrivastava, 2001) allows for comparisons of the same feature, for example dwell time. The designation of suitable ROIs usually will not be a problem if the stimulus results in a clustering of fixations, such as with stimuli that consist of a collection of items (Zelinsky et al., 1997).

In Brandt and Stark's (1997) study, even the stimuli themselves could be used to define ROIs because of the spatial arrangement of subfeatures. Designating different ROIs is more difficult, however, when clustering of fixations does not occur (e.g., when there are no clear separate regions in the stimulus, like a white car in the mist). If it is difficult to designate ROIs, it may even be difficult to determine the number of ROIs. If this is the case, it may be more useful to characterize global densities of fixations in order to investigate aspects of their spatial distribution. Comparing individual scan paths is one thing the researcher can do. Comparing viewing behavior, however, requires comparing sets of scan paths (possibly of different sizes) to one another. One of the major problems in this kind of eye movement research is to distill useful information out of the huge amount of available data. Some important requirements for an objective and quantitative analysis are the independence of (1) stimulus size, (2) number of trials, (3) number of fixations in each trial, and (4) (possibly) number of subjects. In this article, we concentrate on a single property of viewing behavior that meets the requirements mentioned above: the density of fixation locations. This single feature allows for easy comparisons of large data sets (e.g., sets of patients and controls) and is informative about the spatial aspects of viewing behavior. A mathematically elegant procedure to characterize densities is to construct Voronoi diagrams (Voronoi, 1908). The Voronoi diagram is a division of an area into cells that are associated with particular points within that area. A cell is defined by the area taken up by the points that are closer to one particular point than to any other. In our case, the particular points are the fixation locations. The Voronoi diagram constructed for fixation locations has small cells in parts of the stimulus that are fixated frequently and large cells in parts of the stimulus that are fixated occasionally. Distributions of cell sizes therefore correlate to distributions of fixation densities. In the first section of this article, we will explain how Voronoi diagrams can be used to characterize the uniformity of fixation density. In the second section, we apply the Voronoi diagram

method to eye movement data in order to evaluate the outcome. In the third section, we describe simulations performed to get a feeling for the outcomes and requirements of the Voronoi method, and we end by describing our conclusions.

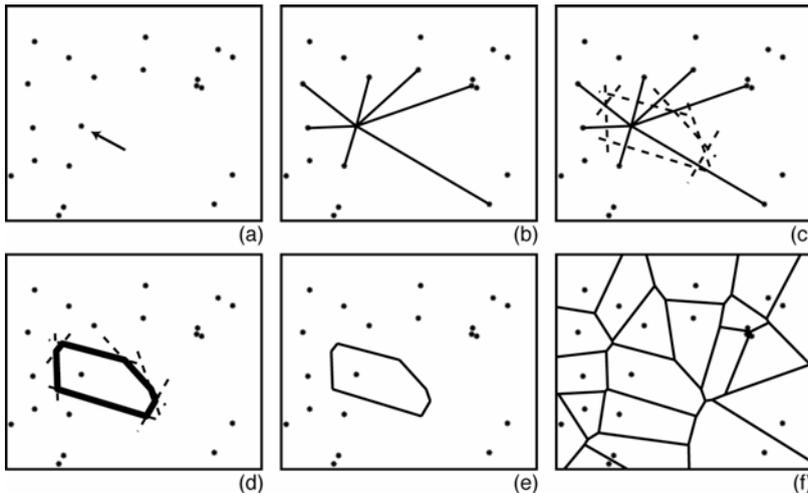


Figure 2.1 Construction of the Voronoi diagram. Pick a single fixation location (a) and draw connecting lines to its neighboring fixation locations (b). Construct a perpendicular line bisecting each connecting line (c). The smallest area enclosed by these perpendicular lines (d) is the Voronoi cell (e). Repeating this procedure for each fixation location yields the Voronoi diagram (f).

The Voronoi method

Areas can be filled with patches in numerous ways. The patches can be regular, as in a mosaic, for example. Or they can be irregular, as in the pattern that emerges when mud is dried and cracked by the sun. The Voronoi tessellation is an example of such an irregular filling of an area.

Construction of Voronoi Diagrams

Figure 2.1 systematically shows in a number of steps the construction of a Voronoi diagram for a set of fixation locations. It depicts the procedure of creating the Voronoi cell that belongs to the first fixation location (panel a): To begin, connecting lines are drawn to all other fixation locations (panel b). For reasons of clarity, we have only drawn the connecting lines that are relevant for this particular cell. The next step is that we compute the perpendicular lines (dashed) that bisect the connecting lines (panel c). The smallest area enclosed by these perpendicular lines is the Voronoi cell for the first fixation location (panels d and e). The procedure is repeated for all fixation locations (panel f). Voronoi cells for fixation locations neighboring the stimulus border are partly bounded by the edge of the stimulus. The sum of all cell areas equals the total stimulus area. Our method focuses on one of the properties of Voronoi diagrams—namely, the size of cells. The size of Voronoi cells depends on the size of the stimulus, the number of fixations per trial, and the location of fixations relative to other fixations. Figure 2.2 shows three hypothetical trials with symmetrical patterns of fixations that differ in these parameters. In this example, we present symmetrical patterns because they are simple and regular, but the same reasoning also applies to irregular patterns. Intuitively, in all three examples of Figure 2.2, the fixations are evenly scattered over the stimulus area. This is also clear from the fact that the Voronoi cells have equal size in each example. In order to match this intuition, the outcomes of the Voronoi method should be equal for all three patterns of fixation locations in Figure 2.2. Neither the size of the stimulus nor the number of fixations should affect the outcome of the method; to achieve this, the size of the cells should be normalized to correct for these two factors. Dividing the individual cell sizes by the mean cell size accomplishes this correction for both the number of fixations and the size of the stimulus, yielding a distribution of normalized cell sizes (dimensionless) with a mean of exactly 1. Because the distribution of normalized cell sizes of every trial has an average value of 1, the summation

of all normalized cell size distributions also has an average value of 1. A wide distribution represents very different cell sizes, and therefore an irregular pattern of fixations. A narrow distribution represents cells of approximately the same size, and therefore a homogeneous distribution of fixations.

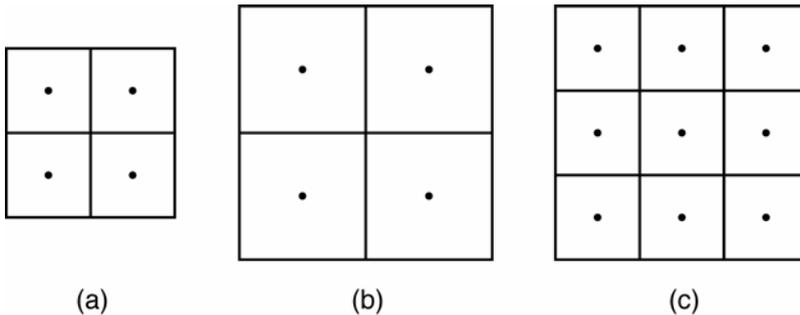


Figure 2.2 Three simple patterns of fixations and their Voronoi diagrams. Using panel b as a reference, panel a has the same number of fixations but a different stimulus size, and panel c has the same stimulus size but a different number of fixations. The three fixation densities, though, are all equally uniform.

Discussion

It may seem like a lot of work to construct Voronoi diagrams for each separate trial, but the Voronoi routine is included in many mathematical software packages (for example, Mathematica and MATLAB). Mathematica includes the option to build bounded Voronoi diagrams like the ones described above (where cells are limited by the stimulus boundaries). If only unbounded Voronoi diagrams are available (e.g., in MATLAB), the cells for all fixation locations can be calculated using the following simple trick:

1. Mirror all fixation locations relative to all stimulus boundaries (which must be straight lines).
2. Construct the unbounded Voronoi diagram for the collection of real and mirrored fixation locations.

3. Remove the cells belonging to the mirrored fixation locations from the analysis.

For the Voronoi method to work properly, there are some requirements for the data. These requirements are not strict, though, and usually they are already met.

1. There must be fixation locations. Since Voronoi diagrams are calculated from discrete points, a scan path that entirely consists of smooth pursuit would be unsuitable for this method.

2. Fixation locations must be unique; that is, a fixation should not be at the exact same location as any other fixation. In most eye movement data, two fixation locations will not fall on the exact same location. This is merely a question of the accuracy of the eye movement recording equipment and analysis method: A higher spatial resolution during recording and more decimal places in the analysis reduce the probability of identical fixation locations. Moreover, fixation locations are usually the result of averaging a number of samples, which further decreases the possibility of identical fixation locations.

3. Fixation locations must be within the boundaries of an area that has to be chosen prior to the analysis. It is logical to choose the field of view relevant to the current task, or otherwise, the smallest region that contains all eye movements is a good choice for setting boundaries. This region should be the same for all trials of the experiment.

Our method depends on only one arbitrary choice— namely, the size (and shape) of the relevant part of the field of view. In most laboratory experiments, these are of course identical to the stimulus boundaries; in daily life, they should be defined by the current task. Imposing this restriction is not an unreasonable thing to do, because in most experiments, and even in daily life, eye movements are restricted to some part of the total field of view (e.g., a computer monitor). The choice of a field of view does not lead to a loss of information, because it can still be used as an indicator for the scale of the

total scan path, and this information is retained in the normalization procedure: A multiplication of the average Voronoi cell size and the number of fixations gives the total stimulus size. The Voronoi method is a purely spatial method. All temporal information included in a scan path is discarded when the scan path is converted to a collection of fixation locations. However, temporal information could possibly be included in the Voronoi method as the third dimension, in addition to the two spatial dimensions. Consequently, the cells of such a three-dimensional Voronoi diagram would themselves be three-dimensional, raising the questions: How should a spatial dimension be scaled to a temporal dimension, and what would the volume of such a three-dimensional cell represent? These issues would have to be investigated thoroughly before the method could be extended to three dimensions in order to include the temporal aspect of scan paths. We will now apply the Voronoi method to an eye movement experiment with three conditions that we expected to yield very different uniformities of fixation density.

Eye movement experiment

To test our Voronoi method, we designed an experiment with three conditions in which we expected to find uniformities of fixation density that were very different from each other. Afterward, this expectation was confirmed by subjective visual evaluation of local fixation densities. We used two tasks—free viewing and visual search. The stimuli were either highly homogeneous or highly structured images.

The first condition (homogeneous search) was designed so that fixation locations were likely to be spread evenly over the stimulus area. The task was to search for an unknown number of targets. The subjects were expected to search the whole stimulus area, because they did not know in advance the total number of targets that were present in the trial. The stimulus was a homogeneous gray field with targets that were slightly darker. In every part of a homogeneously colored search area it is equally difficult to find targets, so the stimulus itself does not favor any part of the stimulus being

fixated more often than other parts. The task in the second condition (structured search) also was to find an unknown number of targets, but the stimulus was a grayscale portrait of a group of people. The search task again required the subject to inspect every part of the stimulus, but the highly structured background made it more difficult to find targets in some parts of the stimulus than in others, for instance because of crowding (Levi, Klein, & Aitsebaomo, 1985; Westheimer & Hauske, 1975). We therefore expected more fixations to occur in the difficult parts of the stimulus, for example parts with large contrast differences. In the third condition (free viewing), observers were not asked to search for targets, but merely to look at pictures. In this condition, there was no need to examine all parts of the stimulus, because observers were free to view the picture as they liked. We expected the structured nature of the stimulus to yield different fixation probabilities for different parts of the stimulus. These different probabilities could be due to differences in the stimuli themselves, such as differences in contrast. The interpretation of the stimulus might also affect its fixation probabilities, though: Interesting parts are more likely to be fixated than parts that are not. The spatial distribution of fixations depends not only on the instruction and the stimulus properties, but also on idiosyncratic behavior of the subjects. For example, the search strategy (Aks, Zelinsky, & Sprott, 2002; Zangemeister et al., 1995) and physical condition (De Gennaro, Ferrara, Curcio, & Bertini, 2001) of the subject play important roles in the properties of eye movements, and consequently in the appearance of a scan path. We therefore expected to find some differences between subjects, but the average distributions of fixations would have to be as follows: (1) In the homogeneous search condition, we expected the fixations to be spread more or less homogeneously in the stimulus area, and therefore we expected cells of approximately the same size. (2) In the structured search condition, we expected the fixations to cluster more in some parts of the stimulus, and we expected cell sizes to be somewhat different. (3) In the free viewing condition, we

expected the fixations to be strongly clustered in interesting parts of the stimulus, so cell sizes would be strongly different from each other in each trial. We applied the Voronoi method to the three conditions to investigate whether the results would be in accordance with the predicted uniformity of fixation density. We expected the skewness of the histograms of the normalized cell sizes to range from low in the homogeneous search condition to high in the free viewing condition.

Method

Apparatus. Two-dimensional orientation of the right eye was measured with an induction coil mounted in a scleral annulus (Skalar Medical, The Netherlands) in an AC magnetic field. This method was first described by Robinson (1963) and refined by Collewijn, Van der Mark, and Jansen (1975). An Apple Macintosh G4/400 equipped with a National Instruments PCI 1200 I/O board presented the stimuli and recorded eye movements and manual responses simultaneously at a sampling rate of 500 Hz. The video card was fast enough to put a stimulus on the screen within one refresh cycle of the monitor. The field generators and lock-in amplifiers were custom built by our technicians and reached the specifications of a Rimmel system or better (Rimmel, 1984). The maximum error in the signal reached 0.05 deg. The data were stored on the computer hard disk for offline analysis. The subjects sat between the field coils (Rimmel Labs, USA; dimensions 1.0 x 1.0 x 1.0 m) in front of a Sony E400 19-in. monitor (distance, 41 cm; resolution, 1024 x 768 pixels; refresh rate, 85 Hz; 256 gray values). Translations and rotations of the head were prevented by a one-axis bite-board. The room was normally lighted during the experiment.

Subjects. Six male (AK, CE, EO, IH, MB, and RW) and 3 female (EL, JZ, and RO) observers participated in all three conditions. None showed any visual or oculomotor pathology other than refraction anomalies. The observers had normal or corrected-to-normal vision. Subjects CE, MB, and RW wore corrective glasses. CE, EO, and IH are the authors

of this study; the other subjects were naive as to the goal of the experiment.

Stimuli. Three conditions were presented to the observers, with no practice trials.

Homogeneous search condition: Search task in a homogeneous background. The search area was the uniformly gray monitor screen; this was a uniform stimulus background because, at all scales, each part was the same as any other part. The targets were crosses that consisted of a horizontal and a vertical line of 1 x 5 pixels. Some pilot experiments were done to determine suitable gray values for both search area and targets. As a result of the pilot experiments, the luminance of the search area was set at 11.0 cd/m² and the luminance of the targets at 9.6 cd/m². Each stimulus contained a random number of targets, from one to five, at random positions. There were 48 trials.

Structured search condition: Search task in a picture. A black-and-white group portrait of 15 people, containing parts with much detail as well as parts with little detail, was used as the stimulus background. This was a structured inhomogeneous stimulus background because it contained objects of different scales. Up to nine targets were distributed over the picture. The targets were white crosses on black squares of 9 x 9 pixels, and there were 24 trials.

Free viewing condition: Looking at pictures. The stimulus backgrounds for this condition were pictures collected from the Internet. The set of 16 pictures consisted of landscapes, cartoons, photographs, art works, and fractals. Subject EO did 23 trials in this condition, and the rest of the observers performed 16.

Procedure. The experiment started with careful placement of the coil on the right eye. The subjects wore the coil for not longer than 30 min, so recording time was limited to about 25 min. Prior to the experiment, a calibration screen was shown to each subject. Five dots had to be fixated one after the other, one in the center of the screen and the others shifted 8° to the left, right, top, and bottom, respectively. In the homogeneous search and structured search conditions, the

subjects were instructed to search for targets until they were confident that they had found them all. The subjects indicated the detection of each target by pressing a button with the right thumb. In the free viewing condition, the task was to look at each stimulus for 30 sec. Each trial ended with a push on another button with the left thumb. Preceding each trial, a fixation dot was shown in the center of the screen, and the stimulus appeared when the left button was pushed.

Data analysis. The velocity signal of eye movements was searched for peak velocities above 10 deg/sec. Each peak was considered a potential indicator of the presence of a saccade. The exact onset of the saccade was determined by going backward in time to the point at which the velocity signal dropped below a level two standard deviations above the average velocity during the stable fixation period before the saccade. The exact offset of the saccade was determined by going forward in time to the point at which the velocity signal dropped below a level two standard deviations above the average velocity during the stable fixation period after the saccade. This method was adopted from an earlier study by Van Der Steen and Bruno (1995). The onsets and offsets of the saccades were used to calculate saccade direction, saccade amplitude, fixation time, and location of fixation. In our further analyses, we rejected fixations shorter than 20 ms and saccades with amplitudes less than 0.1 deg.

Results

Mean saccade amplitudes across all subjects were 9.3, 7.5, and 6.1 degrees for the homogeneous search, structured search, and free viewing conditions, respectively. Standard deviations were 7.6, 5.9, and 5.0 degrees. Mean fixation durations across all subjects were 258, 222, and 270 ms for the homogeneous search, structured search, and free viewing conditions, respectively. Standard deviations were 196, 170, and 225 ms. Subjective visual evaluation of the uniformity of fixation density confirmed our expectations for the three conditions. Viewing pictures yields very inhomogeneous fixation densities, with some parts of the stimulus being

fixated repeatedly, and others not at all. Search in a homogeneous stimulus, on the other hand, yields fixation locations that are spread about evenly over the total stimulus area. For each individual trial, we computed cell sizes. Representative scan paths for individual trials in each condition are shown in Figure 2.3, as well as the Voronoi diagrams. Small cells appear in regions of the stimulus where subjects fixated often, and large cells appear in regions where subjects fixated only a few times. The cells in Figure 2.3a are approximately of the same size, whereas the cells are more differently sized in Figure 2.3c. Fixations seem to be concentrated in parts of the stimulus that are rich in detail, including at the edges and in colorful regions. In Figures 2.3b and 2.3c, for example, the air is fixated only a few times. In the free viewing condition, the grand average number of fixations per trial was 87, resulting in a total of about 1400 fixation locations per subject. In the homogeneous search condition the grand average number of fixations per trial was 58, resulting in a total set of about 1400 fixation locations per subject. In the structured search condition, the grand average number of fixations per trial was 32, resulting in a total of about 1150 fixation locations per subject.



Figure 2.3. Examples of stimuli, scan paths (dashed lines) with fixation locations (white asterisks) and matching Voronoi diagrams (white lines) for each of the three conditions: (a) search in a homogeneous background, (b) search in a photograph, and (c) free viewing pictures.

The skewness measure

The skewness of a data set x can be calculated as follows:

$$skewness(x) = \frac{\sum_{i=1}^N (x_i - \mu)^3}{(N-1)\sigma^3}, \quad (1)$$

where μ is the mean of the data and σ is their standard deviation. Figure 2.4 shows skewness values for the 9 subjects in the three conditions.

Computing the skewness is easy and straightforward, but it has the disadvantage of revealing nothing about the variability of the outcome. A way to obtain a confidence interval is to fit an appropriate probability distribution function to the normalized cell size distribution.

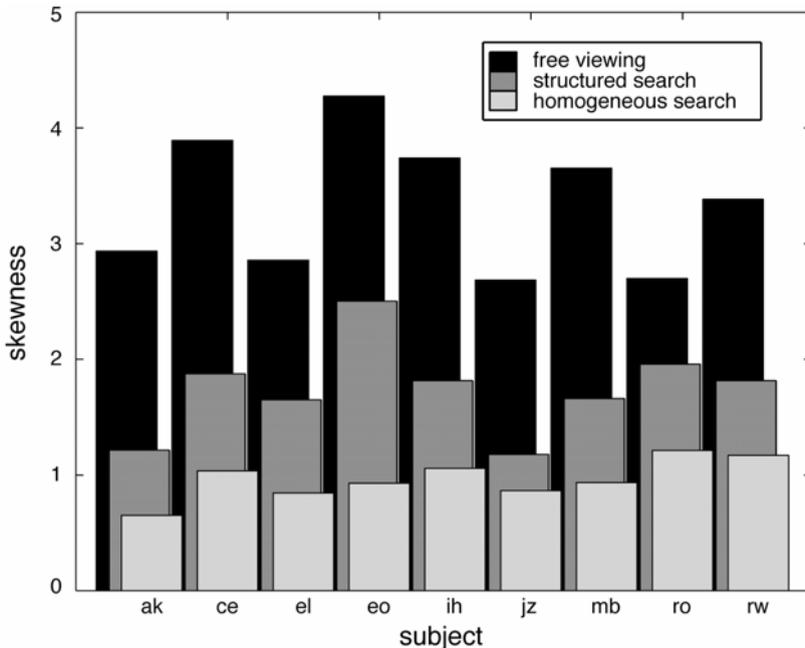


Figure 2.4. Skewness of normalized Voronoi cell sizes for nine subjects in the three conditions.

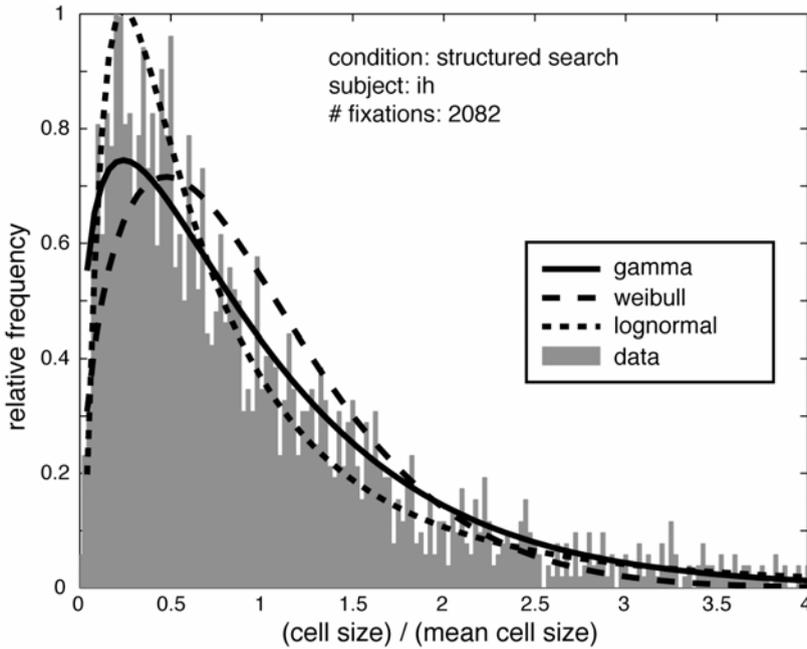


Figure 2.5. Example of a histogram of the normalized cell sizes and the fits of the three proposed distributions.

Fitting the cell size distribution

As can be seen in Figure 2.5 the normalized cell sizes have unimodal distributions of positive real value that are skewed to the right. A number of distributions have these characteristics of unimodality, a positive range of real values and skewness to the right. Some examples are the chi-square, exponential, gamma, lognormal, Rayleigh and Weibull distributions. We will only consider the gamma, lognormal and Weibull distributions since the other three are special cases of either the gamma or the Weibull distribution. The three probability density functions under consideration are:

$$G_{a,b}(x) = \frac{x^{a-1} e^{-x/b}}{\Gamma(a) b^a}, \quad (2)$$

$$W_{\alpha,\beta}(x) = \frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{x}{\alpha}\right)^\beta\right), \quad (3)$$

$$L_{\mu,\sigma}(x) = \exp\left(-\frac{1}{2}\left(\frac{\log(x)-\mu}{\sigma}\right)^2\right), \quad (4)$$

where $G_{a,b}$ is the gamma probability density function, $W_{a,b}$ the Weibull probability density function, $L_{\mu,\sigma}$ the lognormal probability density function, a the gamma shape parameter, b the gamma scale parameter, α the Weibull scale parameter, β the Weibull shape parameter, μ the mean, σ the standard deviation, and Γ the gamma function. In Figure 2.5, the fits of the three different distribution functions are plotted. Subjective visual evaluation of the fits points to the gamma distribution as the most suitable one. An objective measure for goodness of fit is the Kolmogorov-Smirnov test, which tests the probability of data to come from a certain distribution. We applied the Bonferroni correction because we tested the goodness of fit of all three distributions for each data set. The gamma distribution was rejected (total alpha = 0.05) for 11 of the 27 total data sets, whereas the Weibull and lognormal distributions were rejected for 19 of the data sets. Apparently, none of the three proposed distributions fits perfectly to all sets of the data. The literature on Voronoi diagrams indicates that the gamma distribution is an acceptable choice for fitting Voronoi cell sizes: It has proven to yield excellent fits of Voronoi cell size distributions calculated for random Poisson processes in extensive simulations (Okabe, Boots, & Sugihara, 1995). Since we did not find a clear preference for any one of the three distributions, we chose the gamma distribution, in accordance with the literature.

The average value of the gamma probability density function is

$$\mu[G_{a,b}(x)] = ab, \quad (5)$$

in terms of its parameters. Because our data were normalized, by definition the average value of the normalized cell sizes is given by

$$\mu[A_{cell}] = 1, \quad (6)$$

where A_{cell} is the normalized cell size area. The maximum likelihood fitting procedure (which yields the most likely estimates for the parameters) makes use of the following equality:

$$\mu[G_{a,b}(x)] = \mu[A_{cell}], \quad (7)$$

From Equations 5-7, it then follows that $a=1/b$. Therefore, normalization has removed one degree of freedom. Only one fitting parameter remains, and Equation 2 can be rewritten as

$$G_b(x) = \frac{x^{1/b-1} e^{-x/b}}{\Gamma(1/b)b^{1/b}}. \quad (8)$$

Parameter b is positively correlated with the skewness of the distribution: Higher values of b indicate more skewed distributions, corresponding to greater clustering of fixations.

The values for the fitted parameter b according to Equation 8 are plotted in Figure 2.6. The error bars show the 95% confidence intervals. The figure shows three properties of b : First of all, the rank order of b for the three conditions is the same for every subject. This is what we expected for the average distribution across subjects. Second, the rank order of b is what we would expect for the three conditions: b is smallest for the homogeneous search condition and the largest for the free viewing condition. Third, b shows differences between subjects. The variability may reflect idiosyncratic differences in, for example, search strategy or interests. The fact that b is lowest for subject EL in the homogeneous search condition does not necessarily mean that subject EL should also have the lowest values for b in the other two conditions. The homogeneity of the distribution of fixations does not

solely depend on which subject is under investigation; in fact, it is not unlikely that different subjects will react differently under different viewing conditions. In summary, for each subject we found that parameter b was highest in the free viewing condition and lowest in the homogeneous search condition.

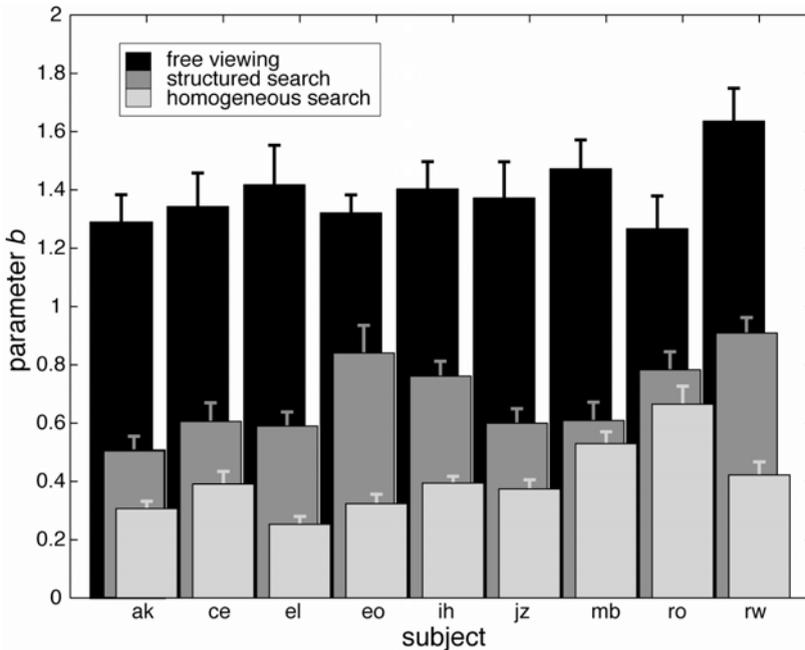


Figure 2.6. Parameter b for nine subjects in the three conditions. The error bars show the upper limit of the 95% confidence intervals.

Discussion

In the previous sections we proposed two measures to describe the uniformity of fixation density: skewness and the gamma fit parameter b . Two advantages of the skewness measure are that it is easy to calculate and independent of the distribution's goodness of fit to any particular probability density function. On the other hand, the gamma fitting yields a confidence interval for parameter b . If the normalized cell

sizes are truly gamma distributed, the measures have the relation

$$\text{skewness} = 2 * \sqrt{b}. \quad (9)$$

The skewness computed indirectly from b , using Equations 8 and 9, differs considerably from the skewness computed directly from the data using Equation 1. This is probably caused by the fact that the normalized cell sizes do not exactly follow a gamma distribution. The difference between the two results may be caused by overestimation of small b values and underestimation of large b values by the fitting algorithm, or by a too-strong influence of extreme values when skewness is calculated with Equation 1. We found the rank order of the skewness measure and parameter b to be consistent for the three conditions and for all subjects. As we expected, there are also some differences between subjects. The rank order of the skewness measure and b confirms our expectations for the outcome of the Voronoi method for the three fixation densities. Therefore, the results of the Voronoi method are in accord with subjective judgments of the uniformity of fixation density. The Voronoi method, however, has two advantages over a subjective judgment of the clustering of fixations: Its results are objective as well as quantitative, and therefore may be more powerful than intuitive judgments in distinguishing between distributions. The present results give us reason to believe that the Voronoi method is a useful tool for the classification of fixation patterns, such as for discriminating the eye movements of patients from those of healthy people.

Simulations

To get a feeling for the possible outcomes of the Voronoi method and the requirements that need to be met in order to apply it to eye movement data, we performed several simulations. Simulations are useful because all conditions are under control, so all properties of a method can be examined systematically. We will only evaluate the results in terms of

parameter b , since the skewness measure has no confidence interval. To investigate the range of possible values for b , we tried to generate sets of fixation locations that yield extreme values for b , as low and as high as possible. The simulation concerned sets of 25 fixation locations. First, in a random set of fixation locations, one fixation was moved so that the value for b was decreased. Changing fixation locations in this way over and over again showed that b can approach 0, although not reach it. The distribution of fixations resulting from this procedure yielded only Voronoi cells of the same size (Figure 2.7a). A histogram of the cell sizes would look like a delta peak, and therefore not skewed. Second, a random set of fixation locations was changed in such a way that b increased. In our simulations, there seemed to be no upper limit to the value of b . As b increased, the distribution yielded many increasingly small cells, as well as some very large ones (Figure 2.7e). A histogram of these cell sizes would be a broad, skewed graph. In Voronoi diagrams from real eye movement data, the Voronoi cells cannot be of zero or infinite size. Therefore, the distribution of cell sizes can never be so skewed that an infinite value for b is found. On the other hand, a value of 0 for b is very unlikely, because all cells would have to be exactly the same size. Figure 2.7 also shows three other values for b and their Voronoi diagrams (panels b–d), corresponding to more common distributions of fixations with intermediate b values. As is clear from Figure 2.7e, $b = 5$ already seems to be an extreme value in terms of real eye movement data. Such a value would mean that the largest part of the stimulus was completely ignored. In conclusion, b values close to 0 correspond to cells of approximately the same size, and therefore fixations that are spread evenly over the total stimulus area. Higher values correspond to a range of different cell sizes: small cell sizes for clustered fixations, and large cell sizes for fixations that are remote from the others. The simulations show that b may take any positive real value. When designing an experiment, it is important to know in advance how many fixations and trials are needed to guarantee a certain confidence interval for parameter b . We investigated

how the accuracy of the fixation distribution obtained from the Voronoi method depends on the mean number of fixations per trial and the number of trials. We simulated fixation locations and investigated how the 95% confidence interval for the fitted parameter b changed for different numbers of trials. We also took into account the mean number of fixations per trial and the uniformity of fixation density. For the fixation density we investigated two conditions: (1) uniform fixation distribution and (2) Gaussian fixation probability around the center of the stimulus with variance of about the stimulus dimensions. The simulated stimulus area was 1024 x 768 pixels. Although the resolution is not of great importance for the result of the simulations, the resolution of 1024 x 768 pixels is widely used for stimuli. Locations falling outside the area of 1024 x 768 pixels were rejected and replaced by newly generated locations. Figure 2.8 shows the number of trials as a function of the mean number of fixations per trial that are needed to obtain 95% confidence intervals that are 20%, 10%, or 5% of the value of the fitted parameter b . For fixation locations from a uniform distribution, the curves in Figure 2.8a can be described by the equation

$$N_t \times N_f = 10^k, \quad (10)$$

where N_t is the required number of trials, N_f the mean number of fixations per trial, and k a constant. For a confidence interval of 20%, $k = 2.8$; for a confidence of 10%, $k = 3.4$; and for a confidence interval of 5%, $k = 4.0$. This relationship between the size of the confidence interval and the value of k means that about 10^3 fixations per condition suffice to obtain a confidence interval of 20%, and about 10^4 per condition to obtain a confidence interval of 5%. For fixation locations from a Gaussian distribution about the center of the stimulus, we find that fewer trials are needed to reach the same accuracy if the average number of fixations per trial is more than about 100. Referring to Figure 2.8 shows that the total numbers of fixations in the eye movement experiment were sufficient to achieve 95% confidence intervals for b smaller than 20% of its

value. For intervals that are 20% of the value of b , a total number of about 10^3 fixations is needed. When standard errors are preferred, note that standard errors are smaller than half of the 95% confidence intervals.

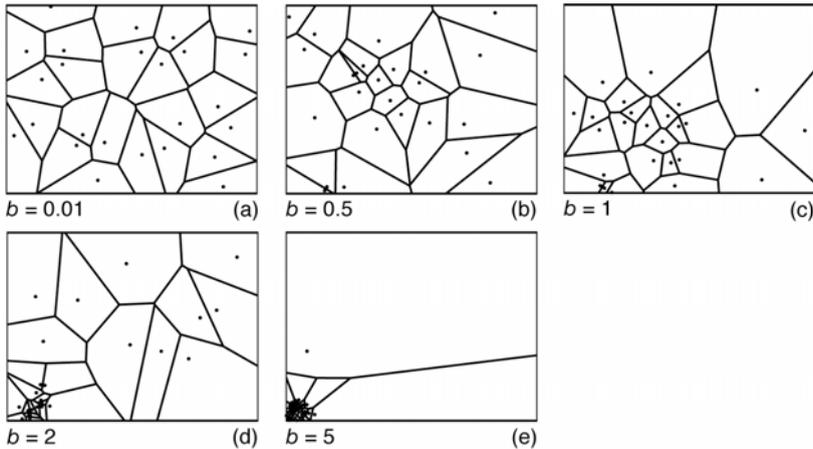


Figure 2.7. Fixation distributions and their Voronoi diagrams for a range of b values. Higher b values go with more grouped fixations.

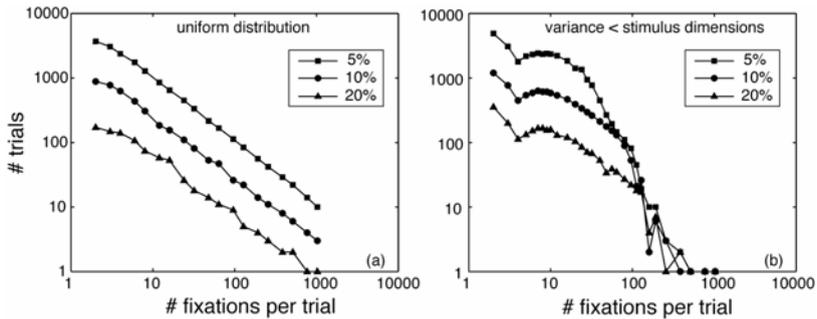


Figure 2.8. The number of trials as a function of the mean number of fixations per trial needed to obtain 95% confidence intervals for parameter b that are 20%, 10%, or 5% of its value. (a) With fixations spread randomly over the stimulus (uniform distribution). (b) With fixations located around the center of the stimulus (Gaussian probability with variance that is smaller than stimulus dimensions).

Conclusions

We developed a method to calculate a quantitative measure for the uniformity of fixation density. This method was tested in three conditions that yielded subjectively different spatial distributions of fixations. On the basis of logical reasoning, we predicted the outcome of the Voronoi method. Agreement of our subjective judgments with the outcome of the Voronoi method indicates that our proposed quantitative measure may be a useful tool for objective description of the spatial features of eye movement data.

The result of our Voronoi method is a single parameter, either the skewness measure or b . Either parameter has a simple interpretation: It is positively related to the clustering of fixations. In order to get accurate results, a large total number of fixations must be analyzed. As a rule of thumb, one should use at least 10^4 fixation locations for high accuracy, whereas 10^3 fixation locations suffice for low accuracy. Another major advantage of the Voronoi method is that voluminous amounts of eye movement data can be analyzed relatively easily. Disadvantages are that just a part of the available information in the data is used and that only one parameter remains. Furthermore, the Voronoi method does not address any temporal aspect of viewing behavior. Although it is not very often used, a simple measure like the nearest-neighbor distance between fixation locations could also provide information about the clustering of fixations and the uniformity of fixation density. The mean nearest-neighbor distance, however, is a measure based on a one-dimensional characteristic, whereas the Voronoi method is based on two-dimensional areas. The Voronoi method has two advantages in comparison with the nearest-neighbor distance. First of all, each fixation location is associated with a unique cell area. In the case of two fixation locations that are very close together, the nearest-neighbor distance will yield the same result for both fixation locations, whereas the Voronoi method will yield different cell areas, since other fixation locations are also taken into account in the computing of the cell areas, not just the nearest fixation location. In short, two-dimensional areas

provide more information than do one-dimensional distances. Even in the case of a one-dimensional stimulus, such as a line of text, this principle still holds: The nearest-neighbor distance for each fixation location is related to only one other fixation location (the nearest), whereas the Voronoi cell size is based on the distance to two neighboring fixation locations, one on either side. The Voronoi method therefore conforms to the number of dimensions of the stimulus: Cell size is expressed in a one-dimensional distance when the stimulus is one-dimensional and in a volume when the stimulus is three-dimensional. The nearest-neighbor distance is always one-dimensional, regardless of the dimensions of the stimulus. Second, each cell in the Voronoi diagram can be uniquely related to one fixation location, and therefore to its corresponding visual span (Bertera & Rayner, 2000; Jacobs, 1986; Lévy-Schoen, O'Regan, Jacobs, & Coeffé, 1984; Pomplun, Reingold, & Shen, 2001; Reingold, Charness, Pomplun, & Stampe, 2001). Each cell area may be an indicator for the size of the visual span at that particular location. A part of a stimulus that has a high information density, for instance as a result of many color and contrast differences, is fixated frequently and has many fixations close to each other. In such a part of the stimulus, the Voronoi cell sizes are generally small, which is consistent with the finding that the visual span is smaller as stimulus density increases (Motter & Belky, 1998). Voronoi diagrams are part of a large family of tessellations that could also be used to construct areas associated with fixation density. For example, Delaunay triangulation (Delaunay, 1934) produces triangles such that no fixation locations are contained within any triangle's perimeter. However, these triangular areas are not as easily interpretable as the cells in a Voronoi diagram. Areas constructed by Delaunay triangulation belong to three fixation locations, because they are triangles, and therefore contain information about three fixation locations. With Delaunay triangulation, each fixation location might relate to any number of triangles, whereas in Voronoi diagrams, each fixation location relates to exactly one cell. The direct

coupling of Voronoi cells with fixation locations also provides an objective way to relate fixation locations to local image content. Each fixation location belongs to a certain part of the stimulus—namely, the part that is contained by its Voronoi cell. Image analysis might be used on each Voronoi cell to try to establish correlations between any frequently used parameter in image analysis, such as contrast or brightness, and characteristics such as fixation duration or the size of the Voronoi cell. Recent research (Vlaskamp, Hooge, & Over, 2003) supports the idea that the Voronoi method is a powerful tool to analyze uniformities of fixation densities. In that study, 6 subjects participated in a search experiment under three stimulus conditions in which the contrast of objects with the background was either consistently high, consistently low, or a mixture of high, medium, and low. In the high- and low-contrast conditions, the fixation density across the stimulus was highly uniform, resulting in (low) values for b ranging from 0.15 to 0.22. In the mixed-contrast condition, the fixation density was less uniform across the stimulus, resulting in values for b ranging from 0.20 to 0.29. For each subject, b was higher in the mixed-contrast condition than in the other two conditions. The uniformity of fixation density was practically the same in the high- and low-contrast conditions, although the number of fixations per trial differed considerably. Thus, even with target-absent stimuli, in which subjects had to search the entire stimulus to realize that there was no target, the Voronoi method yielded results that were congruent with the visual evaluations and was able to distinguish between different uniformities of fixation density: The densities were less uniform in the mixed-contrast condition than in either the high- or the low-contrast condition. Visual evaluation required a substantial amount of time and was subjective, whereas the Voronoi method could relatively easily quantify uniformities.

References

- Aks, D. J., Zelinsky, G. J., & Sprott, J. C. (2002). Memory Across Eye-Movements: 1/f Dynamic in Visual Search. *Nonlinear Dynamics, Psychology, and Life Sciences*, 6(1), 1-25.
- Bertera, J. H., & Rayner, K. (2000). Eye movements and the span of the effective stimulus in visual search. *Perception & Psychophysics*, 62(3), 576-585.
- Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, 9(1), 27-38.
- Choi, Y. S., Mosley, A. D., & Stark, L. W. (1995). String editing analysis of human visual search. *Optometry & Vision Science*, 72(7), 439-451.
- Collewijn, H., Van der Mark, F., & Jansen, T. C. (1975). Precise recording of human eye movements. *Vision Research*, 14, 447-450.
- De Gennaro, L., Ferrara, M., Curcio, G., & Bertini, M. (2001). Visual search performance across 40 h of continuous wakefulness: Measures of speed and accuracy and relation with oculomotor performance. *Physiology & Behavior*, 74(1-2), 197-204.
- De Lucio, J. V., Mohamadian, M., Ruiz, J. P., Banayas, J., & Bernaldez, F. G. (1996). Visual landscape exploration as revealed by eye movement tracking. *Landscape and Urban Planning*, 34, 135-142.
- Delaunay, B. (1934). Sur la sphere vide. *Bulletin of Academy of Science USSR*, VII, 793-800.
- Gbadamosi, J., & Zangemeister, W. H. (2001). Visual imagery in hemianopic patients. *Journal of Cognitive Neuroscience*, 13(7), 855-866.
- Hooge, I. T., & Erkelens, C. J. (1999). Peripheral vision and oculomotor control during visual search. *Vision Research*, 39(8), 1567-1575.

- Jacobs, A. M. (1986). Eye-movement control in visual search: how direct is visual span control? *Perception & Psychophysics*, *39*(1), 47-58.
- Laeng, B., & Teodorescu, D.-S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science*, *26*, 207—231.
- Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, *25*(7), 963-977.
- Lévy-Schoen, A., O'Regan, J. K., Jacobs, A. M., & Coeffé, C. (1984). The relation between visibility span and eye movements in various scanning tasks. In A. G. Gale & F. Johnson (Eds.), *Theoretical and applied aspects of eye movement research* (pp. 133-142). Amsterdam: North Holland.
- Motter, B. C., & Belky, E. J. (1998). The zone of focal attention during active visual search. *Vision Research*, *38*(7), 1007-1022.
- Noton, D., & Stark, L. (1971a). Scanpaths in eye movements during pattern perception. *Science*, *171*(968), 308-311.
- Noton, D., & Stark, L. (1971b). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, *11*(9), 929-942.
- Okabe, A., Boots, B. N., & Sugihara, K. (1995). *Spatial tessellations: concepts and applications of Voronoi diagrams*: Wiley.
- Pelz, J., Hayhoe, M., & Loeber, R. (2001). The coordination of eye, head, and hand movements in a natural task. *Experimental Brain Research*, *139*(3), 266-277.
- Phillips, M. L., & David, A. S. (1997). Visual scan paths are abnormal in deluded schizophrenics. *Neuropsychologia*, *35*(1), 99-105.
- Pomplun, M., Reingold, E. M., & Shen, J. (2001). Investigating the visual span in comparative search: the effects of task difficulty and divided attention. *Cognition*, *81*(2), B57-67.
- Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players:

- evidence from eye movements. *Psychological Science*, 12(1), 48-55.
- Rommel, R. S. (1984). An inexpensive eye movement monitor using the scleral search coil technique. *IEEE Transactions on Biomedical Engineering*, 31, 388-390.
- Robinson, D. A. (1963). A method of measuring eye movements using a scleral search coil in a magnetic field. *IEEE Transactions on Biomedical Engineering*, 10, 137-145.
- Shinoda, H., Hayhoe, M. M., & Shrivastava, A. (2001). What controls attention in natural environments? *Vision Research*, 41(25-26), 3535-3545.
- Van der Steen, J., & Bruno, P. (1995). Unequal amplitude saccades produced by aniseikonic patterns: effects of viewing distance. *Vision Research*, 35(23-24), 3459-3471.
- Vlaskamp, B. N. S., Hooge, I. T. C., & Over, E. A. B. (2003). Local contrast affects fixation density in saccadic search. *Perception*, 32 supplement, 109B.
- Voronoi, G. (1908). Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième mémoire: Recherches sur les parallélogrammes primitifs. *Journal für die Reine Angewandte Mathematik*, 134, 198-287.
- Westheimer, G., & Hauske, G. (1975). Temporal and spatial interference with vernier acuity. *Vision Research*, 15, 1137-1141.
- Williams, L. G. (1966). The effect of target specification on objects fixated during visual search. *Perception & Psychophysics*, 1, 315-318.
- Yarbus, A. (1967). *Eye movements and vision*. New York: Plenum Press.
- Zangemeister, W. H., Sherman, K., & Stark, L. (1995). Evidence for a global scanpath strategy in viewing abstract compared with realistic images. *Neuropsychologia*, 33(8), 1009-1025.
- Zelinsky, G. J., Rao, R. P. N., Hayhoe, M. M., & Ballard, D. H. (1997). Eye movements reveal the spatio-temporal

dynamics of visual search. *Psychological Science*, 8, 448-453.

Zihl, J., & Hebel, N. (1997). Patterns of oculomotor scanning in patients with unilateral posterior parietal or frontal lobe damage. *Neuropsychologia*, 35(6), 893-906.

Chapter 3. Saccade direction in visual search: top-down or bottom-up?

Introduction

People make many saccadic eye movements, on average about three per second. The goal of these saccadic eye movements is to collect visual information with the fovea, the central part of the retina, which is most suited for acquisition of detailed visual information because it is the part of the retina that has the highest density of photoreceptors. The acquired visual information can be used for interaction with the environment, such as communicating with other people, for example in advertising, traffic signs etc. To improve the effectiveness of communication with the help of visual information, one may want to know where people look to acquire the visual information. To this purpose, we would like to have a model that can be used to predict where people are going to look. Some interesting applications of such a model are possible in advertising and medical image enhancement for example. If it is predicted that people would look at uninformative regions for too long, the information should be presented differently. This may for example prevent costly advertising that turns out to be ineffective.

For the prediction of direction and amplitude of saccadic eye movements, it is useful to know what factors may influence whereto saccades are made. There are at least two types of factors. The first type consists of (top-down) factors that depend on the intentions of the observer. These factors may be expected to have influence on saccade direction because of, for example, knowledge about the environment, reading direction (English versus Hebrew), etc. Such factors may form a bias in eye movement behavior. The second type is formed by (bottom-up) factors that depend on the stimulus; features such as color, contrast, luminance, orientation, etc.. For example, highly salient locations often

evoke involuntary eye movements. Theeuwes and colleagues (Theeuwes, Kramer, Hahn, Irwin, & Zelinsky, 1999) showed that top-down control could not prevent the eyes to move towards an abrupt visual onset.

There are several methods that attempt to predict where people look to acquire visual information. They mainly differ in what type of factors they are based on. One method, purely based on bottom-up factors, is the method of saliency maps (Itti & Koch, 2000; Koch & Ullman, 1985). Saliency maps are the combination of several low-level feature maps that are calculated with computer algorithms for each image that is to be examined. In the saliency maps, saliency is quantified for each location in the display. To mimic scan paths, a simulated saccade is made to the most salient location in the map. After the saccade, the saliency at the current fixation location is suppressed, and the next saccade is made to the location with the highest saliency after suppression. Successive salient locations are not necessarily in the vicinity of each other. A disadvantage of this model is that it thus may produce very large saccades, but human observers do not frequently make very large saccades. When human observers do make such large saccades, it is often to reposition the eyes in another region of the field of view, not because the target location of the saccade is salient.

Another method based on bottom-up factors was developed by Najemnik and Geisler (2005). They derived the ideal Bayesian observer for search tasks in which a target is embedded within a random background that has the spectral characteristics of natural scenes. Their ideal searcher used precise knowledge about the statistics of the scenes in which the target was embedded, and about its own visual system, to make eye movements that gain the most information about target location. They found that humans achieved nearly optimal search performance, which may suggest that observer-dependent factors play a less important part than stimulus-dependent factors in visual search.

A third method for predicting where people look to acquire visual information is mostly a top-down method.

Noton and Stark (1971a; 1971b) showed that people use fixed scan paths to look at patterns of checkerboard squares. A scan path appeared repeatedly during initial viewing of a pattern and the same scan path appeared when the pattern was presented to the observer later. The scan path is characteristic for a given observer viewing a certain pattern, but varies with observer and pattern. The scan path also appeared during imagery of a previously viewed stimulus (Brandt & Stark, 1997; Gbadamosi & Zangemeister, 2001). The predictive value of the method is limited because it can only predict eye movement patterns of a particular observer looking again at a particular stimulus.

The aim of this chapter is to answer the question how much observer-dependent and stimulus-dependent factors contribute to making a saccade in a certain direction. For the answer to this question it is necessary to analyze the influence of observer-dependent and stimulus-dependent factors separately. In order to do so, we developed a new method for the prediction of eye movement directions that is capable of capturing top-down as well as bottom-up characteristics. The method needs a set of recorded eye movement data to work on, like the scan path method. Eye movement data reflect the interaction between stimulus and observer, and therefore a method based on eye movement data does not only show the influence of stimulus factors, but also captures characteristics due to observer-dependent factors. In this sense, our method adds to the work of Itti and Koch (2000) and to the work of Najemnik and Geisler (2005). Itti and Koch (2000) did not record eye movements, and therefore their method mainly assesses the influence of stimulus characteristics, rather than individual observer characteristics. The ideal observer developed by Najemnik and Geisler (2005) includes information about the visual system in addition to the characteristics of the stimuli. However, since this information cannot vary with different observers there is essentially no influence of observer-dependent factors in this model either.

Our approach is to construct maps that indicate the most likely saccade direction for each position in the viewing

area and that provide a measure for the likelihood of that direction. To analyze the influence of observer-dependent and stimulus-dependent factors on saccade direction separately, we would like to have two kinds of maps at our disposal:

- A stimulus-averaged map. This map is constructed by combining the eye movements made by a single observer in a large number of stimulus displays. Such a map identifies patterns that are specific for an observer. It is assumed that, because of the combination operation, the influence of stimulus content averages out. Therefore, in a search experiment for example, a requirement for target locations is that they are homogeneously distributed over the stimulus area. The stimulus-averaged map may for example show that an observer often makes upward saccades when he is fixating the left side of stimuli.
- An observer-averaged map. This map is constructed by combining the eye movements of a large number of observers in a specific stimulus. Such a map identifies patterns that are specific for a particular stimulus. The observer-averaged map may for example show that the stimulus under analysis evokes upward saccades when observers are fixating the left side of the stimulus. It is assumed that the influence of factors that vary in individual observers averages out when the number of observers is sufficiently large. However, some observer-dependent factors cannot be averaged out, namely those that are common to all or many observers. For example, a possible remaining effect of reading direction may be expected when all observers are Hebrew readers, as opposed to when all observers are Western Europeans.

It is also possible to analyze the effect of all remaining factors that are common to all observers and stimuli, such as the oculomotor system and the geometry or shape of the viewing area. These factors are not necessarily only top-down or bottom-up. Anyhow, they have in common that their effect on

saccade direction is identical in each trial. To analyze the combined effect of all these factors, a map is constructed that is the result of the combination of all recorded eye movements (of all observers and all stimuli).

We recently developed a new method for the quantification of the uniformity of fixation density (Over, Hooge, & Erkelens, 2006) that is based on Voronoi diagrams (Voronoi, 1908). The method for the construction of the above-mentioned direction-coded maps introduced here is also based on Voronoi diagrams. The Voronoi diagram of a set of fixations yields a Voronoi cell for each fixation location. Voronoi diagrams are independent of choices, such as which area to code for saccade direction and what scales to use. It does not use thresholds or parameters that must be set to their optimum value. The maps are computed as vector plots that indicate the most likely saccade direction for each position in the search area. The length of the vector indicates the likelihood. Comparisons between vector maps are easily made by vector subtraction. This subtraction can be used to compare experimentally obtained maps with models, for example a model of left-to-right and top-to-bottom search. Whereas saliency maps predict the location of the next fixation, our maps provide a way to predict the direction of a saccade starting from any location in the viewing area. What our method adds is that we can determine which of the above-mentioned types of factors contributes most to saccade direction. A possible objection to the methods based on the work of Itti and Koch (2000) and Najemnik and Geisler (2005) is that the role of observer-dependent factors is underestimated. Based on common sense, one may expect a more prominent role for observer-dependent factors. Our method provides a means to verify this expectation. The top-down component can certainly be important when different tasks are compared to each other, because Yarbus (1967) showed that the task at hand strongly influences eye movement patterns.

Method: direction coding of Voronoi diagrams

The bounded Voronoi diagram is constructed from the fixation locations in a trial (Figure 3.1a). Each trial thus yields a Voronoi diagram. Most software packages contain a routine to compute Voronoi diagrams (Matlab, Mathematica etc.). For more detailed information about the construction of Voronoi diagrams see Voronoi (1908), and for Voronoi diagrams on the basis of fixation locations see Over et al. (2006). A grid is constructed for the viewing area. Suitable grid points are for example the pixel positions of the stimulus. For each cell in the diagrams, all grid points within that cell are then assigned a vector that has length one and the direction of the saccade that leaves the cell. We chose the vector to have length one because we wanted all saccades to contribute equally to the direction-coded map, independent of their respective amplitude. The cell that corresponds to the final fixation does not have a saccade leaving it, so all grid points within that cell are not assigned any vector. The result is a direction-coded Voronoi diagram (Figure 3.1b).

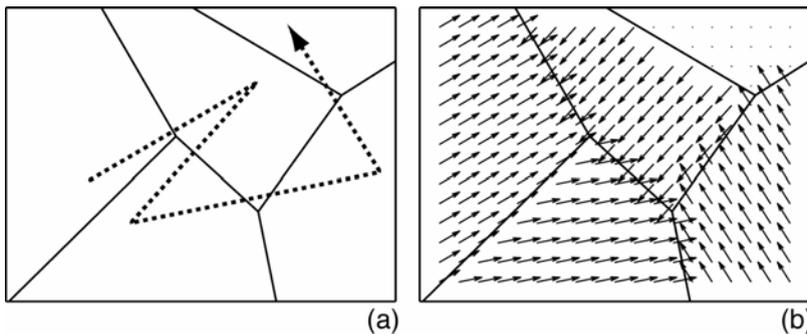


Figure 3.1. (a) Scan path (dashed lines) and the corresponding bounded Voronoi diagram (solid lines) (b) Direction-coded Voronoi diagram. Each grid point in the cell is assigned the direction of the saccade leaving that cell. All vectors have length 1 and have one of the grid points as their starting point. The cell without vectors corresponds to the final fixation.

It is relatively simple to combine for example the data of all stimuli for one observer. It is namely possible to combine several direction-coded Voronoi diagrams by vector addition. The resulting vectors are then scaled to the number of Voronoi diagrams, by dividing the resulting vector lengths by the number of Voronoi diagrams. These calculations yield a new direction-coded map, with all lengths of the resulting vectors between one (all vectors in the individual direction-coded Voronoi diagrams in the same direction) and zero (all vectors in the individual direction-coded Voronoi diagrams canceling each other out). See Figure 3.2 for a schematic example and a direction-coded Voronoi diagram from real eye movement data. The length of each resulting vector relates to the distribution of saccade directions from that position in the viewing area; zero length means a homogeneous distribution of saccade directions (in other words: equal probability for all directions), and length one means all saccade directions in the indicated direction.

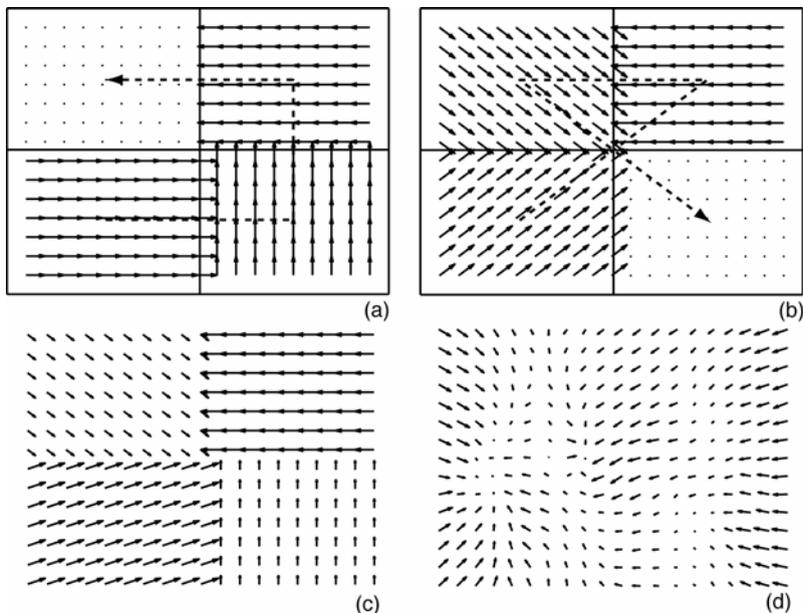


Figure 3.2. (a) and (b): Direction-coded Voronoi diagrams for two hypothetical trials. All vector lengths are 1, the length of a vector between

nearest-neighbor grid points. (c): The combination of the direction-coded Voronoi diagrams in (a) and (b). In the upper right quadrant, vectors in (a) and (b) are in the same direction, and therefore vector length in (c) is 1. In the lower left quadrant the vectors in (a) and (b) are not in the same direction, and therefore the vector direction in (c) is the average direction of (a) and (b), and vector length is smaller than 1. In the other two quadrants, there are only vectors in (a) or (b), and therefore vector length in (c) is 0.5. (d) A real example of a direction-coded map. Eye movement data from 21 observers searching in one stimulus are combined. Vector length 1 is the length of a vector between nearest-neighbor grid points, analogous to (a), (b) and (c). If a grid point has a vector with length 1, then the vectors in all the combined maps were in the same direction at that grid point.

The result of this combination of direction-coded Voronoi diagrams (see Figure 3.2d for a real example) can be analyzed similarly as vector fields are analyzed. One type of analysis of vector fields is the identification of attractors and repellers. (In a gravity field, attractors would be the valleys and repellers would be the hills.) Attractors and repellers, and their strengths, can be calculated from the divergence field. Divergence in two dimensions is defined as:

$$\nabla \cdot (\mathbf{A}) = \frac{\partial \mathbf{A}_x}{\partial x} + \frac{\partial \mathbf{A}_y}{\partial y} \quad 1),$$

where \mathbf{A} is a continuously differentiable vector field, and ∇ is the nabla operator. The divergence of a vector field thus yields a scalar field. Divergence is a linear operator:

$$\nabla \cdot (c\mathbf{A} + d\mathbf{B}) = c\nabla \cdot (\mathbf{A}) + d\nabla \cdot (\mathbf{B}) \quad 2),$$

where \mathbf{A} and \mathbf{B} are continuously differentiable vector fields, and c and d are constants. Linearity is a property that allows for straightforward comparisons of divergences: divergence of a difference equals the difference of the divergences. The sign of the divergence shows whether the field is an attractor or a repeller at a particular location, negative values indicating an attractor, and positive values indicating a repeller. The

magnitude of the divergence is a measure for the strength of the attractor or repellor.

In a visual search task, targets may act as attractors because the majority of final saccades in successful trials are directed towards the target. The strength of the attractor also depends on the directions of other saccades made in the trial. A highly conspicuous target is expected to yield a strong negative divergence at the location of the target because a conspicuous target is a strong attractor for saccades. An inconspicuous target may cause only a slightly negative divergence, because other locations in the stimulus may attract saccades more strongly than an inconspicuous target does. A repellor indicates that the vector field diverges. A candidate for a repellor is the center of the viewing area. If an observer always starts the trial around the center of the viewing area, and he makes his first saccade in an arbitrary direction away from the center, then the saccade direction field diverges at the location of the center of the viewing area.

Divergence is one quantitative measure for properties of a vector field. Another quantitative measure is the sum of the vector lengths. The sum of the vector lengths is a measure for the strength of the field. It is also a useful measure to quantify the difference between two vector maps. A field of difference vectors can be calculated by subtracting the vectors of two direction-coded maps at each grid point. If the two direction-coded maps are identical, then the difference vectors at all grid points have zero length, and the sum of the difference vector lengths is also zero. The maximum difference between two direction-coded maps is the number of grid points times the maximum length of a difference vector, which is two. This is the case if all vectors in both maps have length one and all vector directions in one map are opposite to the directions in the other. The difference vector at each grid point then has length two; hence the sum of all the difference vector lengths is twice the number of grid points.

Visual search experiment

Methods

Eye movements during this search experiment were recorded with an Eyelink system at 250 Hz. Stimuli were presented on a 22-inch LaCie monitor at a frame rate of 85 Hz in a darkened room. With a screen distance of 50 cm, the stimuli subtended 45 degrees horizontally and 33 degrees vertically. Twenty-one observers participated. None showed any visual or oculomotor pathology other than refraction anomalies. Observer IH is one of the authors of this chapter. The other observers were naïve as to the goal of the experiment. Informed consent was obtained from all observers prior to participation in the experiments.

Prior to the experiment, the Eyelink system was properly adjusted and a calibration screen was shown to the observer. Nine dots had to be fixated subsequently, one in the centre of the screen and the others in the corners of the screen and halfway between two corners. Preceding each trial, a fixation dot was shown in the centre of the screen and the stimulus appeared when the space bar was pressed. The task was to find the target as quickly as possible. Observers pressed the space bar again when they fixated on the target. TNO Human Factors in Soesterberg, the Netherlands, provided the original 44 scenes (6144 x 4096 pixels, 16.7 million colors) used in this experiment. For more detailed information about the scenes see Toet, Bijl, Kooi, and Valeton (1998). We selected 136 cutouts of the scenes. The cutouts consisted of 1600 x 1067 pixels. In each of the 136 stimuli one of nine possible military vehicles was present. Target locations were rather homogeneously distributed (Figure 3.3). This was necessary to meet the requirements for being able to average out stimulus content, as mentioned in the introduction. Timeout occurred after thirty seconds, and observers did not receive feedback about their performance.



Figure 3.3. Superimposed silhouettes of the targets in the 136 stimuli.

Analysis

Saccade direction, saccade amplitude, fixation duration and the location of fixations were determined. We rejected fixations shorter than 20 ms and saccades with amplitudes smaller than 0.1 degrees for further analysis. Fixation locations were used to calculate the Voronoi diagrams, and saccade directions were used to code the cells in each diagram as described in the method section. For reasons of clarity, we present the direction-coded maps on 20 by 15 grids in this chapter.

Results

Search times. In total, we recorded eye movements in 2856 trials (21 observers x 136 stimuli). None of the observers fixated outside the stimulus area. For all trials we computed the direction-coded Voronoi diagrams. Search times were also recorded for each trial. They ranged from 0.46 to 29.92 seconds. The response for target location was either correct or false, or there was no response because the trial was timed out

after thirty seconds. Percentage correct ranged from 65% to 97% between subjects. On average, percentage correct was 78% +/- 7%. The other trials resulted in false alarms (18%) or were timed out after thirty seconds (4%). Percentage correct and the wide range of search times indicate that stimuli varied considerably in difficulty. Some targets were easy to find, others were not found within thirty seconds. Eye movements recorded in trials that were timed out, and in false-alarm trials were nevertheless used in the analyses of saccade direction.

Overall field of saccade direction. Our first analysis was to combine the direction-coded diagrams of all recorded trials. Figure 3.4 shows the direction-coded map based on all stimuli and observers. The pattern of saccade directions that emerges is the result of factors that are common in all stimuli and all observers. It appears that there is a single, horizontally stretched attractor somewhat below the center of the viewing area. All vectors are directed towards this attractor. The borders of the viewing area have a repelling effect. This is easily explained because targets could only be found within the stimulus area (looking outside the stimulus area is a waste of time). The fact that the attractor is located below the center of the viewing area may be due to the presence of perspective in the stimuli. The foreground was less difficult to search than the background, because due to perspective, targets in the foreground are generally larger than targets in the background. The stimuli that contained large targets in the upper half of the stimuli were stimuli with a minor perspective difference between foreground and background.

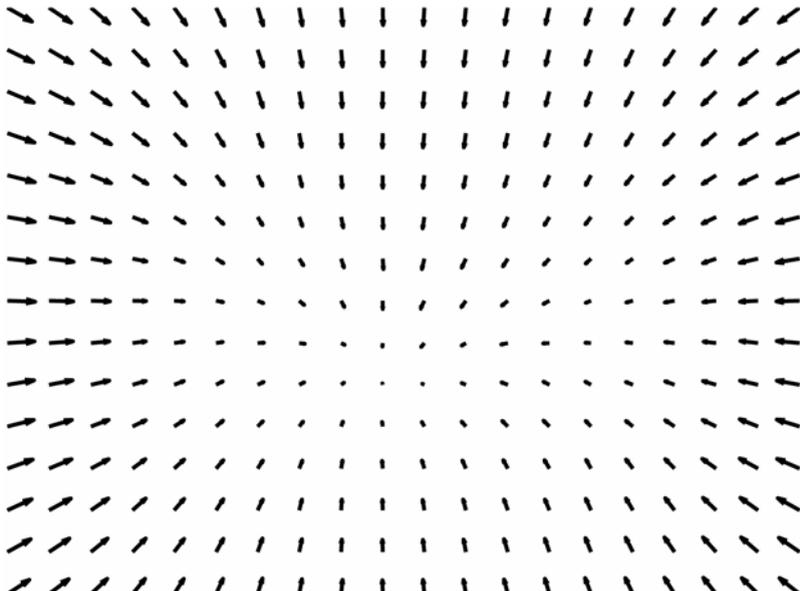


Figure 3.4. Direction-coded map for all experimental data combined (all observers and all stimuli). Vector length 1 is the length of a vector between nearest-neighbor grid points. If a grid point has a vector with length 1 (the length of a vector between nearest-neighbor grid points), then the vectors at that grid point in the combined maps all pointed in the same direction (see also Figure 3.2).

Saccade direction per stimulus. For the analysis of saccade direction patterns that are specific for a stimulus, we computed 136 stimulus-specific maps by combining the direction-coded Voronoi diagrams of the 21 observers per stimulus (see method section, Figure 3.2). As we mentioned in the method section, we expected a conspicuous target to attract saccades more strongly than an inconspicuous target. Indeed, this was a clear result from our data: conspicuous targets yielded one strong attractor (Figure 3.5c) and inconspicuous targets yielded a number of weak attractors and repellers (Figure 3.5f). There were no strong repellers (i.e., large positive divergences) in any of the direction-coded maps per stimulus. The divergence was generally more negative than positive, which is also clear from the range of values of the divergence in all maps: the minimum value of the

divergence in all maps per stimulus was -0.55 , and the maximum $+0.25$.

Conspicuous targets on average result in shorter search times than inconspicuous targets (Engel, 1977). It is therefore reasonable to state that, when average search time is short, the chance that the target was conspicuous is higher than the chance that the target was inconspicuous. If we use search time as an indirect measure for target conspicuity, it is possible to compute the correlation between divergence of the direction-coded maps and target conspicuity. The divergence of direction-coded maps at the target position correlated with median search time per stimulus (the time at which 50% of the observers had correctly identified the target): the correlation coefficient was 0.60. Figure 3.6a shows how median search time per stimulus related to the divergence of the stimulus-specific maps at the location of the target. Stronger attractors generally yielded shorter search times. The relationship was nonlinear, but nonetheless clearly present.

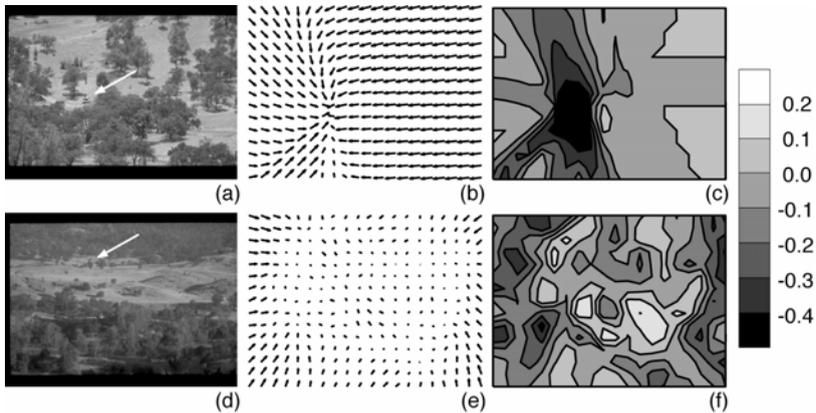


Figure 3.5. (a) Stimulus with a conspicuous target (median search time: 0.67 s). The arrow points to the target. (b) Direction-coded map for that stimulus. (c) Divergence of the direction-coded map. (d-f) As (a-c), but for an inconspicuous target. Less than 50% of the observers found this target within 30 seconds. The color bar at the right belongs to panels (c) and (f). Dark regions (negative divergence) indicate attractors, light regions (positive divergence) indicate repellors.

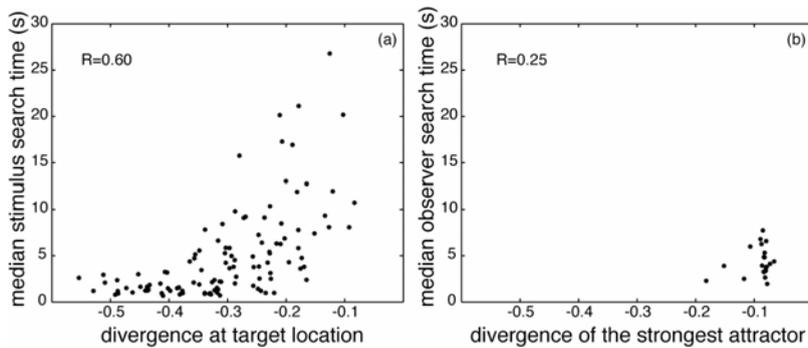


Figure 3.6. (a) Median search time per stimulus as a function of divergence at the location of the target. Data for all 136 stimulus-specific maps are plotted. Correlation coefficient is 0.60. (b) Median search time per observer as a function of minimum divergence. Data for all 21 observer-specific maps are plotted. Correlation coefficient is 0.25.

Saccade direction per observer. For the analysis of patterns that are specific for an observer, we obtained 21 observer-specific maps by combining the 136 direction-coded Voronoi diagrams of each observer. The minimum of the divergence in all maps per observer was -0.18 , and the maximum $+0.03$. This range between minimum and maximum divergence was only about 1/4 of the range between minimum and maximum divergence in the maps per stimulus. The correlation between minimum divergence per observer and median search time per observer (the time at which an observer has correctly identified 50% of the targets) was 0.25, which was less than the correlation between the divergence at the target position and median search time per stimulus. Figure 3.6b shows how median search time per observer related to the minimum divergence of the observer-specific maps. The correlation was only weak; for many observers the minimum divergence was about -0.08 , but their median search times differed considerably.

We also analyzed the strengths of the vector fields represented by the direction-coded maps. Figure 3.7 shows the field strengths compared to the field that is represented by the direction-coded map for all trials (Figure 3.4). The vector fields per stimulus were stronger (i.e. the sum of the vector

lengths was larger) than the vector fields per observer, and the range of strengths was larger for fields per stimulus than for fields per observer. This means that the fields per observer had higher resemblance to the field for all trials than the fields per stimulus.

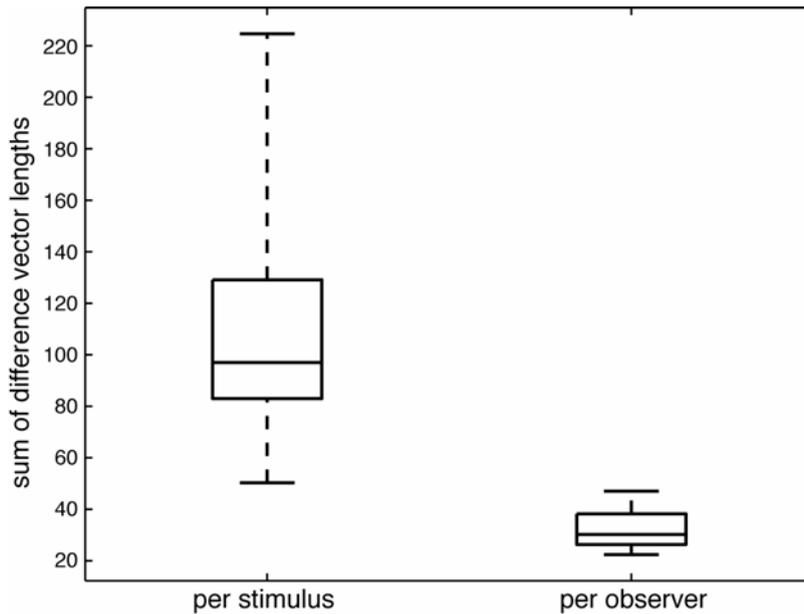


Figure 3.7. Boxplot of the sum of difference vector lengths for direction-coded maps combined per stimulus compared to direction-coded maps combined per observer. The box extends from the 25th percentile to the 75th percentile, with a center line at the median value. The whiskers indicate the extreme values.

A simple model of the generation of saccade directions. Although the abovementioned results indicate a clear difference between the influence of stimulus-dependent and observer-dependent factors on saccade direction, there is also a way to analyze the overall generation of saccade direction in the eye movement system. This analysis is based on the results presented in Figure 3.4: the overall field of saccade direction. The field depicted in Figure 3.4 resembles a pattern of saccade directions that may be expected of a

generator that randomly selects new fixation locations within the viewing area, and so that each location is equally likely to be selected. This is because, with such a generator, the chance of a rightward saccade from a fixation in the left side of the stimulus is also higher than the chance of a leftward saccade. We therefore compared the observed saccade direction behavior to the theoretical saccade direction behavior for two models of random saccade generation: a model of a system with one random generator (a system that randomly selects a new fixation location in the viewing area), and a model of a system with two random generators (one for the amplitude, and the other for the direction of the saccade). The model of a system with two random generators repeats the selection of saccade amplitude and direction until the fixation location that results from executing that particular saccade falls within the search area. Direction-coded maps produced by the two models are shown in Figures 3.8a and 3.8b. For comparison of the observed saccade direction behavior with the simulated saccade directions of the two models we used the quantitative measure of the sum of difference vector lengths (see method section). The sum of the lengths of the difference vectors between observed and theoretical behavior for the model of the system using two random generators was 40, and the sum of the lengths of the difference vectors between observed and theoretical behavior for the model of the system using one random generator was 72. Thus, the difference of the observed average saccade direction behavior with theoretical behavior produced by the model of the system using two random generators was smaller than with theoretical behavior produced by the model of the system using one random generator. However, not even the second model of the system using random saccade generation can explain all of the observed average saccade direction behavior since the sum of the difference vector lengths is different from zero. Two clear differences remain between recorded saccade direction behavior and the saccade directions produced by the model: 1) vectors of zero length for the recorded data are located below the center of the viewing area (possibly due to perspective in

the stimuli), and 2) the total divergence in the recorded data is less negative than in the theoretical direction-coded map of the model of the system using two random generators.

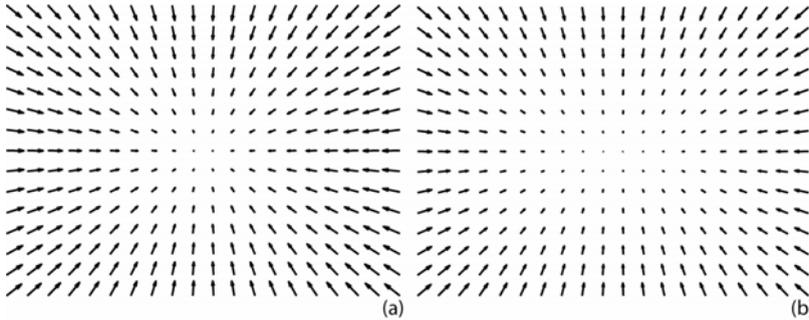


Figure 3.8. (a) Theoretical direction-coded map of the model of the system using one random generator for selection of the next fixation location. (b) Theoretical direction-coded map of the model of the system using two random generators, one for saccade direction, the other for saccade amplitude. Vector length 1 is the length of a vector between nearest-neighbor grid points. If a grid point has a vector with length 1, then all vectors at that grid point in the maps that were combined pointed in the same direction (see also Figure 3.2).

Discussion

In order to answer the question how much observer-dependent and stimulus-dependent factors contribute to making a saccade in a certain direction, we first developed a method to identify average saccade direction at each location in the viewing area. Our new method is based on recorded eye movement data, and thus indirectly captures both stimulus-dependent and observer-dependent factors. We used this method to quantify the contributions of observer-dependent and stimulus-dependent factors to saccade direction. The method yields direction-coded maps in the shape of vector fields. It is possible to do analyses with the help of this method that could not be done before. For example, we computed the divergences of the vector fields, which are indicators for repellers and attractors (for example, targets) in

the field of view. A major advantage of our method is that it is an objective method because it does not depend on arbitrary choices. It is scale and shape independent, which makes it possible to compare datasets of different experiments, for example differing in task or stimulus.

The answer to the question how much observer-dependent and stimulus-dependent factors contribute to making a saccade in a certain direction is that stimulus-dependent factors play a more important part than observer-dependent factors. This conclusion is based on three results. The first result is that the range of values for the divergence of the direction-coded maps was larger for stimulus-dependent factors than for observer-dependent factors. The second result is that the correlation between divergence of the direction-coded maps and median search times was higher for stimulus-dependent factors than for observer-dependent factors. The third result shows that the strengths of the vector fields were higher and the range of strengths was larger for stimulus-dependent factors compared to observer-dependent factors.

It may therefore be hardly worthwhile to train observers in specific search strategies in natural search tasks. This was shown earlier by Widdel and Kaster (1981). They trained one group of observers to use a systematic strategy, and compared their search times to the search times of a group of untrained observers. In the condition of high-density displays, which has the most resemblance to search in natural stimuli, they found no difference in search times between the two groups. This result thus also suggests that observer-dependent factors play a lesser role than stimulus-dependent factors.

On the other hand, there is evidence in the literature that observer-dependent factors play a significant role in eye movement behavior (Andrews & Coppola, 1999). Andrews and Coppola showed that eye movement parameters were highly correlated between different active tasks, such as search and reading, but were uncorrelated with eye movement parameters during passive inspection of scenes. This suggests a stronger influence of observer-dependent (endogenous)

factors in active tasks than in passive tasks. Although we used only one single (active) task (namely search, which is a highly automatic task), we still found that observer-dependent factors played a lesser role than stimulus-dependent (exogenous) factors in our experiment.

The effects in the experimentally obtained saccade-direction maps that could not be attributed to observer-dependent or stimulus-dependent but rather to common factors in all stimuli and all observers were mainly explained by a random saccade generator operating in a confined viewing area. The saccade generator with two random generators, one for saccade amplitude and the other for saccade direction, showed higher resemblance to the recorded saccade direction than the saccade generator that randomly picks the next fixation location from all locations in the viewing area. This implies two possibilities. The first is that the boundary condition that a fixation must be within the viewing area is applied only *after* generating a saccade, which seems inefficient because it may happen that many saccades have to be generated until the new fixation falls within the viewing area. Picking one of the locations in the viewing area available for the next fixation seems more plausible. However, this is not supported by the results of our simulations. The second possibility is that there is a strong preference for small saccades over large saccades. The largest saccade that can occur in the stimuli that we used, which is a saccade along the diagonal of the stimulus, has different probabilities to occur in the two models that we used. Say the current fixation location is the top left corner of the viewing area. In the model with one random generator for selection of the next fixation location, the chance of the largest possible saccade is $1/(N-1)$, where N is the number of grid points. With the grids we used, this chance is $1/299$. In the model with two random generators for saccade amplitude and direction, the chance of the largest possible saccade is about $1/600000$. This is an extreme example, but the general idea is that the model with two random generators has higher probability to produce small saccades than the model with one random generator.

Therefore, the second possibility seems more likely; the comparison of observed saccade directions with simulated saccade directions of the two models suggests that observers have a strong preference for small saccades. Some extra analysis teaches us that this is indeed the case: the size of the stimuli was 33 by 45 degrees, while the average saccade amplitudes were in the range of only 5 to 8 degrees.

An important difference between the saccade-prediction methods that resulted from the work of Itti and Koch (2000), and Najemnik and Geisler (2005) and our method is that the influence of observer-dependent factors may play a more prominent role in our method. The methods of Itti and Koch (2000) and Najemnik and Geisler (2005) are both based on stimulus-dependent factors, the type of factors that in our study appeared to contribute most to saccade direction. It therefore seems likely that methods based on stimulus-dependent factors are better suited for eye movement prediction than methods that are based on observer-dependent factors processes. On the other hand, it is unlikely that observer-dependent factors can always be neglected, especially regarding the work of Andrews and Coppola (1999). Since observer-dependent factors do not play a role in the methods of Itti and Koch (2000) and Najemnik and Geisler (2005), it would be interesting to apply these two methods to a set of stimuli and compare the results to the direction-coded maps resulting from our method that is applied to a recorded data set of human eye movements in the same set of stimuli. It is even possible to combine saliency maps and direction-coded maps to obtain a more accurate map for the prediction of saccade directions. If there would be a difference between the direction-coded maps and the results of the other two methods with any set of stimuli, this difference is likely to be caused by observer-dependent factors. When observer-dependent factors indeed play a prominent role, our method provides a way to quantify this role with respect to the role of stimulus-dependent factors.

References

- Andrews, T. J., & Coppola, D. M. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Research*, 39(17), 2947-2953.
- Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, 9(1), 27-38.
- Engel, F. L. (1977). Visual conspicuity, visual search and fixation tendencies of the eye. *Vision Research*, 17(1), 95-108.
- Gbadamosi, J., & Zangemeister, W. H. (2001). Visual imagery in hemianopic patients. *Journal of Cognitive Neuroscience*, 13(7), 855-866.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40(10-12), 1489-1506.
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4(4), 219-227.
- Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, 434(7031), 387-391.
- Noton, D., & Stark, L. (1971a). Scanpaths in eye movements during pattern perception. *Science*, 171(968), 308-311.
- Noton, D., & Stark, L. (1971b). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, 11(9), 929-942.
- Over, E. A. B., Hooge, I. T. C., & Erkelens, C. J. (2006). A quantitative measure for the uniformity of fixation density: the Voronoi method. *Behavior Research Methods*, 38(2), 251-261.
- Theeuwes, J., Kramer, A. F., Hahn, S., Irwin, D. E., & Zelinsky, G. J. (1999). Influence of attentional capture on oculomotor control. *Journal of Experimental*

- Psychology: Human Perception and Performance*, 25(6), 1595-1608.
- Toet, A., Bijl, P., Kooi, F. L., & Valeton, J. M. (1998). *A high-resolution image dataset for testing search and detection models (TNO-TM-98-A020)*. Soesterberg, The Netherlands: TNO Human Factors Research Institute.
- Voronoi, G. (1908). Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième mémoire: Recherches sur les paralléloèdres primitifs. *Journal für die Reine Angewandte Mathematik*, 134, 198–287.
- Widdel, H., & Kaster, J. (1981). Eye movement measurement in the assessment and training of visual performance. In J. Moraal & K. F. Kraiss (Eds.), *Manned systems design, methods, equipment and applications* (pp. 251-270). New York: Plenum.
- Yarbus, A. (1967). *Eye movements and vision*. New York: Plenum Press.

Chapter 4. Search area edges affect visual span size, fixation density and saccade direction

Introduction

The question asked in this chapter is: Do the edges of a search area affect human visual scanning behavior?

When humans read, search or look around, they make many saccades. They make saccades to direct the fovea to potentially interesting parts of the visual environment. In this way they exploit the fact that visual acuity and color sensitivity is highest in the fovea. Sometimes this may even be necessary because optical resolution in eccentric parts of the retina is insufficient to identify small, inconspicuous targets.

In order to search efficiently, only relevant parts of the field of view should be inspected, and irrelevant parts should be ignored. Usually this is not a problem. For instance, many visual search tasks involve areas with well-defined boundaries between relevant and irrelevant parts of the field of view. Examples of bounded search tasks are search for a car on the parking lot, or search for a child in the playground. The search area boundaries in such tasks are often defined by luminance or color contrasts. This is the case when one is looking outside through a black window-frame during daytime; the window-frame then may have both color and luminance contrasts relative to the outside scene. However, this is not always the case. For example, a radiologist may look only at the right side of a Roentgen photo because the patient suffers from pain at that side of his chest. In this case there is no luminance boundary between the left and right side. Most boundaries are related to the task and the observer's knowledge about the target. When searching for a vehicle one does not scan the sky and when one looks for a chair one usually does not fixate on the ceiling.

Industrial designers often apply physical boundaries to facilitate search. Related objects are grouped within a boundary-defined region. A surrounding line may group all switches on the front panel of audio equipment. This grouping makes it easier for the user to control the equipment, and makes it unnecessary to search the whole front panel for the next switch to use. Also in traffic signs, a line often surrounds destinations in the same directions.

Edges are not always beneficial to visual search. They mark search areas but they may as well influence perception. There is an extensive amount of literature on the effect an object may have on the perception of another object. For example, it is possible to measure an eccentricity threshold for letter identification if this letter is presented in the periphery. However, when other letters are placed around the letter that is to be identified the identification is hampered and the eccentricity threshold decreases. This effect is known as lateral masking or crowding (Kooi, Toet, Tripathy, & Levi, 1994; Levi, Klein, & Aitsebaomo, 1985; Townsend, Taylor, & Brown, 1971; Wolford & Chambers, 1983). Even a letter placed at a larger eccentricity than the target's eccentricity hampers perception of the target letter. In a similar way, edges may act as flankers too and hamper perception of nearby targets.

Boundaries often mark search areas. Sometimes they are even put there intentionally to lighten the task of the searcher. It seems that, at a higher cognitive level, boundaries are able to facilitate search. A question that may arise is whether they also affect eye movement generation. Recent research shows that saccade direction relative to previous saccade direction follows a specific distribution, both in search tasks and under free viewing conditions (Hooge, Over, van Wezel, & Frens, 2005). Such a characteristic indicates that saccade direction depends on previous saccade direction, but does it also depend on the search area edge? Theoretically, the most efficient way to completely search an area would be to distribute fixation locations over the search area so that no parts are missed or checked more than once. Such a strategy

would minimize the total number of fixations needed to cover the whole area and thereby also minimize search time because search time is mainly dependent on the number of fixations, not on individual fixation durations (Gould & Carn, 1973; Megaw & Richardson, 1979). A search strategy could never be efficient if eye movements were made outside the search area. Generally, one assumes that eye movements are rarely made outside the search area. Therefore, search area boundaries are expected to have at least some effect on eye movement patterns.

Another reason that search area boundaries may affect eye movement patterns is that perception of objects near boundaries is possibly impaired by the boundaries; the boundaries thereby indirectly influence eye movement generation by degradation of target perception. For further explanation of perception being hampered near edges, see the control experiment section.

In this study we focus on the influence of luminance boundaries on human eye movement behavior. We present two search experiments. The first experiment is a visual search task in luminance-defined areas of three different shapes and random area orientations. We look at both search performance and eye movement behavior. We will introduce a model that aims to explain the findings of this search experiment. The model is then used to predict eye movement behavior in the second search experiment that has three different search area sizes and random area orientations.

Search area shape experiment

Introduction

In the first search experiment we investigated whether search area edges influence eye movement behavior, i.e. the parameters that characterize scanning patterns such as saccade direction, saccade amplitude, fixation duration and fixation density. We encouraged subjects to scan the whole search area by instructing them to continue searching until they had found

all targets. This kind of search is usually called “exhaustive”. The subjects did not know in advance the number of targets present in a stimulus; they were only told that each stimulus could contain zero to seven targets. To minimize the effect of factors other than the edges of the search area we randomized the orientation of the search areas, and we used small targets that were just slightly brighter than the search area.

As pointed out in the general introduction we argued that it is possible that a stimulus edge affects the perception of targets near the edges. We hypothesized that lateral masking by an edge may hamper target perception. Vlaskamp and Hooge (2005) showed in a search experiment with eye movement recording that when a target is laterally masked, a) search time increases, b) saccade amplitude decreases, and c) fixation duration increases. It is known from the literature that, in more difficult search tasks, saccade amplitude decreases and, at the same time, fixation duration increases (Jacobs, 1986). Similarly, in reading, more skilled readers have shorter fixation durations and larger saccade amplitude (Rayner, 1998). The effect of lateral masking can be summarized as a decrease of the visual span size.

Therefore, near the luminance edges of the search area, one may expect three possible effects: 1) increased fixation duration, 2) smaller saccade amplitude, or 3) higher fixation density, or any combination of these effects. On the other hand, when the above-mentioned eye movement parameters are about equal in all parts of the stimulus one may expect to find poorer performance in regions near luminance edges compared to other regions. Here, we define performance as the percentage of targets found during the time that is spent searching.

We were mainly interested in any evidence for differences in search performance between center and near-edge regions, and in evidence for different eye movement parameters in those two regions.

Methods

Apparatus

Two-dimensional orientation of the right eye was measured with an induction coil mounted in a scleral annulus (Skalar Medical, The Netherlands) in an AC magnetic field. This method was first described by Robinson (1963) and refined by Collewijn, Van der Mark, and Jansen (1975). An Apple Macintosh G4/400 equipped with a National Instruments PCI 1200 I/O Board presented the stimuli and recorded eye movements and manual responses simultaneously at a sampling rate of 500 Hz. The video card was fast enough to put a stimulus on the screen within one refresh-cycle of the monitor. The field generators and lock-in amplifiers have been custom built by our technicians and reach specifications of a Rimmel system or better (Rimmel, 1984). The noise level reached 0.05 degrees at most. The data were stored on the computer hard disk for off-line analysis. Subjects sat between the field coils (Rimmel Labs, USA; dimensions 1.0 m by 1.0 m by 1.0 m) in front of a SONY E400 19-inch monitor (distance 41 cm; resolution 1024 x 768 pixels; refresh rate 85 Hz; 256 gray values). Translations and rotations of the head were restricted by a one-axis bite-board.

Subjects

Five male subjects participated in this experiment (AK, EO, IH, MB and RW). None showed any visual or oculomotor pathology other than refraction anomalies. The subjects had normal or corrected to normal vision. Subjects MB and RW wore corrective glasses. Informed consent was obtained from all subjects prior to participation in the experiments.

Stimuli

The search area was a square, a triangle or a circle, all three shapes having equal area sizes (Figure 4.1). The diameter of the circle stimulus was 493 pixels on the screen corresponding to a visual angle of 23 degrees. The sides of the

triangle subtended 31 degrees of visual angle, the sides of the square 20 degrees. The search areas were gray and located in the center of a black background. The orientation of the squares and triangles was chosen randomly in each trial. The targets were 5 x 5 pixel crosses. Their lines were one pixel thick. The luminance of the black region was $<0.28 \text{ cd/m}^2$, the gray region was set at 1.29 cd/m^2 and the luminance of the targets at 2.24 cd/m^2 . For each stimulus, seven random target positions distributed over the full screen were computed. Only the targets that were located inside the gray search area were made visible, the other targets were discarded. Each stimulus therefore contained 0 to 7 visible targets. Each condition consisted of 60 trials. There were no practice trials.

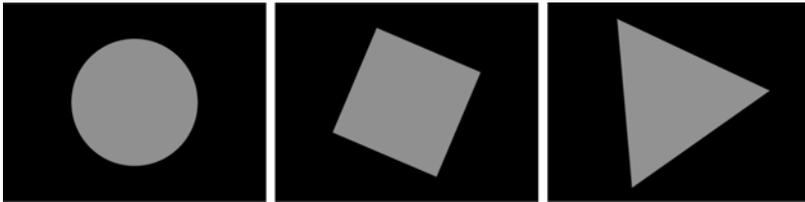


Figure 4.1. Examples of stimuli in the search area shape experiment. Search area sizes were equal and orientations were random. Targets are invisible here due to their small size and low contrast with the background of the search area.

Procedure

Experiments started with careful placement of the coil on the right eye. Recording time was limited to about 25 minutes. Prior to the experiment, a calibration screen was shown to the subject. Five dots had to be fixated successively, one in the center of the screen and the others 8 degrees shifted to the left, right, top and bottom, respectively. Trials of all three conditions were presented in a randomly mixed design. Subjects indicated the detection of each target by pressing a button with the right thumb. Each trial ended by a push on another button with the left thumb. Preceding each trial, a

fixation dot was shown in the center of the screen and the stimulus appeared when the left button was pushed.

Data analysis

The velocity signal of eye movements was searched for peak velocities above 50 deg/s. Each peak was considered a potential indicator of the presence of a saccade. The exact onset of the saccade was determined by going backward in time to the point where the velocity signal dropped below the average velocity plus two standard deviations during the stable fixation period before the saccade. The exact offset of the saccade was determined by going forward in time to the point where the velocity signal dropped below the average velocity plus two standard deviations during the stable fixation period after the saccade. This method was adopted from Van der Steen and Bruno (1995). The onsets and offsets of the saccades were used to calculate saccade direction, saccade amplitude, fixation duration and the location of fixations. We rejected fixations shorter than 20 ms and saccades with amplitudes smaller than 0.1 degrees for further analysis.

In the analyses we made a distinction between the edge region and the center region of the search areas. We did this because we wanted to assess the influence of the luminance edges on performance and eye movement behavior. If there would be any influence of the luminance edges, this influence was very likely to be the largest close to the edges. By dividing the data in two equally large regions (in terms of number of data points) we ensured that both regions had sufficient amounts of data points to perform a reliable analysis. For this purpose, the median distance of fixation locations to the edge was calculated for each combination of five subjects and three conditions (i.e. the search area shapes). The median of all these medians was 2.4 degrees, which we took as the distance from the edge separating the edge region from the center region. In the circle condition, the edge region was 36% of the total area, in the square condition it was 40% and in the triangle condition 45%. These different percentages

are due to the fact that triangles have the largest ratio of perimeter to area, and that circles have the smallest.

Results

Search performance

If the edge acts as a masking flanker that hampers perception of targets nearby we expect search performance in the edge region to be poorer than in the center region. To verify this expectation, cumulative search time curves were obtained for targets in the edge and center region separately (Figure 4.2). For each subject, the curve for the center region lies above the curve for the edge region over the whole range of the plot. This means that, after any search period, the percentage of targets found is higher in the center region than in the edge region. Each subject's performance is therefore better in the center region than in the edge region. Four out of five subjects were able to find more than 95% of the targets. Subject RW found about 75%. If targets were missed, it was more probable that they were located in the edge region than in the center region.

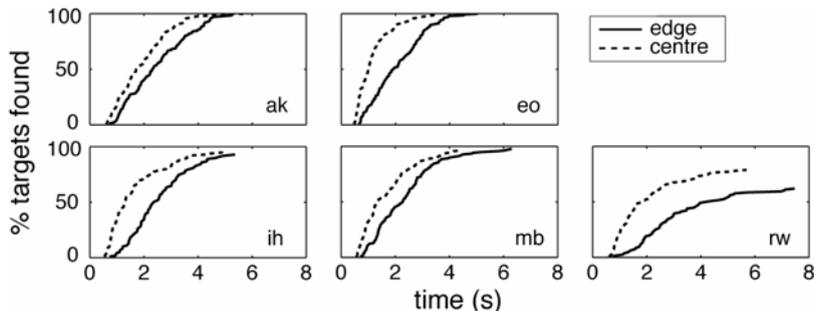


Figure 4.2. Cumulative percentage of targets found as a function of search time in circles, squares and triangles added together for the search area shape experiment. Solid lines: cumulative search time curves for targets in the edge region, dashed lines: cumulative search time curves for targets in the center region.

Fixation duration, fixation density and saccade amplitude

If the edge acts as a masking flanker for targets nearby we expect longer fixation duration, higher fixation density and smaller saccade amplitude in the edge region than in the center region.

Fixation duration was not affected by region [ANOVA, $F(1,4)=0.18$, $p=0.68$] or search area shape [ANOVA, $F(2,8)=0.01$, $p=0.99$], see Figure 4.3a. All subjects showed a significantly higher fixation density in the edge region compared to the center region [t-tests, all $T>4.7$, $p<0.001$], except for subject RW who showed the reversed effect [t-test, $T=-6.9$, $p<0.001$]. The search strategy of subject RW seemed to be inferior because he did not find as many targets as the other four subjects, especially in the edge region. Figure 4.3b shows the mean fixation density, i.e. the number of fixations divided by the total area of each region.

Surprisingly, mean saccade amplitude was larger in the edge region compared to the center region [$F(1,4)=78.25$, $p<0.001$], see Figure 4.3c. This may be explained by the simple fact that saccades can be larger in the edge region than in the center region. A saccade from one side to the opposite side cannot be matched in size by a saccade originating from the center of the search area. Search area shape did not have an effect on saccade amplitude, and interaction was not significant.

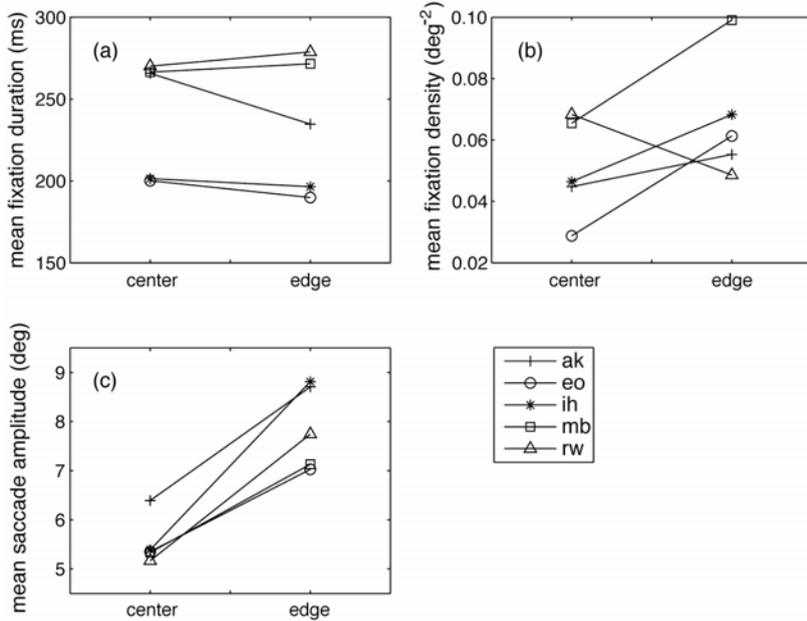


Figure 4.3. (a): Mean fixation duration, (b): mean fixation density and (c): mean saccade amplitude, all as a function of region. Fixation density: all subjects $p < 0.001$.

Saccade direction

Fixation density is generally higher in the edge region as expected when the edge acts as a masker. However, saccade amplitude is larger in the edge region than in the center region. This may imply that the edge region is not scanned with consecutive fixations, or more often than just once. If it is scanned more often, the majority of saccades are likely to be directed along the search area edges. The analysis below addresses the question how saccades are directed within the search area.

Figure 4.4 shows polar plots of the relative frequency of saccade directions with regard to the horizon. In all the polar plots presented in this chapter, histograms of relative frequency are binned in 24 directions of 15 degrees. The outer circles indicate a relative frequency of 10%. In all subjects, except AK, there were no distinct preferred directions with

respect to the horizon. Subject AK seemed to prefer making horizontal eye movements.

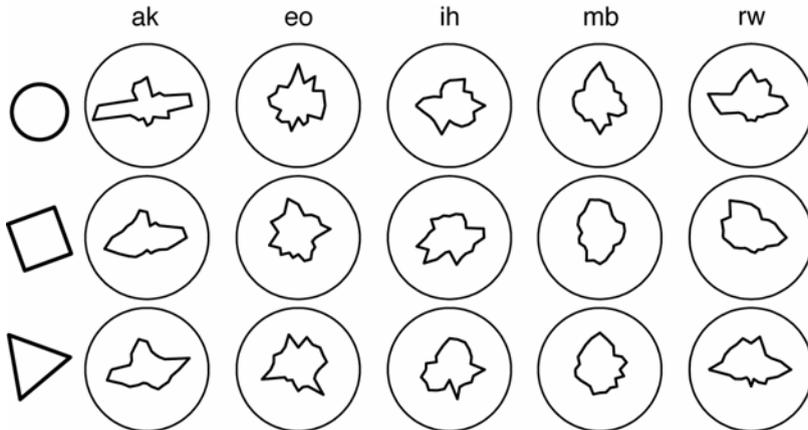


Figure 4.4. Histograms of saccade directions relative to the horizon. Outer circles represent a relative frequency of 10%. Top to bottom: histograms for circle, square, and triangle condition. Bin size: 15 degrees.

While monitoring eye movements from subjects doing the experiment, it appeared that many eye movements were made along the sides of the three search area shapes. This suggests that the influence of the horizon is less important than the orientation of the search area. For purpose of analysis, we determined saccade direction with respect to each of the sides of the search areas. We determined saccade direction with respect to each of the sides because it is not known a priori which of the sides is taken as a reference side. We therefore determined four distributions of saccade direction for the square condition (a square has four sides) and three for the triangle condition (a triangle has three sides). These three or four distributions of saccade direction were then averaged. The results are shown in Figure 4.5.

The circle stimulus has no sides, or one could say an infinite number of sides. It is unlikely that there should be any preference for making a saccade in a particular direction. Therefore, we expect the saccade direction to be

homogeneously distributed. This means that the graph of the relative frequency of the saccade direction in Figure 4.4 would be a circle for the circular search area condition. If the search area is a square, we expect four saccade directions to occur more often than other directions. These four directions consist of two pairs of two parallel sides of the search area that can be followed back and forth. Besides, these directions should be perpendicular to each other. Thus, for the square search area condition the graph of the relative frequency of the saccade direction should have peaks to the right, left, bottom and top. The directions relative to one side of the search area are 0, 90, 180 and 270 degrees. However, from these cross-shaped plots in Figure 4.5 one cannot tell whether saccades are made along the sides or perpendicular to them. The triangular search area condition rules out the hypothesis of saccades being made perpendicular to the sides of the search area. A triangle has no sides perpendicular to each other, so saccades that are made perpendicular to one side are not mixed up with saccades made along another side in the direction histograms. When the majority of saccades would have been made perpendicular to the sides, the peaks in the distribution plot for the triangular search area condition would be at $30 + n \cdot 60$ degrees, $n = [0 \dots 5]$. In fact, the peaks are at $0 + n \cdot 60$ degrees. These directions relate to the three non-parallel sides of the triangle that can be followed in two directions.

The plots seem to represent a kind of signature. The square and triangle condition yield patterns of typical forms, but there are also considerable differences between subjects. For example, some subjects have larger portions of their saccades directed along the edges than others, as is clear from the fact that the peaks in their histograms are larger.

Another feature that can be obtained from the graphs for the triangle condition is the relation between clockwise and counterclockwise scanning. When the six peaks are equally large, clockwise scanning occurred about as often as counterclockwise scanning. This is the case for all subjects. It is not possible to distinguish clockwise from counterclockwise scanning in the square condition because the meaning of each

peak in these plots is ambiguous: it consists of clockwise scanning along one side and counterclockwise scanning along the parallel, opposite side.

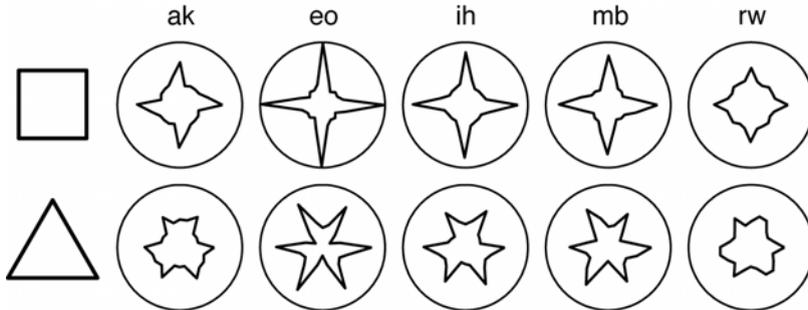


Figure 4.5. Histograms of saccade directions relative to the orientation of the search area. Top row: histograms for the square condition, bottom row: histograms for the triangle condition. Outer circles represent a relative frequency of 10%. Bin size: 15 degrees.

Discussion

We obtained five primary results in this experiment. First of all, search performance was better in the center than in the edge region. Second, fixation duration was not significantly different in edge and center regions. Fixation density was higher in the edge region than in the center region. Counterintuitively, saccade amplitude was larger in the edge region. Lastly, and most strikingly, saccade direction was mainly along the search area edges.

We showed that search area edges affect search performance and eye movement parameters. Usually, higher fixation densities come with smaller saccade amplitudes (Jacobs, 1986). If we understand the influence of edges in visual search we should be able to make a model that mimics the human eye movement behavior recorded in the search area shape experiment. In order to prove that the edge of a search area has a direct influence on eye movement parameters that cannot be neglected we first test models that do not take into

account the hampering effect of luminance edges on perception. In the next section we consider two such models and show that they are inadequate to reproduce human eye movement behavior.

Model 1: Independent or dependent fixations

In this section we are going to build a model starting with the simplest model possible and adding features until the model is able to satisfactorily reproduce human saccade direction distributions.

The simplest model for visual search is the *random search* model (see e.g. Horowitz & Wolfe, 1998), which we will here call the independent fixation model. The independent fixation model's only assumption is that saccade targets are within the search area. The entire search area therefore has uniform fixation probability. There are no requirements for saccade amplitude or direction. Saccade amplitude as well as saccade direction are therefore random, and as a result fixations are also fully independent. Scanning the area in this way is known as sampling with replacement. To simulate sampling with replacement we simply added up all possible saccade directions from each location in the search area with equal weight factors. In this way, the theoretical saccade direction distribution can be computed exactly without doing an infinite number of simulation trials.

Is the independent fixation model sufficient to adequately reproduce saccade directions from the visual search experiment with differently shaped search areas? In other words, is the saccade direction we observed in the search experiment distributed in this particular way due to the geometry of the search area only? Figure 4.6 shows the saccade direction distributions produced by the independent fixation model. For the circle condition, the distribution is uniform. The other two distributions are non-uniform and symmetrical. They are not as peaked as the distributions of the human data, but they are different for different search area

shapes. For the circular and triangular search area a pattern emerges that is similar to the saccade direction distribution of human observers. The squared search area however yields a different pattern. And although the triangular search area yields a similar pattern, the saccade direction distribution is not as peaked as for the human observers. Moreover, fixation densities in the edge and center regions are equal in the simulated data, because fixation probability is identical in edge and center regions. This is also incompatible with the human data from the search area shape experiment, where we found different fixation densities in the edge and center regions. The independent fixation model explains partly why saccades are directed along the search edges: it is an inevitable consequence of the shape of the search areas. However, since the peaks in the saccade direction histograms for the independent fixation model are not as peaked as in the recorded saccade direction histograms, the model needs to be extended.

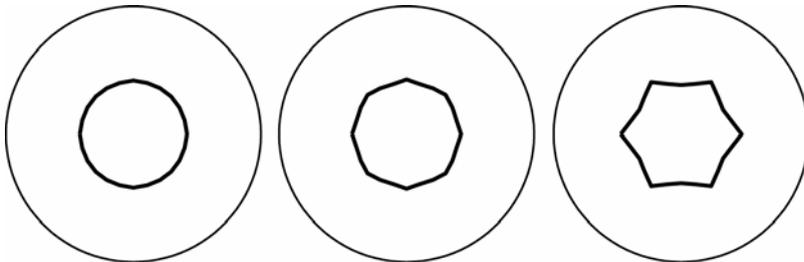


Figure 4.6. Histograms of saccade directions produced by the independent fixation model. From left to right: histograms for circle, square, and triangle condition. Outer circles represent a relative frequency of 10%. Bin size: 15 degrees.

It is not surprising that an independent fixation model cannot perfectly replicate the human visual search data since it basically represents a visual search system that either does not track at all what parts of the search area have been searched or analyzes an infinitely small area during each fixation. Both properties are unrealistic. A more realistic approach would be

sampling without replacement (Gilchrist & Harvey, 2000; Hooge, Stapelkamp, Over, Vlaskamp, & Frens, 2004). To model sampling without replacement it is very useful to introduce the concept of *visual span*. Here, we define the visual span as the area within which a subject can detect a target of unknown position with a certain level of accuracy, in a single fixation (see also Geisler & Chou, 1995).

To make a model for exhaustive search and sampling without replacement, we added the concept of visual span to the random search model and applied the rule that spans are not allowed to overlap each other. This dependent fixation model aims to generate saccades until the whole search area is inspected exactly once. That means that no parts are missed and no parts are fixated more than once. Each trial consists of a certain number of fixations that are needed to search the whole area. The number of fixations depends on the size of the area that can be checked for the presence of a target during a single fixation, the visual span. Thus, a small visual span requires a large number of fixations and, conversely, a large visual span only requires few fixations.

From the literature we know that the size of the visual span in real search tasks mainly depends on idiosyncratic parameters and stimulus properties. For example, the visual span of persons with low vision (e.g. maculopathy, cataract, myopic retinal degeneration) is smaller than the visual span of persons with normal vision (Bullimore & Bailey, 1995; Legge, Ahn, Klitz, & Luebker, 1997). Visual span is also larger for conspicuous stimuli (Legge et al., 1997; Mohammed & Dickinson, 2000; Näsänen, Ojanpää, & Kojo, 2001). In the literature, several ways are used to measure the size of the visual span, varying from a measure based on the mean number of fixation per search trial (Ojanpää, Näsänen, & Kojo, 2002), one based on the number of letters identifiable within a limited time period (Jacobs, 1986), a measure based on the moving window technique (Pomplun, Reingold, & Shen, 2001; Rayner, 1986; Reingold, Charness, Pomplun, & Stampe, 2001), to one based on the number of letters read per second (Legge et al., 1997). We ran the model for several

(realistic) span sizes, and compared the human eye movement data to the simulated data of the best fitting span size. Considered span diameters are 16, 24, 32, 40, 48, 56, and 64% of the diameter of the circular search area in the search area shape experiment. Of course it is possible to extend the simulations with additional span diameters, or even do simulations on a (quasi-) continuous span size scale, but these seven span sizes here suffice to support the conclusions that we want to draw later.

Now that new fixation locations are dependent on the location of previous fixations, because already fixated areas cannot be fixated again, it is impossible to compute a theoretical saccade direction distribution as with the independent fixation model. With the dependent fixation model it is necessary to actually perform the simulation of eye movements. We used the dependent fixation model to simulate at least 5500 saccades for each span size. A simulation trial started with a fixation around the center of the search area, just like the trials the human subjects conducted, and it ended when the new fixation location was unsuccessfully generated 100 times. Generation of a new fixation failed when the new location was outside the search area or when its corresponding visual span overlapped with any of the previously generated visual spans. When the coverage of the area by all generated spans exceeded 80% the model automatically ended the trial because a new span could not be generated without overlap with previously generated spans.

The results for saccade direction distribution of the dependent fixation model are shown in Figure 4.7. Up to a span diameter of 40%, the peaks in the directions of the search area edges are less pronounced than the peaks in the human data plots. Span diameters larger than 40% yield more distinct peaks, but they are not in the direction of the search area edges. In the square condition, the peaks are in the direction of the diagonals of the square, and in the triangle condition they are in the direction of vectors from the center to the vertices of

the triangle. These peaks do not occur in the polar plots of the human data.

It is now possible to determine the most suitable values for the span size parameter per subject by a measure of least squared difference between human and simulated data. The best fit is accomplished by the model parameter for span size that yields the least total of the sum of squared differences in the triangle condition and the sum of squared differences in the square condition. Table 4.1 shows the best fitting parameters.



Figure 4.7. Histograms of saccade directions produced by the dependent fixation model. Top row: histograms for the square condition, bottom row: histograms for the triangle condition. Numbers above the columns denote the size of the visual span in percentages of the diameter of the circle search area. Outer circles represent a relative frequency of 10%. Bin size: 15 degrees.

Table 4.1. Fitting results for the dependent fixation model with constant span size. Best fitting span size of the seven considered sizes, in percentages of the diameter of the circle area in the search experiment. Sum of squared differences of triangle and square condition added.

	ak	eo	ih	mb	rw
Best fitting span size	40%	40%	40%	40%	24%
Sum of squared differences ($\cdot 10^{-3}$)	1.78	11.13	4.35	4.08	0.78

The dependent fixation model seems to fit the human data better than the independent fixation model, but some inaccuracies remain. First of all, the saccade direction distributions cannot be made as peaked as some of the human

data distributions, for none of the span sizes considered here. Second, fixation density in these simulations are equal in the center and edge regions because it solely depends on the size of the visual span. In the human data we found that fixation density was higher in the edge region, and because search was exhaustive it is plausible to assume that the human visual span is smaller in the edge region than in the center region. This may also explain why human performance was poorer in the edge region than in the center region. Before we built this feature into the model we did a visibility control experiment to estimate the ratio of span sizes in center and edge region.

Control experiment: visibility of a target near a luminance edge

Introduction

Target resolvability decreases when other items flank the target item (Bouma, 1970; Toet & Levi, 1992). This phenomenon is known as crowding (or lateral masking). The effect of crowding increases with retinal eccentricity of the target. The distance between target and distracters at which crowding occurs is surprisingly large. Distracters that are separated from the target as far as 0.5 times target eccentricity still evoke crowding, even when the distracter is at a larger eccentricity than the target itself (Bouma, 1970). Vlaskamp and Hooge (2005) showed that, as a result of lateral masking, also in saccadic search the target was less conspicuous (indicated by the eccentricity at which the target was still/became visible) and that the visual span was smaller. The following psychophysical experiment was done to establish whether a luminance edge acts as a mask as well. If it does, visibility eccentricity should be smaller in the vicinity of such an edge.

Methods

The displays (1024 x 768 pixels) consisted of two regions: the left side of the screen was black (281 pixels wide) and the right side was gray (743 pixels wide), see Figure 4.8. A 5 x 5-pixel target cross was presented in all trials. Its lines were one pixel thick. It was located in the gray region at one of twelve horizontal distances from the luminance boundary, and vertically at the center of the screen. The luminance values of the black region, gray region and targets were identical to those used in the search area shape experiment: $<0.28 \text{ cd/m}^2$, 1.29 cd/m^2 and 2.24 cd/m^2 . An Apple Macintosh G4/400 presented the stimuli on a LaCie electron blue 22-inch monitor placed at 48 cm distance from the subjects. A fixation dot (a single white pixel, luminance 16.15 cd/m^2) was present in each trial. It could be located either at the far right of the screen or just to the right of the target cross. The subject moved the fixation dot horizontally with the left/right arrow keys. In case of the fixation dot being close to the target cross, the task was to fixate the dot while moving it to the right until the target cross was just not visible anymore. In case of the fixation dot being at the far right of the screen, the task was to fixate the dot while moving it to the left until the target cross became just visible. At these locations, pushing the down key ended the trial and the eccentricities of these locations to the target were recorded. Bouma (1970) described this method to measure visibility eccentricity. In our visibility experiment, such a method would measure the visibility area rather than the conspicuity area (Engel, 1971), because the location of the target is known when moving the fixation point away from the target. When moving the fixation location towards the target, the exact location of the target is unknown, but it is known that the target is located somewhere in between the fixation point and the luminance edge. Thirteen subjects participated in this experiment. All subjects had normal or corrected to normal vision. Each subject performed 24 trials (12 target-edge distances and two starting points of the fixation dot). The experiment was done in a darkened room, and it started with a

black screen for 15 seconds to allow subjects to get used to the darkness.

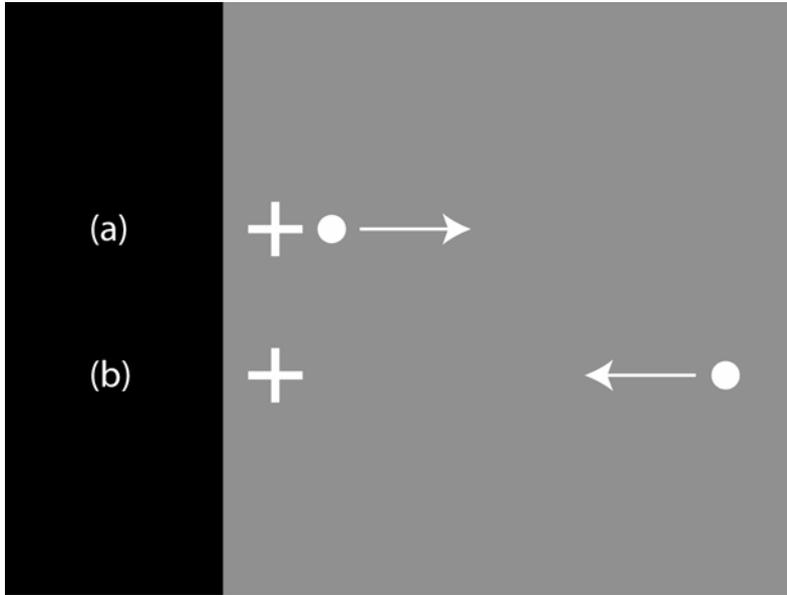


Figure 4.8. Schematic representation of a stimulus in the target visibility experiment. Arrows (not present in the experiment) indicate in which direction the fixation dot had to be moved. Starting position of the fixation dot was either close to the target (a) or at the end of the screen (b). The target (cross) was located at varying distances from the luminance edge. Contrast and object sizes are increased for reasons of clarity. Target and fixation dot were located halfway the height of the screen.

Results

Figure 4.9 shows the mean visibility eccentricity versus target distance to luminance edge of thirteen subjects. Visibility eccentricity increases with distance between target and luminance edge. This increase is significant [ANOVA, $F(1,12)=61.603$, $p<0.001$]. When the distance between the target and the edge exceeds 2.4 degrees, visibility eccentricity reaches an asymptote. At very short distances, less than 0.5

degrees, the visibility eccentricity is about 55% of the asymptotic visibility eccentricity.

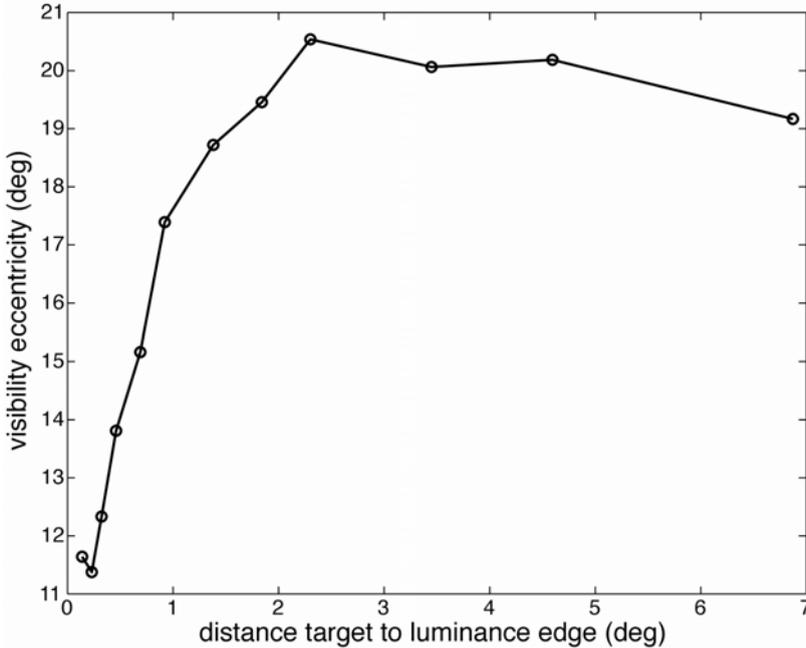


Figure 4.9. Mean visibility eccentricity as a function of distance between target and luminance edge.

Discussion

The main result from this visibility control experiment is that near a luminance edge a target is visible up to a smaller eccentricity than when placed against a homogeneous background. Consequently, the likelihood of a target to be detected is less near a luminance border than when placed in a homogeneous background. To incorporate this psychophysical result into our search model, it has to be interpreted in terms of a visual system property. The interpretation that we give here is that the visibility eccentricity is related to the size of the visual span. In the next section we add the feature of smaller visual spans near a luminance edge to the model.

Model 2: visual span ratio 1:2

We know from the visibility control experiment that visibility eccentricity is smaller for targets near the edges. From the search area shape experiment we know that search performance near the edges is below the performance in the center region. This suggests that in the edge region during each fixation a smaller area is analyzed than in the center region. We therefore built in the model that visual span diameter in the edge region is half of the diameter in the center region, roughly corresponding to the factor 0.55 that we found in the visibility control experiment. The edge region was identical to the one in the search area shape experiment, namely extending to 2.4 degrees from the edge. Other conditions were identical to those for the dependent fixation model with constant span size. An example of a simulation trial is shown in Figure 4.10.

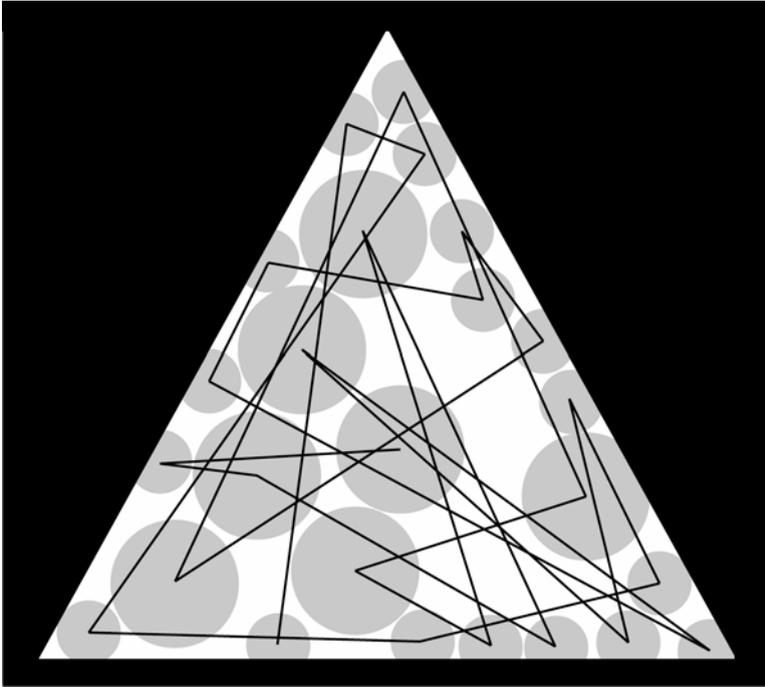


Figure 4.10. Example of a simulation performed by the variable span size model for the triangle condition. Span diameter in the edge region is half the diameter in the center region. Gray circles represent visual spans and black lines represent saccades from one fixation location to the next.

Simulated saccade directions were determined for the seven span sizes just as we did for the human eye movement data before. In Figure 4.11 the simulated saccade direction data are represented in similar graphs as the human eye movement data. It is clear that, for suitable span sizes, the patterns are more similar to the human saccade direction distributions (Figure 4.5) than the dependent fixation model with constant span size (Figure 4.7).

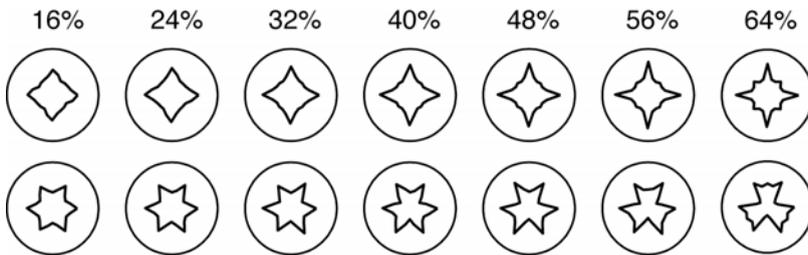


Figure 4.11. Histograms of simulated saccade directions when span size is smaller near the edges. Top row: histograms for the square condition, bottom row: histograms for the triangle condition. Numbers above the columns denote the size of the visual span in percentages of the diameter of the circle search area in the search area shape experiment. Outer circles represent a relative frequency of 10%. Bin size: 15 degrees.

Table 4.2. Fitting results for the variable span size model. Best fitting span size of the seven considered sizes, in percentages of the diameter of the circle area in the search area shape experiment. Sum of squared differences of triangle and square condition added.

	ak	eo	ih	mb	rw
Best fitting span size	56%	56%	56%	48%	16%
Sum of squared differences (*10 ⁻³)	1.13	5.57	1.30	1.29	1.00

We fitted model data to the human data (Table 4.2), similarly as we did before for the dependent fixation model with constant span size. Fits of the variable span size model to the human data are better than for the dependent fixation model with constant span size. This is also proven by the fact that the sum of the squared differences is higher for the dependent fixation model with constant span size than for the model with variable span size for four out of five subjects (compare Tables 4.1 and 4.2). This means that the variable span size model with smaller span size near the search area edges yields better fits than the dependent fixation model with constant span size. Moreover, the two models yield different fixation densities in the edge region. In the fixed-span simulation the fixation density in the edge region is equal to the density in the center region. In the simulations with the variable span size model fixation density near the edges is

higher (see also Figure 4.10), as is also the case in the human data. This is additional evidence that reducing span size near the edges is an essential part of the model. In general, the model with edge-center span size ratio 1:2 can reproduce the pattern in the saccade direction of human data fairly well, but it still cannot replicate it exactly for all subjects. The model reproduces the most important aspects of human saccade direction distributions since the differences between human data and model data are typically small.

Our model is not perfect but it should be good enough to make predictions for other experiments. If the components of the model, namely random search using sampling without replacement and variable span size, are the major mechanisms underlying eye movement patterns during visual search in restricted search areas, the model should be able to predict saccade direction distributions in experimental conditions where the influence of edges is varied. Intuition tells us that the influence of edges in small search areas is bigger than in large search areas. In large search areas edges become less influential whereas in very small areas the entire area may be considered the edge region. Figure 4.12 shows the predictions for three triangular area sizes that we will use in the last search experiment. We used the span size that seems most representative for all subjects, namely the span size that most often yielded the best fit (see Table 4.2). This span size had a diameter that is 56% of the diameter of the circular search area in the search area shape experiment. It is clear that for constant span size, larger search areas yield fewer saccades in the direction of the search area edges. Smaller search areas also evoke saccades in the directions of the center to each of the vertices. Based on the outcome of the model, we expect to find fewer saccades in the directions parallel to the search area edges in larger areas, and more saccades along the edges in smaller areas. In addition, in smaller areas there may also be more saccades in the directions from the center to one of the vertices.

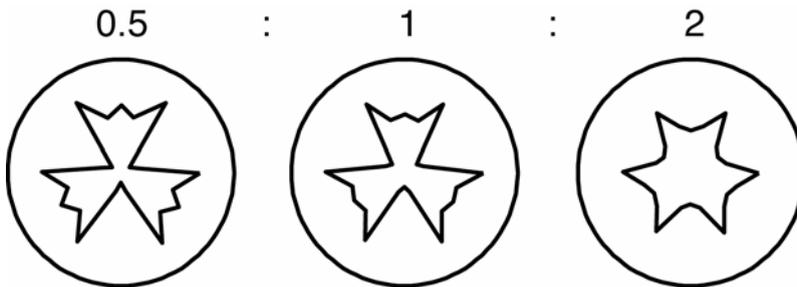


Figure 4.12. Predictions for saccade direction distribution for 56% visual span size in differently sized triangle search areas. Left to right: distributions for small to large search areas. The smallest area size was half the size of the original triangle in the area shape experiment, the medium triangle was identical to the original triangle and the largest triangle was twice as large as the original triangle. Outer circles represent a relative frequency of 10%. Bin size: 15 degrees.

Search area size experiment

Introduction

In this experiment we varied the size of the triangular search area, and compared the human saccade direction distributions to the distributions as predicted by the model.

Methods

Eye movements during this experiment were recorded with an Eyelink system at 250 Hz. Stimuli were presented on a 22-inch LaCie monitor at 85 Hz in a darkened room. Five male subjects participated (BV, EO, IH, RS and TK). None showed any visual or oculomotor pathology other than refraction anomalies. The subjects had normal vision. Subjects EO and IH also participated in the search area shape experiment; the other subjects were naïve as to the goal of the experiment. Informed consent was obtained from all subjects prior to participation in the experiments.

The search area was a triangle, whose orientation and size were randomized. With a screen distance of 31 cm, the sides of the triangle subtended 22, 31 or 44 degrees of visual

angle. The medium sized triangle with sides of 31 degrees was identical to the one used in the search area shape experiment. There were 120 trials for each area size.

Prior to the experiment, the Eyelink system was adjusted and a calibration screen was shown to the subject. Nine dots had to be fixated in random order, one in the center of the screen and the others in the corners of the screen and halfway between two corners. Preceding each trial, a fixation dot was shown in the center of the screen and the stimulus appeared when the space bar was pushed. Subjects indicated the detection of each target by pressing the right arrow button. Mean target density in the three conditions was equal to the mean density in the search area shape experiment. Subjects did not know the number of targets in each stimulus in advance. The trial ended by pushing the left arrow button. Other experimental circumstances were identical to the search area shape experiment. Data analysis was also as in the search area shape experiment.

Results

Although we changed the size of the search area, we applied the same distance to the luminance edge as in the previous experiment as the separation between edge region and center region. This is based on the findings of the target visibility experiment; there we found an eccentricity of 2.4 degrees to the edge wherein visual detection was reduced. The edge regions for the small, medium and large triangles were 57%, 45% and 34% of the total search area, respectively. Target size (in degrees) in this experiment was identical to target size in the search area shape experiment.

Search performance

As in the search area shape experiment, the cumulative search time curves for targets in the edge and center region are plotted in Figure 4.13. For each subject, the curve for the center region again lies above the curve for the edge region over the whole range of the plot. Therefore also in this search experiment, performance in the edge region is below

performance in the center region. Four out of five subjects were able to find more than 95% of the targets. Subject BV found about 80%. If targets were missed, it was more probable that they were located in the edge region than in the center region.

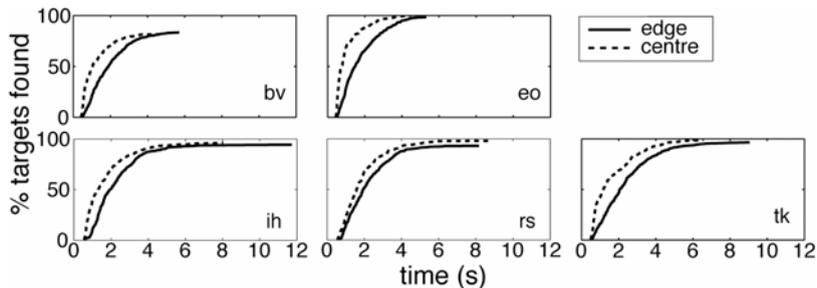


Figure 4.13. Cumulative percentage of targets found as a function of search time in triangles of all sizes added together for the search area size experiment. Solid lines: cumulative search time curves for targets in the edge region, dashed lines: cumulative search time curves for targets in the center region.

Fixation duration, fixation density and saccade amplitude

As in the previous search experiment, there was no significant effect of region on fixation duration [ANOVA, $F(1,4)=0.38$, $p=0.57$]. Neither was fixation duration affected by search area size [ANOVA, $F(2,8)=2.48$, $p=0.14$]. Three subjects (EO, IH and RS) showed a significantly higher fixation density in the edge region compared to the center region [t-tests, all $T>3.5$, $p<0.001$]. Subject TK showed no effect [t-test, $T=-1.6$, $p=0.11$], and subject BV showed the reversed effect [t-test, $T=-2.1$, $p=0.03$]. Figure 4.14a shows the mean fixation duration and Figure 4.14b the mean fixation density, i.e. the number of fixations divided by the total search area, both as a function of region. As in the previous search experiment, the subject (in this experiment BV) who had higher fixation density in the center region than in the edge

region could not find as many targets as the other subjects. It seems that BV underestimated the difficulty of finding targets in the edge region, like subject RW did in the search area shape experiment.

Mean saccade amplitude was larger in larger search areas [ANOVA, $F(2,8)=16.24$, $p<0.01$] and was larger in the edge region compared to the center region [ANOVA, $F(1,4)=49.67$, $p<0.01$], see Figure 4.14c. The last result might be explained by the fact that the maximum possible saccade amplitude influences the mean saccade amplitude. Because it is possible to make larger saccades in the edge region than in the center region, it is in the edge region that larger mean saccade amplitudes are found.

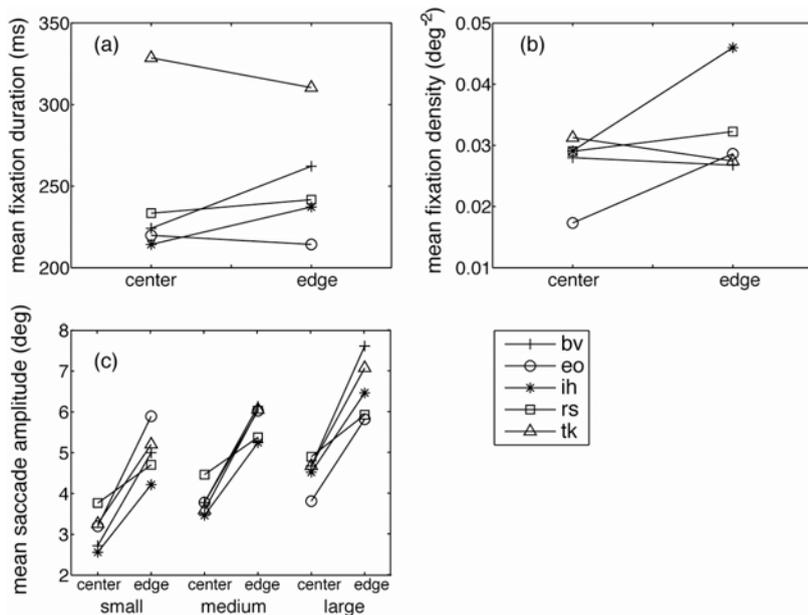


Figure 4.14. (a) Mean fixation duration and (b) mean fixation density, both as a function of region. (c) Mean saccade amplitude as a function of region and area size.

Saccade direction

Plots for saccade direction corrected for stimulus orientation (analogous to Figure 4.5) are shown in Figure 4.15. The distributions of saccade directions show an idiosyncratic pattern. It is as if they are some kind of signature. The fact that subject RS preferred clockwise scanning whereas TK preferred counterclockwise scanning shows this most clearly.

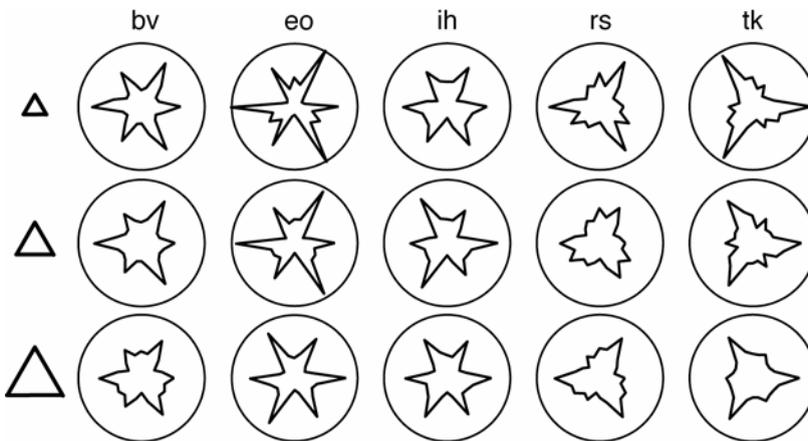


Figure 4.15. Saccade directions relative to the orientation of the search area, for small (top row), medium and large (bottom row) triangle search areas. Outer circles represent a relative frequency of 10%. Bin size: 15 degrees.

The measure of squared differences can now also be used to quantify how peaked the distributions of saccade directions are. Theoretically, the distribution that has the smallest peaks (namely no peaks at all) is the circular uniform distribution. Thus, by computing the squared differences between the uniform distribution and the measured distributions of saccade directions, a measure of preference for directions along the edges is obtained. Figure 4.16 shows the results for this analysis for individual subjects. Analysis of variance teaches that saccade direction is generally more along

the search area edges when the area is smaller [ANOVA, $F(2,8)=6.93, p<0.05$].

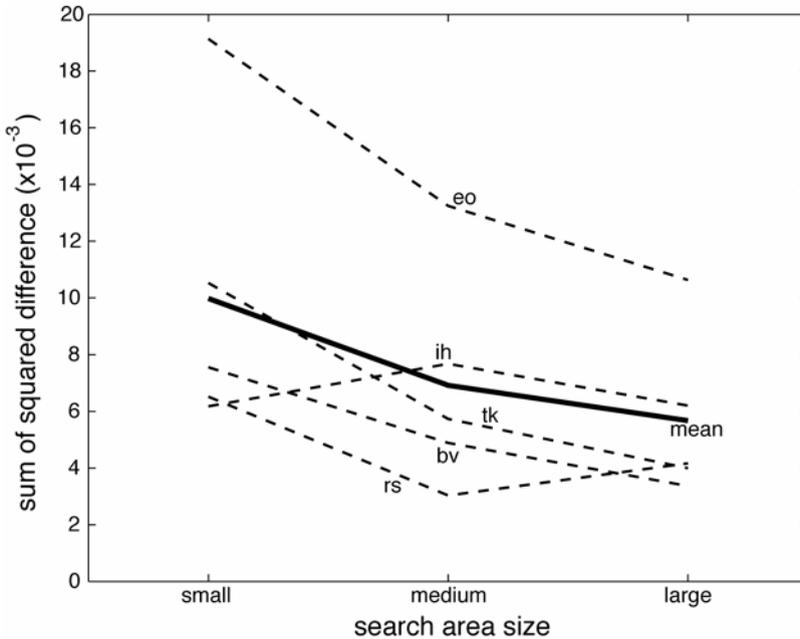


Figure 4.16. Sum of squared difference for saccade direction distributions compared to circular uniform distribution.

Discussion

The model predicted less distinct peaks in the saccade direction distributions for larger search areas. It is clear that the saccade direction distributions for the experimental data have smaller peaks in the direction along the search area edges when the area is larger (Figures 4.15 and 4.16). The prediction of the model was therefore confirmed. It may be concluded that the model incorporates the essential modules to mimic the common aspects of human eye movement behavior for visual search in areas of different shape and size.

Two subjects (EO and IH) who participated in the search area shape experiment also participated in the search

area size experiment. The medium triangle in area size experiment was identical to the triangle in the area shape experiment. One would therefore expect that for the medium triangle the saccade direction distributions in Figure 4.15 are (nearly) identical to the saccade direction distributions for the triangle condition in Figure 4.5. This is not the case. In the search area size experiment, they both seem to have a preference for clockwise or counterclockwise scanning. If this preference would not have been there (take the average of the plot in Figure 4.15 and the same plot mirrored in the vertical axis), the distributions in both experiments would have been much more similar.

General discussion

This study consists of several components. We performed two visual search experiments, one psychophysical experiment, and we constructed a model for two-dimensional visual search. The model, based on dependent fixations and variable visual span size, describes a possible mechanism underlying the eye movement behavior observed in the two search experiments. The psychophysical experiment was performed to find justification for the use of variable span size in the model.

In the search area shape experiment we found that search performance was poorer near the edges, although fixation density was higher there. At any time after the onset of a trial, on average more targets in the center of the search area were found than near the edges. Additionally, more targets were missed in the edge region than in the center region. The psychophysical visibility experiment provided evidence that the visual span was smaller near luminance edges. This fact explains why fixation density was higher in the edge regions. Also, saccade directions were mainly in the directions of the edges in the search area shape experiment. We built a model that can satisfactorily reproduce human saccade direction distributions and used the model to predict human saccade direction distributions in conditions where the

influence of the edges was varied. The prediction was that in large search areas saccade direction would be less often along the search area edges than in small search areas. This prediction was confirmed, which is an indication that the final model incorporates the essential components that determine eye movement generation in differently shaped and sized areas.

The basis of the model is that saccade direction is picked randomly from a uniform distribution, and that fixations must be within the search area. Simulated saccade direction distributions in differently shaped areas showed that directions along the search area edges occurred slightly more often than other directions, and in the square search area there was also a slight preference for the diagonal directions. Fixation densities were uniform, which was in disagreement with the results of the search experiment, and saccade direction distributions produced by the model were not as peaked as those found in the experimental data. Therefore, a new assumption was added, namely that the new fixation location is dependent of the previous fixation locations. This is a reasonable assumption because this kind of behavior is frequently encountered (Gilchrist & Harvey, 2000; Hooge et al., 2004). In the model we introduced the visual span, which should not overlap with previous spans. The preference for directions along the edges became stronger, but still not as strong as recorded from human subjects. Also, fixation densities in edge and center region were still uniform. The last modification of the model was that span size depends on the distance between fixation location and the nearest edge. If this distance is small, the size of the visual span should be small too. After including this feature, model fixation densities were in agreement with human fixation densities (higher near the edges) and the preference for saccades in the direction of the search area edges became even stronger, and therefore more like the preference observed for humans. The model in its final form was able to accurately predict human eye movement behavior in a search experiment where the influence of the edge was manipulated. This proves that the

model can explain general eye movement behavior during search in restricted areas, with a minimum of required mechanisms.

We chose to keep the model as simple as possible, but it is easy to improve the model's performance. A gradual decrease of visual span size towards the edges, not necessarily limited to the factor of two, seems more realistic. Also, it is likely that at least some overlap occurs between spans, because otherwise it is impossible to scan the entire search area. If one desires to reproduce individual scanning behavior exactly, it is also inevitable to introduce some biases, for example for clockwise or counterclockwise scanning.

How do fixation duration and saccade amplitude relate to stimulus content? Moffitt (1980) differentiates dense and sparse stimuli. In dense stimuli, both saccade amplitude and fixation duration may be adjusted to achieve optimal performance. In sparse stimuli, saccade amplitude is determined by the inter element distances, so only fixation duration is adjustable. Question remains, what is a sparse stimulus and what is a dense one? Jacobs (1986) found that saccade amplitude increased and fixation duration decreased with target conspicuity. In a recent study, Vlaskamp and colleagues (Vlaskamp, Over, & Hooge, 2005) found that higher target conspicuity mainly resulted in larger saccade amplitudes, and at the same time slightly shorter fixation durations. Hooge and Erkelens (1996) found that fixation duration increased when target-distracter similarity was higher. All these studies have in common that when the target is less conspicuous, saccade amplitude decreases, fixation duration increases, or a combination of these two effects occurs. Here, we manipulated local conspicuity of the target but found no difference in fixation duration, and larger mean amplitude in the region where the targets are less conspicuous. On the basis of the literature cited above, we would have expected longer fixation durations, smaller saccade amplitudes, or a combination of these effects. A possible reason for this difference is that our stimuli did not consist of discrete objects that could be fixated separately, but instead

had a homogeneous background. This difference also does not mean that the present study is inconsistent with the literature. The consistency appears from the visual span size: a smaller visual span near the edges is necessary to explain three observed features: 1) the higher fixation density in the edge region, 2) worse search performance in the edge region, and 3) the preference for saccade directions along the edges.

From the literature it appears that saccade amplitude is related to the size of the visual span. In one-dimensional search tasks and reading, saccade amplitude can be used safely as an estimator for span size (see e.g. Jacobs, 1986). The assumption is that search proceeds in one direction and that no parts of the stimulus are skipped or checked double. Jacobs (1986) found that 80% of the variance of saccade amplitude is explained by the size of the visual span. Also in two-dimensional search, saccade amplitude is positively correlated to the size of the visual span. Good examples are studies that use the moving window technique (e.g. Pomplun et al., 2001; Rayner, 1986; Reingold et al., 2001). In the present study however, saccade amplitude is not a suitable estimator for visual span size. We found that mean saccade amplitude was larger in the edge region compared to the center region, and this is not in accordance with the finding that visual span is smaller near the edges. This means that search probably does not proceed in one direction in our experiments. The reason for larger mean saccade amplitude in the edge region is that, in the edge region, the maximum possible saccade amplitude is much larger than the maximum possible amplitude in the center region. In addition to saccade amplitude and fixation duration, we therefore propose a third correlate for the conspicuity of targets: fixation density. A requirement for proper functioning of this measure is that search is exhaustive.

On the one hand, edges have a global beneficial property for visual search: they mark the area wherein the target is to be found. For example, a bookcase affords structure to the arrangement of the numerous books in it. On the other hand, edges also have a local detrimental property, namely the masking effect they have on targets nearby. In

experiments with stimuli presented on a monitor, the effect of the monitor housing should not be ignored a priori. For example, in a study by Bijl and colleagues (Bijl, Kooi, & van Dorresteijn, 1997) it is shown that search times for single targets are generally longer when the target is nearer to a monitor edge. In experiments that use grids, such as those by Hooge and Erkelens (1996; 1998), the elements at the outer grid points are in fact themselves the edge of the grid, implying that analyses should be performed separately for these elements and elements in the center of the grid.

It would be nice to make use of the global beneficial property, and to minimize the local detrimental property. In the search area size experiment, target perception in the large areas suffers less from the negative influence of the edges because the relative influence of the edges is less than in the smaller areas. However, in industrial design for example it may be inconvenient to just scale whatever product. It may perhaps make more sense not to use as extreme contrast edges as the luminance edges in our experiments. A moderate edge may for instance be realized by different colors, or simply by a smaller difference in luminance. Another possibility is to put the important items in the center of an area.

References

- Bijl, P., Kooi, F. L., & van Dorresteijn, M. J. (1997). *Visual search performance for realistic target imagery from the DISSTAF field trials (TNO-TM-97-A055)*. Soesterberg, The Netherlands: TNO Human Factors Research Institute.
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*(241), 177-178.
- Bullimore, M. A., & Bailey, I. L. (1995). Reading and eye movements in age-related maculopathy. *Optometry & Vision Science*, *72*(2), 125-138.
- Collewijn, H., Van der Mark, F., & Jansen, T. C. (1975). Precise recording of human eye movements. *Vision Research*, *14*, 447-450.
- Engel, F. L. (1971). Visual conspicuity, directed attention and retinal locus. *Vision Research*, *11*(6), 563-576.
- Geisler, W. S., & Chou, K. L. (1995). Separation of low-level and high-level factors in complex tasks: visual search. *Psychological Review*, *102*(2), 356-378.
- Gilchrist, I. D., & Harvey, M. (2000). Refixation frequency and memory mechanisms in visual search. *Current Biology*, *10*(19), 1209-1212.
- Gould, J. D., & Carn, R. (1973). Visual search, complex backgrounds, mental counters and eye movements. *Perception & Psychophysics*, *14*, 125-132.
- Hooge, I. T. C., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Perception & Psychophysics*, *58*(7), 969-976.
- Hooge, I. T. C., & Erkelens, C. J. (1998). Adjustment of fixation duration in visual search. *Vision Research*, *38*(9), 1295-1302.
- Hooge, I. T. C., Over, E. A. B., van Wezel, R. J. A., & Frens, M. A. (2005). Inhibition of return is not a foraging facilitator in saccadic search and free viewing. *Vision Research*, *45*(14), 1901-1908.
- Hooge, I. T. C., Stapelkamp, M. J., Over, E. A. B., Vlaskamp, B. N. S., & Frens, M. A. (2004). On the role of

- attentional inhibition and memory during visual search. *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, 2, 2053-2058.
- Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature*, 394(6693), 575-577.
- Jacobs, A. M. (1986). Eye-movement control in visual search: how direct is visual span control? *Perception & Psychophysics*, 39(1), 47-58.
- Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*, 8(2), 255-279.
- Legge, G. E., Ahn, S. J., Klitz, T. S., & Luebker, A. (1997). Psychophysics of reading--XVI. The visual span in normal and low vision. *Vision Research*, 37(14), 1999-2010.
- Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, 25(7), 963-977.
- Megaw, E. D., & Richardson, J. (1979). Target uncertainty and visual scanning strategies. *Human Factors*, 21(3), 303-315.
- Moffitt, K. (1980). Evaluation of the fixation duration in visual search. *Perception & Psychophysics*, 27(4), 370-372.
- Mohammed, Z., & Dickinson, C. M. (2000). The inter-relationship between magnification, field of view and contrast reserve: the effect on reading performance. *Ophthalmic and Physiological Optics*, 20(6), 464-472.
- Näsänen, R., Ojanpää, H., & Kojo, I. (2001). Effect of stimulus contrast on performance and eye movements in visual search. *Vision Research*, 41(14), 1817-1824.
- Ojanpää, H., Näsänen, R., & Kojo, I. (2002). Eye movements in the visual search of word lists. *Vision Research*, 42(12), 1499-1512.
- Pomplun, M., Reingold, E. M., & Shen, J. (2001). Peripheral and parafoveal cueing and masking effects on saccadic

- selectivity in a gaze-contingent window paradigm. *Vision Research*, 41(21), 2757-2769.
- Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, 41(2), 211-236.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422.
- Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: evidence from eye movements. *Psychological Science*, 12(1), 48-55.
- Rommel, R. S. (1984). An inexpensive eye movement monitor using the scleral search coil technique. *IEEE Transactions on Biomedical Engineering*, 31, 388-390.
- Robinson, D. A. (1963). A method of measuring eye movements using a scleral search coil in a magnetic field. *IEEE Transactions on Biomedical Engineering*, 10, 137-145.
- Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, 32(7), 1349-1357.
- Townsend, J. T., Taylor, S. G., & Brown, D. R. (1971). Lateral masking for letters with unlimited viewing time. *Perception & Psychophysics*, 10(5), 375-378.
- Van der Steen, J., & Bruno, P. (1995). Unequal amplitude saccades produced by aniseikonic patterns: effects of viewing distance. *Vision Research*, 35(23-24), 3459-3471.
- Vlaskamp, B. N. S., & Hooge, I. T. C. (2005). Crowding degrades saccadic search performance. *Vision Research*, DOI: 10.1016/j.visres.2005.04.006
- Vlaskamp, B. N. S., Over, E. A. B., & Hooge, I. T. C. (2005). Saccadic search performance: the effect of element spacing. *Experimental Brain Research*, DOI: 10.1007/s00221-005-0032-z

Wolford, G., & Chambers, L. (1983). Lateral masking as a function of spacing. *Perception & Psychophysics*, 33(2), 129-138.

Chapter 5. Coarse-to-fine eye movement strategy in visual search

Introduction

During visual search, eye movement parameters depend on a number of stimulus properties and idiosyncratic factors. Jacobs (1986) and Jacobs and O'Regan (1987) showed a relation between visual span (the area that is analyzed during one fixation) and saccade amplitude. Saccade amplitude increases when the target is more salient. How fixation duration is affected has been investigated too. Hooge and Erkelens (1996; 1998) showed that fixation duration increases if target and distracters are made more similar. They showed this in sparse stimuli with elements arranged in an imaginary hexagonal grid. In a recent experiment, Vlaskamp and colleagues (Vlaskamp, Over, & Hooge, 2005) showed that, with smaller target-distracter dissimilarity, not only fixation duration increases but also saccade amplitude decreases. This last finding corresponds to the result of Jacobs (1986). Therefore, a clear relation exists between target conspicuity and eye movement parameters: When the target is less conspicuous, saccade amplitude decreases and fixation duration increases. A conspicuous target can be detected at a larger distance from the point of fixation than an inconspicuous target. At equal eccentricities, a conspicuous target is detected faster than an inconspicuous target.

Eye movement parameters and search performance do not only depend on stimulus properties; they also depend on idiosyncratic factors. One of these factors is, for example, age. Mean fixation duration increases with age (Harris, Hainline, Abramov, Lemerise, & Camenzuli, 1988; Jernajczyk, Sobanska, Czerwosz, & Szatkowska, 2005). Other examples of idiosyncratic factors influencing eye movement parameters are the presence of disorders such as macular degeneration (Bullimore & Bailey, 1995) and acuity (Johnston, 1965).

Most of the factors mentioned above have been studied extensively, especially the influence of stimulus properties. However, a factor in saccadic search that is not as thoroughly investigated is the influence of the subject's knowledge of target conspicuity. The majority of eye movement literature on visual search involves search with known target conspicuity (for reviews see Davis & Palmer, 2004; Rayner, 1998; Viviani, 1990; Wolfe, 1998). Megaw and Richardson (1979) studied the effect of target uncertainty on search times in experiments where subjects were relatively free to choose their search strategies. When the target is one of several possible items, target conspicuity is at least partially unknown. Megaw and Richardson showed that mean search time did not increase with target uncertainty in their study, and neither did mean fixation duration. However, their analysis was not very sophisticated. Their results for mean fixation duration were based on fixations during the first full scan of the display only. Moreover, a fixation could actually consist of several fixations since they treated successive fixations on the same item as one fixation. They reported no other aspects of eye movements than mean fixation duration. Scinto, Pillalamarri, and Karsh (1986) also investigated aspects of target uncertainty during visual search. They used 'textons' (items consisting of identical spatial frequencies) arranged in a grid. The target textons were '10's among 'S's. The three conditions varied in the number of textons that formed the target. The target group consisted of 3, 6 or 9 '10's. The number of target textons in each trial was unknown. Their main questions were whether subjects use cognitive strategies, and if they did whether the experimenter could modify these strategies. They found 1) no evidence for cognitive strategies, 2) no evidence for external influence on search strategy, and 3) an increase in fixation duration and at the same time a decrease in saccade amplitude as search continued to be unsuccessful. This last finding does not "suggest any global or overall systematic search strategy, but rather only very general regulative adjustment of ocular behavior applicable to many visual information-processing tasks". They did seem to consider this finding to be the result

of the fact that target conspicuity is at least partially unknown because the size of the target texton group was variable. We expect that knowledge of target conspicuity may be an important factor for choosing particular search strategies, and thereby also for the setting of eye movement parameters. In this chapter we therefore ask the question: How does eye movement strategy, as reflected in eye movement parameters such as fixation duration and saccade amplitude, depend on the subject's knowledge of target conspicuity?

To study the effect of knowledge about target conspicuity on human eye movement behavior we compared two search conditions: search with varying target conspicuity and search with constant target conspicuity. What is to be expected for the search strategies in these two conditions and how can we expect these strategies to be reflected in the eye movement parameters fixation duration and saccade amplitude? In the following, we assume the existence of an optimal strategy with respect to minimizing search time and maximizing accuracy. We further assume that subjects attempt to use this optimal strategy as good as they can.

1) *Constant target conspicuity.* When target conspicuity is equal in each trial, the optimal eye movement parameters are also equal in each trial. Once the conspicuity of the targets is known, subjects can determine the optimal settings for saccade amplitude and fixation duration and store them in memory. These settings that lead to fastest and most accurate target detection can be used again in all following trials. Target conspicuity is constant over trials in, for example, a blocked-design experiment with identical target and distracters. When these objects are arranged in a regular grid, object locations are identical in each trial. The only unknown parameter then is the target's location in the grid. In the first trials, appropriate settings for fixation duration and saccade amplitude can be determined, and these can be used in all following trials.

2) *Varying target conspicuity.* If target conspicuity is unknown in advance, and it is uncertain whether the target can be found easily, it makes sense to first take the chance that the

target is conspicuous. If the target is conspicuous, search can be fast and may take only little effort, because the target can be found with only a few, short fixations and large saccades. If the target appears to be inconspicuous, short fixations and large saccades do not suffice and eye movement parameters should be changed so that it becomes possible to find a less conspicuous target. Finding an inconspicuous target can be accomplished by deploying smaller saccade amplitudes and fixations that last longer. Thus, more effort has to be put into the search. If we further assume that the transition from relatively effortless search to effortful search is not a step between two discrete states of search, but rather a gradual change, then the time course of search can be described as the result of a coarse-to-fine process (see below). If the visual system uses a coarse-to-fine process, then at first analysis of visual information is fast at a coarse spatial scale, and later analysis is slower at finer spatial scales. At the beginning of each search trial, fixation duration is short and saccade amplitude is large. When the target is not found, fixation duration will increase and saccade amplitude will decrease, adapting to the apparent difficulty of the stimulus at hand (see also Scinto et al., 1986). We will refer to this hypothesis as the coarse-to-fine hypothesis. “Coarse” here refers to the eye movement parameter settings that are optimal for conspicuous targets and “fine” refers to eye movement parameter settings that are optimal for inconspicuous targets.

Coarse-to-fine algorithms are widely used, for example in computer vision (Das & Ahuja, 1996), 3d shape registration (Tarel & Boujemaa, 1999), object identification (Gu & Tjahjadi, 2000), and face detection (Fleuret & Geman, 2001; Schneiderman & Kanade, 2000). Coarse-to-fine algorithms are efficient, especially when iterations are involved (Atiquzzaman, 1999). A computer model designed to generate eye movements during visual search based on the coarse-to-fine algorithm would use short fixation duration and large saccade amplitude at the start of each trial. If the target is not found, the model should infer that the target is not so salient that it can be found using only minimal resources, namely by

large saccades and short fixations. It should therefore increase fixation duration and decrease saccade amplitude. Longer fixation durations and smaller saccade amplitudes yield possibilities for more detailed visual analysis. Smaller saccades lead to a smaller area being analyzed in the same computation time. Longer fixations lead to more computation time spent in the same area. In this way, the model's visual search proceeds at first at a coarse scale and later at increasingly finer scales.

Also in the visual system a coarse-to-fine approach is used. In stereopsis for example, it is generally accepted that the visual system works in a coarse-to-fine manner (Marr & Poggio, 1979; Poggio & Poggio, 1984), so that the probability of false matches is minimized (Mallot, Gillner, & Arndt, 1996). The oculomotor system is another example of a system that uses the coarse-to-fine principle. Saccades often fall short, especially when they are large (Kapoula, 1985). Kapoula and Robinson (1986) suggest this may be a "lazy" strategy: if fixation near the intended location yields the information sought, no correction saccade has to be made, thereby saving energy. However, if the sought information cannot be found at this location, an additional saccade has to be made. The initial saccade, which is on a coarse scale, is then followed by a corrective saccade on a finer scale. As mentioned before, Scinto et al. (1986) found coarse-to-fine time courses of fixation duration and saccade amplitude during a visual search task.

To answer the question how subjects adapt their saccade amplitude and fixation duration to the demands of the search task we compared fixation duration and saccade amplitude from two search experiments. The first experiment was a search task for military vehicles in natural, rural scenes. We assumed it met the conditions required for varying target conspicuity. The target was a possibly camouflaged military vehicle of unknown type. It was photographed from an unknown distance and thus its size was also unknown. The subjects therefore did not know the exact projection of the target on the retina. The second experiment was published

recently (Vlaskamp et al., 2005) and was assumed to meet the conditions required for constant target conspicuity. We performed new analyses on this data set.

Search experiment with varying target conspicuity

In this experiment we investigate temporal changes in fixation duration and saccade amplitude when subjects do not know target conspicuity in advance. We use a database of complex natural images, each containing one military vehicle.

Methods

Apparatus. Eye movements were recorded using an SMI Eyelink I system. A camera, attached to a headband, was placed in front of the left eye of the subject. Although viewing was binocular, only movements of the left eye were recorded. Stimuli were presented with Matlab for Mac OS 9 on a LaCie 22'' monitor (1600 x 1200 pixels). Eye movement recording was controlled by means of the Eyelink Toolbox for Matlab (Cornelissen, Peters, & Palmer, 2002). The eye movement data were analyzed off-line.

Stimulus. TNO Human Factors in Soesterberg (the Netherlands) provided the original 44 scenes (6144 x 4096 pixels, 16.7 million colors) used in this experiment. For more detailed information about the scenes see Toet, Bijl, Kooi, and Valeton (1998). We selected 136 cutouts of the scenes. The cutouts consisted of 1600 x 1067 pixels. Each cutout contained one of nine possible military vehicles.

Procedure. Subjects sat at a distance of 50 cm from the monitor. A chinrest restricted head movements. The room was darkened. The 136 cutouts were presented in two sessions. The second session followed within fifteen minutes after completing the first session. Each session started with the calibration of the eye tracking system. Each trial started with a drift correction of the eye tracker based on a single fixation in

order to maintain accurate eye movement recording throughout the session. The stimulus appeared immediately after drift correction. Search always started in the center of the search display since the fixation point for drift correction was presented in the center of the screen. Subjects searched the display until they found the target. The trial was ended if they did not find the target within 30 seconds. If subjects did find the target, they were instructed to maintain their eyes fixated on the target and to press the space bar to terminate the trial.

Eye movement analysis. Saccades were detected with a velocity threshold of 30 deg/s. Following saccade detection our Matlab program searched back and forth until the velocity was two standard deviations higher than the velocity during fixation, as in Van der Steen and Bruno (1995). Saccades with amplitudes smaller than 0.5 degrees were discarded. If a small saccade was removed, the fixation before and the one after this saccade were added together. Fixations shorter than 20 ms were also discarded. The last saccade before and the first saccade after such a removed fixation were combined to form one new saccade. The last fixation, indicating the target position, was excluded from analysis.

Subjects. Twenty-one civilian subjects participated in the coarse-to-fine experiment. One of the authors (IH) served as a subject. The other subjects were naïve with respect to the goals of this experiment. All subjects had normal vision. The subjects gave their informed consent. The experiment was conducted in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki.

Results

Search times. In total, we recorded eye movements in 2856 trials (21 observers x 136 stimuli). None of the observers fixated outside the stimulus area. Search times were recorded for each trial. They ranged from 0.46 s to 29.92 s. The response for target location was either correct or false, or there was no response because the trial was timed out after thirty seconds. Percentage correct ranged from 65% to 97% between subjects. On average, percentage correct was 78% +/- 7%. The

other trials resulted in false alarms (18%) or were timed out after thirty seconds (4%). Percentage correct and the wide range of search times indicate that stimuli considerably varied in difficulty. Some targets were easy to find, others were not found within thirty seconds. Eye movements recorded in trials that were timed out, and in false-alarm trials were nevertheless used in the analyses.

We analyzed fixation duration and saccade amplitude in all trials combined. Means and standard errors of the means for both fixation duration and saccade amplitude were computed as a function of their ordinal number in the trial. We analyzed only ordinal numbers of fixations and saccades for which a large amount of data was available. Here, we analyzed fixation duration and saccade amplitude up to ordinal number 88. Our (arbitrary) choice for the amount of 88 was that at least five percent of the total number of trials should contain this amount of fixations and saccades. Figure 5.1 shows the time course of the analyzed fixation duration and saccade amplitude. The numbers of fixation durations and saccade amplitudes decrease with ordinal number: for example, there is a first fixation in every trial, but not all trials have as much as 50 fixations because the target may already have been found with fewer fixations. The decreasing amount of fixation durations and saccade amplitudes with ordinal number is one of the reasons for increasing standard error of the means with ordinal number.

Fixation duration. We expected to find increasing fixation duration as a function of ordinal fixation number, corresponding to the coarse-to-fine search strategy. Mean duration of the first fixation was 215 ms. From the second fixation, with a mean duration of 173 ms, mean duration increased by 43% to 248 ms after about 30 fixations, and by 46% to 252 ms after about 80 fixations. The maximum difference between the second and later fixation durations was about 80 ms. Summarizing, fixation duration generally increased with trial duration (Figure 5.1a). Antes (1974) and Unema, Pannasch, Joos, and Velichkovsky (2005) also found this increase during picture viewing. A difference between our

analysis and the analyses of Antes and Unema et al. is that we set apart the first fixation. The duration of the first fixation is reported here as the time between stimulus onset and the start of the first saccade. However, it is likely that most first fixations lasted longer than this time period, because subjects were already fixating the centre of the screen before stimulus onset. This also means that the input of visual information during the first fixation was not constant. In daily life, visual input is usually relatively constant during a fixation since the visual environment does not suddenly change. We therefore believe that the first fixation stems from a different distribution than the successive fixations, and thus the first fixation should be analyzed separately from the successive fixations. We found a significantly longer duration for the first fixation. This result has been reported before (Hooge & Erkelens, 1996; Van Loon, Hooge, & Van den Berg, 2002). Antes analyzed fixations in tenths of the total number of fixations, and Unema et al. analyzed fixation duration as a function of elapsed time. Thus, in those two experiments the first fixation is taken together with other fixations in the first bin. Therefore, if the first fixation duration would be longer than the next few, this effect would be cancelled out by consecutive fixations in the first bin.

Saccade amplitude. We expected to find decreasing saccade amplitudes as a function of ordinal saccade number, corresponding to the coarse-to-fine search strategy. Mean amplitude of the first saccade was 5.9 deg. Mean amplitude slightly increased from the second saccade (7.7 deg) to the ninth saccade (8.1 deg). From the ninth saccade, mean amplitude decreased by 26% to 6.0 deg after about 30 saccades and by 35% to 5.3 deg after about 80 saccades.

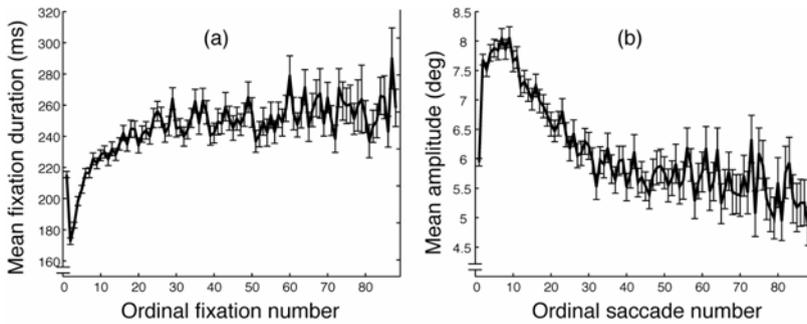


Figure 5.1. Mean fixation duration as a function of ordinal fixation number (a), and mean saccade amplitude as a function of ordinal saccade number (b), for varying target conspicuity. Error bars depict standard errors of the means. Right axes indicate the normalized values: the means per ordinal number divided by the grand average.

Both fixation duration and saccade amplitude showed time courses that correspond to the coarse-to-fine strategy. However, the strategy may not be the only cause for the coarse-to-fine time courses. It is possible that the size of the targets caused some of the coarse-to-fine characteristics found in the time courses of fixation duration and saccade amplitude. Large targets are likely to be found sooner than small targets. However, the correlation between target size in pixels and median search time per subject (the time at which a subject had found 50% of the targets) was -0.10 . This indicates only a minor contribution of target size to the coarse-to-fine characteristics. To see whether these coarse-to-fine time courses were due to the fact that there were different target conspicuities, we split the data set in two equally large parts. One part contained the trials having less than twelve fixations (mostly trials with conspicuous targets); the other part contained the other trials (mostly inconspicuous targets). In both parts, the coarse-to-fine time courses of fixation duration as well as saccade amplitude were present. This fact yielded evidence for only a minor or even absent influence of target conspicuity on the coarse-to-fine time courses. Summarizing, fixation duration and saccade amplitude showed coarse-to-fine time courses. The effects in both time courses were larger than the standard errors of the means.

Search experiment with constant target conspicuity

In this experiment we investigate temporal changes in fixation duration and saccade amplitude when subjects know target conspicuity in advance. We use a eye movement database from a recently published experiment (Vlaskamp et al., 2005) of search for a closed square among squares with a gap in one of the four edges.

Methods

Apparatus. An Apple G3 generated the stimuli, which were presented on a Sony Trinitron 19'' monitor (1024 x 768 pixels). Otherwise, the apparatus was identical to the apparatus in the experiment with varying target conspicuity.

Stimulus. Each search display contained one closed square (the target) among squares with a gap in one of the four edges (the distracters). The elements consisted of lines with a width of 1/12 of their length (0.475 deg). The size of all elements (target and distracters) was 0.57 deg x 0.57 deg. Elements were white and the background was black. All elements were placed on an imaginary hexagonal grid of 37 deg x 31 deg. The target was randomly positioned at one of the grid locations.

The experiment consisted of twelve conditions: four element spacings combined with three gap sizes of the distracters. The four different element spacings were the result of putting different numbers of elements (36, 64, 100, and 144) in the search display while keeping the display size constant. Accordingly, the minimum distance between elements (center-to-center) in a display was 7.1, 5.2, 4.1 or 3.4 degrees. The three gap sizes measured 0.09, 0.19 and 0.28 degrees. All gap sizes were large enough so that the target could be clearly discerned from a distracter when fixated. Distracters in a single condition all had the same gap size.

Procedure. Conditions differed in their number of trials. The numbers of trials were determined at 100, 75, 50 and 25 for the 7.1, 5.2, 4.1 and 3.4 degrees conditions respectively. These numbers were chosen so that a more or less equal number of eye movements would be acquired for each condition. The conditions were presented in blocks. The first 10% of trials in each condition were considered practice trials. Subjects sat at a distance of 41 cm from the monitor. Further, the procedure was identical to the procedure in the experiment with varying target conspicuity.

Eye movement analysis. Eye movement analysis was identical to analysis in the experiment with varying target conspicuity, except for the value of some parameters: 1) saccades were detected with a velocity threshold of 50°/s, 2) saccades with amplitudes smaller than 0.1 deg were removed from the analysis, and 3) fixations shorter than 50 ms were removed from further analysis.

Subjects. Five subjects participated in all conditions. Three of the subjects are also authors (BV, EO and IH, all male). The other two subjects (one male and one female) were naïve with respect to the goals of this experiment. All subjects were between 19 and 35 years old and had normal vision. The subjects gave their informed consent. The experiment was conducted in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki.

Results

In order to investigate common trends in all conditions we normalized fixation duration and saccade amplitude to their means per condition. We did this because it was shown in a previous study (Vlaskamp et al., 2005) that these conditions yielded different mean saccade amplitude and fixation duration per condition. Fixation duration means ranged from 149 ms (medium gap and largest inter-element distance) to 170 ms (smallest gap and inter-element distance). Saccade amplitude means ranged from 5.16 degrees (smallest gap and inter-element distance) to 10.95 degrees (largest gap and inter-element distance). The total amount of trials was

1250 (5 subjects x (100+75+50+25) stimuli). As in the experiment with varying target conspicuity, we analyzed only ordinal numbers of fixations and saccades for which at least five percent of the total number of trials contained data. Five percent of the total 1250 trials consisted of more than 131 saccades. Figure 5.2 shows the time course of fixation duration and saccade amplitude up to fixation and saccade number 131.

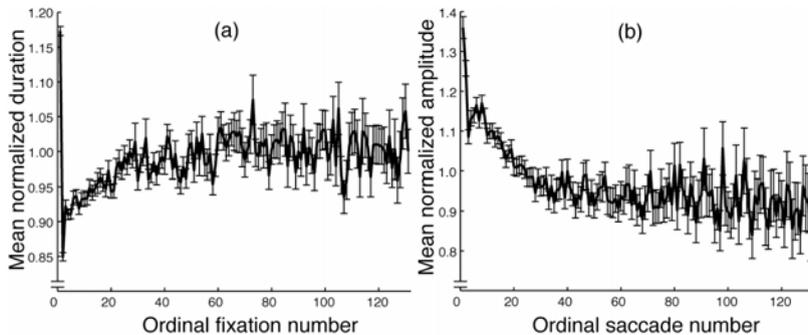


Figure 5.2. Mean normalized fixation duration as a function of ordinal fixation number (a), and mean normalized saccade amplitude as a function of ordinal saccade number (b), for constant target conspicuity. Error bars depict standard errors of the means.

Fixation duration. We expected to find constant fixation duration as a function of ordinal fixation number, corresponding to the constant-parameter hypothesis. As we argued before, the first fixation should be analyzed separately from the successive fixations. Mean normalized duration of the first fixation was 1.17. Thus, as we also found in the experiment with varying target conspicuity, the first fixation has a relatively long duration. From the second fixation, with a mean normalized duration of 0.85, mean normalized duration increased by 16% to 0.99 after about 30 fixations and by 18% to 1.0 after about 80 fixations. Fixation duration was therefore not constant as had been hypothesized but instead generally increased with each next fixation within a trial.

Saccade amplitude. We expected to find constant saccade amplitude as a function of ordinal fixation number, corresponding to the constant-parameter hypothesis. Mean normalized amplitude of the first saccade was 1.36, and of the second saccade 1.26. From the third saccade (1.09) to the eighth saccade (1.17) mean normalized amplitude slightly increased. From the eighth saccade mean normalized amplitude decreased by 17% to 0.97 after about 30 saccades and by 23% to 0.96 after about 80 saccades. A possible explanation for the fact that the first and second saccades are relatively large is that they are repositioning saccades. Subjects might for example start the search with the element in the upper left corner of the screen. Their first saccade would then be about half the size of the diagonal of the screen.

To see whether these coarse-to-fine time courses were possibly due to the fact that we combined twelve conditions with different target conspicuities (a circle target among C's with large gaps is more conspicuous than among C's with small gaps), we also analyzed the data per condition. Linear regression of mean fixation duration showed twelve positive slopes (Table 5.1) and linear regression of mean saccade amplitude showed twelve negative slopes (Table 5.2). We therefore conclude that also when target conspicuity was known in advance, both fixation duration and saccade amplitude showed coarse-to-fine time courses. The coarse-to-fine effects in both time courses were larger than the standard errors of the means. However, the effects were smaller than in the experiment with varying target conspicuity, especially in the case of fixation duration.

Table 5.1. Slope coefficients of linear regression per condition for mean fixation duration as a function of ordinal fixation number. Conditions differed in gap size of the distracters or inter element distance. All coefficients in 10^{-3} ms.

inter-element distance gap size	7.1°	5.2°	4.1°	3.4°
0.09°	2.10	0.98	0.71	0.44
0.19°	5.78	1.80	0.15	0.25
0.28°	5.04	3.98	0.03	0.44

Table 5.2. Slope coefficients of linear regression per condition for mean saccade amplitude as a function of ordinal saccade number. Conditions differed in gap size of the distracters or inter element distance. All coefficients in 10^{-3} degrees.

inter-element distance gap size	7.1°	5.2°	4.1°	3.4°
0.09°	-1.18	-1.16	-0.86	-1.02
0.19°	-6.31	-2.98	-1.23	-2.39
0.28°	-4.83	-6.94	-3.29	-2.66

Discussion

In this study, two experiments were conducted to explore the influence on eye movement parameters of target conspicuity being either constant or varied between trials. The first experiment was a search experiment for (possibly camouflaged) military vehicles in natural backgrounds. The type, size and orientation of the targets and the rural background were unknown to the subjects in advance. In the second experiment the exact appearance of the target was known, as were its surroundings (a fixed grid of distracters). In both experiments we examined the time courses of fixation duration and saccade amplitude. It was expected that the first experiment (with varying target conspicuity) would show coarse-to-fine time courses for both parameters. In the second experiment (with constant target conspicuity) constant fixation duration and saccade amplitude were expected.

In the experiment with varying target conspicuity, we found a coarse-to-fine strategy as predicted by Scinto et al.

(1986). They proposed that the increase of fixation duration and decrease of saccade amplitude with viewing time should be considered as a strategic adaptation to the demands of the task. The second experiment of our study showed that this is clearly not the case. If Scinto's proposal had been correct, we would have found constant fixation duration and constant saccade amplitude in the experiment with constant target conspicuity; there was no need to adapt in this experiment since target conspicuity was exactly known in advance of each trial. What we found in the experiment with constant target conspicuity however, was also a coarse-to-fine time course, although to a lesser extent, especially for fixation duration (compare Figures 5.1 and 5.2). Thus, even in search conditions that were fully known a priori, the coarse-to-fine strategy in eye movement behavior was nevertheless present. The reason that not many other studies reported coarse-to-fine time courses of saccade amplitude and fixation duration is that most of those studies usually involve fairly short trials (<10 saccades, see e.g. Hooge & Erkelens, 1996; Zelinsky, Rao, Hayhoe, & Ballard, 1997), which makes it impossible to identify effects that appear only after many saccades.

Coarse-to-fine time courses of fixation duration and saccade amplitude are not only found in visual search. A number of studies have shown that fixation durations increase and saccade amplitudes decrease during free picture viewing. A classical example has been described by Buswell (1935). Unema et al. (2005) found coarse-to-fine time courses of fixation duration and saccade amplitude during picture viewing too, but offered no explanation for the nature of these time courses. Antes (1974) also investigated the time courses of fixation duration and saccade amplitude during free viewing. Although his analysis was slightly different, his results are comparable to our results: fixation duration increased and saccade amplitude decreased with viewing time. He suggested that the amount of "informativeness" of fixated parts of the stimulus might influence fixation duration: less informative parts are fixated late in the time course, and may require longer visual processing time. Our second experiment

with constant target conspicuity yielded results that contradict this suggestion, since all parts in those stimuli are equally informative.

One possible reason for deploying coarse-to-fine eye movement behavior, even when this type of behavior is suboptimal, is that in daily life the conspicuity of objects that have to be found is usually not exactly known in advance. The reason for this is that the visual environment is much more complex than the stimuli of a laboratory experiment. On average, it will then be profitable to use a coarse-to-fine eye movement strategy since in the majority of cases that strategy will yield the best results. An object with high conspicuity can be found quickly and with minimal effort, and if the object has low conspicuity only little time is lost with searching for a possibly easy to find target. In the experiment with constant target conspicuity, subjects knew in advance of each trial the exact conspicuity of the target. This means that the conditions allowed the use of optimal settings for their fixation duration and saccade amplitude. However, they did not do that. Araujo, Kowler, and Pavel (2001) also found that (five out of six) subjects used suboptimal strategies for making eye movements in visual search. Subjects were cued to the location that contained the target with high probability (80%), but the cue was often ignored. Araujo et al. (2001) proposed four reasons for this suboptimal behavior: 1) decision strategies that failed to recognize the significance of probability cues, 2) built-in preferences to minimize effortful saccadic planning, 3) attraction of attention and saccades to nearby locations, and 4) initiation of saccades while attention remained divided. All four reasons act to facilitate an initial rapid scanning of a series of locations, which is how coarse-to-fine search starts.

Another reason for deploying coarse-to-fine eye movement behavior is that there may be a coarse-to-fine basis within the visual system. There is physiological evidence for the connection between fast-to-slow signals and coarse-to-fine structures in the visual system. Weng, Yeh, Stoelzel, and Alonso (2005) showed that large receptive fields have shorter

response latencies than small receptive fields. The first signals that arrive in the visual cortex are therefore “coarse” signals.

Our evidence for omnipresent coarse-to-fine time courses of fixation duration and saccade amplitude may have consequences for the interpretation of adaptation and learning studies. The effects measured in these studies should probably be corrected for the “adaptation” or “learning” that we found to be inherently present. The term “learning” implies an active improvement of behavior. However, part of such an improvement could also be due to a built-in coarse-to-fine mechanism. The coarse-to-fine eye movement behavior further shows that it may be incorrect to report only means of fixation duration and saccade amplitude in eye movement studies. We showed that means also depend on the number of fixations in a trial, because mean fixation duration increased and mean saccade amplitude decreased with viewing time. Concluding, considering 1) the fact that we did not find constant fixation duration or constant saccade amplitude when target conspicuity was known in advance of each trial, 2) the presence of coarse-to-fine eye movement strategy in plain viewing, and 3) some physiological evidence, we suggest that the visual system makes use of an intrinsic coarse-to-fine mechanism.

References

- Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, 103(1), 62-70.
- Araujo, C., Kowler, E., & Pavel, M. (2001). Eye movements during visual search: the costs of choosing the optimal path. *Vision Research*, 41(25-26), 3613-3625.
- Atiquzzaman, M. (1999). Coarse-to-Fine Search Technique to Detect Circles in Images. *International Journal of Advanced Manufacturing Technology*, 15, 96-102.
- Bullimore, M. A., & Bailey, I. L. (1995). Reading and eye movements in age-related maculopathy. *Optometry & Vision Science*, 72(2), 125-138.
- Buswell, G. (1935). *How people look at pictures*. Chicago: University of Chicago Press.
- Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The EyeLink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior, Research Methods, Instruments, & Computers*, 34(34), 613-617.
- Das, S., & Ahuja, N. (1996). Active surface estimation: integrating coarse-to-fine image acquisition and estimation from multiple cues. *Artificial Intelligence*, 83(2), 241-266.
- Davis, E. T., & Palmer, J. (2004). Visual search and attention: an overview. *Spatial Vision*, 17(4-5), 249-255.
- Fleuret, F., & Geman, D. (2001). Coarse-to-Fine Face Detection. *International Journal of Computer Vision*, 41(1-2), 85-107.
- Gu, Y.-H., & Tjahjadi, T. (2000). Coarse-to-fine planar object identification using invariant curve features and B-spline modeling. *Pattern Recognition*, 33(9), 1411-1422.
- Harris, C. M., Hainline, L., Abramov, I., Lemerise, E., & Camenzuli, C. (1988). The distribution of fixation durations in infants and naive adults. *Vision Research*, 28(3), 419-432.
- Hooge, I. T. C., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Perception & Psychophysics*, 58(7), 969-976.
- Hooge, I. T. C., & Erkelens, C. J. (1998). Adjustment of fixation duration in visual search. *Vision Research*, 38(9), 1295-1302.

- Jacobs, A. M. (1986). Eye-movement control in visual search: how direct is visual span control? *Perception & Psychophysics*, 39(1), 47-58.
- Jacobs, A. M., & O'Regan, J. K. (1987). Spatial and/or temporal adjustments of scanning behavior to visibility changes. *Acta Psychologica*, 65(2), 133-146.
- Jernajczyk, W., Sobanska, A., Czerwosz, L., & Szatkowska, E. (2005). The influence of age and gender on the latency of eye movement in healthy humans. *Journal of Physiology and Pharmacology*, 56 Suppl 4, 93-97.
- Johnston, D. M. (1965). Search performance as a function of peripheral acuity. *Human Factors*, 7(6), 527-535.
- Kapoula, Z. (1985). Evidence for a range effect in the saccadic system. *Vision Research*, 25(8), 1155-1157.
- Kapoula, Z., & Robinson, D. A. (1986). Saccadic undershoot is not inevitable: saccades can be accurate. *Vision Research*, 26(5), 735-743.
- Mallot, H. A., Gillner, S., & Arndt, P. A. (1996). Is correspondence search in human stereo vision a coarse-to-fine process? *Biological Cybernetics*, 74(2), 95-106.
- Marr, D., & Poggio, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society B: Biological Sciences*, 204(1156), 301-328.
- Megaw, E. D., & Richardson, J. (1979). Target uncertainty and visual scanning strategies. *Human Factors*, 21(3), 303-315.
- Poggio, G. F., & Poggio, T. (1984). The analysis of stereopsis. *Annual Review of Neuroscience*, 7, 379-412.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422.
- Schneiderman, H., & Kanade, T. (2000). A Statistical Method for 3D Object Detection Applied to Faces and Cars. *2000 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1746-1751.
- Scinto, L. F., Pillalamarri, R., & Karsh, R. (1986). Cognitive strategies for visual search. *Acta Psychologica*, 62(3), 263-292.
- Tarel, J.-P., & Boujemaa, N. (1999). A Coarse to Fine 3D Registration Method Based on Robust Fuzzy Clustering. *Computer Vision and Image Understanding*, 73(1), 14-28.
- Toet, A., Bijl, P., Kooi, F. L., & Valetton, J. M. (1998). A high-resolution image dataset for testing search and detection

- models (TNO-TM-98-A020)*. Soesterberg, The Netherlands: TNO Human Factors Research Institute.
- Unema, P. J. A., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, 12(3), 473-494.
- Van der Steen, J., & Bruno, P. (1995). Unequal amplitude saccades produced by aniseikonic patterns: effects of viewing distance. *Vision Research*, 35(23-24), 3459-3471.
- Van Loon, E. M., Hooge, I. T. C., & Van den Berg, A. V. (2002). The timing of sequences of saccades in visual search. *Proceedings of the Royal Society B: Biological Sciences*, 269(1500), 1571-1579.
- Viviani, P. (1990). Eye movements in visual search: Cognitive, perceptual and motor control aspects. In E. Kowler (Ed.), *Reviews of Oculomotor Research: Eye Movements and Their Role in Visual and Cognitive Processes*. (Vol. 4, pp. 353-393). Amsterdam: Elsevier.
- Vlaskamp, B. N. S., Over, E. A. B., & Hooge, I. T. C. (2005). Saccadic search performance: the effect of element spacing. *Experimental Brain Research*, DOI: 10.1007/s00221-00005-00032-z.
- Weng, C., Yeh, C. I., Stoelzel, C. R., & Alonso, J. M. (2005). Receptive field size and response latency are correlated within the cat visual thalamus. *Journal of Neurophysiology*, 93(6), 3537-3547.
- Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention*. London UK: University College London Press.
- Zelinsky, G. J., Rao, R. P. N., Hayhoe, M. M., & Ballard, D. H. (1997). Eye movements reveal the spatio-temporal dynamics of visual search. *Psychological Science*, 8, 448-453.

Chapter 6. Summary

This thesis deals with the quantitative analysis of saccadic search strategy. The goal of the research presented was twofold:

- 1) To quantify overall characteristics of fixation location and saccade direction. These two eye movement parameters are usually not reported quantitatively, probably due to the lack of suitable methods to calculate them. Chapters 2 and 3 provide two new methods, and show that these methods are useful in the quantitative analysis of saccadic search strategy.
- 2) To identify search strategies, with the use of a quantitative description of eye movement parameters (amplitude and direction of saccades, and location and duration of fixations). In chapters 4 and 5 two different strategies are described.

Below are the summaries of the chapters in this thesis.

Quantitative description of saccade direction and fixation location

The first aim of this thesis was to provide methods for the quantification of saccade direction and fixation location. Chapter 2 provides a method to quantify a general property of fixation locations. We proposed a quantitative measure based on Voronoi diagrams for the characterization of the uniformity of fixation density. The method does not depend on arbitrary choices and can therefore be applied to many kinds of databases of fixation locations. Due to the robustness of the method, it is also possible to compare different databases of eye movement recordings: the method is independent of the number of fixations in a trial, and the size of the stimuli. We investigated how normalized cell-size distributions were related to the homogeneity of fixation densities. Two possible measures of the result of the Voronoi method were discussed.

The first was the fitted parameter of the gamma distribution; the second was the skewness of the normalized cell sizes. Both showed good correlation with subjective visual evaluation of the uniformity of fixation densities. Not only are these two measures objective and quantitative, they also have a simple intuitive meaning: they may be thought of as indicating the clustering of fixations. We showed that during a visual search task, a structured (natural) background leads to higher clustering of fixations compared to a homogeneous background. In addition, in natural stimuli, a search task leads to higher clustering of fixations than the instruction to freely view the stimuli.

Chapter 3 provides a method to identify the overall field of saccade directions in the viewing area. We extended the Voronoi method of chapter 2 so that it became possible to create vector maps. These maps indicate the preferred saccade direction for each position in the viewing area. Calculating the divergence of the vector field enabled quantitative analysis of attractors and repellers in the maps. The sum of the vector lengths in the maps was a quantitative measure for the field strength. We used these measures to quantify the influence of observer-dependent and stimulus-dependent factors on saccade direction in a search task with natural scenes. We identified attractors and repellers in the direction-coded maps for stimuli and observers. The results showed that the influence of stimulus-dependent factors appeared to be larger than the influence of observer-dependent factors. A random saccade generator with two random generators for the direction and amplitude of a saccade operating in a confined viewing area mainly explained the remaining effects in the experimentally obtained saccade-direction maps. Since observer-dependent factors played a lesser role than stimulus-dependent factors with respect to the preferred saccade directions, it seems that the training of observers to use specific search strategies may have limited added value.

Quantitative description of eye movement strategies

The second aim of this thesis was the investigation of strategies for eye movements during visual search. In chapter 4 we showed that the border of the search area played a role in the search strategy. In a search experiment in differently shaped areas we measured that search performance was poorer near the search area luminance edges. Fixation density, however, was higher in the edge region, and saccade direction was mainly along the edges of the search areas. Simple models for independent and dependent fixations did not account for these results about saccade direction and fixation density. In a target visibility experiment we established that the visibility of targets near a luminance edge is less than the visibility of targets against a homogeneous background. We interpreted this as the visual span being smaller near luminance edges than in homogeneous backgrounds. This feature was added to the simple dependent fixation model. The model was used to predict saccade direction distributions for search in areas of different size. Our model predicted that fewer saccades are made along the search area edges in large search areas than in small search areas. The data collected for search in areas of different sizes confirmed the prediction by the model. We concluded that search area edges affected eye movement strategy both at sensory level and cognitive level. The sensory-level influence was due to the fact that fixations usually fall within the search area. Search area edges hampered target perception, so that fixation density had to be higher in the edge regions in order to achieve similar search performance as in the center regions. A cognitive-level influence of the edges appeared from the fact that saccades were made more in the direction along the edges than in other directions. Another factor in cognitive-level strategy appeared from subjects preferring clockwise or counterclockwise scanning. The magnitude of the influences depended on the shape and size of the search area.

In chapter 5 we investigated the effect of a priori knowledge about target conspicuity on saccade amplitude and

fixation duration. The visual system may use fixation duration and saccadic amplitude as optimizers for visual search performance. The optimal value for fixation duration lies between zero, where visual information is not collected or processed at all, and infinity, where just one fixation occurs, and therefore not all the available visual information elsewhere can be collected and processed. The optimal value for saccade amplitude also lies between zero, where subsequent fixations would be at the exact same position, and infinity, which is physically impossible for the human eye. We expected to find (nearly) optimal settings for saccade amplitude and fixation duration in a search task with known target conspicuity, and sub-optimal settings in a search task with unknown target conspicuity. First, we analyzed saccades and fixations in a search experiment for (possibly camouflaged) military vehicles of unknown type and size. Target conspicuity was unknown in each trial. We found that saccade amplitude decreased and fixation duration increased gradually as a function of the serial order of saccades and fixations. These changes are the mark of a coarse-to-fine search strategy, in which a stimulus is at first scanned quickly, and progressively better scrutinized to find the target. In addition we analyzed saccades and fixations recorded during a search experiment in which the only unknown was the location of the target (target conspicuity was known a priori). In this experiment too, we unexpectedly found decreasing saccade amplitude and increasing fixation duration as a function of the serial order, although to a lesser extent. The fact that target conspicuity was known only slightly attenuated the coarse-to-fine time courses of fixation duration and saccade amplitude. Therefore, whether target conspicuity was known or unknown appeared to have minor influence on the time courses of fixation duration and saccade amplitude. We hypothesized an intrinsic coarse-to-fine strategy for visual search that is even used when such a strategy is suboptimal.

Concluding, chapters 2 and 3 provide two new quantitative measures to describe the spatial distribution of

fixation locations and the overall field of saccade directions, two parameters that could not be described quantitatively. With these two new measures, the whole set of four most important eye movement parameters (saccade direction and amplitude, and fixation location and duration) can now be described quantitatively. In chapters 4 and 5 two strategies are described which are used in search tasks. It was shown that stimulus factors seem to be the leading factors in settings of eye movement parameters during visual search (chapter 3), although also cognitive factors might influence eye movement behavior (chapter 4), and that saccade amplitude and fixation duration show coarse-to-fine time courses (chapter 5). Considering the fact that the visual system is an extensively trained system, it must be assumed that, on average, it yields optimal search performance. It then follows that target conspicuity in daily life is usually unknown, which is not illogical. Subjects use a more or less default strategy, that takes into account that some targets are more difficult to find than others, and that can be adjusted to stimulus characteristics. How it is adjusted depends on how eye movement parameters should be set to achieve optimal information gathering, such that of all visited locations just enough information is gathered to decide whether a target is located there. Because the visual span (the area around the fixation location where a target can be detected) is not a discrete entity, but rather an area gradually increasing with fixation duration, and defined by a threshold that changes with target conspicuity, there always is the issue of overlap between different visual spans. Overlap can never be avoided when the entire search area needs to be searched, since the visual span is a circle, or circle-like. On the other hand, overlap should be avoided, since areas that are already scanned don't need to be scanned more than once. The default strategy, however, seems to be quite autonomous, and omnipresent.

Hoofdstuk 7. Samenvatting

Dit proefschrift behandelt de kwantitatieve analyse van saccadische zoekstrategieën. Het doel van dit onderzoek was tweeledig:

1) Het kwantificeren van algemene kenmerken van fixatiepositie en saccaderichting. Deze twee oogbewegingsparameters worden gewoonlijk niet kwantitatief vermeld, waarschijnlijk vanwege het gebrek aan geschikte methodes om ze te berekenen. Hoofdstukken 2 en 3 beschrijven twee nieuwe methodes, en tonen aan dat deze methodes bruikbaar zijn in de kwantitatieve analyse van saccadische zoekstrategieën.

2) Het identificeren van zoekstrategieën, door middel van een kwantitatieve beschrijving van de oogbewegingsparameters (amplitude en richting van saccades, en positie en duur van fixaties). In hoofdstukken 4 en 5 worden twee verschillende strategieën beschreven. Hieronder staan de samenvattingen van de hoofdstukken in dit proefschrift.

Kwantitatieve beschrijving van saccaderichting en fixatiepositie

Het eerste doel van dit proefschrift was om methodes te ontwikkelen voor de kwantitatieve maten voor saccaderichting en fixatiepositie. Hoofdstuk 2 beschrijft een methode om een algemeen kenmerk van fixatieposities te kwantificeren. Wij introduceerden een kwantitatieve maat, gebaseerd op Voronoi diagrammen, voor de uniformiteit van de fixatiedichtheid. De methode hangt niet van willekeurige keuzes af en kan daarom op vele soorten gegevensbestanden van fixatieposities worden toegepast. Omdat de methode zo robuust is, is het ook mogelijk om verschillende gegevensbestanden te vergelijken: de methode is onafhankelijk van het aantal fixaties in een proef, en van de grootte van de stimuli. Wij onderzochten hoe het verband was tussen verdelingen van genormaliseerde celgroottes in

Voronoi diagrammen en de uniformiteit van de fixatiedichtheid. Twee mogelijke maten voor het resultaat van de Voronoi methode werden besproken. De eerste was een fitparameter van de gamma verdeling; de tweede was de scheefheid van de genormaliseerde celgrootte. Allebei toonden goede correlatie met subjectieve visuele evaluatie van de uniformiteit van de fixatiedichtheid. Niet alleen zijn deze twee maten objectief en kwantitatief, ze hebben ook een eenvoudige intuïtieve betekenis: ze zijn een maat voor de groepering van fixaties. In een visuele zoektaak toonden wij verder aan dat een gestructureerde (natuurlijke) achtergrond leidt tot een sterkere groepering van fixaties dan een homogene achtergrond. Bovendien leidt, in natuurlijke stimuli, een zoektaak tot een sterkere groepering van fixaties dan de instructie om de stimuli te bekijken zoals je zelf wilt.

Hoofdstuk 3 beschrijft een methode om het algemene veld van saccaderichtingen in het te bekijken gebied te identificeren. Wij breidden de Voronoi methode van hoofdstuk 2 uit zodat het mogelijk werd om vectorvelden te creëren. Deze vectorvelden geven een beeld van de meest voorkomende saccaderichting voor elke positie in het te bekijken gebied. Het berekenen van de divergentie van de vectorvelden maakte kwantitatieve analyse van attractors en repellers mogelijk. De som van de vectorlengtes in de velden was een kwantitatieve maat voor de totale veldsterkte. Wij gebruikten deze maten om de invloed van waarnemer-afhankelijke en stimulus-afhankelijke factoren op de saccaderichting in een zoektaak met natuurlijke scènes te kwantificeren. Wij identificeerden attractors en repellers in de richting-gecodeerde velden voor zowel stimuli als waarnemers. De resultaten toonden aan dat de invloed van stimulus-afhankelijke factoren groter scheen te zijn dan de invloed van waarnemer-afhankelijke factoren. Een model voor een saccadegenerator met twee random generators voor de richting en de amplitude van een saccade resulteerde in een aannemelijke benadering van de experimenteel verkregen saccade-richtingsvelden. Aangezien de waarnemer-afhankelijke factoren een kleinere rol speelden dan stimulus-

afhankelijke factoren, lijkt het erop dat de training van waarnemers om specifieke zoekstrategieën te gebruiken slechts een beperkte toegevoegde waarde kan hebben.

Kwantitatieve beschrijving van oogbewegingsstrategieën

Het tweede doel van dit proefschrift was het onderzoeken van oogbewegingsstrategieën tijdens visueel zoeken. In hoofdstuk 4 staat dat de rand van het zoekgebied een rol speelde in de zoekstrategie. In een zoekexperiment in gebieden van verschillende vorm bleek dat de zoekprestaties slechter waren dichtbij de randen van het zoekgebied, randen die bestonden uit een verschil in helderheid tussen het zoekgebied en zijn omgeving. Toch was de fixatiedichtheid hoger in het randgebied, en waren saccades hoofdzakelijk gericht langs de randen van de zoekgebieden. Simpele modellen voor onafhankelijke en afhankelijke fixaties konden deze resultaten met betrekking tot saccaderichting en fixatiedichtheid niet verklaren. In een controle-experiment stelden wij vast dat de opvallendheid van een doel dichtbij een rand minder is dan de opvallendheid van een doel tegen een homogene achtergrond. Wij interpreteerden dit als een kleinere visuele spanwijdte dichtbij de randen dan in homogene achtergronden. Deze eigenschap werd toegevoegd aan het model voor afhankelijke fixaties. Het aangepaste model werd vervolgens gebruikt om de verdeling van saccaderichtingen te voorspellen voor zoeken in gebieden van verschillende grootte. Ons model voorspelde dat minder saccades langs de randen van het zoekgebied worden gemaakt in grote zoekgebieden dan in kleine zoekgebieden. De resultaten van het zoekexperiment in driehoeken van verschillende grootte bevestigden de voorspelling van het model. Wij concludeerden dat de randen van het zoekgebied de oogbewegingsstrategie zowel op sensorisch niveau als op cognitief niveau beïnvloedden. De invloed op sensorisch niveau was toe te schrijven aan het feit dat fixaties gewoonlijk binnen het zoekgebied vallen. De randen van het zoekgebied

belemmerden doelwaarneming, zodat de fixatiedichtheid in de randgebieden hoger moest zijn om even goede zoekprestaties als in de centrumgebieden te bereiken. Een invloed op cognitief niveau van de randen bleek uit het feit dat saccades meer in de richting langs de randen dan in andere richtingen werden gemaakt. Een andere invloed op cognitief niveau was dat sommige proefpersonen met de wijzers van de klok mee oogbewegingen maken, en anderen tegen de wijzers van de klok in. De grootte van al deze invloeden hing van de vorm en de grootte van het zoekgebied af.

In hoofdstuk 5 onderzochten wij het effect van a priori kennis over opvallendheid van het doel op saccadeamplitude en fixatieduur. Het visuele systeem zou fixatieduur en saccadeamplitude kunnen gebruiken om visuele zoekprestaties te optimaliseren. De optimale waarde voor fixatieduur ligt tussen nul, waarbij helemaal geen visuele informatie wordt verzameld of verwerkt, en oneindig, waarbij slechts één fixatie voorkomt, en daarom niet alle op andere lokaties beschikbare visuele informatie kan worden verzameld en verwerkt. De optimale waarde voor saccadeamplitude ligt ook tussen nul, waar volgende fixaties op precies dezelfde positie zouden zijn, en oneindig, wat voor het menselijk oog fysiek onmogelijk is. Wij verwachtten (bijna) optimale instellingen voor saccadeamplitude en fixatieduur in een zoektaak met bekende opvallendheid van het doel, en suboptimale instellingen in een zoektaak met onbekende opvallendheid van het doel te vinden. Eerst analyseerden wij saccades en fixaties in een zoekexperiment naar (soms gecamoufleerde) militaire voertuigen van onbekend type en grootte. Voor elke stimulus verscheen was de opvallendheid van het doel onbekend. Wij vonden dat saccadeamplitude geleidelijk kleiner werd en fixatieduur geleidelijk steeg als functie van saccade- en fixatienummer. Deze veranderingen zijn het teken van een “grof-naar-fijne” zoekstrategie, waarin een stimulus aanvankelijk vluchtig wordt afgetast, en later meer in detail doorzocht om het doel te vinden. Vervolgens analyseerden wij saccades en fixaties tijdens een zoekexperiment waarin de enige onbekende de positie van het doel was (de

opvallendheid van het doel was a priori wel bekend). Hoewel wij optimale instellingen verwachtten, vonden wij in dit experiment ook kleiner wordende saccadeamplitude en stijgende fixatieduur als functie van saccade- en fixatienummer, zij het in mindere mate. Het feit dat de opvallendheid van het doel bekend was dempte de grof-naar-fijne tijdsverlopen van fixatieduur en saccadeamplitude slechts weinig. Dus scheen het feit of de opvallendheid van het doel bekend of onbekend was een niet zo grote invloed op de tijdsverlopen van fixatieduur en saccadeamplitude te hebben. Wij stelden als hypothese dat een intrinsieke grof-naar-fijne strategie in visueel zoeken wordt gebruikt, zelfs wanneer een dergelijke strategie suboptimaal is.

Concluderend, hoofdstukken 2 en 3 beschrijven twee nieuwe kwantitatieve maten voor de ruimtelijke verdeling van fixatieposities en het algemene veld van saccaderichtingen, twee parameters die tot op heden niet kwantitatief konden worden beschreven. Met deze twee nieuwe maten kan de gehele reeks van de vier belangrijkste oogbewegingsparameters (richting en amplitude van saccades, en positie en duur van fixaties) nu kwantitatief worden beschreven. In hoofdstukken 4 en 5 worden twee strategieën beschreven die in zoektaken worden gebruikt. Het bleek dat stimulusfactoren de belangrijkste factoren zijn voor het instellen van de oogbewegingsparameters tijdens visueel zoeken (hoofdstuk 3), hoewel ook cognitieve factoren het oogbewegingsgedrag zouden kunnen beïnvloeden (hoofdstuk 4), en dat saccadeamplitude en fixatieduur grof-naar-fijne tijdsverlopen vertonen (hoofdstuk 5). Gezien het feit dat het visuele systeem een zeer goed getraind systeem is moet worden verondersteld dat het, gemiddeld gezien, optimale zoekprestaties levert. Dit zou betekenen dat de opvallendheid van een doel in het dagelijkse leven gewoonlijk onbekend is, wat logisch is. Proefpersonen gebruiken een min of meer standaard strategie, die rekening houdt met het feit dat sommige doelen moeilijker zijn te vinden dan andere doelen, en die aan stimuluseigenschappen van dat moment kan

worden aangepast. Hoe die strategie wordt aangepast hangt af van hoe de oogbewegingsparameters zouden moeten worden ingesteld om een optimaal resultaat te bereiken, zodanig dat van alle bezochte posities net genoeg informatie wordt verzameld om te besluiten of zich daar een doel bevindt. Omdat de visuele spanwijdte (het gebied rond de fixatiepositie waarin een doel kan worden ontdekt) geen constante entiteit is, maar meer een gebied dat geleidelijk groeit met fixatieduur, en is bepaald door een drempel die met de opvallendheid van het doel verandert, is er altijd de kwestie van overlapping tussen verschillende visuele spanwijdten. De overlapping kan nooit worden vermeden wanneer het volledige zoekgebied moet worden doorzocht, aangezien de visuele spanwijdte een cirkel, of cirkelvormig is. Toch zou die overlapping wel zoveel mogelijk moeten worden vermeden, aangezien gebieden die reeds zijn bekeken niet meer dan eens hoeven te worden bekeken. De standaardstrategie schijnt echter vrij autonoom en alom aanwezig te zijn.

Publications

- Hooge, I. T. C., Over, E. A. B., van Wezel, R. J. A., & Frens, M. A. (2005). Inhibition of return is not a foraging facilitator in saccadic search and free viewing. *Vision Research*, *45*(14), 1901-1908.
- Hooge, I. T. C., Stapelkamp, M. J., Over, E. A. B., Vlaskamp, B. N. S., & Frens, M. A. (2004). On the role of attentional inhibition and memory during visual search. *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, *2*, 2053-2058.
- Hooge, I. T. C., Vlaskamp, B. N. S., & Over, E. A. B. (2007). Saccadic Search: on the duration of a fixation. In R. van Gompel, M. Fischer, W. Murray, and R. Hill (Eds), *Eye Movement Research: Insights into Mind and Brain*. Oxford: Elsevier.
- Over, E. A. B., Hooge, I. T. C., & Erkelens, C. J. (2006). A quantitative measure for the uniformity of fixation density: the Voronoi method. *Behavior Research Methods*, *38*(2), 251-261.
- Over, E. A. B., Hooge, I. T. C., & Erkelens, C. J. (submitted). Saccade direction in visual search: stimulus dependent or observer dependent?
- Over, E. A. B., Hooge, I. T. C., & Erkelens, C. J. (submitted). Search area edges affect visual span size, fixation density and saccade direction.
- Over, E. A. B., Hooge, I. T. C., Vlaskamp, B. N. S., & Erkelens, C. J. (in press). Coarse-to-fine eye movement strategy in visual search. *Vision Research*.
- Vlaskamp, B. N. S., Over, E. A. B., & Hooge, I. T. C. (2005). Saccadic search performance: the effect of element spacing. *Experimental Brain Research*, *167*(2), 246-259.

Abstracts and conference proceedings

- Frens, M. A., Vlaskamp, B. N. S., Over, E. A. B., & Hooge, I. T. C. (2005). The distribution of fixation time in saccadic search. *Washington, DC: Society for Neuroscience* (Program No. 859.8).
- Hooge, I. T. C., Vlaskamp, B. N. S., & Over, E. A. B. (2002). Saccadic search: the relation between fixation duration and saccade amplitude. *Perception, 31*(Supplement), 101a.
- Hooge, I. T. C., Vlaskamp, B. N. S., & Over, E. A. B. (2003). Saccadic Search: direct or indirect control of fixation duration? *European Conference on Eye Movements, 12*.
- Hooge, I. T. C., Vlaskamp, B. N. S., Over, E. A. B., & Frens, M. A. (2005). The distribution of fixation time in saccadic search. *European Conference on Eye Movements, 13*.
- Over, E. A. B. (2003). Visual search: saccade parameters depend on the shape of the search area. *Journal of Vision, 3*(9), 123.
- Over, E. A. B., & Erkelens, C. J. (2004). Prediction of eye-movement direction in natural stimuli. *Perception, 33*(Supplement), 134B.
- Over, E. A. B., Hooge, I. T. C., & Erkelens, C. J. (2002). A quantitative measure for the spatial distribution of fixations. *Perception, 32*(Supplement), 182A.
- Vlaskamp, B. N. S., Hooge, I. T. C., & Over, E. A. B. (2002). Stimulus density does not affect search performance. *Perception, 31*(Supplement), 174A.
- Vlaskamp, B. N. S., Hooge, I. T. C., & Over, E. A. B. (2003). Local contrast affects fixation density in saccadic search. *Perception, 32*(Supplement), 109B.

Curriculum Vitae

Eelco Over was born on May 7, 1976, in Utrecht, the Netherlands. After completing his secondary education (gymnasium) at the Christelijk Gymnasium Utrecht, he studied Applied Physics at Delft University of Technology. In 2000 he obtained his masters degree, and in 2001 he started his PhD study at the department Physics of Man, Universiteit Utrecht. In 2005, he was hired to develop an eye movement analysis program for a start-up company. Since 2006, he is a post-doc at the department of Neuroscience, ErasmusMC, Rotterdam.